École doctorale 129 : Sciences de l'Environnement

THÈSE

pour obtenir le grade de docteur délivré par

Sorbonne Université

Spécialité doctorale "Sciences de l'Environnement"

qui sera
présentée et soutenue publiquement par

Anna Lea ALBRIGHT

le 18 mars 2022

The trade-wind boundary layer and climate sensitivity

Directrice de thèse : **Sandrine BONY**
Co-Directeur de thèse : **Bjorn STEVENS**

Jury

Sandrine BONY	Laboratoire de Météorologie Dynamique, CNRS	Directrice
Florent BRIENT	Sorbonne Université	Examinateur
Francis CODRON	Sorbonne Université	Examinateur
Kerry EMANUEL	Massachusetts Institute of Technology	Rapporteur
Cathy HOHENEGGER	Max Planck Institute for Meteorology	Rapporteur
Marie LOTHON	Laboratoire d'Aerologie, CNRS	Examinatrice
Bjorn STEVENS	Max Planck Institute for Meteorology	Co-Directeur

Laboratoire de Météorologie Dynamique
UMR 8539, 4 Place Jussieu, 75005 Paris, France

Abstract

The response of clouds in the trades to warming remains uncertain, raising the specter of a high climate sensitivity. Decreases in cloud fraction are thought to relate to couplings among convective mixing, turbulence, radiation, and the large-scale environment. The EUREC^{4}A (*Elucidating the role of cloud-circulation coupling in climate*) field campaign made extensive measurements that allow for deeper physical understanding and the first process-based constraint on the trade cumulus feedback, as described in this thesis, in two parts.

The first part (Chapters 2–5) uses EUREC^{4}A observations to improve understanding of the characteristic vertical structure of trade-wind air and the processes that determine it. The second part (Chapters 6–7) applies an improved physical understanding of the trade-wind boundary layer to the evaluation of trade cumulus feedbacks. Ideas developed in these chapters support new conceptual models of the structure of the lower trade-wind atmosphere and a more active role of clouds in maintaining this structure, and show little evidence for a strong trade cumulus feedback to warming.

Chapt. 3 calculates clear-sky radiative profiles from 2580 *in situ* soundings launched during EUREC^{4}A, which are then used to observationally close subcloud layer moisture and energy budgets in Chapt. 4. Chapt. 4 shows that mixed layer theory, evaluated with EUREC^{4}A observations and with uncertain parameters constrained in a Bayesian approach, provides a closed description of subcloud layer thermodynamic variability. Monthly-mean residuals are 3.6 Wm^{-2} for moisture and and 2.9 Wm^{-2} for energy, and synoptic residuals are small and unbiased. Mixed layer theory is therefore a useful framework for characterizing subcloud layer variability and the processes controlling it. Surface wind speed variability is found to influence the subcloud layer depth and fluxes, yet thermodynamic variability above the subcloud layer top emerges as the primary control on subcloud layer moisture and temperature variability. Observed thermodynamic profiles and effective entrainment parameters constrained in the Bayesian methodology show evidence of an about 150 m thick transition layer separating the well-mixed part of the subcloud layer from the subcloud layer top.

Contrary to previous theory, Chapt. 5 shows that the canonical structure of strong jumps at the subcloud layer top is rare and only occurs in large cloud-free areas. A population of small clouds is shown to be responsible for smoothing vertical gradients over the transition layer. These findings suggest a new conceptual picture that the dissipation of small clouds modulates the transition layer structure. This analysis allows for positing an interplay between shallow and deeper trade-wind convection and a more active role for clouds in mixing processes that determine the subcloud layer structure.

Chapt. 6 traces how, in a representative subset of CMIP6 models, differing trade-wind cloud radiative responses can explain about 70% of the spread in global cloud feedbacks, and differences in trade-wind cloud responses discriminate between high and low climate sensitivity models. Chapt. 7 presents analysis supporting a weak trade cumulus feedback. Observations support a positive relationship between cloudi-

ness and convective mixing, which is opposite to the negative relationship thought to underlie strongly positive trade cumulus feedbacks. Three metrics related to the couplings among clouds, convective mixing, and relative humidity are presented that can be applied to both observations and climate models and yield a probabilistic estimate of the trade cumulus feedback.

Acknowledgments

The last three and a half years as a graduate student at the Laboratory of Dynamic Meteorology in Paris have been a wonderful adventure. I would like to warmly thank Sandrine Bony for sharing her brilliance and joy of doing science, and for the independence she gave me in developing my research projects. When I sent her an email one crisp fall day in Cambridge, asking how to apply to her group, I could not have imagined how much of the world she would open to me, both concretely[1] and more metaphorically. With the constraints of the covid-19 pandemic, our group meetings moved outside to underneath the canopy of the Lebanese cedar in the Jardin des Plantes[2], which felt a fitting parallel for her poetic, yet strong and guiding influence.

A heartfelt thank you also goes to Bjorn Stevens. I first met Bjorn whilst swiftly walking to a panel discussion he was giving at AGU in New Orleans in 2017. This feeling of dynamism[3] has always remained. I have never met someone who could develop such good metaphors, so quickly — which symbolizes, for me, at once his expansive knowledge, quick wit, and ability to distill complex ideas to their essence. I feel grateful to have joined such a positive and collaborative research environment that Sandrine and Bjorn brought to life, in particular around the EUREC^{4}A field campaign.

Many thanks to fellow cloud enthusiasts — you made the last years such a lot of fun. Ben Fildier and Ludovic Touzé-Peiffer, thank you for being the best roommates in Barbados and outreach team, or 'équipe de choc'. Our radiative profiles project helped me through the first covid-19 lockdown, and Ben's initiative to grow tomatoes on the ENS rooftop was a reminder to pause and enjoy small delights. Raphaela Vogel and Jessica Vial helped me process and work with *in situ* observations and GCM output, sometimes while canoeing through Hamburg's canals or hiking in Barbados.

[1] e.g., moving to Paris and traveling to Boulder, Barcelona, Mykonos, Barbados, Hamburg, and the Provence.

[2] `https://commons.wikimedia.org/wiki/File:P1260407_Paris_V_jardin_des_Plantes_cedre_du_Liban_1734_rwk.jpg`

[3] Meetings included but were not limited to: meetings while cycling, meetings in the MPI 'birch forest', and group chats while swimming in the ocean around Barbados.

Thank you to Hauke Schulz for smooth sailing on the Schlei to balance less-smooth days writing these chapters. Un grand merci und herzlichen Dank also to Caroline Muller, Max Popp, Nicolas Rochetin, Marius Léna, Addisu Semie, Geet George, Tobias Kölling, Theresa Mieslinger, Marc Prange, Jule Radtke, Theresa Lang, and Julia Windmiller for knowledge and happy moments shared – in Paris and Hamburg, or as neighbors on Coverly Drive in Barbados.

A few years ago, in 2016, it was a stroke of luck to walk into Mike McElroy's class, 'Energy and Climate: Vision for the Future' one snowy January day. This class and an ensuing research project with Mike led me to pursue a master's in Earth and Planetary Sciences during my fourth year at Harvard. A year of research with Peter Huybers after graduation taught me the ropes of scientific research and writing. His creativity and ability to meaningfully contribute to a vast arc of questions have been a steady source of inspiration over the past few years. I am grateful to have continued to work together throughout my PhD.

Financially, this thesis was largely supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement #694768).

Contents

Chapter 1

Why look to the trades?

> "We were still sailing in the blue zone of the trade winds. And it was every day, every day, every night, the same regular breath, warm, exquisite to breathe; and the same transparent sea, and the same small white clouds, speckled, calmly passing across the deep sky; ..." –Pierre Loti, *My Brother Yves*, 1883.

> "Climate change is a question of our world's destiny — it will determine the well-being of everyone on Earth". –Angela Merkel, 2017.

1.1 A defining challenge of our time

Climate prediction is a defining challenge of our time. It is long-established that greenhouse gases lead to a warming of the atmosphere (e.g., Arrhenius, 1908; Charney et al., 1979). A central open question, which motivates much of the research in this thesis, is what sets the speed and intensity of warming at the global scale. How much and how quickly will global temperatures rise due to greenhouse gas forcing? A subsequent question is what effects are produced by a certain amount of warming. Heating provides the energy to drive winds, evaporate moisture, form rain, and melt ice. Global warming is thus expected to influence extreme weather events, such as storms, floods, and droughts, and increase sea level. These regional changes have attendant consequences for human health, water resources, agriculture, migration, and biodiversity. As expressed by former German chancellor Angela Merkel, climate

change is a question of Earth's destiny and will affect the well-being of everyone on Earth.

Economically, climate change is an externality that is unprecedentedly large and uncertain (Tol, 2009). Estimates of the social cost of carbon range in magnitude (e.g., Stern, 2006; Weitzman, 2011; Nordhaus, 2018), yet they all project large consequences. These estimates, however, likely remain an incomplete representation of the possible effects of climate change, in particular heavy-tailed structural uncertainties associated with low-probability, high-impact scenarios (e.g., Weitzman, 2011, 2014; Wagner and Weitzman, 2016). There is thus a need for advancing knowledge about how much the Earth will warm and the implications of a given level of warming.

A primary lens for studying future climate is General Circulation Models (GCMs), also called global climate models. In 1963, Joseph Smagorinsky published a seminal paper describing numerical experiments using the primitive equations, a set of fluid equations that describe global atmospheric flows (Smagorinsky, 1963). This first GCM was based on the premise that studying the atmospheric circulation required a model capable of resolving heat transport from the equator to the poles. Smagorinsky's work paved the way to a vast research effort to couple the atmospheric circulation to different physical processes, such as moist convection and cloud formation, which have long been appreciated as important for the energetics of the atmosphere (e.g., Arakawa and Schubert, 1974; Stevens and Bony, 2013).

GCMs are a powerful tool for understanding how the climate system works, yet they remain an imprecise lens into the future. These models must make a number of approximations and simplifications, in particular to represent small-scale processes below the 100–200 km resolution of typical GCMs. Most processes related to moist convection and cloud formation cannot, for instance, be simulated explicitly and are instead represented indirectly by so-called parametrizations, which relate subgrid-scale processes to the large-scale variables explicitly resolved by the model (e.g., Siebesma et al., 2020). Representing tropical diabatic responses, in particular, relies heavily upon such parametrizations, whereas in the mid-latitudes, much of the energy and moisture transport is accomplished by baroclinic eddies that are better

resolved by GCMs (Smagorinsky, 1963; Stevens and Bony, 2013; Siebesma et al., 2020). Beyond these structural uncertainties, there are additional parametric uncertainties that result from inadequately sampling the high-dimensional parameter spaces of these models; a model with only 20 free parameters is, for instance, already a hypercube with around one million corners (Carslaw et al., 2018).

Due to these uncertainties, GCMs lay out a broad swath of futures in response to greenhouse gas forcing. A common metric for quantifying the response of the climate system to carbon dioxide (CO_2) radiative forcing is the equilibrium climate sensitivity (ECS). ECS is defined as the amount of surface warming at equilibrium following a doubling of atmospheric CO_2 concentrations. Knutti et al. (2017) write that ECS "has reached almost iconic status as the single number that describes how severe climate change will be". This idealized metric also serves as a way to compare model predictions.

Despite a long history of recognizing the influence of carbon dioxide on Earth's temperature, ECS has proven stubbornly challenging to estimate. In the 1820s, the French physicist Joseph Fourier suggested that the Earth traps heat (Fleming, 1999). In 1896, in developing a theory to explain the Ice Ages, Swedish physicist and chemist Svante Arrhenius first estimated how increases in atmospheric CO_2 are responsible for increases in Earth's temperature. In 1908, Arrhenius wrote, "Although the sea, by absorbing carbonic acid [resulting from carbon dioxide interacting with water], acts as a regulator of huge capacity, which takes up about five-sixths of the produced carbonic acid, we yet recognize that the slight percentage of carbonic acid in the atmosphere may by the advances of industry be changed to a noticeable degree in the course of a few centuries" (Arrhenius, 1908). Arrhenius was the first person to recognize the potential for anthropogenic CO_2 to change Earth's temperature and estimated a temperature increase of 4°C in response to a CO_2 doubling (Arrhenius, 1908). In the 1960s, Charles Keeling presented observational evidence of increasing atmospheric CO_2 concentrations due to human activities, relative to preindustrial levels trapped in ice cores. Climate models developed beginning in the 1960s provided a physical-numerical framework for studying the relationship between radiative forcing by CO_2

and surface temperature (e.g., Smagorinsky, 1963; Manabe and Strickler, 1964; Manabe and Wetherald, 1967). In 1979, the United States National Academy of Sciences convened a group of scientists, led by Jule Charney, to assess the potential influences of CO_2 on climate. The ensuing report, "Carbon Dioxide and Climate: A Scientific Assessment" became known as the 'Charney Report' and popularized the concept of equilibrium climate sensitivity. Drawing on results from two early climate models and physical insights, the Charney report put forward a best estimate of ECS of 3°C and range from 1.5–4.5°C (Charney et al., 1979), known as the 'Charney range'.

Tightening the Charney range of ECS is one of climate science's most enduring problems. The most recent assessment of climate sensitivity, across multiple lines of evidence gives a narrower range of 2.6–3.9°C (Sherwood et al., 2020). Notably, the tighter upper bound is not directly constrained by the latest generation of GCMs in the Coupled Model Intercomparison Project phase 6 (CMIP6) whose range is 1.8–5.6°C (Zelinka et al., 2020; Meehl et al., 2020). The upper bound is instead constrained by process-based and paleoclimate constraints, though paleoclimate constraints also use GCMs (Sherwood et al., 2020). Despite the apparent tighter upper bound, there are still open questions regarding the influence of changing spatial patterns of warming ('pattern effect'), long timescale feedbacks, and cloud feedbacks, which could still allow for a more positive upper bound on ECS (Sherwood et al., 2020).

The increase in complexity of GCMs has expanded the number of physical processes that can be studied with these tools, yet uncertainties in climate sensitivity remain similar to those that emerged from early modeling studies. In his landmark study, Smagorinsky concluded, "In pursuing the objective to generalize theoretical models we must ask ourselves whether greater detail in formulating the contributing processes is warranted by truncation errors, by sensitivity of the results to detail, by the resulting increase in computational complexity and time, and by ignorance of the way these processes really work. Very often this cannot be determined in advance, but must wait for computational experiments to be performed". Stevens and Bony (2013) respond, writing, "Although the drive to complexity has not reduced key un-

certainties, it has addressed Smagorinsky's question as to what level of process detail is necessary to understand the general circulation. There is now ample evidence that an inadequate representation of clouds and moist convection, or more generally the coupling between atmospheric water and circulation, is the main limitation in current representations of the climate system". The Charney report already discussed that a primary obstacle to better climate predictions was uncertainty in the radiative feedbacks associated with clouds (e.g., Charney et al., 1979; Bony et al., 2013b). These obstacles remain in CMIP6, with the response of clouds and convection still representing a primary uncertainty in constraining climate sensitivity (e.g., Zelinka et al., 2020; Meehl et al., 2020). This chain of research casts into relief that in addition to increasing the complexity of models, it is equally important to deepen physical, process understanding using observations and simpler modeling frameworks.

1.2 Climate questions traced to the trades

In the novel, *My Brother Yves*, the French naval officer and writer, Pierre Loti, writes about a journey in the trade winds, as cited in the opening quotation. The trade-wind zones extend from about 10–30° north and south of the equator. In both hemispheres, these zones are dominated by the trades, which are steady easterly winds with an equatorward component, blowing from the subtropical high to the equatorial low.

The etymology of the trade winds in different languages gives insight to their nature and perception. In German the winds are known as *Passatwinde*, coming from Portuguese, *passar*, or Italian *passata*, both meaning to 'go through', highlighting their steady movement. In French the trades are *les vents d'alizés*, similar to Portuguese, *ventos alísios*. The French name is speculated to relate to the Latin root *lixare*, which was used in medieval French to qualify the "smooth, delicate, gentle character of these measured winds that blow regularly" (Bellec, 2003). The serenity and steadiness of this etymology and Loti's description appear at odds with the urgency of climate change. How can such a calm and steady region play a role in determining how much the Earth will warm?

1.2.1 Role of trades in climate

A first answer comes by way of their large statistical weight: trade-wind regions are the most frequent tropical regime, and trade-wind cumuli are, moreover, the most frequent cloud type on Earth (e.g., Norris, 1998; Bony and Dufresne, 2005). Trade-cumulus regions cover approximately 20% of the Earth (e.g., Myers et al., 2021). Small changes in these regions with warming, such as associated with their cloudiness, therefore have a large global impact. Their geographic extent provides an intrinsic motivation to better understand the trades.

In the vertical, trade-wind air exhibits a characteristic layered structure (Riehl et al., 1951; Malkus, 1958; Augstein et al., 1974; Stevens, 2006). A robust vertical structure also allows for developing conceptual models, which are useful frameworks for developing process understanding (e.g., Stevens, 2006). In its most general form, the trade-wind atmosphere is divided into a shallow, moist layer typically extending to 2–3 km in height, topped by a much drier free troposphere. Near the equator, deep convection forms in response to the destabilizing effect of radiative cooling from water vapor (e.g., Emanuel et al., 1994). This deep, precipitating convection effectively dehydrates the atmosphere, such that air subsiding in the subtropics is very dry. Large-scale subsidence in the subtropics also suppresses deep convection, explaining why the trade-wind moist layer does not grow beyond few kilometer depth.

This shallow, moist layer is often referred to as the 'planetary boundary layer' and is itself divided into additional layers. Previous descriptions of these layers, primarily from Malkus (1958) and Augstein et al. (1974), are briefly summarized:

1. *surface layer*: the lowest 50–100 m over the ocean that are characterized by mechanical turbulence.

2. *mixed layer*: a neutrally stable, vertically well-mixed layer from the top of the surface layer up to approximately 600 m.

3. *transition layer*: a layer of approximately 100 m thickness that is thought to separate dry convection and mechanical mixing below from cloud convection

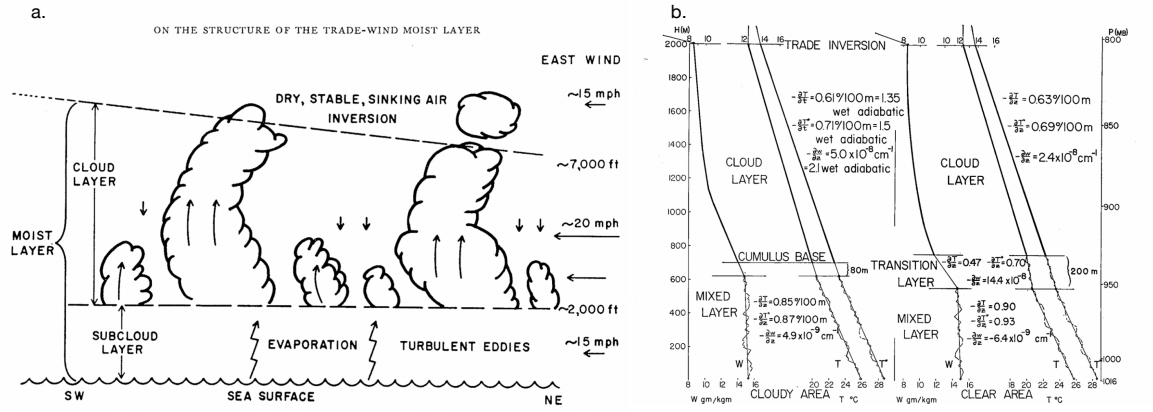


Figure 1-1: Reproducing two figures from Malkus (1958), one of the first papers to analyze the vertical structure of the trade-wind atmosphere. Panel (a) plots a cartoon of the vertical levels in the moist layer, from the sea surface to the cloud top, overlain by a dry, subsiding free troposphere, and (b) mean sounding profiles in cloudy (nine soundings) and clear areas (sixteen soundings) for the mixing ratio, temperature, virtual temperature. These profiles are annotated with estimates of the different vertical levels introduced above: mixed layer, transition layer, cloud layer, and inversion layer.

above the transition layer top. Together the mixed layer and transition layer are often referred to as the *subcloud layer*.

4. *cloud layer*: a layer wherein shallow trade cumuli are embedded. This layer typically has decreasing specific humidity and a temperature gradient that is slightly stronger than the moist adiabat. Cloudiness typically peaks around cloud-base and again near the trade-wind inversion. Trade-wind cumuli are shallow in height because they are capped by subsiding air and a stable inversion layer.

5. *trade-wind inversion layer*: a layer of strongly decreasing humidity ('hydrolapse') and increasing temperature around 2–3 km in height, which separates the moist layer from the dry free troposphere. The mean height of this interfacial zone is determined by the opposing balance between compression due to large-scale subsidence and expansion due to convection.

In the following sections, three vignettes are discussed that describe the implications

of this characteristic vertical structure and motivate a closer examination of the trade-wind planetary boundary layer.

Link in the global energy budget

One motivation to study the trade-wind planetary boundary layer is its role as an early link in the global energy budget (e.g., Riehl et al., 1951; Malkus, 1958). The trade-wind subcloud layer couples the surface to the cloud layer, and in doing so, regulates the import of energy and moisture from the ocean (e.g., Malkus, 1958; LeMone and Pennell, 1976; Stevens, 2007). Subcloud layer moisture and temperature variability control moist static energy variability, which influences the convective potential and thus cloudiness (e.g., Emanuel, 1986; Neelin et al., 1987; Lindzen and Hou, 1988; Emanuel, 1993). Clouds, however, also influence the subcloud layer. Clouds bring down drier air from aloft, which can then be entrained into the subcloud layer (Stevens, 2006). They can also influence the radiative budget of the subcloud layer and of the surface.

The export of latent heat from the trades to the equator is accomplished by advection of moisture by the trades. The trades can be interpreted in a number of ways. They can be interpreted, for example, as the return flow of the Hadley circulation, or, from a boundary layer perspective, as the result of the boundary layer momentum budget that is driven by pressure gradients induced by sea-surface temperature differences (Lindzen and Nigam, 1987). After being exported to equatorial regions, latent heat that originally accumulated in the trade-wind boundary layer has numerous downstream effects. The advection of water into the deep tropics, favors high values of boundary layer moist static energy and thus deep convection (Emanuel et al., 1994) there rather than in the trade wind regions (Oueslati and Bellon, 2013; Popp and Silvers, 2017). The deep convection in the deep tropics causes divergence near the tropopause and a return flow away from the equator that exports energy poleward, thus helping to maintain an approximate thermal equilibrium globally (e.g., Riehl et al., 1951; Malkus, 1958; Pierrehumbert, 1995). Outside the tropics, the poleward energy transport is dominated by eddies. These eddies transport latent heat poleward

at mid-latitude thus influencing large-scale circulations and the global hydrological cycle (e.g., Riehl, 1954; Heckley, 1985; Tiedtke, 1989)

Given its role in global energy transports, primarily by funneling latent heat to the tropics, subtle thermodynamic variations in the trades can have significant global reverberations (e.g., Malkus, 1958; Augstein et al., 1974; Tiedtke, 1989). These considerations motivate an investigation of the vertical structure of the trade-wind atmosphere, as well as its thermodynamic variability and vertical mixing processes.

Earth's 'radiator fins'

The vertical structure of trade-wind air also has implications for radiative processes. Fig. 1-2 shows that Earth is a water planet. It depicts upper-atmospheric water vapor absorption above about 4 km observed from the Meteosat satellite in the water vapor channel. Areas with a dry upper atmosphere are rare globally but emerge in the subtropics (e.g., Spencer and Braswell, 1997; Soden, 1998).

The trade-wind atmosphere is effective at cooling to space because of the large transmissivity of the dry free troposphere. Strong radiant energy loss to space occurs from the top of the shallow moist layer, whereas the dry atmosphere above does not return much downwelling radiation. The loss of infrared energy to space increases nonlinearly with decreases in relative humidity (Spencer and Braswell, 1997; Soden, 1998). For instance, a free troposphere with 10% mean relative humidity radiates about 320 Wm^{-2}, whereas an atmosphere with 30% relative humidity radiates only 300 Wm^{-2} (e.g., Fig. 2 in Spencer and Braswell (1997)). The key to allowing the tropics to lose its heat — the, on average, 400 Wm^{-2} of solar radiation that pour into the tropics from the sun — is thus to have vast dry areas. Evocatively, Pierrehumbert (1995) calls the subtropics the 'radiator fins', in contrast with the equatorial regions being like a 'furnace' or a 'hothouse'. Through their dryness, the trades thus increase the efficiency of radiant heat loss, stabilize tropical temperatures, and reduce the amount of energy that must be exported to the poles.

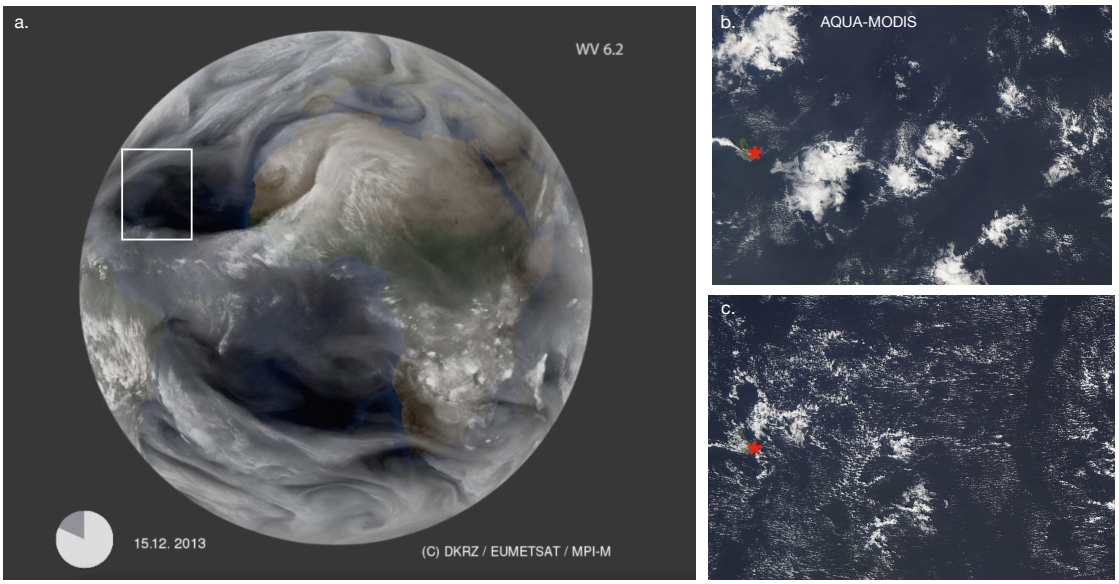


Figure 1-2: (a) the 6.2 micron METEOSAT channel, which measures atmospheric water vapor absorption above about 600 hPa. Darker shades indicate lower humidity and thus greater transparency to emitted longwave radiation. One such area with a dry free troposphere is highlighted by a white box. (b) Two AQUA MODIS GOES-E views of a trade-wind cloud field east of Barbados (indicated by a red star) on February 2, 2020 and (c) February 9, 2020.

A less cloudy future?

Besides its dry free troposphere, the trades further cool tropical temperatures through the presence of shallow cumuli, which have a negative radiative effect (e.g., Hartmann et al., 1992). Regarding these shallow trade-wind clouds, the question is not how much longwave radiation do the subtropics lose to space. It is, rather, how much shortwave energy do the subtropics absorb, with more shallow clouds implying a larger albedo and less heat gain, and vice versa. Trade-wind cumuli are, as introduced previously, the most frequent cloud type on Earth (e.g., Norris, 1998; Bony et al., 2017). The 'small white clouds', as described by Loti, are not as spectacular as cumulonimbus clouds that can grow higher than the tallest mountains or the expansive decks of stratocumulus clouds off western continental coasts. Yet they constitute the 'silent', or humble majority of Earth's clouds whose changes are greatly amplified by their frequency.

Climate change experiments with GCMs show that there is a propensity of trade-wind cloudiness, in particular at its base, to decrease with warming, raising the specter of high ECS values. In a pioneering study, Bony and Dufresne (2005) showed that cloud radiative responses in GCMs diverged most in shallow cumulus regimes, and these responses discriminate between high and low climate sensitivity models (Fig. 1-3a). This analysis also highlights how climate models, despite their known deficiencies, allow us to form hypotheses — for instance, that the sensitivity of trade-wind clouds to environmental conditions is a large source of uncertainty in global cloud feedbacks and ECS (Bony and Dufresne, 2005). Ensuing research has further supported the Bony and Dufresne (2005) findings, such as using different CMIP ensembles (e.g., Webb et al., 2006; Vial et al., 2013; Zelinka et al., 2020) and more idealized aquaplanet configurations (e.g., Medeiros et al., 2008). Fig. 1-2 shows two trade-wind cloud fields from the GOES-E satellite. This uncertainty can be rephrased pictorially: how will the views in Fig. 1-2 change with warming, towards a cloudier or less cloudy future in the trades?

Process-based studies, on the other hand, suggest that trade-wind clouds are less

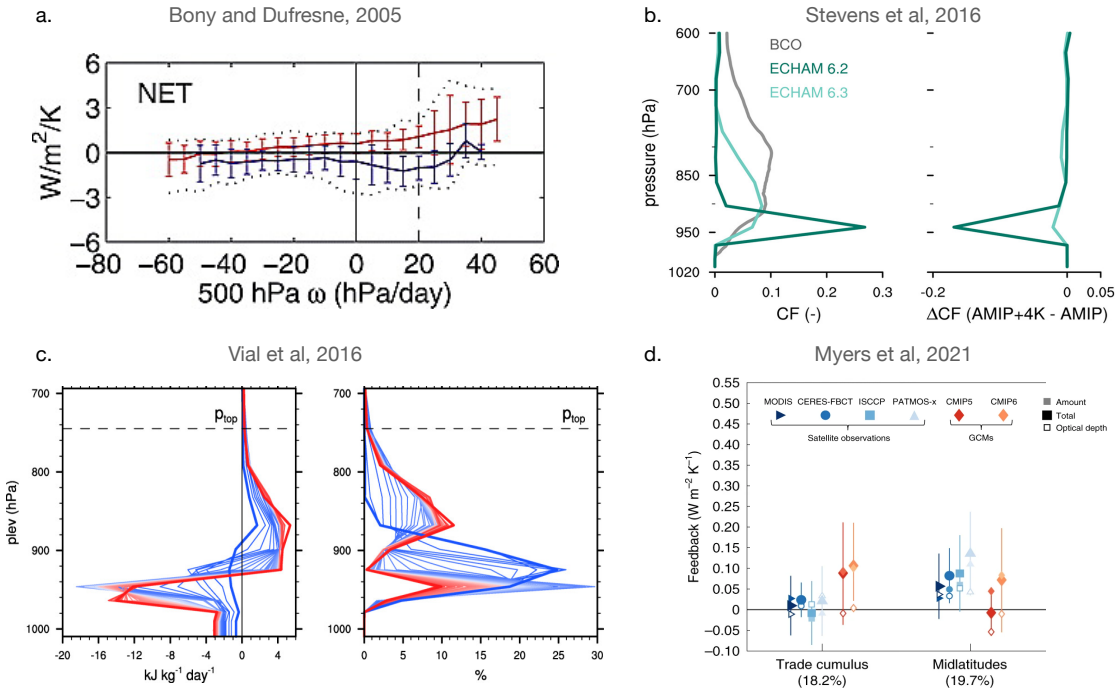


Figure 1-3: Figures illustrating the discrepancy between a large sensitivity of cloud-base cloudiness to changing environmental conditions in GCMs (a. ,b. , c.) and greater resilience in satellites (d) and LES (not shown). (a) Reproduction of Fig. 2 from Bony and Dufresne (2005) showing the change in the net cloud radiative effect with warming as a function of dynamical regimes. Red (blue) points correspond to models predicting a positive (negative) tropical-mean cloud radiative response. Markers show the mean, lines the standard deviation, and dotted lines the range of 15 coupled GCMs. (b) Reproduction of Fig. 1 in Stevens et al. (2016) showing cloud fraction (left) and change in cloud fraction (right) for two model versions differing in a single shallow convective mixing parameter. Also shown is an observed cloud fraction profile at the Barbados Cloud Observatory (grey). (c) Reproduction of Fig. 1 in Vial et al. (2016) showing the vertical distribution of the moist static energy flux convergence due to convection (left, in %) and cloud fraction (right, in %) for a range of convective mixing strengths, with minimum in thick blue, and maximum in thick red. (d) Adapted from Fig. 3 in Myers et al. (2021) showing observationally-constrained marine low cloud feedbacks scaled by regime frequency globally. Error bars for observations span 90% confidence intervals estimated from observational uncertainty and inter-model uncertainty in cloud-controlling factors. Bars for CMIP models span the range of simulated feedbacks associated with changes in low cloud properties.

sensitive to changing environmental conditions than GCMs predict. In large-eddy simulations (LES), trade-wind clouds at cloud-base are resilient to changing environmental conditions both in the current climate (Nuijens et al., 2015b) and idealized climate change experiments (Rieck et al., 2012; Blossey et al., 2013; Bretherton, 2015; Vogel et al., 2016). These process-based studies suggest a neutral or slightly positive trade-wind cloud feedback. Two recent studies, Myers et al. (2021) and Cesana and Del Genio (2021), use satellite measurements and also find evidence for a near-zero trade-wind cloud feedback (Fig. 1-3d.). Myers et al. (2021) also find that trade cumulus feedbacks estimated using CMIP models are markedly higher than their satellite-constrained estimate. In CMIP6 relative to CMIP5, a too-negative midlatitude cloud feedback, compared to satellite constraints, was corrected (e.g., Zelinka et al., 2020). This correction is argued to have unmasked a compensating too-positive cloud feedback in the trades, which then led to much higher ECS values in CMIP6, outside the Charney range (Myers et al., 2021), again emphasizing the importance of testing the credibility of strongly positive trade cumulus feedbacks simulated by GCMs.

There is thus a discrepancy to resolve — why are trade-wind clouds in GCMs much more sensitive to changing environmental conditions than in LES or satellite-derived studies? On the one hand, there are reasons to interpret results from LES and satellite-based studies with caution. In LES, the cloud amount is not linked to the large-scale circulations in which the clouds form, and cloud fraction is known to be sensitive to microphysics scheme, resolution, or domain size (e.g., Vogel et al., 2016; Vial et al., 2017; Radtke et al., 2021). Cloud fraction estimates from satellites were shown to diverge widely across passive remote sensors (Stubenrauch et al., 2013). Another consideration is that due to their coarse spatial resolution, satellite retrievals could, in principle, underestimate the number of cloudy pixels and overestimate the clear-sky fraction (e.g., Mieslinger et al., 2019, 2021); this would lead to an underestimate of the cloud feedback, which is proportional to the difference between the all-sky and clear-sky cloud radiative effect. The satellite-based studies also rely upon reanalysis data to estimate environmental conditions, yet these reanalysis data

are known to exhibit biases in the trades (e.g., Stevens et al., 2016). More fundamentally, the question remains why these lines of evidence suggest a small trade cumulus feedback. That is, the physical mechanism explaining the resilience of trade cumulus cloud fraction to changing environmental conditions remains unknown.

Growing evidence suggests that the diversity of GCM trade cumulus feedbacks can, on the other hand, be traced to a specific idea: how models represent lower-tropospheric vertical mixing (e.g., Sherwood et al., 2014; Bony et al., 2017). Fig. 1-3b,c. highlight two case studies from GCMs that differ in the strength of convective mixing, with large implications for changes in cloudiness and ECS. The general idea, which will be discussed in Chapt. 7, is that increased convective mixing brings dry air down from the free troposphere, drying the large-scale environment near cloud-base and decreasing cloud-base cloudiness. Greater mixing therefore reduces cloudiness at cloud base. In GCMs, differences in this vertical mixing are thought to result from different representations of the interplay among surface turbulent fluxes, convective mixing, and low-level radiative effects (e.g., Tomassini et al., 2014; Vial et al., 2016).

These processes are thought to explain large differences in the trade-wind cloud feedback, yet they have never been tested observationally. During the EUREC^{4}A (*Elucidating the role of cloud-circulation coupling in climate*) field campaign, we collected data to measure these interplays, such as among cloudiness, shallow convective mixing, radiative heating, and the large-scale environment, in order to constrain cloud feedbacks (Bony et al., 2017; Stevens et al., 2021). The campaign took place in January and February 2020 in the downstream trades of the North Atlantic. The original aim of EUREC^{4}A to constrain cloud feedbacks opened a broader set of questions to improve fundamental understanding of the trades. The Charney report stated that climate prediction may be "expected to improve gradually as greater scientific understanding is acquired and faster computers are built" (Charney et al., 1979). Since this report, faster computers have been built, but there are still numerous open questions regarding scientific understanding of the trades, several of which this thesis, using EUREC^{4}A data, tries to answer.

1.3 Thesis outline

From a methodological perspective, this thesis has two parts. In the first part (Chapters 2–5), *in situ* EUREC^{4}A observations are collected (Chapt. 2) and analyzed to deepen physical understanding of the characteristic vertical structure of trade-wind air and the processes that maintain it. Chapt. 4 closes subcloud layer moisture and energy budgets for the first time with *in situ* observations, using radiative heating profiles calculated and studied in Chapt. 3 and bulk models of surface and entrainment fluxes constrained in a Bayesian approach. The EUREC^{4}A observations also provide a basis for reconceptualizing the transition layer, which contrasts with previous views based on theory for cloud-free boundary layers (Chapt. 5).

Better understanding how the trades behave in the present-day using EUREC^{4}A observations allows for better evaluating how trade-wind clouds will change in the future. In the second part (Chapters 6–7), this improved physical understanding is applied to the evaluation of general circulation models. Chapt. 6 traces how, in CMIP6, uncertainties in trade cumulus feedbacks are strongly associated with uncertainties in the global cloud feedback and, to a lesser extent, in equilibrium climate sensitivity. Chapt. 7 shows analysis supporting the first process-based constraint on the trade cumulus cloud feedback. This constraint draws upon extensive EUREC^{4}A observations, mixed layer theory as evaluated in Chapt. 4, and output from ten GCMs. Conclusions and perspectives following from ideas presented in this thesis are discussed in Chapt. 8. Chapt. 9 presents additional materials.

1.4 A door in the sky — learning to see clouds

Clouds, while remaining the subject of rich scientific inquiry, have also long fascinated artists and writers. One can argue that there are a number of parallels between observing art and observing nature. Both involve noticing and connecting particular elements with the entirety. Art and nature evoke a sense of beauty and wonder that are thought-provoking. Seeing artwork, like observing nature, is enriched by

Figure 1-4: Artistic of clouds and turbulence from different time periods and cultures: (a) knight on horseback hidden in the clouds Andrea Mantegna's (1456–1459) painting of St. Sebastian, (b) figure hiding in the clouds in a 13$^{\text{th}}$ century fresco by Giotto di Bondone, (c) sketch of turbulence by Leonardo da Vinci (1452–1519), (d) *Starry Night* by Vincent van Gogh (1853–1890), (e) L'Embellie, or The Upswing by René Magritte (1898–1967), (f) the *The Great Wave off Kanagawa* by Hokusai (1760–1849), and (g) *La corde sensible*, or *The Heartstring* by René Magritte.

being with people of various backgrounds, as each person brings different perspectives and experiences to bear. Works of art, such as poetry, painting, or dance, also foster concision and intensity in communicating ideas, which is not unlike a scientific equation or explanation.

Looking at art, moreover, sharpens our senses and encourages close examination. The precise observer Andrea Mantegna's (1456–1459) painting of St. Sebastian is rewarded with a knight on horseback hidden in the turbulent clouds (Fig. 1-4a). Similarly, if looking closely at a 13$^{\text{th}}$ century fresco by Florentine artist, Giotto di Bondone, a profile of a smirking devil, or alternatively the artist himself, emerges in the vaporous clouds (Fig. 1-4b). In both cases, once you see the hidden figure a first time, you cannot fail to see it the following times. These vignettes raise the question of how we learn to see, as one often sees what one is conditioned to see. How often do we miss surprises hidden in plain sight? One way that EUREC^{4}A avoids potential blind spots is by bringing together a vast community. By merging differing, complementary perspectives, one can hope to overcome our blind spots and discover some of the surprises that have remained hidden in the clouds.

These cloud mysteries have long fascinated the human imagination. Art has often depicted depicted clouds and turbulent motion, as illustrated in Fig. 1-4. Painter

René Magritte writes about his research on "the problem of the cloud": "I feel a drive to paint a cloud, perhaps a hundred" (Magritte et al., 1992). Indeed, his painting L'Embellie, or The Upswing, (1962) depicts a door in the sky that opens onto a field of shallow clouds (Fig. 1-4e). The Barbadian poet, Kamau Brathwaite (1930–2020), uses the same visual language as Magritte, writing, 'from under the clouds where I write the first poem ... a door opening in the sky" (Brathwaite, 2005). To return to the questions introduced in this chapter, Magritte and Brathwaite's images of a door in the sky evoke the history of climate model analysis of the trades. These model analyses revealed limitations in our physical understanding about shallow convective processes in the trades, which represent bottlenecks for climate predictions. That is, they have shown us the door to deeper physical understanding and improved climate prediction. Now, EUREC^{4}A has allowed us to walk through this door. The German physicist, Max Planck, writes about the scientific process of formulating questions and then posing them to nature that is at the heart of EUREC^{4}A: "An experiment is a question which science poses to Nature, and a measurement is the recording of Nature's answer. But before an experiment can be performed, it must be planned — the question to Nature must be formulated before being posed" (Planck, 2014). The 'recordings' of nature's answers from EUREC^{4}A invite us to deepen our physical understanding, all the while being filled with a sense of wonder and curiosity about these questions in the sky.

Chapter 2

Can EUREC^{4}A answer climate questions?

At the time of EUREC^{4}A, the scientific study of the trades had progressed to the point of formulating specific questions that, if tested observationally, could lead to a step-change in understanding, as discussed in Chapt. 1. Before transitioning to the research chapters, this chapter discusses additional motivations for organizing the EUREC^{4}A field campaign (Sec. 2.1) and presents the primary measurement platforms (Sec. 2.2.1). An argument is also given why these measurements are amenable to answering climate questions (Sec. 2.2.2). I had the opportunity to take part in EUREC^{4}A and briefly describe my research flight experiences (Sec. 2.3) and contributions to the campaign (Sec. 2.4).

2.1 Why organize EUREC^{4}A?

The EUREC^{4}A (*Elucidating the role of cloud-circulation coupling in climate*) field study took place in January and February 2020 in the downstream Atlantic trades, east of Barbados (see Fig. 2-1 for experiment location). As introduced in Chapt. 1, it is the most extensive field campaign ever performed in the trades and measured the interplay among clouds, circulations, and the large-scale environment (Bony et al., 2017; Stevens et al., 2021). EUREC^{4}A was originally conceived to observationally test trade cumulus feedback mechanisms and to provide benchmark measurements for a new generation of models and satellite observations (Bony et al., 2017; Stevens et al., 2021). These aims require jointly characterizing the micro- and macrophysical

environment of clouds, and the cloud field itself. Yet the question arises, why organize such a campaign, given the large number of previous field campaigns and the wealth of data from satellites?

Advancing understanding on the questions raised in Chapt. 1 requires observational data, but such data is lacking in the trades. The last major field studies focusing on the trades took place about 50 years ago, before climate change was considered a salient issue. They include the *Atlantic Expedition* from September to October 1965 (Augstein et al., 1973), the *Atlantic Tradewind EXperiment*, ATEX, in February 1969 (Augstein et al., 1974), the *Barbados Oceanographic and Meteorological Experiment*, BOMEX, from May to July 1969 (Holland, 1970), and the Puerto-Rico Experiment in December 1972 (LeMone and Pennell, 1976). These campaigns took place at the advent of the satellite era, so they were not yet accompanied by observations from space. Despite occurring half a century ago, these studies continue to act as reference data sets for studying and simulating cloudiness the trades (e.g., Stevens et al., 2001; Siebesma et al., 2003).

More recent field campaigns focused on different cloud regimes and other questions than jointly characterizing clouds and their large-scale environment. Some field studies sampled primarily stratocumulus regimes (e.g., Stevens et al., 2003; Bretherton et al., 2004; Wood et al., 2011; Russell et al., 2013), or regions of deeper convection (e.g., Betts, 1974; Johnson et al., 1999; Raymond et al., 2003). Other field campaigns made measurements in the trades, such as *Rain in shallow cumulus over the ocean* (RICO Rauber et al., 2007), *Cloud System Evolution in the Trades* (CSET Albrecht et al., 2019) and *Organization of Tropical East Pacific Convection* (OTREC Fuchs-Stone et al., 2020), yet their scientific motivations were targeted towards more specific physical processes, such as precipitation and the stratocumulus-to-cumulus transition.

Since early field campaigns in the trades, the expansion of satellite measurements has provided global views of Earth and extensive new data. Yet as alluded to in Chapt. 1, the coarse resolution of passive remote sensing measurements leads to biases in their representation of vertical moisture profile features, such as the sharp

decrease at the planetary boundary layer top or elevated moisture layers (e.g., Loeb et al., 2009; Chepfer et al., 2010; Stevens et al., 2017; Prange et al., 2021), especially in the lowest three kilometers, corresponding to the weakest absorption lines (Chazette et al., 2014). Regarding cloudiness, biases in cloud detection among different passive remote sensors lead to large discrepancies in estimates of trade-wind cloud fraction and related quantities (Stubenrauch et al., 2013). The highest-resolution satellite measurements better resolve trade-wind clouds (e.g. the ASTER instrument has up to 15 m spatial resolution), yet these high-resolution sensors have limited temporal sampling (e.g., Mieslinger et al., 2019, 2021). Cloudiness is often studied in mesoscale large-eddy simulations, yet cloud fraction and cloud organization are sensitive to resolution, domain size, and microphysics scheme (e.g., Bretherton and Blossey, 2017a; Vogel et al., 2016; Radtke et al., 2021).

Given these limitations, *in situ* observations are critical for testing ideas about trade-wind cumuli. To begin filling this measurement gap, in 2010, the Max Planck Institute for Meteorology and the Caribbean Institute for Meteorology and Hydrology established the Barbados Cloud Observatory (BCO) on the windward side of Barbados, at the easternmost point of the island (Stevens et al., 2016). The BCO intercepts air that has been undisturbed by land influence since the western coast of Africa. Other long-term, ground-based measurement stations are located in the extratropics, such as Cloudnet] (Illingworth et al., 2007) and the U.S. Department of Energy's ARM (Atmospheric Radiation Measurement) climate research facilities (e.g., Moran et al., 1998; Long et al., 2013). The BCO is unique, however, in being the only long-term observational site in the trades. The BCO was also an anchor for two Next-Generation Aircraft Remote Sensing for Validation Studies airborne field campaigns (NARVAL and NARVAL2), held in December 2013 and August 2016 in preparation for EUREC^{4}A Stevens et al. (2016); Konow et al. (2019).

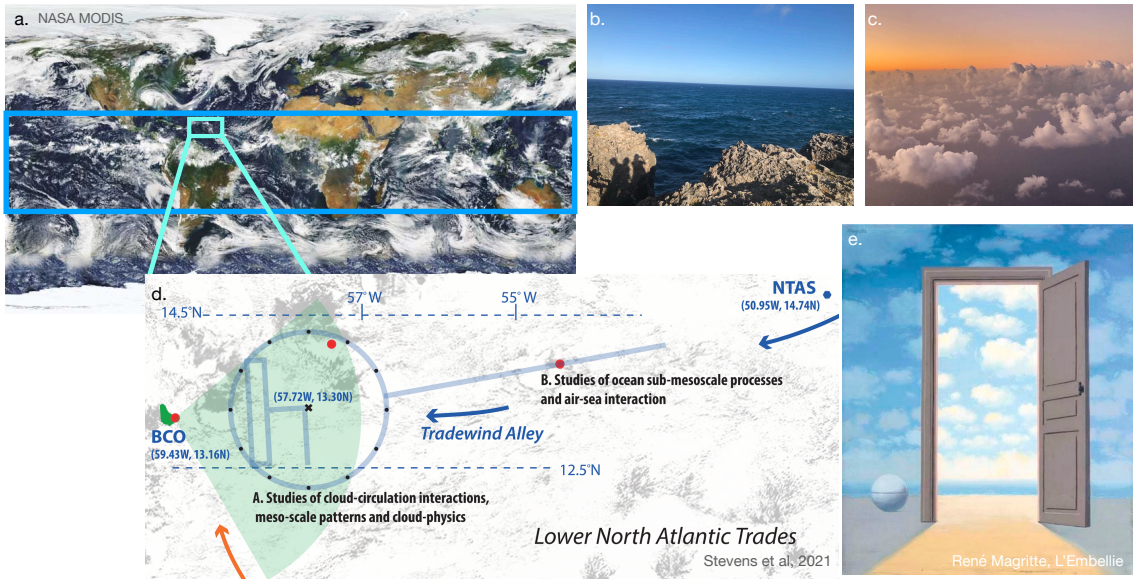


Figure 2-1: Location of EUREC^{4}A field study in the downstream North Atlantic trades, east of Barbados, as shown in (a) and (d). (b) shows the view from the Barbados Cloud Observatory in a rare cloud-free moment, and (c) depicts a view from flying on the WP-3D aircraft. In (e) is René Magritte's painting, *L'embellie* or *The Upswing* (1962), as if opening the door to the trade winds that blow from the east, as illustrated schematically in (d).

2.2 Building a cloud laboratory in the trades

Growing from these observational foundations, EUREC^{4}A brought together four ships, five aircraft, the BCO, a C-band rain radar, and a multitude of uncrewed aerial and seagoing systems, including fixed-wing aircraft, quadcopters, drifters, buoys, underwater gliders, and Saildrones (Stevens et al., 2021). EUREC^{4}A also launched an extensive sounding array consisting of 1068 dropsondes and 1512 radiosondes. This 'symphony' of measurements can be thought of building a transient cloud laboratory in the trades. This cloud laboratory can then be used to observationally test questions, such as those raised in the previous chapter.

2.2.1 Primary measurement platforms

In the following section, the four measurement platforms whose data are used throughout the following chapters are briefly outlined. Further information is given in later chapters regarding particular analyses. The high-flying aircraft, HALO (High Altitude and Long Range Research Aircraft) (Konow et al., 2021) characterized clouds and the environment from approximately 9–10 km, both with remote-sensing instruments and by launching approximately 800 dropsondes around the perimeter of a fixed, mesoscale circle (Stevens et al., 2021). It followed a circular flight pattern known as the 'EUREC^{4}A circle' (Fig. 2-2). This EUREC^{4}A circle was centered at 13.3° N, 57.7° W and had an about 220 km diameter, roughly comparable to the meso-β scale from Orlanski (1975) or the size of a typical GCM grid box. A second aircraft, the ATR-42, primarily followed a 120x15 km rectangular, or racetrack, flight pattern at low-levels within the EUREC^{4}A circle (Fig. 2-2) (Bony et al., 2022). One particularity of the ATR-42 was its sideways-pointing lidar and radars that, notably, measured the cloud fraction at cloud base. Flight legs below cloud base also characterized turbulent structures, microphysics (aerosols and clouds), and thermodynamic variability in the subcloud layer and near-to-surface. A third aircraft, the WP-3D Orion, augmented this sampling of clouds and their environment with additional dropsondes and remote sensing (Pincus et al., 2021). Within the EUREC^{4}A circle, a research vessel, the R/V Meteor, complemented the three aircraft by providing sea temperature estimates, surface-based remote sensing, and surface flux measurements along an about 200 km transect in the same domain, from 12.5–14.5° N along the 57.255° W meridian (Fig. 2-2). Data from the Barbados Cloud Observatory provide additional data.

2.2.2 Relevance for climate questions

A common denominator of the platforms described in Sec. 2.2.1 is that they undertook intense, unbiased statistical sampling. That is, they did not 'cloud chase', or seek out specific meteorological conditions, as did other platforms, such as the Twin Otter

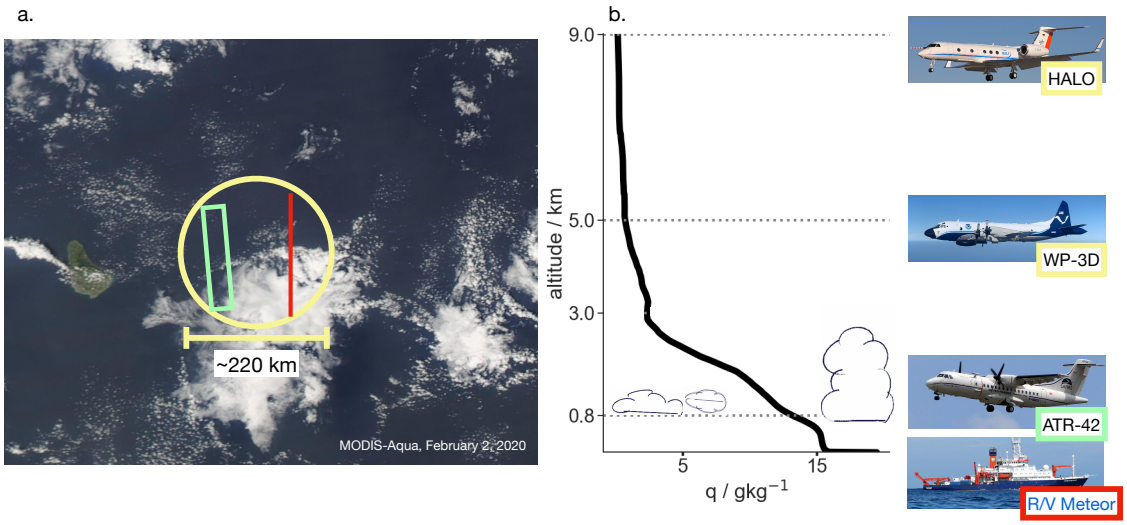


Figure 2-2: (a) MODIS-AQUA satellite image overlain with approximate tracks for HALO and the WP-3D in the EUREC^{4}A circle (yellow), ATR-42 rectangular pattern (green), and R/V Meteor ship tracks (red), and (b) the campaign-mean specific humidity profile annotated with the approximate flight levels of the different aircraft, noting that HALO dropsondes yield data from 9–10 km and WP-3D dropsondes from about 5–7 km downwards.

aircraft that was also part of the campaign. Instead, the platforms followed prescribed flight patterns – the EUREC^{4}A circle for HALO, and to a lesser extent, the WP-3D, the racetrack for the ATR-42, and the transect for the R/V Meteor. These measurements therefore yield unbiased realizations of the large-scale environment. As a result, measurements from the different platforms are assumed to be consistent and comparable, despite sampling at slightly offset locations and times.

For the relevance of EUREC^{4}A measurements to climate questions, the measurements need to fulfill two additional conditions beyond being unbiased samples. First, measurements must be representative of the broader trade wind regions, such that inferences from EUREC^{4}A generalize to the trades as a whole. Second, inferences from short timescales must be relevant for longer climate timescales.

For the first condition, Medeiros and Nuijens (2016b) show that in GCMs and observations (from the BCO and Calipso), shallow cumulus clouds near Barbados exhibit similar structure and variability to clouds across the trade wind belt as a whole. Medeiros and Nuijens (2016b) also show that errors simulated by GCMs

near Barbados – regarding the vertical structure of clouds, the planetary boundary layer, covariance with environmental conditions, and turbulent and convective mixing processes that give rise to clouds – are similar to errors simulated in other trade-wind regions. Rasp et al. (2020) find, similarly, that patterns of shallow cumulus cloud organization present near Barbados are also present through the trades as a whole. These studies lay an empirical foundation that clouds near Barbados are representative of trade-wind clouds across the tropics.

The second condition is that short timescale variability is informative of variability expressed on longer scales. Clouds and many associated processes, such as turbulence, entrainment, and convective mixing, are 'fast physics'. Over the past two decades, numerous studies have shown that climatological biases in GCMs, in particular regarding moist convection, are already evident on shorter timescales, even within a few days of model simulation, motivating the evaluation of GCMs in 'weather mode' (e.g., Rodwell and Palmer, 2007; Williams et al., 2013). Regarding changes in tropical circulation and precipitation associated with the direct adjustments to CO_2, Bony et al. (2013a) show that about half of the 30-year mean change from a quadrupling of CO_2 occurs within five days, showing that these changes rely upon fast physical processes. Based on these considerations, it therefore appears a reasonable ansatz that EUREC^{4}A observations can provide a window into understanding these fast physical processes, which can then inform longer-term responses. Indeed, in Chapt. 7, the relationships among convective mixing, cloudiness, and relative humidity are shown to be similar at the three-hourly and monthly timescales in GCM output. These results indicate that three-hourly processes are informative for explaining longer timescale variability. Vial et al, in prep also demonstrate that variability on the diurnal timescales is representative of variability of longer, climatological timescales. On slightly longer time scales, previous studies also indicate that cloud feedbacks are similar whether calculated from inter-annual, or longer timescale variability (e.g., Zhou et al., 2015; Sherwood et al., 2020).

2.3 Flying

I had the remarkable opportunity to take part in the EUREC^{4}A campaign in January and February 2020. The *in situ* data that I analyze in the following chapters is thus data that I helped collect. The benefits of being immersed in the natural environment one wishes to study is perhaps best expressed by the renowned American mathematician and meteorologist, Edward Lorenz. Regarding atmospheric circulations, Lorenz wrote, "Before one can make any serious attempt to explain the circulation of the atmosphere, he must become familiar with the circulation which he wishes to explain... experience suggests that the investigator who attempts to deduce the atmospheric circulation without first observing it is placing himself at a considerable disadvantage" (Lorenz, 1967). This sentiment certainly holds for observing and making an attempt to better understand processes in the trades. There is also a unique spirit of discovery, joy, and camaraderie in participating in a field campaign that brings together people from numerous institutions and countries in a shared pursuit.

While in Barbados, I took part in two research flights: one on the ATR-42, and another on the WP-3D, known informally as the 'hurricane hunter' or 'Miss Piggy'. These were my first research aircraft experiences and gave me different perspectives of the trade-wind clouds. The ATR flew through clouds near their base around 800 m, and the WP-3D flew higher, around 5–6 km on average. On the French ATR-42, I sat next to Marie Lothon who was already analyzing turbulence data in real-time and initiated me as a first-timer into the workings of a flight. On this day, February, 5, 2020, there were many clouds at cloud base (Fig. 2-3a,d). On other days, the meteorological situation was quite different, and given the unbiased (vs. cloud-chasing) nature of the sampling, the crew found themselves flying through long stretches of nearly clear-sky. After hours of flat lines on different measurement devices designed to measure clouds and humidity, one can understand when Nicolas Rochetin, a French scientist, exclaimed, "Holy mackerel, we missed the stratiform zone again!" ("Saperlipopette, on a encore raté la zone stratiforme !"). Even when the skies were clear, a rapid descent down to 60 m provided an adrenaline rush. During this flight leg, it felt

Figure 2-3: Photos from flying in the ATR-42 (left) near (a,d) the maximum in cloud-base cloudiness and (e) near the surface around 60 m, and (right, b,c,f) in the WP-3D, depicting wisps of clouds detraining condensate at cloud top (b,c).

as though the airplane was skimming along the ocean surface. One was close enough to spot wave crests breaking into whitecaps in the strong wind and the sargassum seaweed that is so troublesome for Barbadian tourism. Within the ATR aircraft, it was still possible to appreciate the coordination of the EUREC^{4}A experiment. The online *Planet* interface allowed for tracking, for instance, the Twin Otter aircraft that zigzagged around, chasing clouds, and the Poldirad rain radar, which we had helped build a few days earlier and that now measured precipitation. One could also track how the HALO aircraft flew above us in circles. On our headphones, the French pilots joked about the Germans who throw "800 lithium batteries in the sea". (That is, about 800 dropsondes launched from HALO.) The concurrence of the ATR-42 and HALO flight patterns (Fig. 2-2) and their unbiased sampling suggest that their measurements should be coherent. Indeed, after the campaign, strong agreement was found between the two platforms across the campaign, such as regarding their thermodynamic variability at different heights, indicating that measurements indeed reflect large-scale, rather than local variability (Bony et al., 2022).

My flight on NOAA's hurricane hunter was also a poetic, fascinating experience. The hurricane hunter, as its name implies, normally flies in much more trying conditions. The evening before my flight, Kerry Emanuel was telling us about the first and only crash of a hurricane hunter plane during Hurricane Janet in 1955, and another harrowing, near-crash during Hurricane Hugo in 1989. In the pre-flight meeting prior to our take-off, unusually strong wind shear was forecast for the flight area. We were then instructed to put on safety suits whose pockets contained essentials in case of an emergency, such as a flashlight, knife, and warming blanket. These aspects certainly added to the mystique and excitement, tinged with very slight fear, when stepping on board for the flight.

The WP-3D flight on February 9, 2020 was an all-night flight during full moon. Even deep in the night, the moonlight illuminated a panorama of cloud formations: billowing cumuli, cloud arcs, decks of clouds, veil clouds that appear like flying carpets, and a handful of extremely narrow cloud towers that grew to 7–8 km. I sat in the cockpit during take-off and parts of the flight. Despite the illusion of gliding

across a tactile or well-defined surface of the clouds, there is no such obvious boundary between clouds and their environment. Encountering the cloud edge is, of course, not like walking through a door, but rather like encountering a gradient of water vapor that would be easier to observe in infrared light. Stevens et al. (2017) expresses it evocatively and succinctly: "If it did not have to condense to become visible, water vapor would fuel the fascination of many more scientists. Imagine seeing with the naked eyes how elevated layers of water vapor, and its radiative effects, engender shallow circulations, or how pockets of humidity surround and socialize cumulus convection. Imagination is indeed necessary because water vapor's mysteries arise as much from its visible transparency as from the opulence of its infrared opacity".

One frequently-occurring cloud formation was small wisps of clouds that detached from cloud tops and then drifted and dissipated into the drier environment (Fig. 2-3b,c). This detrainment moistens and cools the surrounding environment. Observing these clouds out the window was the initial spark for proposing a new conceptual model of the transition layer in Chapt. 5.

2.4 Organizing outreach

Before concluding this chapter and transitioning to the research chapters, I summarize two contributions to the campaign. The first contribution is outreach work that we organized in collaboration with Barbadian colleagues Rebecca Chewitt-Lucas, Branden Spooner, Shanice Whitehall from the Caribbean Institute for Meteorology and Hydrology (CIMH), and French colleagues, Benjamin Fildier and Ludovic Touzé-Peiffer. Ulrike Kirchner and Dörte de Graaf from the Max Planck Institute for Meteorology also helped us organize outreach. The outreach initiatives included a two-day scientific symposium, about 20 visits to local schools, a weekly seminar series with local institutes, and an open house for the broader community to visit different measurement platforms. The second contribution was calculating clear-sky, aerosol-free atmospheric radiative profiles from 2580 soundings (1068 dropsondes and 1512 radiosondes) launched during the campaign (Albright et al., 2021a). This analysis is

presented in Chapt. 3.

The research symposium, 'From BOMEX to EUREC^{4}A', brought together 24 local and visiting scientists and about 100 participants to consider the advances since BOMEX, a major field campaign that took place 50 years before EUREC^{4}A (Fig. 2-4). While BOMEX had many similarities with EUREC^{4}A, involving several ships and research aircraft, the main difference lies in their objectives. For BOMEX, the main objective was studying the exchange of moisture, heat, and momentum between the ocean and the atmosphere. At the time, climate change was not yet a topical issue, so the main goal of the campaign was to improve weather forecasting. For EUREC^{4}A, by contrast, better understanding and constraining the magnitude of future climate change is at the heart of the campaign.

Two 'alumni' of the BOMEX campaign, 50 years before, spoke during the symposium: Clyde Outram, former director of Civil Aviation on Barbados, and Pat Callender, then the head of the Barbadian airport. They vividly recounted their experiences, even bringing an insignia from BOMEX that they had kept for five decades. Callender raised further points of contrast between BOMEX and EUREC^{4}A, describing how he had to negotiate on behalf of air traffic controllers seeking compensation for managing United States aircraft during the BOMEX campaign, and local scientists were not often co-authors on the papers published with BOMEX data. EUREC^{4}A, on the other hand, is a Barbadian-German-French-American initiative with strong Barbadian involvement and expertise.

Together with the Barbadian Ministry of Education, we also organized visits to about 20 local elementary and secondary schools (Fig. 2-4). The Barbadian government has strongly invested in education, and our efforts sought to reinforce their environmental education. We interactively explored the scientific method by conducting small experiments using simple materials. Using commonly-found materials made it easier for students to recreate the experiments at home. The experiments included creating a cloud in a bottle, observing surface tension, and asking how stratification affects mixing and overturning (e.g., by observing whether an ice cube melts more quickly in fresh or salty water). During the school visits, we also tried to vi-

Figure 2-4: Photographs from outreach activities organized, including a two-day symposium, 'From BOMEX to EUREC^{4}A (a) and a school visit where Ludovic Touzé-Peiffer and I perform experiments with local students (b,c).

sualize questions asked during EUREC^{4}A and benefited from cloud animations and presenting tips provided by Tim Cronin and Pier Siebesma. We also organized an outreach day, wherein a number of facilities opened to the general public, such as for the launching of radiosondes and drones. Our outreach activities are summarized in the EUREC^{4}A overview paper, Stevens et al. (2021).

Chapter 3

Atmospheric radiative profiles during EUREC^{4}A

This chapter presents the method to calculate atmospheric radiative profiles from EUREC^{4}A soundings and presents initial results. Calculating these radiative profiles is a necessary step for analyses in later chapters, namely closing subcloud layer thermodynamic budgets (Chapt. 4) and reconceptualizing the transition layer (Chapt. 5).

3.1 Abstract

The couplings among clouds, convection, and circulation in trade-wind regimes remain a fundamental puzzle that limits our ability to constrain future climate change. Radiative heating plays an important role in these couplings. Here we calculate clear-sky radiative profiles from 2580 in situ soundings (1068 dropsondes and 1512 radiosondes) collected during the EUREC^{4}A field campaign, which took place in the downstream trades of the western tropical Atlantic in January-February 2020. We describe the method used to calculate these cloud-free, aerosol-free radiative profiles. We then present preliminary results sampling variability at multiple scales, from the variability across all soundings to groupings by diurnal cycle and mesoscale organization, as well as individual soundings associated with elevated moisture layers. We also perform an uncertainty assessment and find that the errors resulting from uncertainties in observed sounding profiles, and ERA5 reanalysis employed as upper and lower boundary conditions are small. The present radiative profile data set

This chapter was published in *Earth System Science Data* in February 2021 (Albright et al., 2021a).

can provide important additional detail missing from calculations based on passive remote sensing and aid in understanding the interplay of radiative heating with dynamic and thermodynamic variability in the trades. The data set can also be used to investigate the role of low-level radiative cooling gradients in generating shallow circulations. All data are archived and freely available for public access on AERIS at `https://doi.org/10.25326/78`.

3.2 Introduction

The EUREC^{4}A field campaign, which took place in January and February 2020 in the downstream trades of the western tropical Atlantic, was designed to elucidate the couplings among clouds, convection, and circulation in trade-wind regimes and understand the role of this interplay in climate change (Bony et al., 2017). Shallow trade-wind clouds cover large parts of tropical oceans (Wood, 2012), yet their response to warming remains largely unknown, and uncertainty in shallow convective processes are the cause for large uncertainties in climate projections (Bony and Dufresne, 2005; Vial et al., 2013; Sherwood et al., 2014; Zelinka et al., 2020). Among all physical processes involved in shallow convection, atmospheric radiative cooling emerges as key to the coupling between low-level circulations and convection. Understanding the dynamics driven by variations in radiative heating rates, and potential relationship to the mesoscale organization of clear and cloudy regions, was one motivation for the campaign (Bony et al., 2017).

A characteristic feature of the trade-wind vertical moisture profile is a sharp humidity gradient between the moist marine boundary layer and dry, subsiding free troposphere around two kilometers Riehl et al. (1951); Malkus (1958). This characteristic vertical moisture structure has important implications for radiative cooling profiles, but it is difficult to observe with satellite remote sensing (Stevens et al., 2017). Indeed, moisture profile features, such as the sharp decreases in moisture at the top of the marine boundary layer or elevated moisture layers, are smaller than typical weighting functions of even hyperspectral instruments (e.g. Maddy and Barnet, 2008;

Schmit et al., 2009; Menzel et al., 2018), especially in the lowest three kilometers, corresponding to the weakest absorption lines (Chazette et al., 2014). The lack of informative observations means that the vertical profile of water vapor in large-scale atmospheric analyses do not represent the fine-scale moisture structure indicated by soundings (Pincus et al., 2017). Errors in the vertical moisture structure estimated from passive remote sensing produce corresponding errors in radiative cooling profiles computed from retrievals and/or analyses, making *in situ* soundings especially valuable.

Here we calculate radiative profiles from 2580 in situ soundings (1068 dropsondes and 1512 radiosondes) collected during EUREC^{4}A, whose network of observations provided extensive sampling of the tropical trade-wind environment. Similar studies have been conducted over continents as part of the Atmospheric Radiation Measurement program (Kato et al., 1997; Mlawer et al., 1998), over the western Pacific warm-pool region as part of the TOGA COARE (Coupled Ocean–Atmosphere Response Experiment) (Guichard et al., 2000), and over the western tropical Atlantic, albeit focused on transported Saharan dust layers (Gutleben et al., 2019). The present radiative profiles have the potential to complement and further what can be learned from calculations based on passive remote sensing. In addition, this data set may help in understanding how low-level gradients in radiative cooling fuel shallow circulations, as observed to emerge in remote sensing and large eddy simulations (L'Ecuyer et al., 2008; Stephens et al., 2012; Seifert et al., 2015). These shallow circulations are speculated to influence the mesoscale spatial organization of shallow convection, a question at the core of EUREC^{4}A (Bony et al., 2020; Stevens et al., 2020b).

In Sec. 3.3, we describe the data, the radiative transfer code, and the procedure underlying the calculation of the radiative profiles. We then present initial results to open the discussion on questions that could be explored with these radiative profiles (Sec. 3.4). Lastly, we perform an uncertainty assessment (Sec. 3.5) and find that errors resulting from uncertainties in the sea surface skin temperature, *in situ* soundings, and ERA5 reanalysis used as boundary conditions are modest.

3.3 Data and methods

3.3.1 Radiosonde and dropsonde data

From January 8 to February 19, over 2500 atmospheric soundings were conducted using dropsondes and radiosondes over the western tropical Atlantic ocean south and east of Barbados. As the sondes fall or ascend, their simple autonomous sensors, equipped with a GPS receiver, measure the vertical profiles of pressure, temperature, relative humidity, and instantaneous horizontal wind. To calculate radiative profiles, we employ level-3 data, which have been interpolated into a common altitude grid with 10 meter spacing (Stephan et al., 2020; George et al., 2021). We select dropsondes and radiosondes that have measurements on more than ten atmospheric levels in total. This filter suffices to remove failed soundings and results in an input data set consisting of 1068 atmospheric profiles from dropsondes and 1436 profiles from radiosondes. The minimum and maximum levels z_{min} and z_{max} measured by each sonde are reported in the final data set.

Figure 3-1a shows the geographic and temporal distributions of the sondes used to calculate the radiative profiles. Radiosondes were launched from a network of one land station and four research vessels, within a region ranging from 51–60°W to 6–16°N. On land, radiosondes were launched from the Barbados Cloud Observatory (BCO), located on a promontory on the easternmost point of Barbados called Deebles Point (13.16°N, 59.43°W), where it is exposed to relatively undisturbed easterly trade-winds. The fleet of four research vessels includes the French research vessel L'Atalante, two German research vessels Maria S. Merian (MS-Merian) and Meteor, and the American research vessel from the National Oceanic and Atmospheric Administration (NOAA) Ronald H. Brown (RH-Brown). Dropsondes were launched from both the German High Altitude and Long Range Research Aircraft (HALO) and the United States Lockheed WP-3D Orion from NOAA (WP-3D). HALO typically flew at an altitude of 30,000 ft (approximately 9 km), following a circular flight pattern with 90 km radius centered at 13.3°N, 57.7°W. When launching sondes, the WP-3D flew

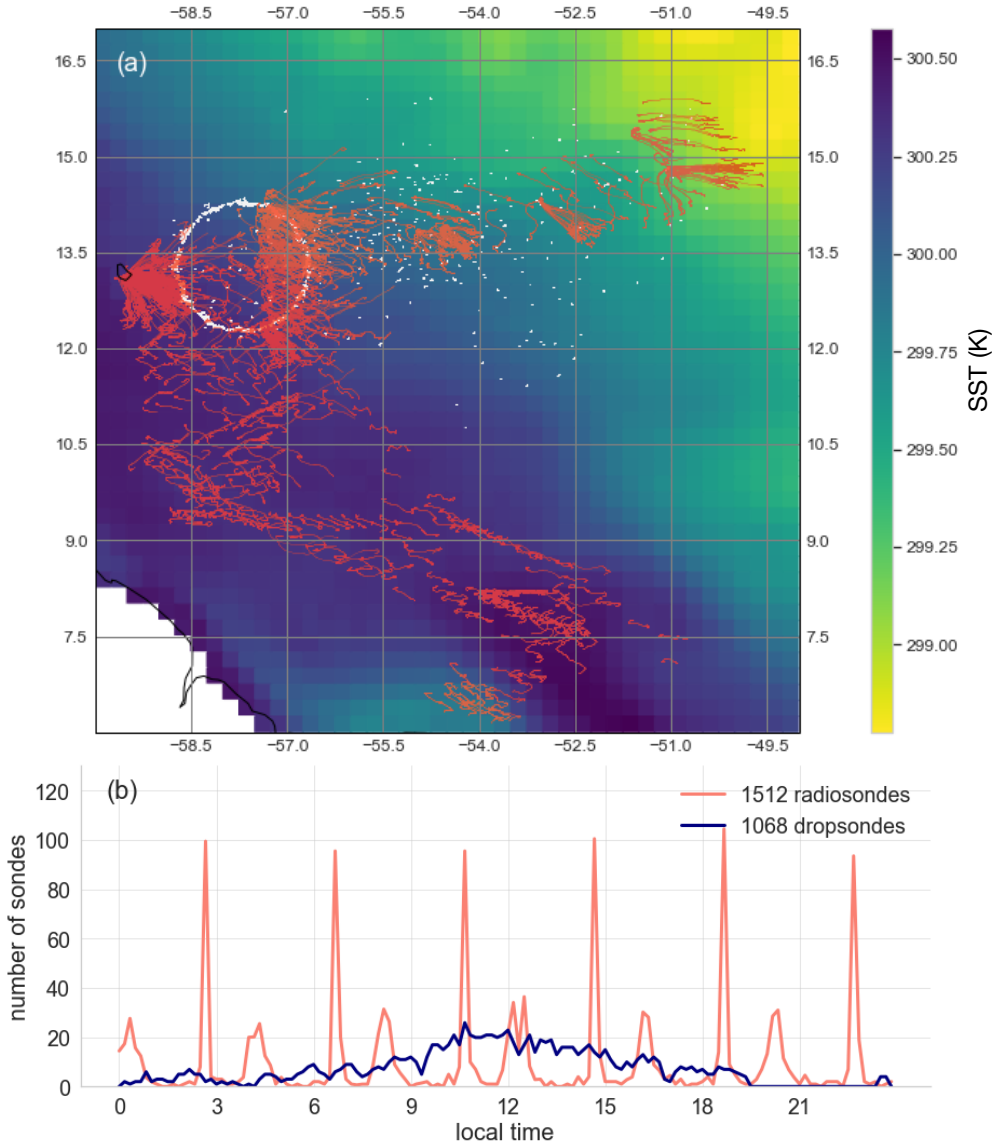


Figure 3-1: (a) The EUREC^{4}A sounding network: 1068 soundings from dropsondes (white) and 1512 from radiosondes (coral). We employ 810 dropsondes launched from HALO and 258 dropsondes from the WP-3D to calculate radiative profiles, as well as 325, 344, 156, 377, and 310 radiosondes launched from L'Atalante, BCO, MS-Merian, Meteor, and RH-Brown, respectively. Background colors show sea surface skin temperature (SST_{skin}) from ERA5 reanalysis at 0.25° resolution averaged over January and February. (b) The diurnal distribution of the 1068 dropsondes (blue) and 1512 radiosondes (coral) with sonde launch time are binned in 10 min intervals.

at 24,000 ft (approximately 7 km), releasing sondes along both linear and circular patterns in the region covered by HALO, as well as further to the east close to the nominal position of the RH-Brown.

Radiosondes were launched every four hours, daily from January 8–February 19, 2020, approximately synchronously from each platform. Given that the time-lag between ascending and descending radiosondes is on the order of hours, and that there is substantial horizontal drift between the ascent and descent, we chose to compute separate radiative profiles for ascending and descending radiosondes. For dropsondes, HALO flight takeoffs were staggered at 5, 8, and 11 am local time, with flights lasting approximately eight hours, yielding roughly 72 sondes per flight. The WP-3D undertook three night flights, which allows for a better characterization of the diurnal cycle, together with the radiosondes launched during the night (Figure 3-1b).

We refer the reader to Stephan et al. (2020) and George et al. (2021) for a complete description of the radiosonde and dropsonde data sets, respectively, and Bony et al. (2017) and Stevens et al. (2021) for an overview of the campaign scientific motivations and measurement strategy.

3.3.2 Radiative transfer calculation

The radiative transfer code used here, RRTMGP (Rapid Radiative Transfer Model for GCMs, Parallel) (Pincus et al., 2019), is a plane-parallel correlated-k two-stream model based on state-of-the-art spectroscopic data for gas and condensate optics. It is based on line parameters from Atmospheric and Environmental Research and the MT_CKD water vapor continuum absorption model (Mlawer et al., 2012). The calculation of radiative profiles from radiosonde and dropsonde data then proceeds in the following way:

1. vertical soundings of temperature, pressure, and water vapor specific humidity at 10 meter resolution are interpolated onto a 1 hPa vertical grid and then merged with temperature and specific humidity from ERA5 reanalyses in the

following manner. Sonde measurements below 40 m are first truncated for all sondes: radiosondes do not provide data in this surface layer because of deck heating effects on ships (Stephan et al., 2020), and we apply the same filter to dropsondes for consistency. The ERA5 profiles at hourly and 0.25° resolution (European Centre for Medium-Range Weather Forecasts, 2017) are linearly interpolated temporally and spatially to the time, latitude, and longitude of the sounding. ERA5 values are used above the highest level measured by each sonde to extend the observed soundings vertically to 0.1hPa and account for the effect of high-altitude thermodynamic variability on the radiative cooling profiles below. To obtain the lower boundary condition, we linearly interpolate the ERA5 sea surface skin temperature (SST_{skin}), also at hourly and 0.25° resolution (European Centre for Medium-Range Weather Forecasts, 2017), to the time, longitude and latitude where the sounding was launched;

2. CO_2 concentrations are set to the present day value of 414 ppm while CH_4, O_3 and N_2O concentrations are taken from the standard tropical atmosphere profile of Garand et al. (2001);

3. the set of resulting profiles is then used as input to RRTMGP to derive upwelling and downwelling clear-sky radiative fluxes in the shortwave and longwave ranges of the spectrum. The calculation uses a spectrally-uniform surface albedo of 0.07 and a spectrally-uniform surface emissivity of 0.98, typical values for tropical oceans.

Dropsondes and radiosondes drift horizontally as they rise and/or fall (Figure 3-1a), which could give slight errors due to aliasing of horizontal moisture variability into vertical variability. This potential error source is less pronounced for dropsondes than radiosondes due to their faster travel speed through the troposphere.

We compute radiative fluxes and heating rates only for the gaseous component of the atmosphere, without explicitly taking into account cloud or aerosol properties. These radiative profiles are therefore clear-sky and aerosol-free. The soundings do, however, capture the water vapor structure, including regions of high humidity in

cloud areas and aerosol layers. Cloud cover in trade-wind regimes is relatively low, between 10% (Nuijens et al., 2015a) and 20% (Medeiros and Nuijens, 2016a) for active clouds, so cloud-free, or clear-sky, profiles are representative of the thermodynamic environment. Taking into account the influence of cloud liquid water would require a number of *ad hoc* assumptions about microphysical and optical properties within clouds (see for instance Guichard et al., 2000). Similarly, we do not directly represent the radiative effect of mineral dust aerosols. The dominant aerosol radiative effect in this region has been shown to result from the covariance of aerosols with water vapor, such that aerosols tend to be associated with elevated moisture layers Gutleben et al. (2019, 2020). Dust aerosol layers are, moreover, more common in the summer than in winter (Lonitz et al., 2015). We leave open the possibility that direct scattering by dust aerosols has an additional role on radiative heating rates, but do not have the coincident data to appropriately address this question for all soundings.

3.4 Preliminary results and discussion

This section includes a first exploration of the data set. We examine radiative variability at different scales – across all soundings, at the diurnal timescale, and according to different patterns of mesoscale organization – as well as in individual profiles showing the influence of sharp vertical moisture gradients on radiative heating rates.

3.4.1 Variability across soundings

A distribution of longwave, shortwave, and net heating rates, as well as large-scale thermodynamic quantities, are shown in Fig. 3-2. Local extrema in the median shortwave, longwave, and net heating rates occur near 2 km (Fig. 3-2d,e,f), associated with the rapid decrease in specific and relative humidity at this level (Fig. 3-2b,c). The top of the planetary boundary layer, or interface between the moist marine boundary layer and dry free troposphere above, is expected to occur around 2km in the trades (Malkus, 1958; Cao et al., 2007; Stevens et al., 2017). The spread in specific and relative humidity is greater than that in temperature, suggesting a strong

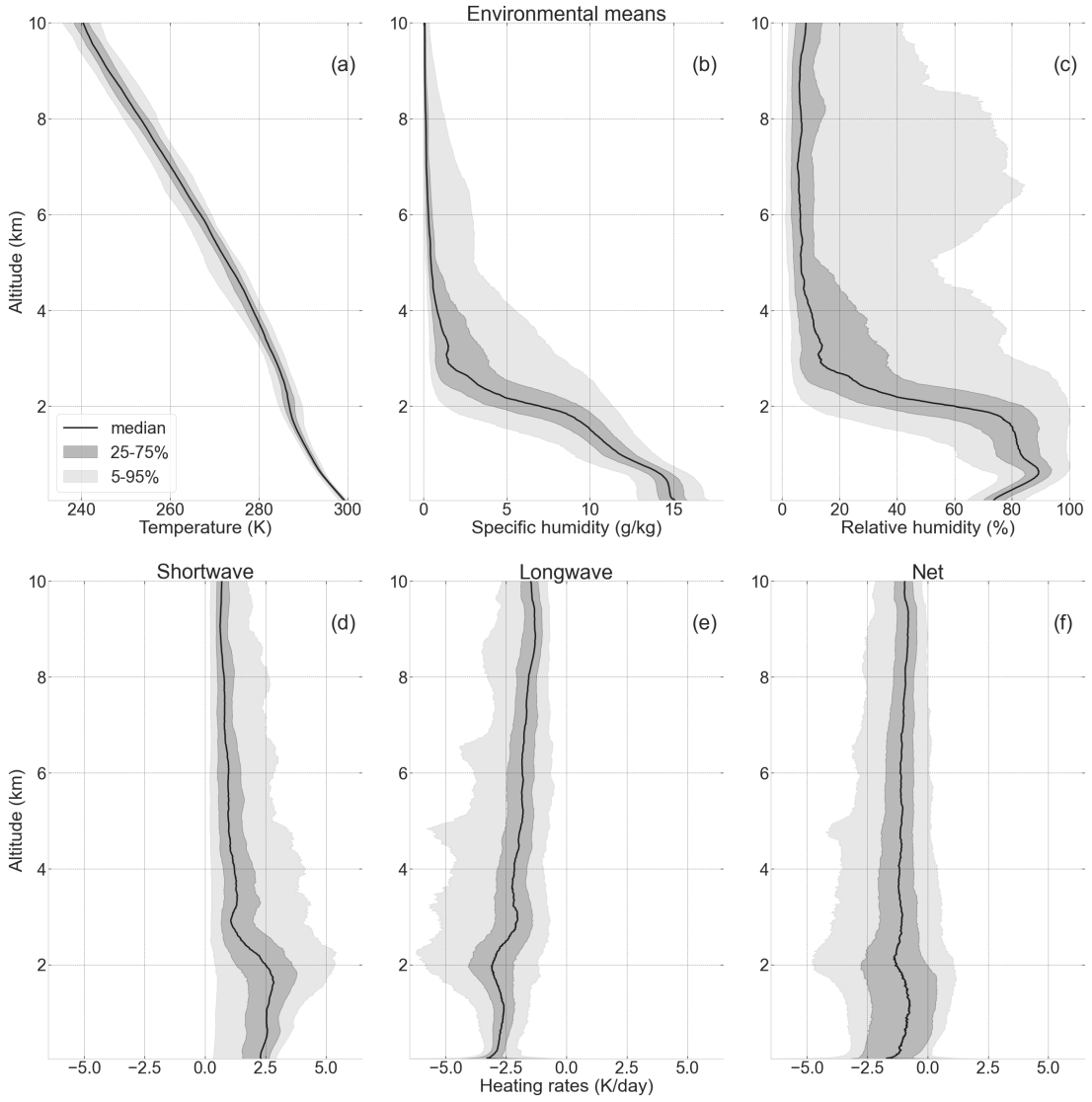


Figure 3-2: Top: Temperature (a), specific humidity (b) and relative humidity (c) (with respect to ice for $T < 0°C$) from EUREC^{4}A dropsonde and radiosonde data. Bottom: Shortwave (d), longwave (e) and net (f) heating rates calculated from EUREC^{4}A dropsonde and radiosonde data using the radiative transfer code RRTMGP. The center traces are the median profiles, and the medium and light grey shadings indicate the 25–75% and 5–95% intervals, respectively. For the shortwave, the median and the interquartile range are calculated using daytime values only.

role for moisture variability on the variability in radiative heating rates. On average, longwave cooling is stronger than shortwave heating, such that net heating rates are largely negative from the surface up to 10 km, with a median value around -1 K/day. Additional local minima in longwave heating are observed around 3 and 5 km between the 5% and 25% quantiles. These local minima could, for instance, correspond to the radiative effect of elevated moisture layers arising from convection detraining moisture at these higher levels, albeit less frequently, or aerosol layers associated with increased water vapor concentrations (Stevens et al., 2017; Wood et al., 2018a,b; O et al., 2018; Gutleben et al., 2019).

We next partition radiative heating variability into its variability in time (e.g. diurnal cycle, day-to-day variability) and regarding the spatial characteristics of the convection field (e.g. the spatial distribution of clear and cloudy regions).

3.4.2 Diurnal cycle and day-to-day variability

Figure 3-3 gives an overview of the diurnal variability of radiative heating, which has been implicated in the diurnal cycle of convection and cloudiness (e.g., Gray and Jacobson Jr, 1977; Randall and Tjemkes, 1991; Ruppert and Johnson, 2016). Shortwave radiative heating follows the solar cycle. Longwave heating rates show less diurnal variability and have approximately the same amplitude (with an opposite sign) as shortwave heating rates during daytime. This compensation between longwave cooling and shortwave heating results in a daytime net heating rate that is slightly positive in the lower 2km. The daytime heating contributes to stabilizing the lower atmosphere, disfavoring convection. At night, strong radiative cooling destabilizes the lower troposphere and strengthens convection. The maximum nighttime longwave cooling occurs slightly above 2 km, with secondary cooling peaks occurring around 4 and 6km. During daytime, the peak in stabilizing radiative heating appears slightly below 2km. This difference in the height of peak radiative heating, albeit of different sign, could reflect differences in the height of the moist, convecting layer over the diurnal cycle: a shallower marine boundary layer during the day that deepens at night (Vial et al., 2019). These considerations highlight the potential for subtle interactions

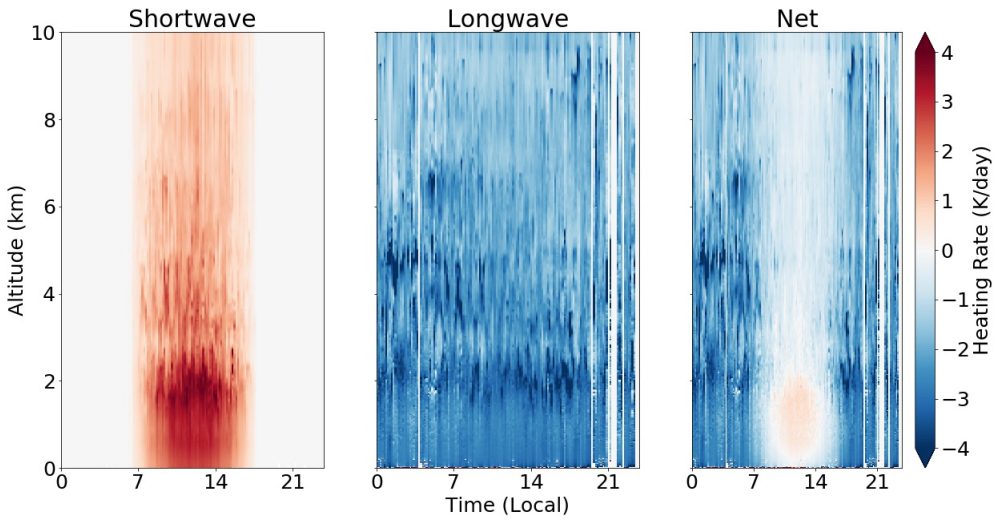


Figure 3-3: Diurnal composite of shortwave (left), longwave (middle), and net (right) clear-sky heating rates binned in 10-minute intervals. Colored shadings indicate heating rates in units of K/day. The data are plotted with respect to local solar time to simplify interpretation of the diurnal cycle. White indicates the absence of data. We note that some variability, such as in the nighttime longwave radiative cooling variability, could result from different numbers of sondes launched throughout the diurnal cycle (as illustrated in Fig. 3-1b).

among radiation, convection, and cloudiness on the diurnal timescale.

Fig. 3-4 shows the day-to-day evolution of the shortwave (top), longwave (middle) and net (bottom) heating rates derived from radiosondes launched at BCO. In the shortwave and net heating rates, the daily stripes are due to zero shortwave heating during the night. In the longwave component alone, the amplitude of the diurnal cycle is less evident. Regarding the day-to-day variability, both in the shortwave and the longwave components, trends in the height-evolution of the radiative heating maxima appear to persist over several days. These trends are likely due to variations in humidity (e.g. Dopplick, 1972; Jeevanjee and Fueglistaler, 2020) and are consistent with the presence of multi-day trends in moisture observed at BCO during the campaign (see Figure 13 in Stevens et al., 2021). At the end of the campaign, the rise in the peak of longwave cooling appears to correspond to the rising location of the interface between the moist, convecting layer below and dry free troposphere above (not shown). The persistence and evolution of radiative heating patterns could be tied to larger-scale synoptic moisture activity or to the evolution of mesoscale organization patterns.

3.4.3 Radiative signatures of mesoscale patterns of cloud organization

We next aggregate radiative heating rates spatially. Fig. 3-5 illustrates four representative cases of the Fish-Gravel-Flower-Sugar classification established previously for mesoscale (20-2,000km) organization patterns of clear and cloudy regions (Bony et al., 2020; Stevens et al., 2020b). These cloud organization patterns were identified visually from satellite imagery and correspond to differences in large-scale environmental conditions (Bony et al., 2020). They are also observed to have different top-of-the-atmosphere radiative effects (Bony et al., 2020). As outlined in Stevens et al. (2020b), Sugar refers to a 'dusting' of small, shallow clouds with low reflectivity and a random spatial distribution. Gravel clouds tend to be deeper than Sugar (up to 3-4km), have little stratiform cloudiness, precipitate, and organize along apparent gust fronts or

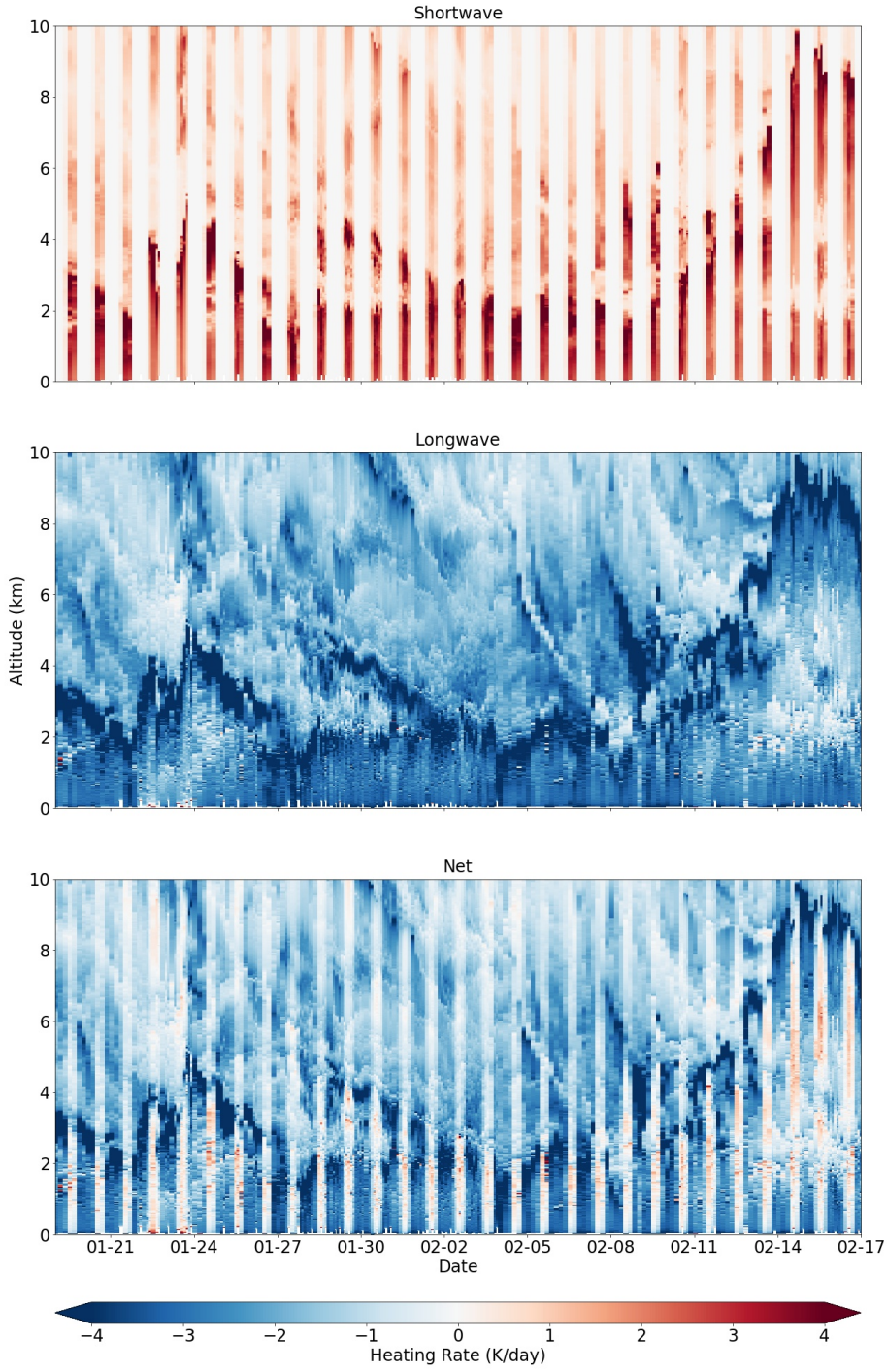


Figure 3-4: Shortwave (top), longwave (middle), and net (bottom) heating rates at BCO during EUREC^{4}A, from January 19 to February 17. The heating rates are calculated from radiosondes launched at BCO. In colors are heating rates with units of K/day. White indicates the absence of data.

cold pools at the 20-200km scales. Fish are skeletal networks (often fishbone-like) of clouds at the 200-2,000km scale with stratiform cloud layers; the Fish pattern is often associated with extratropical intrusions. Flowers are circular features defined by their stratiform cloud elements. Both Fish and Flowers are surrounded by large swaths of clear air.

We choose four days as an example of the large-scale environmental and radiative signature of each pattern, given the spatial pattern observable in the GOES-16 satellite images in the HALO flight path shown by the white circle. We plot daily-mean profiles for temperature, specific humidity, and relative humidity (Fig. 3-5a,b,c), as well as shortwave, longwave, and net radiative heating rates (Fig. 3-5d,e,f). These profiles were calculated from approximately 70 HALO dropsondes launched during the eight-hour flight on each day. We also plot the standard deviation of radiative heating for each flight (Fig. 3-5g,h,i). As a first approximation, the standard deviation of daily radiative heating profiles acts as a proxy for spatial variability in radiative heating rates.

Spatial variability in radiative heating has been shown to drive shallow circulations (e.g. Naumann et al., 2019) and affect convective organization (e.g. Bretherton et al., 2005; Muller and Held, 2012). In this illustrative example, the differences in the mean and standard deviation of the radiative heating rates hint at a role for differences in radiative cooling rates in the onset or maintenance of mesoscale patterns of organization. For instance, the 'Fish' pattern on January 22, 2020 is associated with a moister lower troposphere between 1 and 3km and slightly drier free troposphere above 4km. This vertical moisture distribution may give rise to the observed vertical variability in radiative heating rates, with larger peaks in the mean profile (Fig. 3-5e) and standard deviation (Fig. 3-5h) in radiative heating observable between 2 and 4km, likely corresponding to strong humidity gradients at these levels.

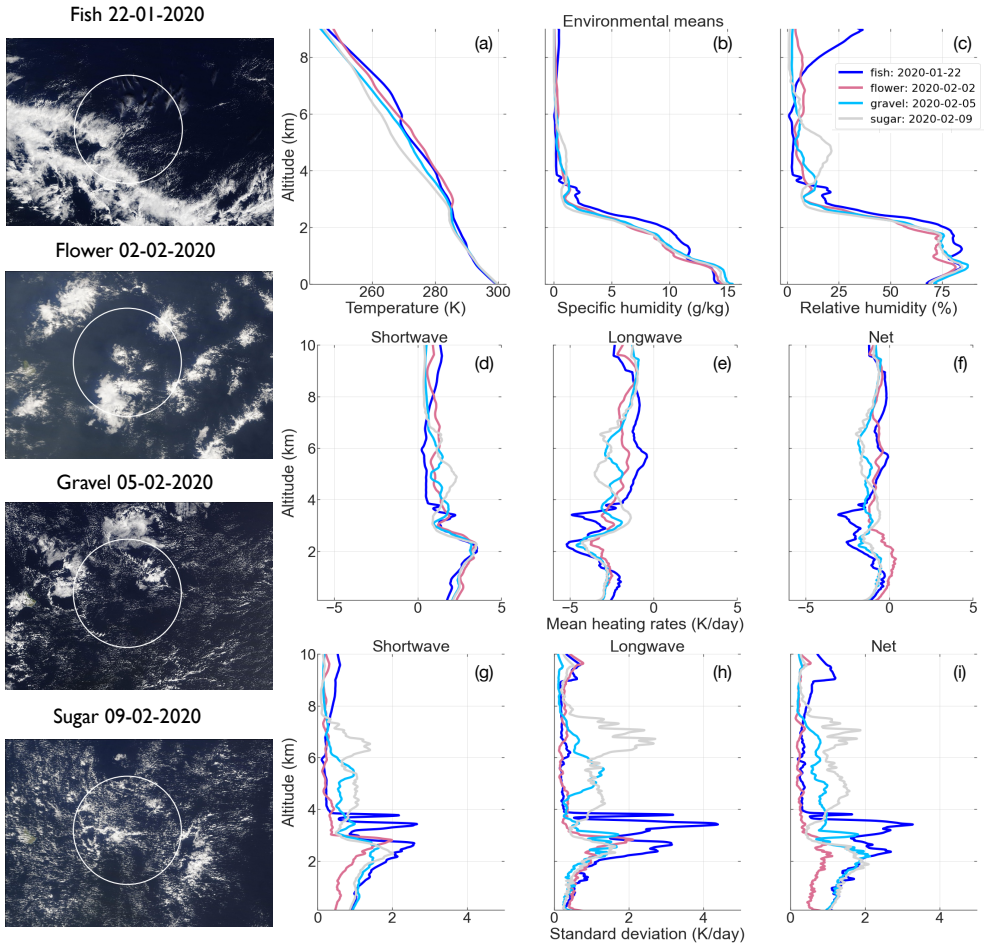


Figure 3-5: Thermodynamic (a-c), daily mean radiative heating (d-f), and daily standard deviation of radiative heating (g-i) profiles classified by mesoscale organization pattern, using a characteristic example of each type as diagnosed from snapshots from GOES-16 infrared channel (left column). This figure employs HALO dropsondes launched in the circular flight pattern (shown by the white circle) on the chosen day, corresponding to roughly 70 dropsondes each. We focus on the spatial extent of the HALO flight pattern because the cloud organization pattern does not necessarily extend across the entire sampling domain Figure 3-1a, nor have the patterns been shown to be scale-invariant.

3.4.4 Effect of sharp moisture gradients on radiative heating profiles

Figure 3-6 highlights the radiative signatures of elevated moisture layers, which can persist for multiple hours at inversion levels (Stevens et al., 2017; Wood et al., 2018a; Gutleben et al., 2019). We focus in detail on two thermodynamic and radiative heating profiles of a particular elevated moisture layer extending to 4 kilometers, alongside GOES-16 images (Fig. 3-6i,j) corresponding to these soundings. This structure persisted for at least four hours on January 24, 2020, and we plot thermodynamic conditions and radiative heating profiles sampled three hours apart, at 12:55 and 15:55 UTC (see Fig. 3-6). A striking feature is the sharp peak in longwave cooling at the top of the moisture layer of nearly -20 K/day at 15:55 UTC, corresponding to the strong humidity gradient, with relative humidity decreasing by nearly 70% in 100 meters (Fig. 3-6c,d).

Although we calculate clear-sky profiles only, the present work could be extended to account for radiative effect of cloud liquid water, which could be used, for instance, to investigate the radiative effect of geometrically- and optically-thin 'veil clouds' persisting at inversion levels (Wood et al., 2018a,b; O et al., 2018), such as those illustrated by the flight photographs (Fig. 3-6a,e). Over global oceans, approximately half of low clouds do not fully attenuate space-borne lidar, suggesting that these optically-thin clouds contribute significantly to total cloud cover estimates (Leahy et al., 2012a) and could have an important radiative impact (e.g., Wood et al., 2018b).

3.5 Uncertainty assessment

To evaluate the robustness of our results and ensure good use of this data set, we performed several uncertainty assessments by perturbing the $\mathrm{SST_{skin}}$, *in situ* moisture data, and ERA profiles used. We also included in the data set the minimum and maximum levels z_{min} and z_{max} measured by each sonde. Unless indicated otherwise, the errors reported below correspond to a subset of profiles with valid data starting

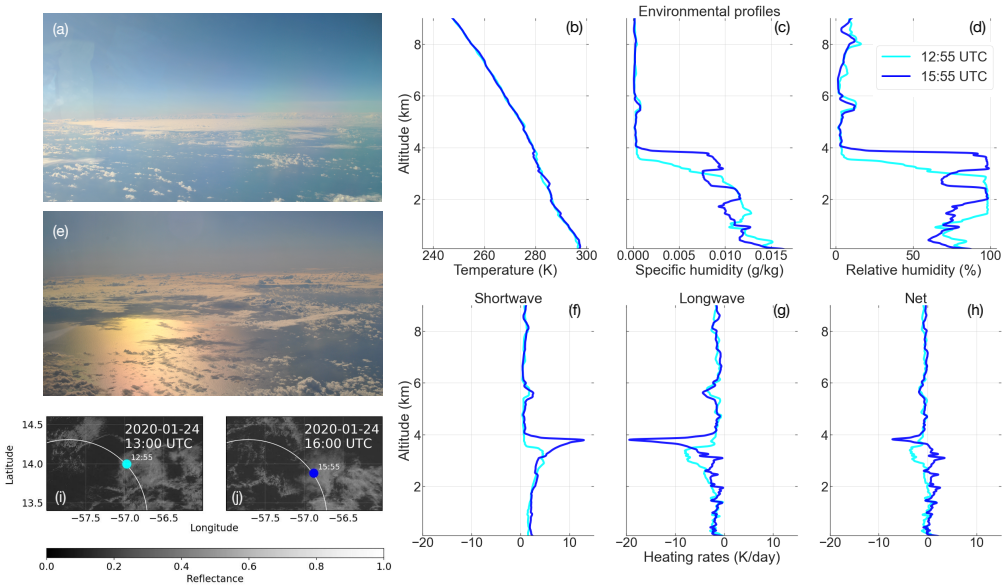


Figure 3-6: Thermodynamic and radiative heating profiles associated with an elevated moisture layer persisting for multiple hours on January 24, 2020 in the HALO flight pattern. Plotted here are the temperature (b), specific humidity (c), relative humidity (d), as well as shortwave, longwave, and net radiative heating rate (f-h) profiles for two soundings sampled three hours apart, at 12:55 and 15:55 UTC. Alongside these profiles are photographs (a,b) taken from the HALO aircraft during the flight and GOES-16 satellite images (i,j), with the dropsonde location and launch time indicated by a circle along the circular flight pattern. Credit for the two flight photographs: J. Vial.

at 40 m (ie. $z_{min} \leq 40$ m) and during daytime, which corresponds to 1314 profiles. The daytime filter was required for relevant calculation of the error in the shortwave, and then kept for consistency for the longwave, but the magnitude of errors in the longwave is not affected by this filter (not shown).

We first test the sensitivity to the ERA5 SST_{skin}. To this end, we perturbed the original SST_{skin} by ± 0.42 K and recalculate all heating rates. This value is chosen as it corresponds to the root-mean-square-error (RMSE) between between ERA5 SST_{skin} and Marine-Atmosphere Emitted Radiance Interferometers (M-AERI) measures taken during a series of cruises in the Carribbean Sea from 2014 to 2019 (Luo and Minnett, 2020). Figure 3-7 shows the RMSE between the original and perturbed radiative profiles (blue curves). In the longwave and net, the effect of the perturbation is strong in the first atmospheric layer, but then decreases rapidly and becomes negligible after a few hundred meters. Except for the first few atmospheric layers, the uncertainty around the SST_{skin} can therefore be safely neglected.

We then investigate the sensitivity to the uncertainty of sounding measurements by perturbing all soundings by a vertically-uniform relative error and redoing all radiative transfer calculations. The manufacturer predicts an uncertainty of ± 0.1 K for the temperature and $\pm 3\%$ for specific humidity (Vaisala, 2018). The temperature uncertainty has virtually no effect on radiative profiles (not shown). The effect of $\pm 3\%$ uncertainty on the specific humidity profiles is shown in Fig. 3-7 in red. The highest RMSE for this specific humidity perturbation occurs in the cloud layer, between 800 m and 2 km, with a magnitude of 0.05 K/day for net radiative heating. A secondary peak with a magnitude of 0.03 K/day is also evident near the inversion, at about 3 km. Given a median radiative heating value of -1 K/day throughout the lower troposphere (Sec. 3.4.1), these errors are roughly 3-5% for the net radiative heating. These maxima likely correspond to the cumulative errors at the altitude of large vertical humidity gradients, which lead to peaks in longwave, and to a lesser extent shortwave heating rates for individual profiles.

Finally, we explore the uncertainty associated with ERA5 temperature and humidity data employed as an upper boundary condition. Similarly to the uncertainty

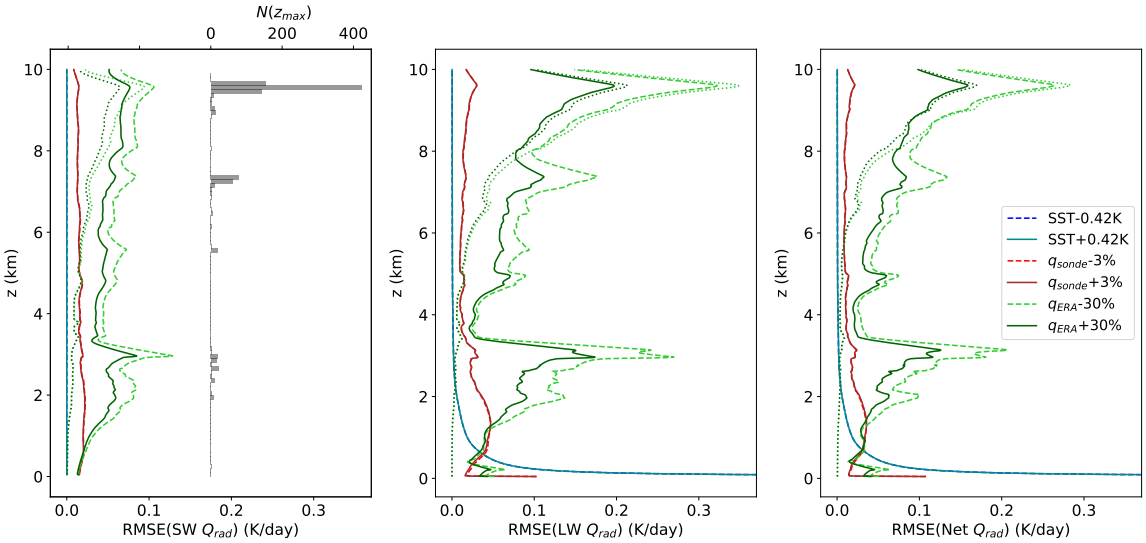


Figure 3-7: Root-mean-square error (RMSE) estimates in shortwave (left), longwave (center) and net heating rates (right) for perturbations in $\mathrm{SST_{skin}}$ (blue), ERA5 humidity profiles (green) and sonde humidity measurements (red) for the 1314 daytime profiles that have valid data starting at 40 m. Dashed curves show negative perturbations, solid curves show positive perturbations and dotted green curves show ERA5 humidity perturbations restrained to the 1117 daytime profiles that have valid data at all levels between 40 m and 8 km. The horizontal grey bars on the left panel show the frequency distribution in the maximum level measured (z_{max}).

analysis for the sounding data, we perturb ERA5 3D fields — used as input to the radiative transfer code — by a uniform relative error. Previous studies have shown that ERA5 reanalyses can present biases of various kinds (Nagarajan and Aiyyer, 2004; Dyroff et al., 2015). We compare ERA5 humidity and temperature data with coincident radiosonde measures to obtain an estimate of ERA5 biases up to 100 hPa. From the surface to 100 hPa, the RMSE in temperature between co-located radiosonde soundings and ERA5 is between 0.3 and 0.7 K, with a mean of 0.5 K, and between 5% (at the surface) and 70% (near the inversion) for the specific humidity, with a mean around 30%.

Fig. 3-7 only shows the effect of the ERA5 specific humidity uncertainty, taken at ±30%, on radiative profiles, as the temperature has once again a negligible influence. The corresponding green curves (respectively dashed and solid) reveal local maxima in the longwave and net heating rates around 3, 7 and 9.5 km. Again given a median radiative heating value of -1K/day throughout the lower troposphere (Sec. 3.4.1), the errors at these local peaks are between 10–30%. These maxima coincide with the modes in the frequency distribution of the highest level z_{max} measured by the soundings, indicated in grey in the left panel. These peaks suggest that the uncertainty arises from the large discontinuities emerging at the ERA5-sounding junction level when perturbing ERA5 humidity profiles. The results suggest that the corresponding uncertainty mainly occurs in the vicinity of the junction levels. This notion is further confirmed by calculating the RMSE only on profiles which have data between 40 m and 8 km (ie. $z_{min} \leq 40$ m and $z_{max} \geq 8$ km, dotted green curve): the remaining 1117 profiles left do not contain vertical discontinuities in humidity in this range, and we see that the remaining upper-tropospheric discontinuities do not affect heating rates in the lowest troposphere.

Overall, the small uncertainty values given with these tests support the robustness of this data set and gives confidence regarding its use for more detailed investigations in the lower troposphere. The uncertainty from sea surface skin temperature is limited to the first few atmosphere layers, and uncertainty from merging with ERA5 specific humidity is largely contained to the sounding-reanalysis junction point. Uncertainty

associated with observed specific humidity profiles produces localized errors in the cloud and inversion layers below 3km, though these errors are approximately 5% or less. We recommend that users carefully compare the magnitude of the signal they analyze with the magnitudes of the errors provided here.

3.6 Conclusions

The first objective of this work is to present the method used to calculate clear-sky, aerosol-free radiative profiles from 2580 radiosonde and dropsonde soundings launched during the EUREC^{4}A field campaign. These radiative profiles are calculated using a state-of-the-art correlated-k model, RRTMGP, in which ERA5 reanalyses provide lower and upper boundary conditions. We then aggregate the radiative heating profiles at multiple scales to examine temporal and spatial variability in trade wind regimes. We find that radiative heating rates in the wintertime trade-wind environment display significant diurnal and day-to-day variability, and we observe hints that this variability may be associated with different types of mesoscale organization. An uncertainty assessment is further conducted to demonstrate that the influence of uncertainties in the sounding data, and upper and lower boundary conditions, is small relative to the magnitude of estimated radiative heating.

These results present a first overview of how this data set could help answer existing research questions, in particular: 1) What is the role played by radiation in the mesoscale organization of shallow convection? (e.g., Seifert and Heus, 2013; Bretherton and Blossey, 2017b) 2) what is the interplay between the diurnal variability in radiative heating, convection, and cloudiness? (e.g., Gray and Jacobson Jr, 1977; Ruppert Jr and O'Neill, 2019; Vial et al., 2019), and 3) what is the influence of clear-sky radiative cooling gradients on atmospheric circulations? (e.g., Gray and Jacobson, 1977; Mapes, 2001; Emanuel et al., 2014; Thompson et al., 2017; Naumann et al., 2019). Such questions regarding the coupling of clouds, convection, and circulations in trade-wind regimes are at the heart of the EUREC^{4}A field campaign, and the radiative profiles presented here complement other EUREC^{4}A observations and

data products in forming a toolbox for these investigations.

Chapter 4

Observed subcloud layer moisture and energy budgets in the trades

4.1 Abstract

The trade-wind subcloud layer is a structural component of the atmosphere. Its thermodynamic variability has long been characterized using simple frameworks, of which mixed layer theory is the simplest kind. Past observations qualitatively support such a description, yet the adequacy of mixed layer theory as a quantitative description has not been tested. Here we use observations collected during the EUREC^{4}A (*Elucidating the role of clouds–circulation coupling in climate*) field campaign to test this framework and evaluate our understanding of the trade-wind subcloud layer. We find evidence for a transition layer separating the mixed layer and subcloud layer tops. The presence of such a finitely-thick transition layer complicates the application of mixed layer theory with its assumptions of no vertical gradients and an infinitesimally-thin 'jump' at the subcloud layer top. This ambiguity introduces effective parameters and motivates their estimation through a Bayesian inversion. Results from this joint inversion further reflect a nonzero depth of entrainment mixing. We find that subcloud layer moisture and energy budgets close for both synoptic variability and a monthly campaign-mean, yielding a campaign-mean residual of 3.6 Wm^{-2} for moisture and and 2.9 Wm^{-2} for energy. Surface wind speed variability influences the subcloud layer depth and fluxes, yet thermodynamic variability above the subcloud layer top

This chapter is under review in the *Journal of the Atmospheric Sciences*.

emerges as the primary control on subcloud layer moisture and temperature variability. Given that mixed layer theory offers a closed description for observations, it offers an appealing framework for evaluating larger-scale models that must parameterize the processes regulating this important part of the atmosphere.

4.2 Introduction

The trade-wind subcloud layer is an important component of the tropical atmosphere. Typically defined as extending from the top of the surface layer (at approximately 50 m) to cloud base (at approximately 700 m) (e.g., Malkus, 1958; Stevens et al., 2017), it couples the surface to the trade-wind cloud layer and, in so doing, regulates the import of energy and moisture from the ocean (e.g., Malkus, 1958; LeMone and Pennell, 1976; Stevens, 2007). Clouds in the trades are rooted in the subcloud layer, as subcloud moisture variability controls moist static energy variability, which influences convective potential and cloudiness (e.g., Emanuel, 1986, 1993). Clouds, however, also influence the subcloud layer, as they mix down dry air from aloft, making it available to be entrained into the subcloud layer (Stevens, 2006). This subtle interplay between mixing processes in the subcloud layer and their interaction with the clouds above governs the magnitude of latent heat export from the trades to the equator (e.g., Malkus, 1958). The trade-wind subcloud layer thus forms an important link in the global thermodynamic budget, transporting latent heat to the equatorial belt where it influences large-scale circulations and the global hydrological cycle (Riehl, 1954; Heckley, 1985; Tiedtke, 1989).

The clouds of the trades provide an additional motivation to study the trade-wind subcloud layer. By virtue of their large spatial extent and thus statistical weight, trade-wind cloud regimes have a large influence on the global energy budget and global dynamics (e.g., Bony et al., 2004). Diverging cloud feedbacks in the trades were also shown to explain differences in climate sensitivity estimates (e.g., Bony and Dufresne, 2005; Webb et al., 2006; Vial et al., 2013), some of which has been shown to relate to how efficiently moisture is exported out of the subcloud layer (Sherwood

et al., 2014).

Given the importance of the trade-wind subcloud layer, it is useful to understand what controls its properties. To aid this understanding, the subcloud layer has long been characterized using simple frameworks. Strong buoyancy- and, to a lesser extent, wind shear-driven turbulence homogenizes subcloud layer thermodynamic variables in the vertical (e.g., Mahrt, 1976; Stull, 2012). Such a well-mixed vertical structure allows for simplification by solving for the vertical integrals, or bulk properties, of the boundary layer. Among these the vertically-integrated models, or bulk models, the mixed layer model is the simplest case. Bulk models have formed the basis of parameterizations in larger-scale models (Arakawa and Schubert, 1974; Deardorff, 1972). Mixed layer models are a special type of bulk model that assume that the vertical structure in conserved quantities is not important for their dynamics. This simplification allows the evolution of the layer to be represented by simply tracking its mean mass and energy budgets, usually as represented by the mean subcloud layer specific humidity and potential temperature (e.g., Betts, 1976; Stevens, 2006). Mixed layer models allow for assessing the relative magnitude of different processes and provide a quantitative map between subcloud properties and their environmental controls.

The adequacy of this mixed layer description of the subcloud layer has, however, only been assessed from relatively few measurements and large-eddy simulations often performed for idealized conditions over small (and usually homogeneous) domains. Past observational studies typically used approximately 100 dropsondes and fixed certain parameters, such as the sea surface temperature and vertical thermodynamic structure above the subcloud layer (Betts, 1976; Betts and Ridgway, 1989). Augmenting such analyses with passive remote sensing remains challenging (e.g., Kalmus et al., 2014). Indeed, subcloud layer moisture variability is poorly quantified by passive remote sensing because observed moisture profile features, in particular vertical gradients at the top of the boundary layer, are at scales much finer than the typical weighting functions of even hyperspectral instruments (e.g., Maddy and Barnet, 2008; Stevens et al., 2017; Pincus et al., 2017). Mixed layer models have been used

effectively to interpret large-eddy simulations (e.g., Neggers et al., 2006; Bellon and Stevens, 2012, 2013; Schalkwijk et al., 2013). The apparent realism of large-eddy simulation can suggest a trustworthy representation of nature, yet these simulations use idealized boundary conditions, employ a range of domain sizes and resolutions, and rarely represent the diversity of mesoscale patterns of convection observed in nature, all of which can influence the resultant moisture fields (e.g., Bony et al., 2017). These limitations render *in situ* observations especially important for evaluating the adequacy of the mixed layer model in encapsulating the main modes of subcloud layer variability and ultimately attributing it to changes in the trade-wind environment.

During the EUREC^{4}A (*Elucidating the role of clouds–circulation coupling in climate*) field campaign, the most comprehensive field study ever performed in the trades (Bony et al., 2017; Stevens et al., 2021), we collected the necessary data to investigate the structure and variability of the subcloud layer. These data allow us to test the assumptions of mixed layer theory as a closed description of the trade-wind subcloud layer. To the extent that mixed layer theory provides an adequate description, it allows us to go a step further and quantify which large-scale environmental conditions influence variability in subcloud layer properties.

Section 5.2 presents the EUREC^{4}A observations, and Section 4.4 defines subcloud layer variability first in terms of five parameters — its height, as well as its means and gradients in specific humidity and potential temperature – and assesses how each parameter contributes to observed day-to-day variability. Section 4 introduces the mixed layer theory framework for assessing the importance of the subcloud layer structure in shaping its bulk variability. To do so we introduce a Bayesian methodology to jointly constrain uncertain parameters related to entrainment. Section 5 then evaluates whether mixed layer theory can explain observed synoptic variability and the monthly campaign-mean balance for moisture and energy. Section 6 examines the relationships among subcloud layer properties and large-scale meteorological conditions.

4.3 EUREC^{4}A field campaign data

EUREC^{4}A field study measurements were made in January and February 2020 in the downstream trades of the North Atlantic, anchored in Barbados. In both models and observations, clouds and the large-scale environment around Barbados were found to be representative of the Atlantic and Pacific trades (Medeiros and Nuijens, 2016b; Rasp et al., 2020), suggesting that many inferences from EUREC^{4}A observations are informative of the global trade-wind regimes as a whole. EUREC^{4}A sets itself apart from previous field campaigns, as, for instance, reviewed by Garstang et al. (2019), through the intense and unbiased sampling; the use of novel observing strategies and exploiting improved and novel instrumentation; and the coincidence of satellite measurements with very high spatial resolution and temporal sampling (Bony et al., 2017; Stevens et al., 2021).

Figure 4-1 shows the geographic distribution of the measurements used in this study. Our core data are 810 dropsondes from the German High Altitude and Long Range Research Aircraft (HALO) launched between January 22, 2020 and February 15, 2020 (Konow et al., 2021). These dropsondes yield vertical profiles of pressure, temperature, and relative humidity with a manufacturer-stated accuracy of 0.4 hPa, 0.1°C, and 2%, respectively. We employ level-3 and level-4 dropsonde data, which have been processed and interpolated into a common altitude grid with 10 m vertical resolution (George et al., 2021). We note that George et al. (2021) identify a dry bias in the HALO dropsondes, which they correct with a multiplicative factor of 1.06 applied to relative humidity and all associated moisture quantities from these sondes. We use these corrected data.

During EUREC^{4}A, dropsonde measurements were distributed along the 'EUREC^{4}A circle'. The EUREC^{4}A circle is defined by a circular flight pattern with an approximately 220-kilometer diameter, centered at 13.3°N, 57.7°W (Fig 4-1). This flight pattern was repeated 69 times, over 12 flights. Typically each flight incorporated two – temporally well separated – periods of circling. A 'circling-mean' is defined as the mean of three 'circle-means' that each average about 12 dropsondes along

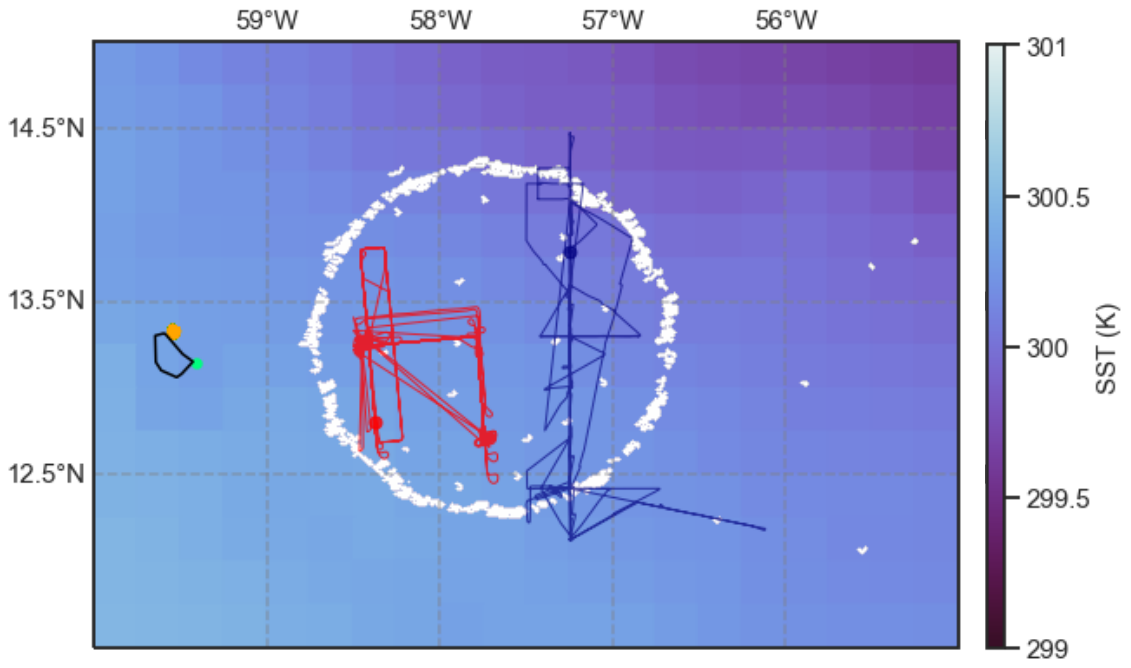


Figure 4-1: Data employed in this study include dropsondes launched in the EUREC^{4}A circle (white; 810 HALO dropsonde soundings), subcloud layer thermodynamic measurements from the ATR-42 aircraft (red tracks), and sea surface temperature and surface flux measurements from the R/V Meteor (navy tracks). We also use data from the Barbados Cloud Observatory (green) and subcloud layer thermodynamic measurements from the remotely-piloted aircraft CU-RAAVEN (orange). Background sea surface temperatures are ERA5 data 0.25° resolution averaged over January and February 2020.

the EUREC^{4}A circle (Table 4.1). Given that measurements did not target specific meteorological conditions they provide unbiased sampling of the trade-wind layer. The structure of the data collected encourages the definition of 69 circle-means, 24 circling-means, and one campaign-mean (Table 4.1).

The French ATR-42 aircraft made additional thermodynamic measurements during 18 flights from January 26 to February 13, 2020. The ATR-42 flew coincident rectangular patterns inside the EUREC^{4}A circle (see Fig. 4-1), which show good agreement with HALO measurements (Bony et al, in prep, Fig. 15). We also use thermodynamic measurements between January 24 to February 15, 2020 below one kilometer from a remotely-piloted aircraft CU-RAAVEN (**?**), as well as eddy covariance surface sensible and latent heat fluxes from the R/V Meteor (Schirmacher et al,

Term	Definition
EUREC^{4}A circle-mean	One circle-mean refers to the mean of 7–12 dropsondes launched over one hour by the HALO aircraft along a circular, 220-kilometer airborne sounding array centered at 13.3°N, 57.7°W at 10.5 km altitude. A dropsonde is launched for every 30 degree change in heading. The spatial scale of the EUREC^{4}A circle characterizes the large-scale environment and corresponds to the size of a typical general circulation model gridbox, or what Orlanski (1975) called the meso-β scale (20-200km).
circling-mean	One circling-mean is defined as the mean of three consecutive circle-means, corresponding to 30–36 consecutive soundings aggregated over 210 minutes. The start time of successive circles was offset by fifteen minutes. The circling-mean scale represents synoptic variability of the large-scale environment.
flight-mean	The mean of all dropsondes launched along the EUREC^{4}A circle during one of eleven research flights of the HALO aircraft for which circles were performed. Each flight lasted approximately eight hours.
campaign-mean	The mean of 810 dropsondes launched from the HALO aircraft between January 22, 2020 and February 15, 2020, or approximately a monthly-mean, spanning the majority of the EUREC^{4}A campaign.

Table 4.1: Defining terminology to describe the sampling strategy in EUREC^{4}A used in this analysis and described in Sec. 5.2. This terminology follows the terms introduced in Stevens et al. (2021). We refer the reader to Konow et al. (2021) for additional information about data from HALO flights.

in prep). Sea surface temperatures are from the R/V Meteor, with these values extrapolated from the R/V Meteor location to the respective dropsonde location based on fixed zonal and meridional sea surface temperature gradients of -0.14 K degree^{-1} of latitude or longitude. These gradients are estimated from two satellite products (GOES-16 ABI and CLS), and ECMWF Reanalysis (ERA5) reanalysis (Hersbach et al., 2020), which agree well over the same spatiotemporal domain. We subtract 0.25 K from Meteor sea surface temperatures measured at few-meter depth to account for the 'cool skin' surface effect (e.g., Fairall et al., 2003). To a lesser extent and solely for purposes of comparison with observations, we use ERA5 reanalysis at 0.25° spatial and hourly temporal resolution for January and February 2020 for

surface sensible and latent heat fluxes and vertically-resolved specific humidity and potential temperature.

4.4 Describing the subcloud layer structure and its variability

To conceptualize the subcloud layer we first consider a representation in terms of five scalar variables: height or depth (h), mean potential temperature ($\overline{\theta}$), mean specific humidity ($\overline{q}$), the potential temperature vertical gradient ($\partial\theta/\partial z$), and the specific humidity vertical gradient ($\partial q/\partial z$), which we refer to jointly as $X_{\mathrm{SCL}} = \{h, \overline{\theta}, \overline{q}, \partial\theta/\partial z, \partial q/\partial z\}$.

4.4.1 Defining the subcloud layer height

We first ask to what extent the subcloud layer height, h, can be defined from thermodynamic profiles. Despite its role as a key vertical length scale, there is no consensus on how to define this height (e.g., Seibert et al., 2000). Different methods applied to a single data set have been shown to yield a wide range of heights (e.g., Liu and Liang, 2010; Beyrich and Leps, 2012; Dai et al., 2014), leading to ambiguity in the basic question of the depth of the trade-wind subcloud layer.

To estimate the subcloud layer depth, we use three subcloud height definitions. The methodology for applying each definition to the data is given in Appendix A, and Table 4.2 summarizes the results. The first method estimates the depth over which there is no vertical gradient in a conserved variable within a threshold (Canut et al., 2012). This 'gradient method' selects the height where a thermodynamic variable exceeds its mean, averaged over the levels below, by a certain threshold ϵ: for instance, for specific humidity, the depth over which $|q(z) - \overline{q}_\rho| \leq \epsilon_q$, where $\overline{q}_\rho$, hereafter $\overline{q}$, is updated at each vertical level and computed as the density-weighted mean from 100 m to z. We apply this method to q, θ, and virtual potential temperature, θ_v, a proxy for buoyancy. The method implicitly defines the sub cloud layer as that layer over which the vertical deviation in the mean structure is less than the expected horizontal

variability within the layer. Empirically, we choose a threshold that is one-third of small-scale variability, estimated as within-day variability from 50-550 m depth. ϵ-values for q, θ, and θ_v are 0.30 gkg^{-1}, 0.15 K, and 0.20 K, respectively. The second method, following Holzworth (1964), estimates the level at which a hypothetical rising parcel of near-surface air, representing a thermal, reaches its level of neutral buoyancy based on the θ_v profile, without any overshoot. The third definition finds the peak in the relative humidity profile, given that relative humidity is expected to maximize at the subcloud layer top if specific humidity is constant and temperature linearly decreases with height.

From this height analysis a conceptual picture emerges of two distinct layers that may be controlled by different mixing processes. We find evidence for a well-mixed layer in q and θ with a mean depth of 570 m, which also corresponds to the distribution of relative humidity maxima (Fig. 4-2, Fig. B1). This layer appears to be homogenized vertically by boundary layer eddies driven by surface fluxes. We call this height the mixed layer top. The buoyancy variable, θ_v, is, however, approximately well-mixed over a deeper layer, to a mean depth of 718 m. This depth also corresponds to the mean lifting condensation level (LCL) of 708 m, corresponding to the mean of LCL values estimated for air masses from 50–300 m. Given its correspondence with the mean LCL, we call this depth the subcloud layer top. The subcloud layer depth also coincides with peaks in cloud base height distributions from the BCO and R/V Meteor ceilometers. These ceilometer distributions do, however, have a substantial fraction of cloud bases below the mean subcloud layer height (not shown).

We associate the intermediate layer, defined by the offset between the mixed layer and subcloud layer top, with the presence of a transition layer, which has often been observed (Malkus, 1958; Augstein et al., 1974) and simulated (e.g. Stevens et al., 2001). Over this interfacial layer, q and θ begin to exhibit vertical gradients, albeit in ways that have compensating effects on buoyancy, so that θ_v gradients are much less pronounced. This offset is also present in individual soundings, demonstrating that it is not a product of averaging over layers that vary in depth. Using a smaller set of dropsondes, Betts and Albrecht (1987) also note a subcloud layer with constant θ_v

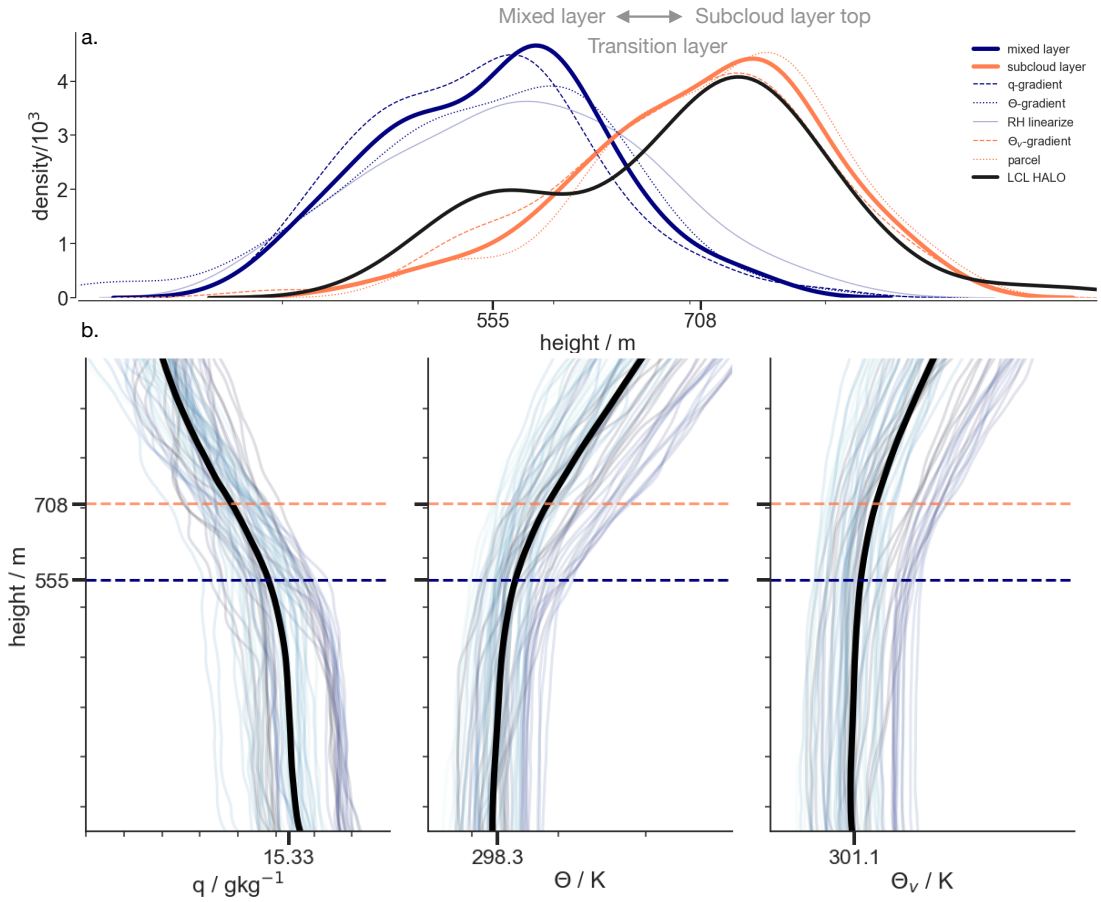


Figure 4-2: (a) distributions of different methods employed to estimate the mixed and subcloud layer heights. We find that three methods based on specific humidity q or potential temperature θ individually and relative humidity (blue curves) correspond to the mixed layer, whereas the θ_v-gradient and parcel method based on θ_v, a proxy for buoyancy (red), and the lifting condensation level (black). For the mixed layer and subcloud layer top distributions, the thicker line is the mean of individual distributions. (b) 69 circle-mean profiles from HALO for q, θ, and θ_v. The black line is the time-mean across all profiles, and colored profiles correspond to time (dark to lighter blue with time). Dotted lines mark the mixed layer height (estimated from the q-gradient method, blue) and subcloud layer height (estimated from the θ_v-method, red). Their difference indicates the transition layer.

(dry virtual adiabat), though their paper does not highlight a transition layer.

Contrary to what is often assumed in the application of mixed layer theory (e.g., Lilly, 1968; Tennekes, 1973), and previously asserted for the transition layer (e.g., Augstein et al., 1974; Albrecht et al., 1979), our measurements do not show the subcloud layer to be a layer defined by a sharp jump in θ_v. Spatial and temporal

variability could smooth vertical profiles and prevent sharp gradients, though we also do not find sharp jumps in most individual sounding profiles. Vertical variability in height is similar for different depths. The mixed layer top, defined with the q-gradient, and subcloud layer top, defined with the θ_v-gradient, for instance, vary coherently, with a Pearson correlation coefficient of r=0.86. While this vertical structure merits further study, for the purposes of our analysis we note that the presence of a transition layer introduces ambiguity in the application of the mixed layer theory. We address this uncertainty through the introduction of effective parameters estimated through a Bayesian approach, as outlined in the next section.

4.4.2 Evidence that vertical thermodynamic gradients are small

To test the assumption that conserved subcloud layer properties can be represented by a single vertical or bulk value, we compare the magnitude of vertical gradients relative to differences about the mean. We calculate the root mean square error (RMSE) from assuming a vertical, perfectly well-mixed profile relative to the observed profile up to the mixed layer top. We then compare this RMSE with one standard deviation of circle-mean data averaged up to the mixed layer top, σ, as $\left[\frac{\sigma - RMSE}{\sigma}\right] \times 100$. For q, the σ value is 83±3.2% across circle-mean data, suggesting that vertical deviations from the mean value are small relative to differences in the mean (see Fig. B1). Similarly, the σ value for potential temperature is 76±6.0% (Fig. B1). A related question is the extent to which the strength of vertical gradients encode differences in variability among subcloud layers. We find rank (Kendall and Spearman) and Pearson correlations of $\partial\theta/\partial z$ and $\partial q/\partial z$ with h, $\overline{\theta}$, and $\overline{q}$ below 0.3. We therefore infer that vertical gradients are small and therefore justify omission from the mixed layer description. We return to the question of whether excluding vertical gradients biases interpretation of subcloud layer moisture and energy budgets in Sec. 4.

4.4.3 Moisture variability is the primary mode of subcloud layer thermodynamic variability

We find that the subcloud layer varies thermodynamically primarily through variability in $\overline{q}$. Note that the height of the subcloud layer is defined using the θ_v-gradient method, and subcloud layer means, $\overline{q}$ and $\overline{\theta}$, are defined as the density-weighted means from 50 m to this height. We compare variability in $\overline{q}$ and $\overline{\theta}$ by diagnosing their contribution to variability in vertical length-scales. In the circle- and circling-mean observations, we find that anomalies in $\overline{q}$ have a Pearson correlation coefficient of -0.97 with anomalies in the LCL and -0.71 with h anomalies, showing that variability in the subcloud layer depth strongly associated with $\overline{q}$ variability. The correlation between anomalies in h and the LCL is 0.77. Anomalies in $\overline{\theta}$, by contrast, have a Pearson correlation coefficient of only -0.5 with anomalies in the LCL and -0.24 with h anomalies, counter intuitively suggesting that higher temperatures are associated with lower LCL heights. Warmer temperatures are also associated with increased humidity, which would lower the LCL, so the weak anticorrelation could reflect the trade-off between temperature and humidity. There is a nearly one-to-one relationship for variance in subcloud layer moisture and moist static energy (MSE), defined as MSE$=C_pT + L_vq + gz$ where C_p is the specific heat at constant pressure, T is the absolute pressure in Kelvins, L_v is the latent heat of vaporization, q is the water vapor specific humidity, g is the gravitational constant, and z is height above the surface. The Pearson correlation coefficient of $\overline{q}$ and mean subcloud-layer mean MSE is r=0.99.

From continuous ERA5 data, we find the large-scale $\overline{q}$ signal decorrelates after approximately two days, with a lag-1 autocorrelation (AR1) coefficient that decreases from 0.98 after one hour to 0.19 after 48 hours (Fig. 4-3d). This decorrelation timescale aligns with the mean gap of two days between EUREC^{4}A flights, such that the EUREC^{4}A data quantify independent realizations of synoptic moisture variability. Performing a fast Fourier transformation on ERA5 $\overline{\theta}$ data shows that its variability is largely diurnal (Fig. 4-3c) and likely driven by variability in radiative

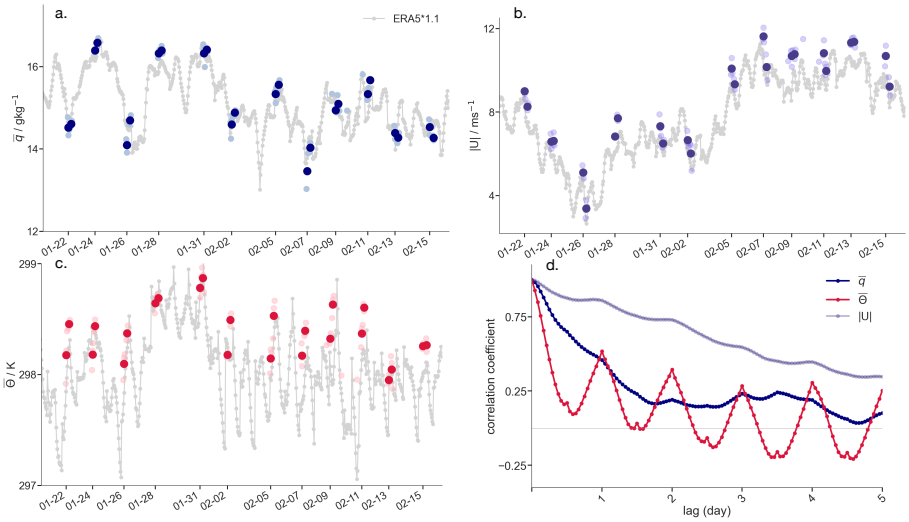


Figure 4-3: Evolution of subcloud layer-mean specific humidity, $\overline{q}$, (a.), 10 m wind speed (b.), and subcloud layer mean potential temperature, $\overline{\theta}$ for circle-mean data (lighter circle) and circling-mean data (darker circle), compared with hourly ERA5 data. Note that ERA5 moisture displays a dry bias (Bock et al., 2021) and ERA5 specific humidity is corrected by 1.1 Its variability is, however, coherent with variability in the *in situ* measurements Panel d. shows lag-1 autocorrelation (AR1) coefficients calculated from hourly ERA5 data for $\overline{q}$ and $\overline{\theta}$, interpolated to the same heights as the *in situ* data, and 10 m wind speed.

heating (Albright et al., 2021a). $\overline{\theta}$ has a strong peak in the power spectral density at a 24-hour period, a signal not seen in $\overline{q}$ (not shown). The ERA5 10 m wind speed is highly autocorrelated, with an autocorrelation coefficient of 0.74 after two days and 0.48 after 8 days (Fig. 4-3d). The wind speed signal decorrelates after ten days with r=0.04, demonstrating the dominance of lower-frequency variability in the wind.

4.5 Mixed layer theory for subcloud layer moisture and energy

Mixed layer theory assumes that the interface between the subcloud layer and the more stratified fluid above is vanishingly thin and that the subcloud layer gradients are negligibly small (e.g., Stevens, 2006). The presence of a transition layer and its gradients complicates this interpretation. EUREC^{4}A observations, however, provide unprecedented data to test the adequacy of this interpretive framework.

4.5.1 Theory and closure assumptions

For a subcloud layer scalar, ϑ, after performing a Reynolds decomposition on the conservation equation ($\frac{\mathrm{D}\vartheta}{\mathrm{D}t} = 0$) and integrating over the depth of the layer, the mixed layer budget of ϑ can be written as,

$$hQ_\vartheta = \overline{w'\vartheta'}\big|_0 - \overline{w'\vartheta'}\big|_1 . \tag{4.1}$$

Here, Q_ϑ is a source term that includes diabatic tendencies (e.g., radiant energy sources, evaporation, precipitation) and the material derivative. The thickness of the layer, h, is $h = z_1 - z_0$. The subscript 0 denotes values at the lower interface of the bulk layer, the ocean-to-subcloud layer interface, subscript 1 denotes values at the upper interface, the subcloud-to-cloud layer interface, and w refers to the vertical velocity. The equation expresses that the vertical divergence of the turbulent flux balances the sum of the non-turbulent processes, denoted by Q_ϑ (e.g., Betts, 1976; Stevens, 2006).

The flux at an interface is given as product of the velocity of the interface relative to the mean flow and a 'jump', at that interface i, such that,

$$\overline{w'\vartheta'}\big|_i = -V_i \Delta_i \vartheta, \tag{4.2}$$

where $\Delta_i\vartheta$ defines the change in ϑ across the interface, from top to bottom, so that $\Delta_1\vartheta = \vartheta_1 - \overline{\vartheta}$ and $\Delta_0\vartheta = \overline{\vartheta} - \vartheta_0$. At the top of the layer, this jump is traditionally assumed to occur over a layer of zero thickness. Such a sharp discontinuity is seen in stratocumulus mixed layers (e.g., Lilly, 1968; Deardorff, 1972) and represented in idealizations of the dry mixed layer (e.g., Tennekes, 1973; Stull, 1976; Albrecht et al., 1979).

As shown in the previous section, h and hence the jumps are poorly defined in the trades. These findings introduce ambiguity in the application of mixed layer theory. We attempt to accommodate this ambiguity by introducing scaling coefficients, C_q and C_θ, in the entrainment flux calculation. These coefficient scale jumps at the top

of the layer and compensate for possible errors in our choice of the subcloud layer height, h, and uncertainty in the depth over which the jumps are computed. To the extent that these coefficients do not vary on a case-to-case basis, they indicate that variability in the transition layer depth is not the primary mechanism regulating mixing into the boundary layer. The jumps $\Delta_1 q$, $\Delta_1 \theta$, and $\Delta_1 \theta_v$ are formulated as,

$$\Delta_1 q = C_{\mathrm{q}}(q_{h+} - \overline{q}_{|h-}) \tag{4.3}$$

$$\Delta_1 \theta = C_{\theta}(\theta_{h+} - \overline{\theta}_{|h-}) \tag{4.4}$$

$$\Delta_1 \theta_v = \Delta_1 \theta + 0.61(\overline{\theta}\Delta_1 q + \overline{q}\Delta_1 \theta) \tag{4.5}$$

The subscript $h+$ refers to the value of q or θ above h, computed as the average from h to h + 100 m. $\overline{q}_{|h-}$ or $\overline{\theta}_{|h-}$ are averages from 50 m to the mixed layer top defined from the linearized relative humidity method (see Appendix A), though it is insensitive to defining the mixed layer top using other methods. The choice of averaging up to the mixed layer top, rather than up to h, to calculate these jumps is motivated by the desire to exclude transition layer air in the average. Due to sub-circling variability in h, excluding transition layer air can best be achieved by a conservative (and therefore lower-altitude) choice of averaging height. An idealization of these jumps is given schematically in Fig. 4-4. This scaling coefficient approach is also similar to the linear mixing-line model that represents cloud layer air mixing into the mixed layer (Betts and Ridgway, 1989). Whereas our formulation is similar to the 'transfer coefficient' approach in Neggers et al. (2006) and Zheng (2019), these studies consider values in the free troposphere whereas we consider values at the subcloud layer top. The structure of the transition layer, as discussed in Sec. 3a, suggests that this layer is actively mixed by local processes, rather than by an overturning mixing that extends to the free troposphere.

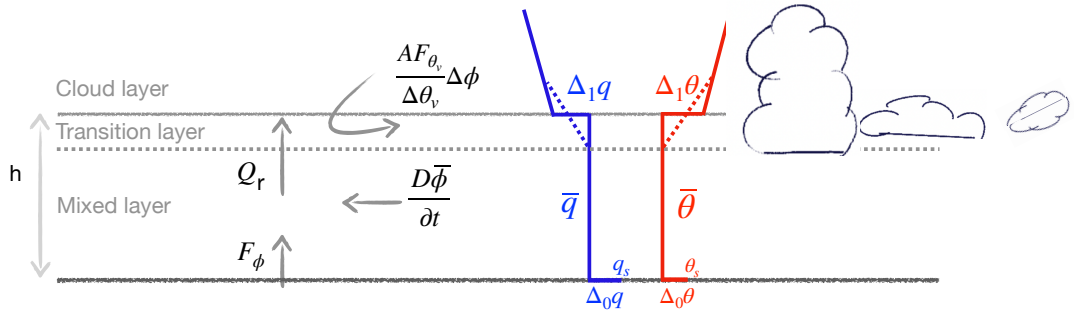


Figure 4-4: Schematic of subcloud layer budgets, as described in Eq. (4.9) and Eq. (4.10). We give a description of the surface fluxes (F_ϑ), entrainment fluxes ($E\Delta\vartheta$), material derivatives, and clear-sky radiative heating term (Q_r) in Sec. 4a. The height h refers to the depth of the subcloud layer, which includes both a well-mixed layer and a transition layer. Also shown are idealized profiles of specific humidity, q, and potential temperature, θ, from foregoing studies, with the dotted lines marking a transition layer gradient.

With these jumps we can rewrite Eq. (4.1) as,

$$hQ_\vartheta = -V_0\Delta_0\vartheta + V_1\Delta_1\vartheta. \tag{4.6}$$

The first term on the left-hand side, $V_0\Delta_0\vartheta$, defines a surface flux wherein the surface exchange velocity, V_0, denotes the product of the 10 m horizontal wind speed, U (wherein easterly is defined as negative), and a dimensionless parameter, C_d, following surface layer similarity theory (e.g., Stevens, 2006). Note that $-V_0\Delta_0\vartheta$ is positive when surface values are larger than near-surface -air values, which is almost always the case. The dimensionless parameter, C_d, depends on the surface roughness, the structure of the surface layer, and the stability of this layer but is generally taken to be constant and equal to 0.0011 (e.g., Deardorff, 1972; Fairall et al., 2003). In this analysis, we set $C_d = 0.0010$, which is slightly smaller than its typical value of 0.0011. Our smaller value compensates for the larger difference between the surface and subcloud layer mean value, compared to the typical difference taken between the surface and 20 m value (e.g., Fairall et al., 2003).

V_1 represents the diabatic growth of the subcloud layer into the overlying fluid

and is taken to equal the entrainment rate, E (Stevens, 2006). Given a transition layer of zero thickness, a common closure for E is to represent the turbulent flux at the subcloud layer top as a fixed fraction of the surface turbulent flux (Lilly, 1968; Stull, 1976). That is, the turbulent entrainment flux is energetically constrained by its surface source, such that,

$$E = -\frac{AV_0\Delta_0\theta_v}{\Delta_1\theta_v}. \tag{4.7}$$

The constant A defines the entrainment 'efficiency' and is between 0 and 1. The jump, $\Delta_1\theta_v$, is positive and models the resistance that entrainment feels when expanding into the overlying, more buoyant fluid.

As demonstrated in Fig. 4-2, the transition layer does, however, have finite thickness, and the above closure is not appropriate. (Include the following either here in the text or in an Appendix) We present two expository examples that motivate the appearance of an effective A parameter, A_e, reflecting a thick interfacial layer. Integrating over a layer above and below h from $h_+ = h + \epsilon$ to $h_- = h - \epsilon$ for some small ϵ yields an expression for E, wherein θ_{v+} is the value at h_+,

$$E = \frac{-AV_0\Delta_0\theta_v}{\Delta_1\theta_v} + \frac{\delta h}{2\Delta_1\theta_v}\left(\frac{d\overline{\theta_v}}{dt} + \frac{d\theta_{v+}}{dt}\right). \tag{4.8}$$

The derivation for Eq. (4.8) is given in Appendix C. Setting these two equations for E equal Eq. (4.7) and Eq. (4.8), we can consider that the second term on the right-hand side of Eq. (4.8) as being absorbed to increase A in Eq. (4.7), which renders A as an effective parameter, A_e that accounts for jumps occurring over finite-thickness.

A similar justification is obtained by specifying of A_e to be the value that gives the correct flux divergence:

1. Modeling the subcloud layer with an infinitely thin transition layer (zero-order flux-jump model), the rate of warming or drying in the subcloud layer mean is given by the flux gradient $\frac{\partial F}{\partial z}$, equivalent to $\frac{F_0(1+A)}{h}$ if $F_\mathrm{h} = -AF_0$.

2. If the flux minimum is not at the top of the subcloud layer, but there is a finite thickness transition layer above the subcloud layer top, then the minimum flux

is not at h, but at a height h - δh. Defining A as the ratio of the minimum to maximum flux, the equation becomes $\frac{F_0(1+A)}{(h-\delta h)}$.

3. Setting these two terms equal and replacing A in 1. with A_e yields $\frac{F_0(1+A_\mathrm{e})}{h} = \frac{F_0(1+A)}{(h-\delta h)}$, or $A_\mathrm{e} = \frac{(1+A)h}{(h-\delta h)}$ -1. As an example, if A=0.2 as often assumed, h=700 m, δh=120 m, then A_e=0.45.

Combining these assumptions, the equilibrium budget for Eq. (4.1) for specific humidity, $\overline{q}$ is thus

$$h\left[\frac{\partial \overline{q}}{\partial t} + \overline{(\vec{u}\cdot\nabla q)}\right] = -C_d U \Delta_0 q - \frac{A_e V_0 \Delta_0 \theta_v}{\Delta_1 \theta_v}\Delta_1 q = 0. \tag{4.9}$$

The $\overline{q}$ balance is between a surface kinematic latent heat flux, $C_d U \Delta_0 q$, which moistens the layer from a saturated ocean surface moisture source, q_s, wherein $\Delta_0 q = \overline{q} - q_s < 0$; an entrainment flux, $\frac{A_e V_0 \Delta_0 \theta_v}{\Delta_1 \theta_v}\Delta_1 q$, which imports drier cloud layer air into the subcloud layer, wherein $\Delta_1 q < 0$; and large-scale horizontal advection, $\vec{u}\cdot\nabla\overline{q}$ and a time-derivative, $\frac{\partial \overline{q}}{\partial t}$, both of which can either moisten or dry the subcloud layer. Note that we neglect phase changes, such as associated with evaporating precipitation.

For potential temperature, $\overline{\theta}$, the equilibrium budget equation is,

$$h\left[\frac{\partial \overline{\theta}}{\partial t} + \overline{(\vec{u}\cdot\nabla\theta)}\right] = -C_d U \Delta_0 \theta - \frac{A_e V_0 \Delta_0 \theta_v}{\Delta_1 \theta_v}\Delta_1 \theta + hQ_\mathrm{r} = 0. \tag{4.10}$$

This energy balance is more difficult to constrain, both because it involves more terms and because the magnitude of individual terms is smaller. Eq. (4.10) includes a surface kinematic sensible heat flux, $C_d U \Delta_0 \theta$, which warms the subcloud layer given $\Delta_0\theta = \overline{\theta} - \theta_s < 0$; an entrainment flux $\frac{-A_e V_0 \Delta_0 \theta_v}{\Delta_1 \theta_v}\Delta_1 \theta$, which brings warmer cloud layer air into the subcloud layer; large-scale horizontal advection $\vec{u}\cdot\nabla\overline{\theta}$, which could either warm or cool the layer; a time-derivative $\frac{\partial \overline{\theta}}{\partial t}$, which is predominantly associated with the diurnal cycle; and a clear-sky radiative heating term $Q_\mathrm{rad,clr}$, another cooling term.

The $\overline{q}$ and $\overline{\theta}$ balances are illustrated schematically in Fig. 4-4.

4.5.2 Observational estimates of terms in mixed layer theory budgets

Here we describe how the different terms in Eq. (4.9) and Eq. (4.10) are calculated from the observations. Clear-sky, aerosol-free radiative heating profiles for EUREC^{4}A dropsondes and radiosonde profiles are calculated in Albright et al. (2021a). Large-scale horizontal moisture advection, $\vec{u} \cdot \nabla \overline{q}$, and potential temperature advection, $\vec{u} \cdot \nabla \overline{\theta}$, are calculated in George et al. (2021) using the regression method from Bony and Stevens (2019). We estimate the time derivatives or storage terms, $\frac{\partial \overline{q}}{\partial t}$ and $\frac{\partial \overline{\theta}}{\partial t}$, as the ordinary least squares regression slope of the three circle-means per circling-mean. Estimating this derivative as the regression slope for the approximately 30–36 individual soundings per circling yields similar results (r=0.82), yet these soundings are more affected by small-scale variability than are circle-means. For uncertainty estimates, we calculate the standard error of the three circle-mean data per circling-mean for all terms in Eq. (4.9) and Eq. (4.10), except for uncertainty on the time-derivative, which we take to be the standard error on the regression slope.

Our bulk estimates of the sensible heat flux are 6.3 ± 2.7 Wm^{-2} and 166 ± 54 Wm^{-2} for the latent heat flux, with the notation denoting the mean and standard deviation. We find good agreement among these bulk surface fluxes and four co-located estimates: bulk estimates from the R/V Meteor, eddy covariance flux measurements from the R/V Meteor, fluxes calculated with the COARE algorithm (Fairall et al., 2003) using HALO dropsonde data, and ERA5 surface fluxes (see Table 2). Pearson correlation coefficients of our bulk sensible heat flux estimates are 0.81, 0.63, 0.82, and 0.72 with these four estimates, respectively. For the latent heat fluxes, the correlations are 0.84, 0.78, 0.92, and 0.95, respectively.

The surface turbulent flux, F_{θ_v}, is used to compute E and is defined as $F_{\theta_v} = F_\theta + \epsilon \overline{\theta} F_q$, where $\epsilon \approx 0.608$ is a thermodynamic constant that depends on the molecular weight of water relative to that of dry air. F_{θ_v} relates to surface buoyancy flux by a factor $\frac{g}{\overline{\theta_v}}$, which then denotes the time rate of change of turbulent kinetic energy production. F_{θ_v} fluxes estimated from mixed layer theory or bulk methods are $16 \pm$

	units	mean	1 s.d.
mixed layer	m	555	79.0
subcloud layer	m	708	83.6
lifting condensation level	m	694	105
transition layer	m	153	52
$\overline{q}$	$\mathrm{gkg^{-1}}$	15.3	0.886
$\overline{q_{\mathrm{sc}}}$	$\mathrm{gkg^{-1}}$	15.1	0.883
$\overline{\theta}$	K	298	0.235
$\overline{\theta_{\mathrm{sc}}}$	K	298	0.253
SST	K	300.3	0.166
θ_s	K	299.2	0.260
q_s	$\mathrm{gkg^{-1}}$	22.59	0.189
$\partial q/\partial z$ (ML)	$\mathrm{gkg^{-1}m^{-1}}$	$-1.06*10^{-3}$	$2.93*10^{-4}$
$\partial \theta/\partial z$ (ML)	$\mathrm{Km^{-1}}$	$4.24*10^{-4}$	$2.91*10^{-4}$
$\partial \theta_v/\partial z$ (ML)	$\mathrm{Km^{-1}}$	$2.33*10^{-4}$	$2.47*10^{-4}$
$\partial q/\partial z$ (TL)	$\mathrm{gkg^{-1}m^{-1}}$	$-6.69*10^{-3}$	$2.11*10^{-3}$
$\partial \theta/\partial z$ (TL)	$\mathrm{Km^{-1}}$	$2.49*10^{-3}$	$6.44*10^{-4}$
$\partial \theta_v/\partial z$ (TL)	$\mathrm{Km^{-1}}$	$1.30*10^{-3}$	$3.24*10^{-4}$
10 m wind speed, $\vec{u}_{10}$	m/s	8.46	2.19
$\overline{\vec{u}\cdot\nabla q}$	$\mathrm{gkg^{-1}ms^{-1}}$	$1.59*10^{-5}$	$2.12*10^{-5}$
$\overline{\vec{u}\cdot\nabla \theta}$	$\mathrm{K\ ms^{-1}}$	$8.51*10^{-3}$	$8.46*10^{-3}$
$Q_{\mathrm{rad,clr}}$	$\mathrm{K\ d^{-1}}$	-0.853	1.01
E	$\mathrm{mm\ s^{-1}}$	20.3	7.74
Δq	$\mathrm{gkg^{-1}}$	-2.36	0.597
$\Delta \theta$	K	0.782	0.115
$\Delta \theta_v$	K	0.359	0.0297
F_{q} bulk, dropsondes	$\mathrm{Wm^{-2}}$	166	56
F_{q} bulk, Meteor	$\mathrm{Wm^{-2}}$	165	48
F_{q} eddy covariance, Meteor	$\mathrm{Wm^{-2}}$	150	51
F_{q} COARE, dropsondes	$\mathrm{Wm^{-2}}$	162	45
F_{q} ERA5	$\mathrm{Wm^{-2}}$	178	49
F_θ bulk, dropsondes	$\mathrm{Wm^{-2}}$	6.3	2.7
F_θ bulk, Meteor	$\mathrm{Wm^{-2}}$	6.5	3.8
F_θ eddy covariance, Meteor	$\mathrm{Wm^{-2}}$	12.4	4.1
F_θ COARE, dropsondes	$\mathrm{Wm^{-2}}$	6.0	6.1
F_θ ERA5	$\mathrm{Wm^{-2}}$	10	4.5

Table 4.2: Campaign-mean and standard deviation of different terms as calculated from the 69 circle-mean data (hourly timescale) located in the 'EUREC^{4}A circle'. Results are qualitatively similar for 24 circling-mean data aggregated to a three-hourly timescale. For the surface flux terms, bulk refers to bulk theory flux formulations, COARE refers to the COARE algorithm, and Meteor refers to the R/V Meteor.

5.2 Wm^{-2}, which agree well with the four other estimates in magnitude and with correlation coefficients around r=0.9.

4.5.3 Bayesian inversion of uncertain entrainment parameters

The most uncertain terms in Eq. (4.9) and Eq. (4.10) relate to the entrainment fluxes: the effective entrainment efficiency, A_e, and the scaling parameters for the jumps, C_q and C_θ. The entrainment exchange velocity, E, and fluxes at the subcloud layer top have long been challenging to measure observationally (e.g., Lenschow et al., 1999; Kawa and Pearson Jr, 1989; Stevens et al., 2003) or estimate from simulations (e.g., Moeng et al., 1999; Bretherton et al., 1999; Vogel et al., 2020).

There are a lack of foregoing constraints on these jumps, and from the trade-wind observations, there are ambiguities associated with how to define the jumps at the upper interface. Given these uncertainties, we constrain the parameters, $\Theta = \{A_e, C_\mathrm{q}, C_\theta\}$, using a Bayesian framework. This approach allows for estimating a joint distribution of parameters Θ that are most likely to explain observed data. The values of C_q, C_θ, and A_e should be physical, and synoptic variability should be explained without having to vary these parameters, as the latter would be indicative of a failed assumption.

The Bayesian approach is similar to other optimization techniques, yet it yields joint posterior distributions and thus provides an estimate of uncertainty for the constrained parameters. Following Bayes rules, we invert for the joint posterior distribution of Θ,

$$P(\Theta \mid y_\mathrm{obs}) \propto P(y_\mathrm{obs} \mid \Theta) \cdot P(\Theta). \tag{4.11}$$

Sampled sets of parameter values, Θ, are used with observed estimates of the other terms to model subcloud layer moisture and temperature budgets, y_obs, following Eq. (4.9) and Eq. (4.10). We model the likelihood, $P(y_\mathrm{obs} \mid \Theta)$, such that the residuals of the subcloud layer moisture and energy budgets are normally distributed around zero with standard deviations σ_q and σ_θ, respectively. The likelihood $P(y_\mathrm{obs} \mid \Theta)$ is thus formulated based on the multivariate Gaussian distribution of the modeled

subcloud layer moisture and temperature budgets,

$$P(y_{\text{obs}} \mid \Theta) \sim N(0, \Sigma^2(\Theta)). \tag{4.12}$$

$$\Sigma^2(\Theta) = \begin{bmatrix} \sigma_q^2 & 0 \\ 0 & \sigma_\theta^2 \end{bmatrix}. \tag{4.13}$$

This likelihood expresses the probability of closing the moisture and energy budgets, or jointly obtaining residuals equal to zero for both budgets given parameters Θ. We do not expect any covariance between the moisture and temperature residuals, and therefore the off-diagonal terms are zero in Eq. (4.13). We assume a standard deviation σ_q of 10^{-8} kgkg^{-1}s^{-1} or 17 Wm^{-2}. For potential temperature, the standard deviation σ_θ is chosen to be 3*10^{-6} Ks^{-1} or 2.5 Wm^{-2}, though our results are insensitive to these choices. The posterior distribution, $P(\Theta \mid y_{\text{obs}})$, then represents the distribution of parameter values that are most likely to close the budgets given observed data.

For the prior distribution $P(\Theta)$ on A_e, we choose a normal prior with mean of 0.2 and standard deviation 0.4, $\mathcal{N}(0.2, 0.4^2)$. A common view of A is the ratio of minimum to maximum buoyancy fluxes, when assuming that the minimum buoyancy flux occurs at the subcloud layer top, and that the transition from the negative, minimum to zero buoyancy flux occurs over an infinitely thin layer. Its value is often taken to be 0.2 (Lilly, 1968; Stull, 1976; Canut et al., 2012). A value of A greater than one is energetically inconsistent, given that the source for entrainment mixing is the surface turbulent fluxes. Obtaining a posterior distribution of A_e that does not exceed one serves as a physical test of the model. We model the prior distributions for C_{q} and C_θ as a normal distribution with mean 1 and standard deviation of 0.5.

Sampling is performed using the Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970), and we run four chains of 60,000 samples. The first 10,000 samples are discarded for each chain, yielding 200,000 samples. Results are consistent among chains, indicating that our model is adequately sampled.

Notably, A_e is well-constrained by the inversion and has a maximum likelihood estimate (MLE) posterior value of 0.43 and a 5–95% credible interval (c.i.) of 0.34–0.53 (Fig. 4-5). Its marginal posterior distribution is similar regardless of the prior distribution. A_e being larger than 0.2 is consistent with it being an effective parameter that reflects the presence of a finitely-thick transition layer as previously discussed. A finitely thick layer is consistent with what is seen in large eddy simulations simulations (e.g., Vanzanten et al., 1999) and direct numerical simulations (e.g., Garcia and Mellado, 2014), albeit in the case of the cloud-free mixed layer.

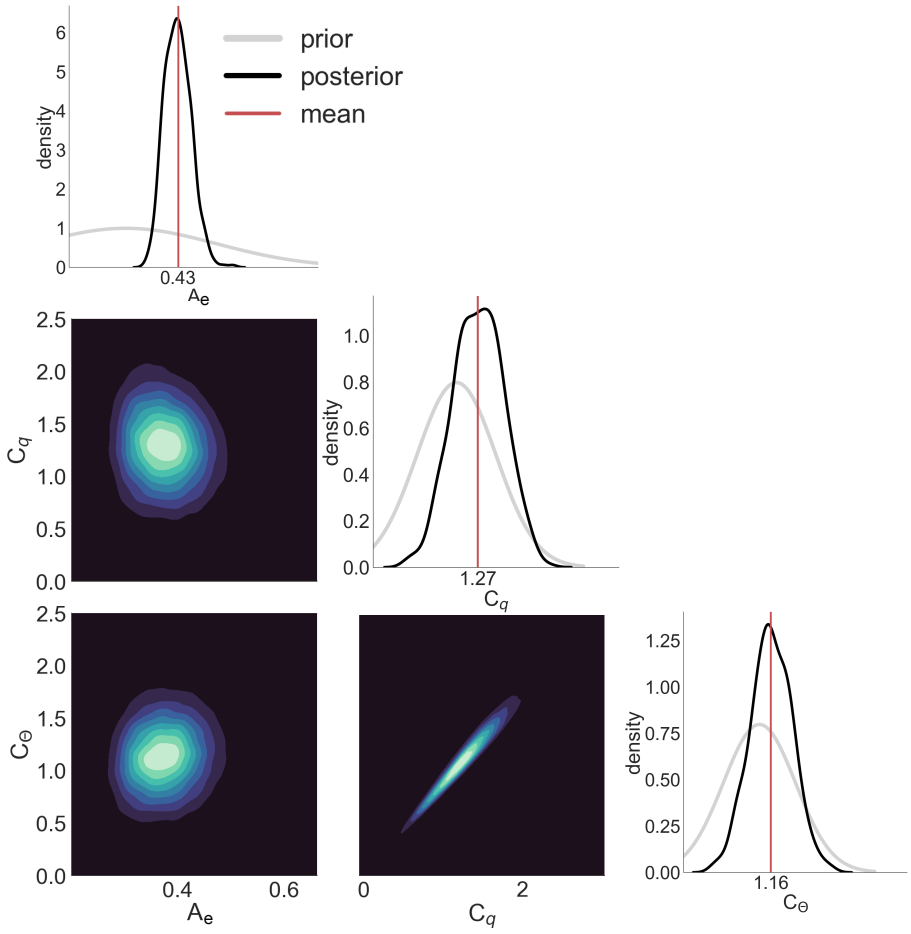


Figure 4-5: Marginal and joint posterior distributions for the uncertain entrainment parameters $\Theta = \{A, C_q, C_\theta\}$. For the marginal posterior distributions (black), the marginal prior distribution (grey) and maximum likelihood estimate, MLE (red), are also shown.

The MLE C_q is 1.27 with a 5-95% c.i. of 0.74–1.89. The MLE C_θ is 1.16 with a 5-95% c.i. of 0.65–1.66 (Fig. 4-5). We multiply the MLE transfer coefficients by

the time-varying values of Δq and $\Delta\theta$ (calculated following Eq. (4.3) and Eq. (4.4) to obtain specific humidity and potential temperature jumps used to calculate the entrainment flux. We also note the strong covariance between C_{q} and C_θ, which have a Pearson correlation coefficient of r=0.97, highlighting the utility of a Bayesian framework that can capture this parameter covariance. The covariance suggests that the same subcloud layer eddies mix moisture and energy, consistent with physical expectation. That the parameters constrained by the Bayesian inversion are physical – namely, an A_e greater than 0.2 and scaling coefficients that strongly covary – acts as a first validation of mixed layer theory.

4.6 Resulting moisture and temperature budgets

Adopting the maximum likelihood estimate parameters from our Bayesian inversion, Fig. 4-6 shows that the bulk theory budgets close to within 3.6 Wm^{-2} for moisture and 2.9 Wm^{-2} for potential temperature for the campaign-mean and can largely explain synoptic variability. For the moisture budget, the campaign-mean residual is 2.2% of the largest term, the surface latent heat flux. Out of 24 circling-means, 20 residuals are unbiased given uncertainty estimates, and the budgets close equally well for circling-means measured during daytime or nighttime (Fig. 4-6a). In the energy balance, 14 out of 24 residuals are unbiased, and as for moisture, the budget holds equally well for day and night (Fig. 4-6b).

Regarding the relative magnitude of physical processes, for the moisture budget, the dominant balance is between surface latent heat flux (166 ± 56 Wm^{-2}) and entrainment drying flux (-128 ± 53 Wm^{-2}), with a secondary role for large-scale moisture advection (-34 ± 41 Wm^{-2}) and the storage term (-1.0 ± 40 Wm^{-2}). Note that the advection term is negative because it is subtracted from the left-hand side in Eq. (4.9). Its absolute value is positive, as the product of negative (easterly) winds and a negative difference (moving west to east), as the advection brings colder, less moist air from east to the west. In the energy budget, the entrainment flux (18 ± 6.5 Wm^{-2}) has roughly twice the magnitude of the other terms, which have a similar

magnitude of 6–8 Wm^{-2} (Fig. 4-6b). Physically this balance says that radiative cooling of the layer is disproportionately balanced by entrainment warming, enabled by moisture flux contributions to the turbulence kinetic energy production, i.e., though the F_q contribution to F_{θ_v}.

One might be tempted to think that the flexibility afforded by the Bayesian framework allows for closing the budgets by construction. A_e, C_q, and C_θ are constants, yet vertical profiles of moisture and potential temperature change across days, such that there is no guarantee that a fixed combination of A_e, C_q, and C_θ allows for budgets to close. We close moisture and energy budgets jointly, which provides a stronger constraint than closing a single budget. That is, in the moisture budget, drying by entrainment balances moistening by surface fluxes, whereas in the energy budget, both entrainment and surface fluxes warm the layer, such that each budget place counteracting constraints on the entrainment rate.

That the budgets close to within these small residuals suggests that knowledge of the mean state in Eq. (4.9) and Eq. (4.10) is sufficient to close the budgets, without knowledge of the vertical thermodynamic gradients or incorporating additional processes. The correlation of residuals with vertical gradients informs whether the omission of vertical gradients is justified. Indeed, correlations of residuals with vertical gradients are small. The correlations of $\overline{\theta}$ residuals with $\partial\theta/\partial z$ and $\partial q/\partial z$ are 0.21 and -0.17, respectively, and the correlations of the $\overline{q}$ residuals with $\partial\theta/\partial z$ and $\partial q/\partial z$ are 0.33 and -0.25, respectively. The ratios of $\partial\theta/\partial z$ and $\partial q/\partial z$ multiplied by $h/2$ to the jumps at the upper interface are also small, with a mean value of 20% for q and 22% for θ, providing further evidence that the influence of vertical gradients on the subcloud layer budgets is small.

Structure in the residuals is also indicative of observational error or missing processes. We use two proxies for precipitation or precipitation-driven downdrafts, whose influences we neglect: cloud top height estimated from the WALES instrument onboard HALO (Konow et al., 2021), mindful that deeper clouds are more likely to precipitate (e.g., Stevens et al., 2016), and a cold pool fraction per circling, wherein a cold pool sounding is defined as having θ_v-gradient height less than 400 m (Touzé-

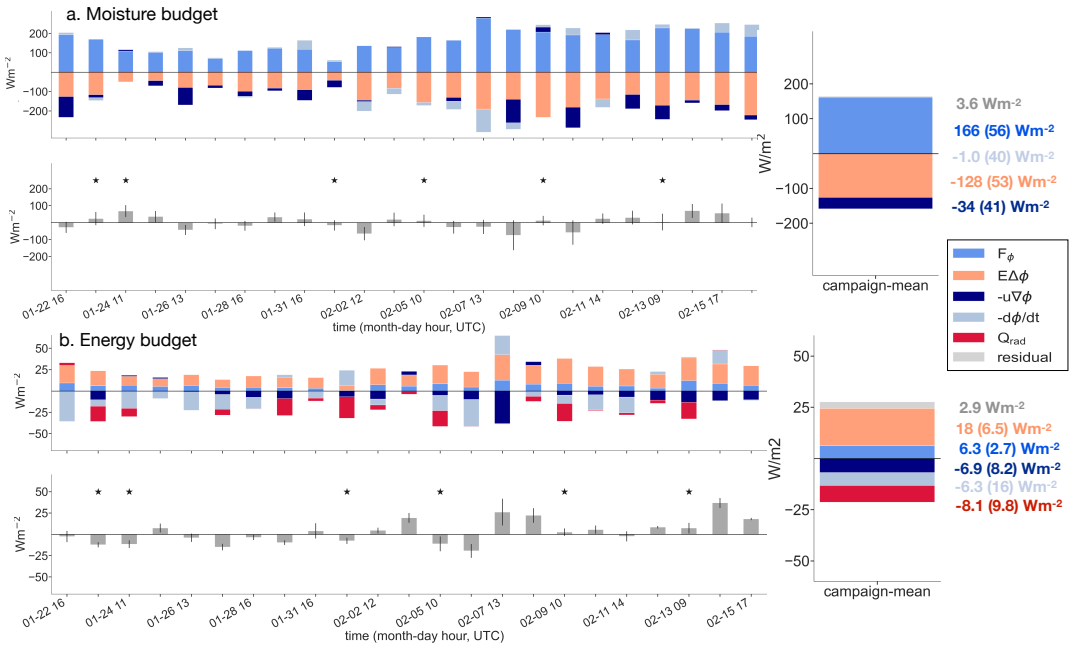


Figure 4-6: Synoptic variation and campaign-mean moisture balance, showing the surface flux (blue), entrainment flux (orange), large-scale horizontal advection (dark blue), the time derivative (light blue), clear-sky radiative cooling (red), and the residual term (grey). Black stars flag circling-means that include sondes launched during the nighttime. Panel a. shows the moisture balance whereas panel b. shows the energy balance. Uncertainties are added in quadrature for the 5-95% Bayesian credible interval on entrainment parameters, as well as one standard deviation for individual terms calculated across the three circle-means per circling-mean.

Peiffer et al., 2021). The residual structure is, however, not correlated with these proxies. For WALES cloud top height, correlations are r=0.19 for $\overline{q}$ residuals and r=0.16 for $\overline{\theta}$ residuals. For the cold pool fraction, these correlations are also small, r=0.23 for $\overline{q}$ residuals and r=0.24 for $\overline{\theta}$ residuals. These weak correlations support our finding that the subcloud layer moisture and energy budgets can close solely by representing small-scale entrainment mixing. These findings that the influence of downdrafts and other coherent structures is relatively small in the trades relative to turbulent entrainment mixing are consistent with Thayer-Calder and Randall (2015), justifying assumptions in many parameterizations, dating back to Arakawa and Schubert (1974).

4.7 How do these subcloud layer properties relate to the large-scale environment?

Surface and entrainment fluxes are strongly associated with U variability, as expected given their structural dependence on the wind speed. Fig. 4-7 relates variability among U, h, surface fluxes, and entrainment fluxes, and the clear-sky radiative cooling, Q_{rad}. A deeper subcloud layer is associated with stronger U, r=0.62, consistent with Nuijens and Stevens (2012). Fixing other parameters at their campaign-mean value and only allowing U to change recovers most variance in surface and entrainment fluxes: 87% of the variance in F_q, 64% of the variance in $E\Delta q$, 74% of the variance in $E\Delta\theta$, though only 22% of the variance in F_θ. If we instead allow only the sea surface temperature to vary, we recover 32% of the variance in F_q, 38% of the variance in $E\Delta q$, 11% of the variance in F_θ, and 35% of the variance in $E\Delta\theta$. The surface wind speed plays a larger role in explaining variability in the fluxes except for $E\Delta\theta$.

Having established that mixed layer theory is a skillful framework, we can further employ it as a physical mapping to diagnose how boundary conditions, such as the surface wind speed, influence $\overline{q}$ and $\overline{\theta}$. Solving for $\overline{q}$ at equilibrium from Eq. (4.9) yields,

$$\overline{q} = \frac{V_0 q_s + E q_+ - h(\frac{\partial q}{\partial t} + \overline{\vec{u} \cdot \nabla q})}{V_0 + E}. \tag{4.14}$$

Similarly, solving for $\overline{\theta}$ at equilibrium yields,

$$\overline{\theta} = \frac{V_0 \theta_s + E\theta_+ + hQ_r - h(\frac{\partial \theta}{\partial t} + \overline{\vec{u} \cdot \nabla \theta})}{V_0 + E}. \tag{4.15}$$

The velocity scale $V_0 = C_d U$, q_+ and θ_+ correspond to values 100 m above the subcloud layer top. E can, moreover, be rewritten as $\frac{A_e F_{\theta_v}}{(\overline{\theta}+0.61[\overline{\theta}(q_+-\overline{q})+\overline{q}(\theta_+-\overline{\theta})]-\overline{\theta_v}}$ as function of q_+ and θ_+.

Predictions of $\overline{q}$ from Eq. (4.14) nearly recover observed $\overline{q}$, with a correlation coefficient r=0.92 (Fig. 4-8a). Predictions of $\overline{\theta}$ from Eq. (4.15) have a correlation

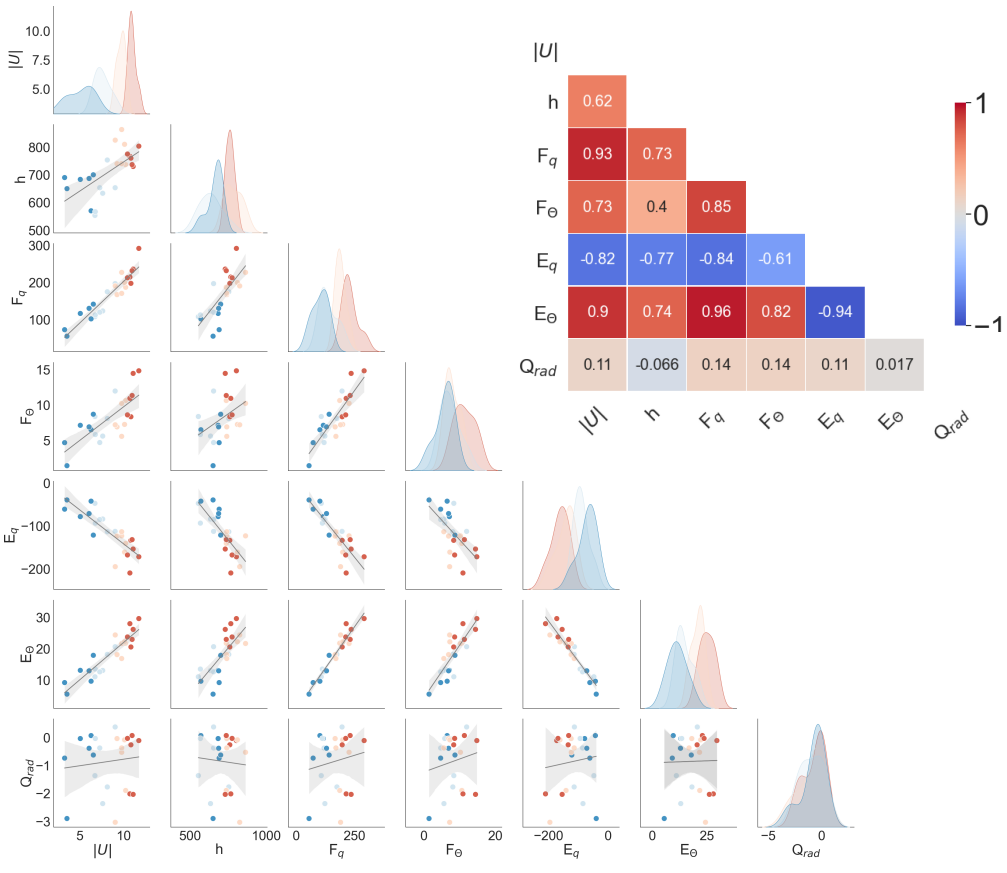


Figure 4-7: Relationships among 10 m wind speed, subcloud layer depth, surface fluxes, and entrainment fluxes. The black line is the central ordinary least squares regression, and the grey shading is the 5–95% confidence interval on the regression. Colors correspond to quartiles of the wind speed with increasing wind speed going from blue to red. The inset shows the Pearson correlation coefficients.

coefficient of r=0.48 with observed $\overline{\theta}$, qualitatively consistent with larger residuals in the energy budget (Fig. 4-8d). Note that if only considering residual terms whose absolute value is less than 5 Wm^{-2}, which occurs 29% of the time, then the correlation of observed and predicted $\overline{\theta}$ becomes r=0.88.

Different boundary conditions, set by the large-scale environment, are U, q_+, θ_+, SST, $\vec{u}\cdot\nabla\overline{q}$, $\vec{u}\cdot\nabla\overline{\theta}$, $\frac{\partial q}{\partial t}$, $\frac{\partial \theta}{\partial t}$, and $\mathrm{Q_r}$. We vary one external parameter at a time, and fix the other parameters at their campaign-mean value, to estimate the predicted $\overline{q}$ or $\overline{\theta}$ values driven just by variability in a single external condition (Fig. 4-8). Perhaps surprisingly, varying only the surface wind speed to predict $\overline{q}$ yields a weak correlation with observed $\overline{q}$ (r=-0.28) or with predicted $\overline{q}$ when allowing all quantities to vary (r=-

0.11), not only the wind. That the net influence of the wind speed is weak results from its opposing influences, both moistening the layer through surface fluxes and drying it through entrainment. Whereas the correlation of the surface wind speed with individual fluxes is strong (Fig. 4-7), the correlation of the wind speed with the sum of the surface moistening flux and entrainment drying flux is weak (r=0.32). Varying only q_+, which also influences the entrainment rate through $\Delta\theta_v$, yields the highest correlation with predicted moisture (r=0.66, Fig. 4-8b) and observed moisture (r=0.54). Varying other factors yields weak correlations (Fig. 4-8c). For thermal energy, we find, similarly, that θ_+ recovers the most variance in $\overline{\theta}$ that is predicted (r=0.43) and observed (r=0.93), whereas the predictive power of other external factors is smaller. If again considering the case of residual terms whose absolute values are less than 5 Wm^{-2}, the correlation of observed $\overline{\theta}$ is r=0.89 when only varying θ_+ to calculate the predicted $\overline{\theta}$ and r=0.93 when only varying q_+.

In summary, variability in the fluxes is very strongly influenced by U variability. Yet because of opposing influences of the surface and entrainment fluxes on $\overline{q}$, the surface wind speed influence on $\overline{q}$ is weak. Instead, knowing the humidity above the subcloud layer, q_+, is the most informative for predicting $\overline{q}$ variability. Subcloud layer moisture and moisture just above are coupled, such that it is not possible with this analysis to infer causality.

Regarding external influences on vertical gradients, we also find weak rank and Pearson correlations (r<0.2) of $\partial q/\partial z$ and $\partial\theta/\partial z$ with the 10 m wind speed or wind shear between different vertical levels. This finding contrasts with Malkus (1958) who found the vertical moisture gradient to be anticorrelated with the surface wind speed, r=-0.68, though the earlier study used many fewer soundings (order-10 vs. order-1000 in our study).

Naumann et al. (2017) and Naumann et al. (2019) suggest that stronger radiative cooling is associated with a smaller h, while Zheng (2019) finds that stronger radiative cooling deepens the subcloud layer. Unfortunately, our observations do not allow us to resolve this discrepancy. We do not find a simple linear relationship between clear-sky radiative cooling and h (Fig. 4-7) or mean thermodynamics (not shown). During

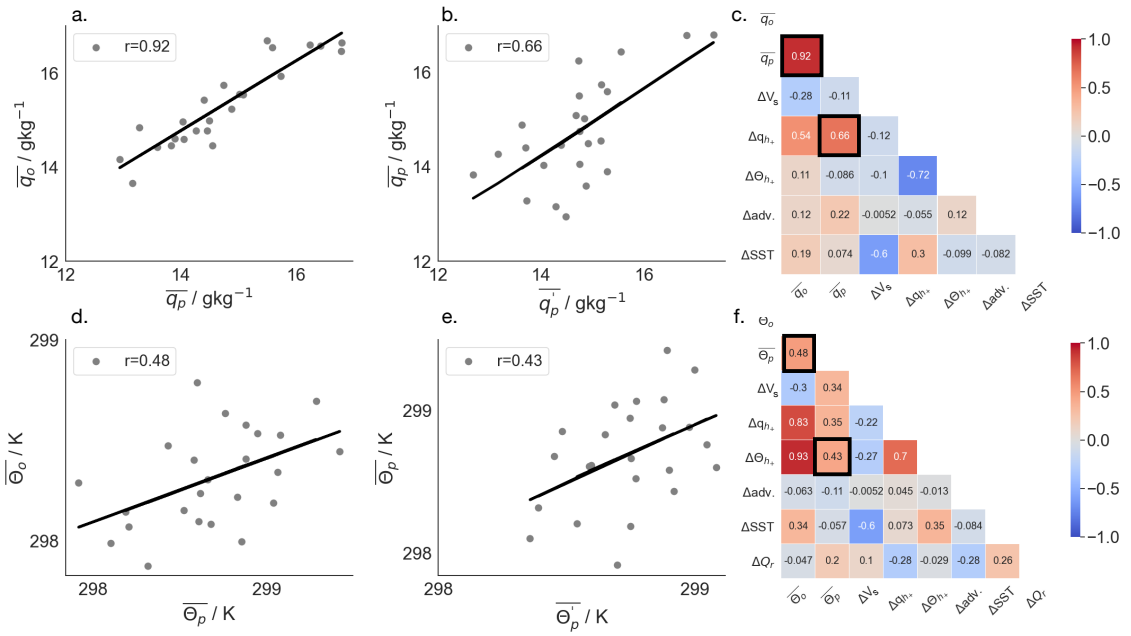


Figure 4-8: a,b,d,e. Scatter plots with linear regression and Pearson correlation coefficient. Panel (a) plots subcloud layer mean specific humidity predicted using mixed layer theory (Eq. (4.14)), $\overline{q_p}$, and observed specific humidity, $\overline{q_o}$, and similarly, panel (d) plots $\overline{\theta_p}$ from Eq. (4.15) and $\overline{\theta}$. Panels (b) and (e) compare predicted values using all terms with predicted values only varying values at h_+. Insets (c) and (f) show Pearson correlation coefficients among subcloud layer mean thermodynamics and meteorological controls. Scatter plots on the left are highlighted by black squares. Large-scale meteorological controls that vary are the near-surface wind (ΔV_s), values at h_+ (e.g. Δq_{h+}), horizontal advection (Δadv.), and sea surface temperatures (ΔSST).

the night, the wind speed tends to increase, deepening the subcloud layer, which could offset a decrease in the depth of this layer due to stronger nighttime radiative cooling. This compensation highlights the difficulty in disentangling the influence of clear-sky radiative cooling on subcloud layer properties when its variability is aliased onto variability in other variables, such as the surface wind speed.

4.8 Conclusions

In this analysis, we quantify thermodynamic variability in the trade-wind subcloud layer and test mixed layer theory using extensive *in situ* observations from the

EUREC^{4}A campaign. In defining the subcloud layer height, we find evidence for a transition layer that separates the subcloud layer top from the mixed layer top. Both heights were not clearly defined in prior literature, and we develop three definitions for calculating these heights from vertical thermodynamic profiles. Vertical gradients are small relative to variability about the mean, providing an initial justification of their omission from mixed layer theory. We find that $\overline{q}$ varies primarily day-to-day and decorrelates after about two days, such that EUREC^{4}A flight data samples nearly-independent realizations of large-scale variability. Anomalies in the subcloud layer depth and LCL are largely associated with anomalies in $\overline{q}$. Given its synoptic variability and influence on subcloud layer vertical length-scales, the primary mode of subcloud layer thermodynamic variability therefore appears to be through $\overline{q}$ variability.

The presence of a transition layer and its vertical gradients introduces ambiguity in the application of mixed layer theory. We address this uncertainty through the introduction of effective parameters estimated through a Bayesian methodology. These entrainment parameters, in particular a mean effective entrainment parameter $A_e = 0.43$ that is greater than the value of 0.2 as often assumed, are consistent reflections of a finitely-thick transition layer. Such a finitely-thick layer with thermodynamic gradients contrasts with foregoing theory based on dry boundary layers that assumes a sharp discontinuity at the layer's upper interface (e.g., Lilly, 1968), but it is consistent with direct numerical simulations (e.g., Garcia and Mellado, 2014) albeit of a dry boundary layer.

Despite the apparent challenges in applying mixed layer theory in the trades, we find that this framework offers a closed description for moisture and energy budgets, with campaign-residuals of 3.6 Wm^{-2} for moisture and 2.9 Wm^{-2} for energy and small, unbiased residuals for synoptic variability. That the budgets close to within these small residuals suggests that knowledge of the mean state through $\overline{q}$, $\overline{\theta}$, and h is sufficient to close the thermodynamic budgets, without having to include vertical thermodynamic gradients. We also find little evidence that closing moisture and energy budgets requires representing additional processes, such as precipitation or

coherent downdrafts.

After showing that mixed layer theory is a useful framework, we use this theory as a mapping between meteorological variables and subcloud layer thermodynamics. Mixed layer theory predictions of $\overline{q}$ have a correlation r=0.92 with observed $\overline{q}$, whereas predictions of $\overline{\theta}$ have a correlation r=0.48 with observed $\overline{\theta}$, consistent with more skillful closure of the subcloud layer moisture budget. We find strong linear relationships among surface wind speed variability and subcloud layer depth, surface fluxes, and entrainment fluxes. Yet due to the compensating influences on $\overline{q}$ through surface moistening and entrainment drying, the net influence of the wind speed on $\overline{q}$ is weak. Instead, only knowing moisture and temperature values above the subcloud layer has the most predictive skill for predicted $\overline{q}$ and $\overline{\theta}$, respectively, because these are the properties of the air mixed into the subcloud layer by entrainment. Clear-sky radiative cooling variability does not appear to exert a primary influence on the subcloud layer depth or fluxes compared to the stronger influence of the surface wind speed.

Given the skill of the mixed layer framework, it would be worthwhile to apply this framework to representations of the trade-wind subcloud layer by a hierarchy of models, from general circulation to storm-resolving models and large-eddy simulations. Variables analyzed could include subcloud layer moisture or energy as in this study, momentum (e.g., Holland and Rasmusson, 1973), or isotopes (e.g., Risi et al., 2020). Quantifying the relative magnitudes of different processes, how well the budgets close, and how individual terms vary according to large-scale environmental conditions would serve as a litmus test for evaluating how well models represent physical processes, such as surface and entrainment fluxes, relative to novel observational anchoring from the EUREC^{4}A field campaign.

4.9 Appendix A: Methodology for various boundary and subcloud layer height estimates

4.9.1 Thermodynamic variable gradient method

The vertical stratification of the tropical atmosphere occurs in all variables, but it is most evident in moisture (Augstein et al., 1974; Stevens et al., 2001). We first define a subcloud layer length scale as the depth over which there is no vertical gradient in specific humidity within a threshold, applying the method from Canut et al. (2012). The method selects the height where the specific humidity becomes greater than the density-weighted mean specific humidity of the levels below by a certain threshold ϵ_q: $|q(z) - \overline{q}| \leq \epsilon_q$, where $\overline{q}$ is updated at each vertical level. We begin at a height of 100 m to minimize the influence of the surface layer. This humidity-jump approach was implicitly adopted in Malkus (1958), and is similar to previous estimates based on discontinuities in observed profiles (e.g. Heffter, 1980; Marsik et al., 1995).

In implementing the q-gradient method, the primary uncertainty is the choice of threshold ϵ_q, which should be large enough not to be biased by small-scale vertical variability, but precise enough to identify the humidity discontinuity at the subcloud layer top. To choose a threshold, we turn to intensive sampling from both the CU-RAAVEN remotely-piloted aircraft and the ATR-42 and HALO aircraft. Empirically, we choose a threshold that is one-third of turbulent, eddy-scale variability, estimated as within-flight variability (compared with day-to-day variability). Calculating the specific humidity standard deviation below 550 m within a three-hourly flight of the CU-RAAVEN suggests a threshold ϵ= 0.3 gkg^{-1}. Calculating the standard deviation in q below 500m from the ATR-42 yields a threshold ϵ= 0.35 gkg^{-1}, and for HALO soundings within one flight, one-third of the standard deviation is ϵ= 0.27 gkg^{-1}. We use ϵ= 0.30 gkg^{-1}. The maximum allowable vertical gradient in the boundary layer is thus 0.03 gkg^{-1}m^{-1}, given a 10m grid spacing. This threshold allows for a certain moisture gradient, or deviation from a perfectly well-mixed profile, noted previously for both the trades and other environments (Malkus, 1958; Mahrt, 1976;

Dai et al., 2014) and shown by our analyses. Across the HALO dropsonde soundings, this 0.30 gkg^{-1} threshold corresponds to a 10% difference between mean air in the cloud and subcloud layer, when averaging air masses between 1000–1200m and 100-300m depths. We evaluate this height method and empirically-chosen gradient in Sec. 4.4. Heights from the q-gradient method are 546±82 m, with values denoting the mean and standard deviation across the 69 circle-mean data.

An advantage of this threshold definition is its straightforward application to other thermodynamic variables like θ and θ_v: $|\theta(z) - \overline{\theta}| \geq \epsilon_\theta$. We use thresholds 0.15 K for θ and 0.20 K for θ_v. The threshold of 0.2K for θ_v is also employed in Touzé-Peiffer et al. (2021). These thresholds are similarly chosen from the CU-RAAVEN, ATR-42, and HALO sounding data as one-third of one standard deviation within flights. These thresholds correspond to 10% of differences between cloud and boundary layer air (estimated conservatively as the 1000-1200m minus 100-400m layer-means) for θ and θ_v, respectively. Heights from the θ-gradient are 549±97 m and 697±94 m for the θ-gradient, with values denoting the mean and standard deviation across the 69 circle-mean data.

4.9.2 Parcel method

Next, we examine the parcel method, also referred to as the 'Holzworth method', as introduced by Holzworth (1964), which estimates the level at which a hypothetical rising parcel of surface air, representing a thermal, reaches its level of neutral buoyancy. We compute the level of neutral buoyancy where θ_v surface parcels intersect a background profile fitted to the cloud layer θ_v profile determined by linear regression. Surface air is defined as 0-50m values; choosing 0-90m air affects the height by O(1%). We calculate the cloud layer θ_v profile from 100m above the height determined from the q-gradient method to the first inversion base height, defined where the static stability first exceeds 0.1 K/hPa, similar to a definition given in Bony and Stevens (2019).

This parcel method can be viewed as a simplification of the Richardson-number method that neglects the shear contribution (e.g., Seibert et al., 2000; Dai et al.,

2014; Zhang et al., 2014). Although the Richardson and gradient Richardson number methods are related to the generation and consumption of mixed layer turbulence and diagnose flow stability (e.g., Garratt, 1994; Stull, 2012), we do not employ this method due to the considerable uncertainty underlying choices in its free parameters (e.g., Zilitinkevich and Baklanov, 2002; Seidel et al., 2012). Heights from the parcel method are 719±85 m.

4.9.3 Linearized relative humidity profile

A third type of definition involves the relative humidity profiles. The relative humidity increases throughout the subcloud layer (e.g., Nuijens et al., 2015c), as the specific humidity q remains largely constant while temperature decreases. In practice, spurious peaks in relative humidity in our circle-mean profiles could arise from the spatial averaging of multiple soundings. For instance, a circle-mean could average between profiles falling through a cloud close to the subcloud layer top, saturated in relative humidity, and drier profiles elsewhere along the EUREC^{4}A circle flight path.

To circumvent this bias, we introduce a linearization of the relative humidity profile. We find the first local maximum in relative humidity above 300m and then linearize the relative humidity profile, by ordinary least squares regression, from 50m above the surface to 50m above this first local relative humidity maximum. We then find all local relative humidity maxima below one kilometer and choose the height that minimizes the relative humidity difference between the observed and linearized profiles. Heights from the relative humidity maximum method are 571±96 m.

4.10 Appendix B: Vertical thermodynamic profiles

In Figure B, we show vertical profiles of specific humidity, q, and potential temperature, θ, to further illustrate the vertical structure as described in Sec. 4.4.

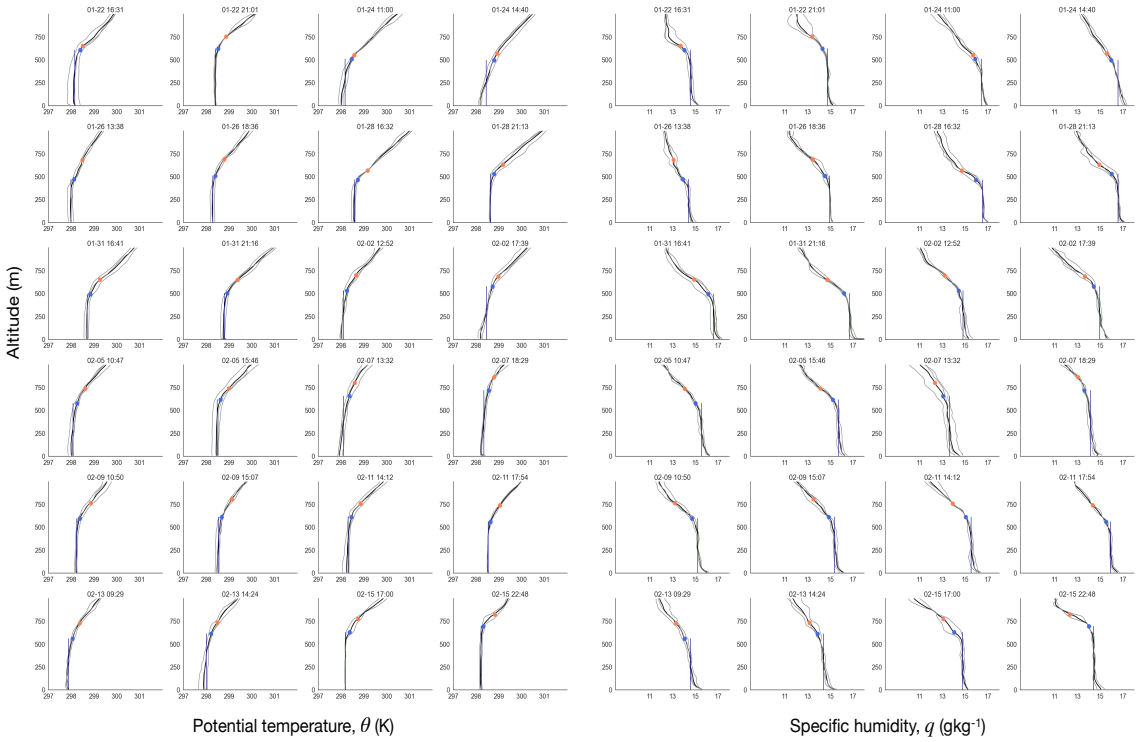


Figure 4-9: The 24 vertical circling-mean profiles (black) of specific humidity, q and potential temperature, θ, together with the three circle-mean profiles (grey) averaged to estimate each circling-mean. Blue dots correspond to the mixed layer top, estimated with the q-gradient definition, whereas red dots denote the subcloud layer top, estimated with the θ_v-gradient definition. The vertical navy line denotes the mixed layer-mean value and demonstrates that the majority of circling-mean profiles have a vertically well-mixed layer.

4.11 Appendix C: Derivation for entrainment efficiency parameter, A_e

For a subcloud layer scalar, ϑ, Eq. (4.8) results from integrating the following equation over a thin interfacial layer with lower and upper boundaries h$_-$ and h$_+$ and layer-thickness $\delta h = h_+ - h_-$,

$$\int_{h_-}^{h_+} \frac{\partial \overline{\vartheta}}{\partial t}\, dz = -\int_{h_-}^{h_+} \frac{\partial}{\partial z}\overline{w'\vartheta'}\, dz. \tag{4.16}$$

Applying the Leibniz integral rule for differentiation under integration yields,

$$\frac{d}{dt}\left[\langle\overline{\vartheta}\rangle_{\delta h}\delta h\right] - \frac{dh_+}{dt}\vartheta_{h+} + \frac{dh_-}{dt}\vartheta_{h-} = -\overline{w'\vartheta'}|h_+ + \overline{w'\vartheta'}|h_- \tag{4.17}$$

Assuming that turbulence vanishes at h_+ implies that $\overline{w'\vartheta'}|h_+ = 0$. Heights h_+ and h_- are offset from h by a constant, ϵ, such that h_+=h+ϵ and h_-=h-ϵ, which means that $\frac{dh_+}{dt} = \frac{dh_-}{dt} = \frac{dh}{dt}$. Assuming that the layer is well-mixed implies that $q_{h-} = \overline{q}$. With these assumptions, Eq. (4.17) becomes,

$$\frac{d}{dt}\left[\langle\overline{\vartheta}\rangle_{\delta h}\delta h\right] - \frac{dh}{dt}\Delta\vartheta = \overline{w'\vartheta'}|h_- \tag{4.18}$$

If δh is constant and $\langle\overline{\vartheta}\rangle_{\delta h}$ is approximately equal to $\frac{\overline{\vartheta}+\vartheta_+}{2}$, with $\overline{\vartheta}$ equaling the mixed layer mean value then Eq. (4.18) becomes,

$$\frac{\delta h}{2}\left[\frac{d\overline{\vartheta}}{dt} + \frac{d\vartheta_+}{dt}\right] - \frac{dh}{dt}\Delta\vartheta = \overline{w'\vartheta'}|h_- \tag{4.19}$$

Note that if assuming that the interfacial layer has zero-thickness, e.g. $\delta h = 0$, Eq. (4.19) becomes,

$$-\frac{dh}{dt}\Delta\vartheta = -E\Delta\vartheta = \overline{w'\vartheta'}|h_-, \tag{4.20}$$

where the growth of the layer, $\frac{dh}{dt}$ is considered the entrainment rate, E.

Replacing ϑ with θ_v yields Eq. (4.8),

$$\frac{\delta h}{2}\left[\frac{d\overline{\theta_v}}{dt} + \frac{d\theta_{v+}}{dt}\right] - \frac{dh}{dt}\Delta\theta_v = \overline{w'\theta_v'}|h_- \tag{4.21}$$

Rearranging to solve for $E = \frac{dh}{dt}$ yields and adopting the formulation for the flux, $\overline{w'\theta_v'}|h_-$, given in Eq. (4.2) and Eq. (4.7) yields,

$$E = \frac{-AV_0\Delta_0\theta_v}{\Delta_1\theta_v} + \frac{\delta h}{2\Delta_1\theta_v}\left(\frac{d\overline{\theta_v}}{dt} + \frac{d\theta_{v+}}{dt}\right). \tag{4.22}$$

The effective A_e can be interpreted as absorbing the second term on the right-hand side in Eq. (4.22).

Chapter 5

A new conceptual picture of the transition layer

5.1 Introduction

The transition layer in the trades has long been observed (e.g., Malkus, 1958; Augstein et al., 1974; Yin and Albrecht, 2000) and simulated (e.g., Stevens et al., 2001), but previous studies have not investigated its origins. This layer is often associated with an approximately 200 m deep layer between the well-mixed part of the subcloud layer (around 500 m depth) and the subcloud layer top (around 700 m depth) (e.g., Malkus, 1958; Betts, 1976; Augstein et al., 1974; Arakawa and Schubert, 1974; Yin and Albrecht, 2000). The transition layer is typically identified from thermodynamic soundings by its vertical gradients that extend over a certain depth and are stronger than those in the cloud layer above or the mixed layer below. Its top is also often associated with the mean cloud base level estimated from the lifting condensation level (LCL) of surface parcels (e.g., Malkus, 1958; Augstein et al., 1974).

From these recognitions, previous studies developed a conceptual view that that the transition layer is predominantly a feature of cloud-free regions, and clouds, if they occur, are rooted at the transition layer top. Malkus (1958), for instance, examines vertical gradients to conclude that the transition layer was always (100%) present in clear-sky regions and generally (55%) absent in cloudy regions. She also proposes that the transition layer is thicker in clear areas (200 m) than in cloudy areas (80 m),

This chapter is in preparation for submission to the *Journal of the Atmospheric Sciences*.

as illustrated schematically in Fig. 5-1a. Augstein et al. (1974) goes a step further to outline a qualitative scheme (his Fig. 12) that the transition layer is maintained by dry convection and mechanically-driven turbulence, whereas moist convective processes play a role above the transition layer top.

The inferences of strong vertical gradients over the transition layer, moreover, established a conceptual picture of the transition layer as a barrier to convection, or a valve that regulates subcloud to cloud layer transports (e.g., Augstein et al., 1974; Yin and Albrecht, 2000; Neggers et al., 2006). In this view, buoyant updrafts must be sufficiently energetic to overcome the stable transition layer, similar to the notion of convective inhibition. Earlier, Ooyama (1971) presented the 'cumulus dispatcher function', which represents the probability that a buoyant updraft in a statistical ensemble successfully forms a cloud. The probability of success depends on the large-scale environment near cloud base – that is, on environmental conditions in the transition layer. Neggers et al. (2006) expresses a similar idea of the transition layer influencing the coupling between subcloud and cloud layer processes, noting that "the bulk subcloud-layer properties that determine the saturation characteristics of the transition layer, and thus the area fraction of shallow cumulus, act as a regulator or valve on the moist convective transport".

In the intervening years from these early observational studies, a different interpretation arose that the transition layer could be modeled as an infinitely-thin layer with 'jumps', or abrupt discontinuities, in analogy with the cloud-free convective boundary layer (e.g., Lilly, 1968; Stull, 1976; Stevens, 2006). A schematic of a cloud-free boundary layer (Fig. 5-1b) and results from a direct numerical simulation of a dry boundary layer (Fig. 5-1c) highlight a layer that is well-mixed by turbulence, topped by an abrupt discontinuity. Neglecting cloudiness was rationalized by the small cloud-base cloud fraction in the trades. Indeed, trade-wind cloud fraction near cloud-base was estimated to be about 3-4% as by the ATR-42 aircraft during EUREC^{4}A (Bony et al., 2022), such that this assumption appears reasonable.

Closure schemes for the entrainment rate based on cloud-free layers are features of cumulus parametrizations in the seminal Arakawa and Schubert (1974) study, as

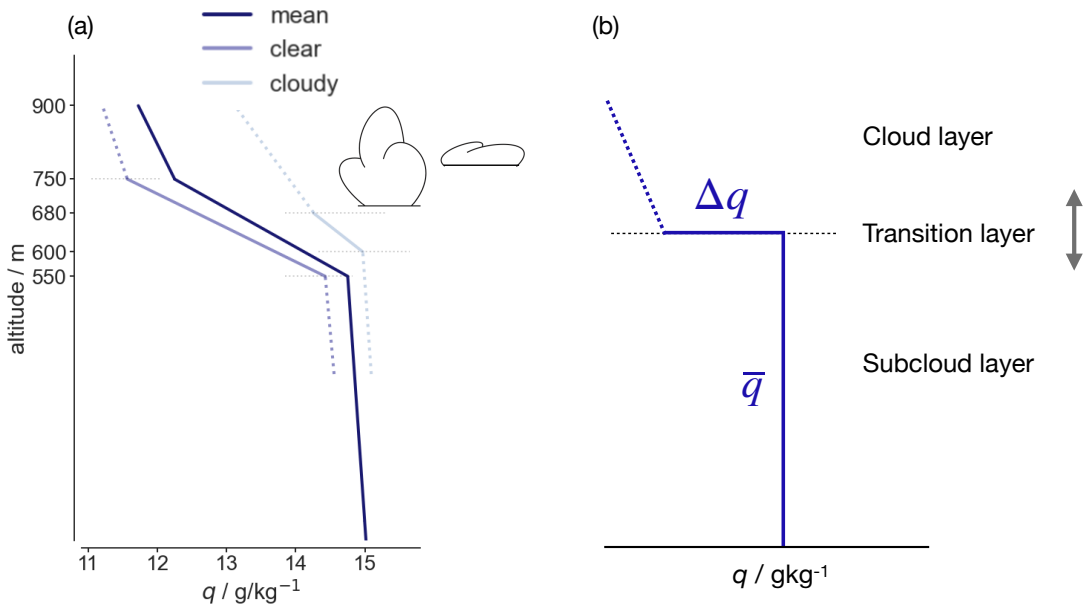
well as the Betts (1973) and Albrecht et al. (1979) representations of shallow cumulus fields. The conceptual view of entrainment in a cloud-free boundary layer is that overshooting plumes entrain filaments of more stratified overlying air into the turbulent layer (e.g., Bretherton, 1997). Such a formulation suggests that clouds are not explicitly agents of subcloud layer turbulent mixing, and turbulence and any associated entrainment are instead generated by surface fluxes.

During EUREC^{4}A, we collected data to better understand the characteristic vertical structure of the trades, as introduced in Chapt. 1. In Chapt. 4, the trade-wind subcloud layer is shown to have a nuanced vertical structure that includes a finite-depth transition layer between the well-mixed part of the subcloud layer and subcloud layer top. That is, extensive soundings from EUREC^{4}A show that most thermodynamic profiles do not exhibit the jump that is expected from cloud-free boundary layers and used as the conceptual basis for mixed layer theory.

There are different reasons a finite-thickness transition layer with smooth vertical gradients, as opposed to a jump, could be evident in thermodynamic profiles. Averaging individual profiles that show a jump, albeit at different heights, would yield a finite-depth layer with vertical gradients simply as a product of averaging. Yet the transition layer and lack of a clear discontinuity at the boundary layer top exist in individual soundings, as will be discussed in this chapter, disproving this first idea. A second hypothesis is that a transition layer exists as a structural component of the atmosphere and has finite thickness. The question then arises, what processes maintain such a transition layer in the trades? This chapter provides a first answer to this question.

5.2 EUREC^{4}A data

The primary data are similar to those used in the previous chapter, in particular 810 dropsondes from the High Altitude and Long Range Research Aircraft (HALO) launched between January 22, 2020 and February 15, 2020 (Konow et al., 2021). These dropsondes provide vertical profiles of pressure, temperature, and relative hu-



(c) Reproduced from Garcia, Mellado, 2014

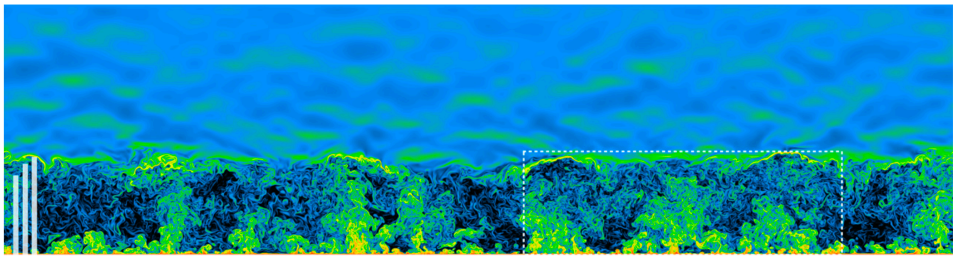


FIG. 2. Vertical cross section of the CBL showing the logarithms of the magnitude of the buoyancy gradient $N^{-2}|\nabla b|$ for case Re100 at the final time $z_{enc}/L_0 = 18$. Colors black, blue, yellow, and red correspond, respectively, to values varying between 10^{-1} and 10^2 in powers of 10. Color scale is as Fig. 6. The tops of the three white vertical bars indicate the location of the different CBL height definitions, from left to right: the encroachment height z_{enc}, the flux-based height $z_{i,f}$, and the gradient-based height $z_{i,g}$.

Figure 5-1: (a) Representing specific humidity profiles from Malkus (1958) with the heights and gradients given in her study. The mean profile (dark blue) is the weighted average between 16 clear-sky (averaged in light blue) and 9 cloudy (averaged in medium blue) soundings. According to her view, clouds are rooted at the transition layer top. (b) An idealized profile of specific humidity, similar to idealized profiles such as in Albrecht et al. (1979) for moist static energy. That the infinitely-thin transition layer can vary in height is indicated by the grey arrow. (c) Figure reproduced from Garcia and Mellado (2014), showing results from a direct numerical simulation of a dry convective boundary layer. Shown is the logarithm of the buoyancy gradient, which acts similarly to the θ_v gradient. Colors correspond to increasing values (e.g., from black to blue to green).

midity, which have been processed and interpolated into a common altitude grid with 10 m vertical resolution (George et al., 2021). As described in Chapt. 4, dropsonde measurements were distributed along the 'EUREC^{4}A circle'. The EUREC^{4}A circle is defined by a circular flight pattern with an approximately 220-kilometer diameter, centered at 13.3°N, 57.7°W. This flight pattern was repeated 69 times, over 12 flights. Typically each flight incorporated two – temporally well separated – periods of circling. A 'circling-mean' is defined as the mean of three 'circle-means', which each average about 12 dropsondes along the EUREC^{4}A circle. Given that measurements did not target specific meteorological conditions they provide unbiased sampling. The structure of the data collected encourages the definition of 69 circle-means, 24 circling-means, and one campaign-mean value.

We also use ceilometer cloud base height estimates from the R/V Meteor and the Barbados Cloud Observatory (BCO) at 10-second resolution from January 19, 2020 to February 19, 2020. Data during night times that the HALO aircraft did not fly are dropped. Cloud base heights vary, and to estimate the base of clouds forming from updrafts within the subcloud layer, ceilometer data between 350 and 1000 m are analyzed. These data span the range of the mixed layer lifting condensation levels. We bin data into three-hourly segments and select the most frequently-sampled value as a representative cloud base. Typically, the first peak corresponds to the distribution peak. In the cases where they differ, we select the first peak that is within 50% of the absolute distribution peak. The first peak is chosen because the first decile is biased by rain, whereas higher deciles increasingly reflect cloud side detection from sheared convection or decaying cloud fragments that are not indicative of cloud base (Nuijens et al., 2014). Example three-hourly cloud base distributions are given in Fig. 5-2 to illustrate the methodology. In the following analysis, the ceilometer cloud base height distribution refers to the aggregate of distribution peaks from three-hourly cloud base height data.

Cloud *top* height data are taken from the WALES (*WAter vapor Lidar Experiment in Space*) instrument, a water vapour differential absorption lidar. This lidar operates at four wavelengths around 935 nm to measure water vapor mixing ratio

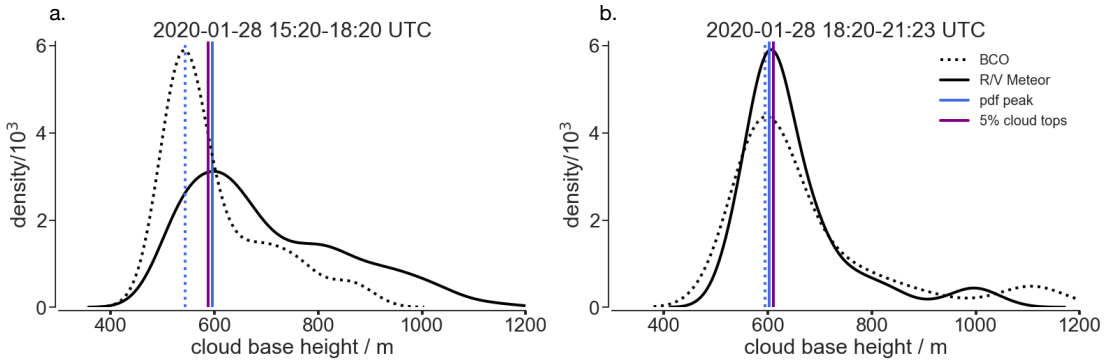


Figure 5-2: Example three-hourly cloud base height distributions from the R/V Meteor (solid) and the Barbados Cloud Observatory ceilometers (dotted), annotated with the first distribution peak (blue vertical line). Also shown is the 5% cloud top height estimate from WALES (purple vertical line) for the same time interval.

profiles below HALO (Wirth et al., 2009; Konow et al., 2021). An additional High Spectral Resolution Lidar (HSRL) channel at 532 nm allows for the retrieval of the atmospheric transmission. The backscatter data from the HSRL has a resolution of 40 m in the horizontal (corresponding to temporal resolution of 5 Hz and a typical 200 m/s flight speed) and 15 m in the vertical. Data products include a cloud flag inferred from the lidar backscatter ratio at 532 nm, cloud top height with both a precision and accuracy of about 10 m, and optical depth between the cloud top and subcloud layer top (Konow et al., 2021). By way of comparison, even high-resolution satellite retrievals from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) instrument have a vertical uncertainty of 250–500 m on cloud top height estimates, similar to the uncertainty in Mieslinger et al. (2019). Cloud top height estimates from these lidar data with their approximately 10 m accuracy are thus particularly valuable.

5.3 A different conceptual picture emerges

5.3.1 Transition layer depth

The expectation from cloud-free boundary layers is a single layer that varies in height (e.g., as illustrated in Fig. 5-1b). Sampling the undulation of this layer would result in a single Gaussian height distribution. To test this idea, we perform a vertical length-scale analysis of the convective boundary layer. The methods for calculating various heights are described in the previous chapter and its appendices. Contrary to expectations from cloud-free boundary layers, Fig. 5-3a shows that two height distributions exist even in individual soundings. A bimodal distribution is also evident in more aggregated data, such as EUREC^{4}A circle-mean and circling-mean data (see Fig. 4-2).

We associate the first distribution with a well-mixed layer in q and θ, having a mean depth of 500 m for individual soundings and 570 m for circle-mean data (Fig. 5-3). This layer appears to be homogenized by surface-flux driven eddies. We call this height the mixed layer top. The buoyancy variable, θ_v, is, however, approximately well-mixed to a mean depth of 710 in individual soundings and 708 m in circle-mean data. We call this depth the subcloud layer top. These heights are generally consistent with what Malkus (1958) and Augstein et al. (1974) identified as the mixed layer and subcloud layer top depths.

We hypothesize that the offset between the mixed layer and subcloud layer top indicates the presence of a transition layer, over which thermodynamic variables begin to have vertical gradients. Using individual dropsondes and taking the difference between the top of the subcloud layer and mixed layer distributions, we find a transition layer thickness of 180±207 m, with the values denoting the mean and standard deviation. The mean value appears to converge towards a mean depth of 150 m, and, as expected, the standard deviation decreases with increasing levels of aggregation from individual sondes to circle- and circling-mean data (151±77 m for circle-mean and 152±50 m for circling-mean data). The transition layer therefore appears to have

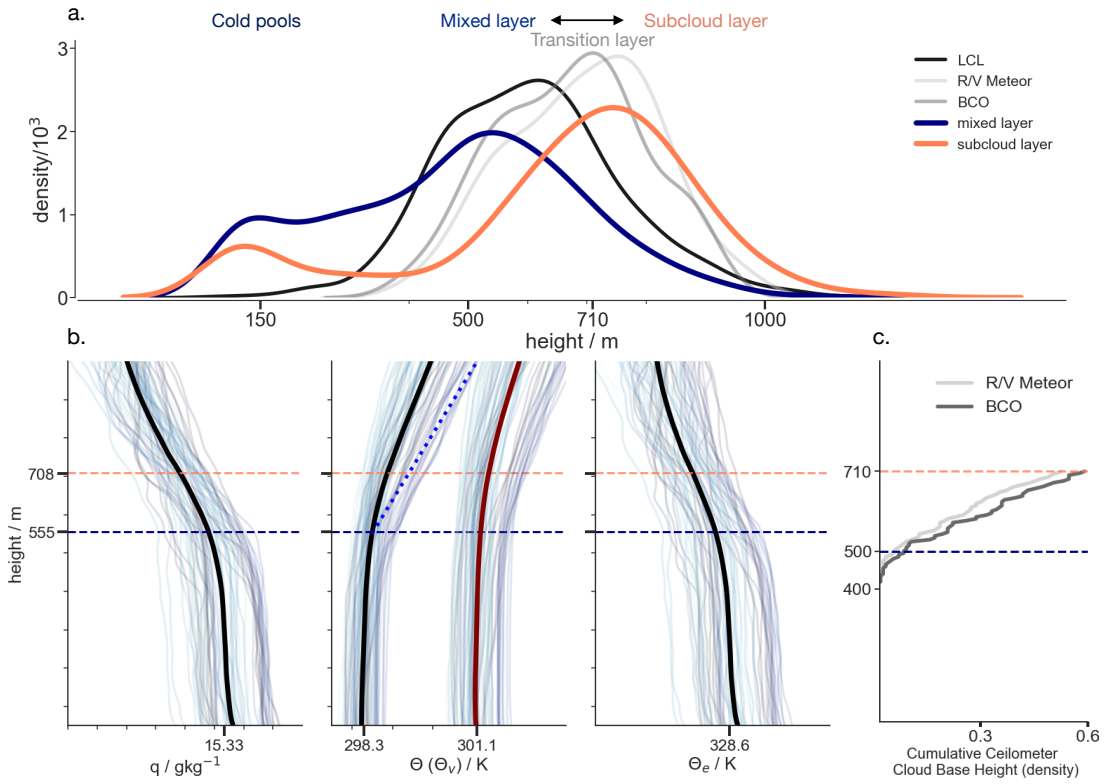


Figure 5-3: (a) distributions of different methods employed to estimate the mixed and subcloud layer heights for 810 individual dropsondes. We find that three methods based on specific humidity q or potential temperature θ individually and relative humidity (averaged in the blue curve) correspond to the mixed layer, whereas the θ_v-gradient and parcel method based on θ_v, a proxy for buoyancy (averaged in the orange curve) correspond to what is often called the subcloud layer height. Also shown are the lifting condensation level averaged from 100–300 m air parcels (black) and ceilometer cloud base height estimates (grey). Cold pool soundings correspond to the distribution around 150 m height. (b) 69 circle-mean profiles from HALO for q, θ, θ_v, and θ_e. The black line is the time-mean across all profiles, and colored profiles correspond to time (moving from darker to lighter blue). Dotted lines mark the mixed layer height calculated using circle-mean data (blue) and subcloud layer height (orange). Their difference indicates the presence of a transition layer. Note that the mixed layer and subcloud layer height values in panel b. are calculated from circle-mean profiles, explaining the difference in heights with panel a. (c) Empirical cumulative distribution function (CDF) of aggregated ceilometer cloud base height distributions measured by the R/V Meteor and at the BCO. In this aggregated distribution, each cloud base height value corresponds to first distribution-peak from three-hourly distributions, corresponding to the blue lines in Fig. 5-2 and as described in Sec. 5.2.

about 150–200 m depth, within the range given in previous observational studies (e.g., Malkus, 1958; Augstein et al., 1974; Yin and Albrecht, 2000), but contrasting with the thin transition layer view from modeling.

5.3.2 Less stable layer than in previous studies

Observed thermodynamic profiles exhibit smooth vertical gradients over the transition layer (Fig. 5-3b). The θ_v (buoyancy) gradients, however, appear relatively weak because q and θ vertical gradients have compensating effects on θ_v, as also pointed out by Yin and Albrecht (2000). Previous conceptualizations of the transition layer, however, posit that the transition layer acts as a barrier or cap to convection (Sec. 5.1), which would suggest a relatively strong θ_v gradient.

The transition layer gradients given Malkus (1958) and Augstein et al. (1974) are a useful point of comparison. Fig. 5-4 shows that transition layer θ_v gradients observed during EUREC^{4}A are, on average, weaker than those in Malkus (1958) and Augstein et al. (1974). Specific humidity gradients are similar, and the weaker θ_v gradient in EUREC^{4}A is driven by a weaker θ gradient. We speculate that the reason for the greater stability implied by the mean profiles in Malkus (1958) and Augstein et al. (1974) is their smaller sample size and that their sampling was not entirely unbiased, compared to the extensive, unbiased sampling in EUREC^{4}A. Malkus (1958), for instance, launch 16 out of their 25 soundings into very clear-sky regions, whereas the other nine soundings explicitly targeted active cloud cores. Augstein et al. (1974) analyze a larger set of soundings from field campaigns in 1965 and 1969. He, however, removes soundings wherein a transition layer was not apparent, which could bias results towards stronger transition layer gradients. During EUREC^{4}A, strong gradients with the magnitude of those in Malkus (1958) and Augstein et al. (1974) are seen, but they occur infrequently (Fig. 5-4).

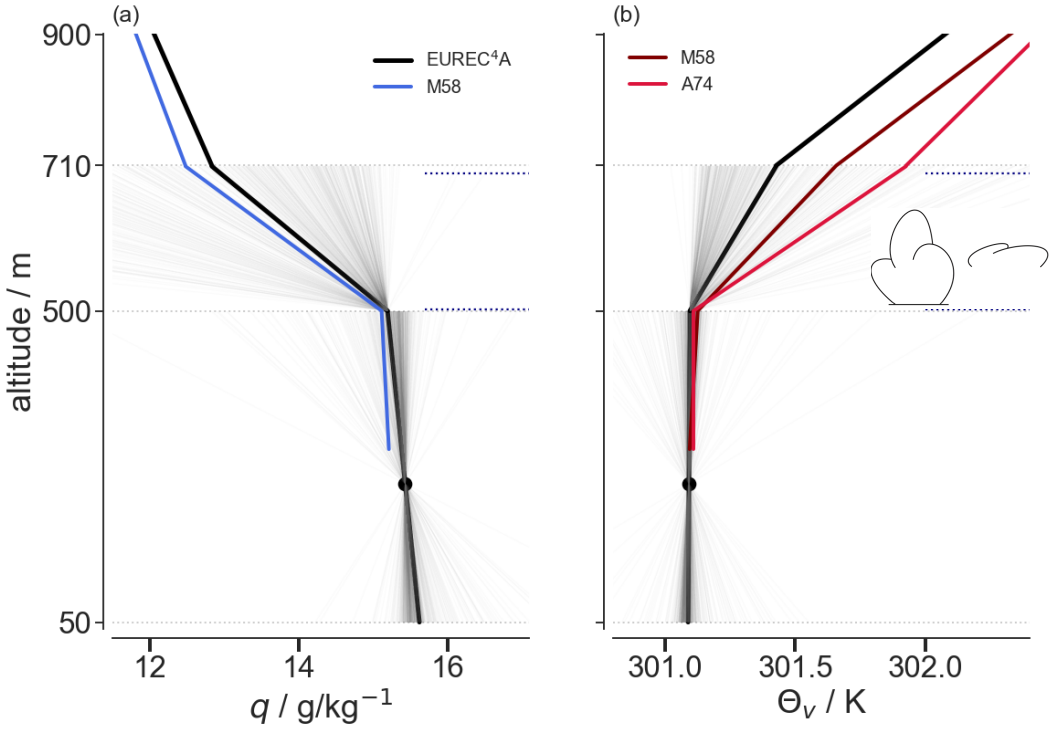


Figure 5-4: Composite profiles of (a.) specific humidity, q, and (b.) virtual potential temperature, θ_v. The profiles are constructed from the mean mixed layer value, mean gradients over the mixed layer, transition layer, and lower cloud layer, and mean heights of the mixed layer and transition layer top. Heights are estimated from EUREC^{4}A sounding data and are generally consistent with those in Malkus (1958) and Augstein et al. (1974). EUREC^{4}A data are plotted in black, both the campaign-mean (thick black line) and individual dropsondes (thin black lines), with the mixed layer mean value indicated by the black dot. Colorful profiles use vertical gradients from previous studies, but with the mean value adjusted to the mean mixed layer value in EUREC^{4}A to better compare transition layer gradients. M58 corresponds to Malkus (1958), and A74 refers to Augstein et al. (1974). θ_v gradients are given in Augstein et al. (1974), but for Malkus (1958), θ_v values are calculated from observed temperature, mixing ratio, and pressure profiles in her Fig. 7. Based on our view, clouds are shown as rooted at the mixed layer top, rather than the transition layer top as in Augstein et al. (1974), but for Malkus (1958). Also shown are the 25$^{\text{th}}$- and 75$^{\text{th}}$-percentile of lifting condensation levels (LCL) calculated from individual dropsondes, averaging LCL values from 100–300 m air parcels (blue dotted lines).

5.3.3 Clouds rooted within transition layer

Another difference with Malkus (1958) and Augstein et al. (1974) is that we find that clouds are already rooted within the transition layer, instead of at its top. Decreasing θ_e over the transition layer (Fig. 5-3b), a necessary but not sufficient condition for a conditionally unstable layer, hints at a potential role for convective processes in this layer.

Examining ceilometer cloud base estimates, about 60% of cloud bases occur below what is typically called the subcloud layer top height around 700 m. A cumulative distribution of these cloud base heights is given in Fig. 5-3c, showing that the R/V Meteor has 61% of cloud bases below 710 m, and the BCO has 55% of cloud bases below 710 m. Below 500 m, approximately the mixed layer top height, 9% clouds at BCO and 6% of clouds measured by the R/V Meteor ceilometer already have their bases.

We therefore conjecture that the transition layer is populated by shallow clouds, and that cloud-mediated mixing processes shape its structure. While many clouds may continue to grow above the transition layer, a fraction of clouds both form and dissipate within the transition layer. When these clouds form, they warm and dry the ambient environment and when they dissipate, they cool and moisten the large-scale environment, such that this air takes on properties that more closely resemble air in the mixed layer. Such cloud-driven processes could 'precondition' the surrounding air and reduces the work to entrain more-buoyant air into the mixed layer. Viewed from the point of thermals and cloud formation, reducing the $\Delta\theta_v$ jump by cooling and moistening ambient air reduces the barrier to convection and thus the updraft velocity required by a thermal to continue ascending and form a cloud. We thus hypothesize that the structure of the transition layer is an important way in which the cloudy boundary layer differs from a cloud-free, or dry boundary layer. This hypothesis can directly be tested using EUREC^{4}A observations.

5.4 Physical origins of transition layer structure

We first test the idea that the presence of clouds changes the transition layer structure through a 'denial of mechanism' approach. A distinction is made between a clear convective boundary layer, or a convective boundary layer without clouds at its top over scales much larger than the depth of the layer (e.g., cloud-free over scales greater than 10 km), as compared to cloud-topped convective boundary layers. Large cloud-free areas, rather than simply clear-sky areas in between clouds, are selected because the area between clouds is still materially influenced by cloud condensate detrainment and mixing. We identify such large cloud-free patches in two ways: first, by eye, from GOES-E satellite movies at one one-minute resolution overlain with dropsonde locations and times (Bony et al., 2022); and second, in Sec. 5.4.3, using the cloud flag product from WALES lidar as introduced in Sec. 5.2.

Fig. 5-5 illustrates two case studies for January 31, 2020 and February 2, 2020 with large cloud-free areas. In each case, GOES-E satellite movies are used to identify one dropsonde from a clear convective boundary layer and another sonde that is more influenced by clouds. The cloud-free sondes exhibit the characteristic structure of a clear convective boundary layer, with a well-mixed layer topped by a jump. In the non-clear-sky case, a discontinuity is not visible, and instead there are smoother vertical gradients. The influence of clouds on the surrounding environment is also suggested by the bottom dropsonde profile in Fig. 5-5a, which does not necessarily fall within a cloud, but nonetheless does not display a canonical clear-sky profile.

The expected difference between clear and cloudy profiles holds in these two cases, but how well does the distinction generalize across the 810 HALO soundings? To this end, we find the distribution of the maximum vertical gradient (over 10 m) in all dropsonde vertical profiles of specific humidity, q, between 100–900 m. This metric captures the strongest jump that is evident in an observed boundary layer moisture profile. Results are similar for different choices of lower and upper bound. The lower bound of 100 m is chosen to avoid possible surface layer gradients, and 900 m acts as a conservative estimate of the mixed layer top. The vertical layering of

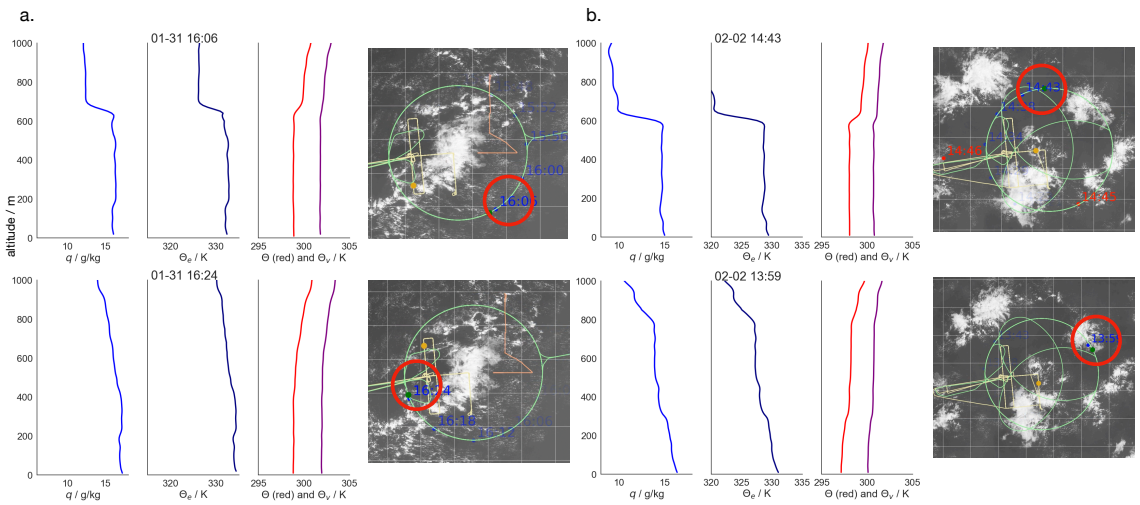


Figure 5-5: Two pairs of sondes launched in large clear-sky patches (top panels) and areas influenced by cloudiness (bottom panels), and their corresponding GOES-E satellite images with the sonde time (listed above the profiles) and sonde location (circled in red). Sondes are launched on January 31, 2020 (a.) and February 2, 2020 (b.). Profiles are specific humidity, q (medium blue), equivalent potential temperature, θ_e (dark blue), potential temperature, θ (red), and potential temperature, θ_v (purple). Clear-sky profiles exhibit a stronger jump at the mixed layer top than do profiles influenced by clouds, which tend to have smoother vertical thermodynamic gradients.

the atmosphere is present in numerous variables, but particularly evident in moisture (e.g., Augstein et al., 1974), motivating our choice of q, though results are similar for other thermodynamic variables.

Fig. 5-6 shows the majority of soundings have small values of this first difference metric, corresponding to smooth gradients at the mixed layer top. The 25$^{\text{th}}$-percentile and median values of the maximum vertical gradient are 0.17 and 0.25 gkg^{-1} compared to a standard deviation of 1.06 gkg^{-1} for q averaged from 100–500 m. As described in Sec. 4.7, mixed layer theory allows for predicting the mean subcloud layer specific humidity, $\overline{q}$, by varying only certain factors, but keeping other parameters fixed at their campaign-mean value. The predicted $\overline{q}$ from only varying the 10 m wind speed and surface saturation specific humidity, indicative of variations in the surface latent heat flux, has a standard deviation of 0.61 gkg^{-1} across the campaign. These comparisons indicate that the median gradient value at the boundary layer top

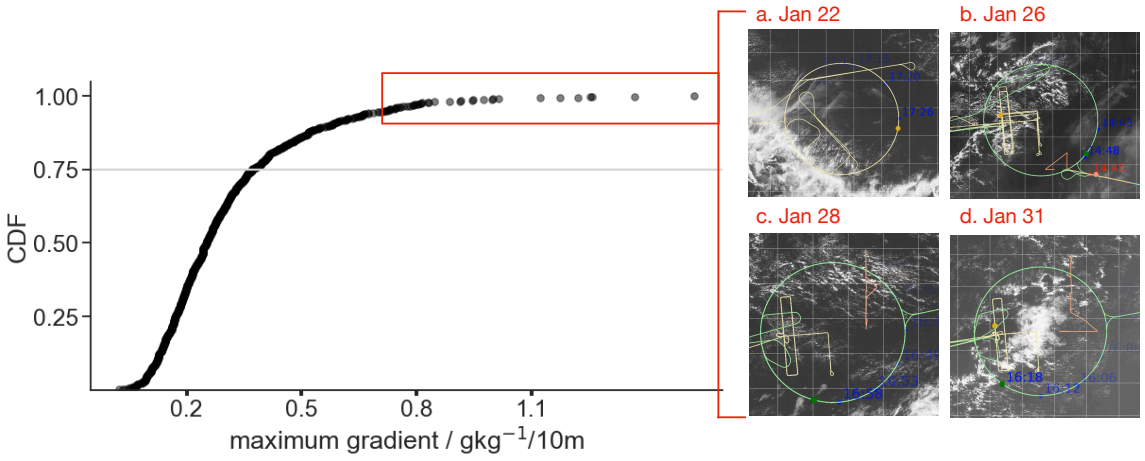


Figure 5-6: Empirical cumulative distribution function (CDF) for vertical gradients (across 10 m) calculated for specific humidity, q, profiles between 100–900 m. The red box highlights the largest-5% vertical gradients. On the right are four satellite snapshots from GOES-E corresponding to days most frequently represented in these largest-5% values and illustrating large clear-sky swaths.

is significantly smaller than variations about the mean, even when the variations are driven only by surface fluxes.

Examining vertical gradients above the 95$^{\text{th}}$-percentile, we find that larger values systematically occur in large clear-sky areas. Fig. 5-6 shows satellite images for four days whose sondes make up 56% of the largest-5% vertical gradient values. On these days, large cloud-free areas frequently extend across the EUREC^{4}A circle. This analysis provides an initial indication of an association between sharp gradients in thermodynamic profiles and large, cloud-free areas. That is, the cloud-free convective boundary layer exhibiting an abrupt discontinuity at its top appears to be the exception that makes the rule, rather than the typical vertical structure.

5.4.1 What maintains cloud-free regions?

A subsequent question is what physical processes create such clear skies. Subsiding branches of shallow circulations are expected to be associated with large dry areas (e.g., Naumann et al., 2019). Qualitatively, sondes with gradient values exceeding the 95$^{\text{th}}$-percentile tend to be associated with mesoscale subsidence (not shown). This vertical velocity is calculated at the circle-scale rather than at the scale of individual

dropsondes, making it difficult to associate individual sondes within a large-scale vertical motion value.

Cold pools could, conceivably, give rise to canonical dry layer profiles as illustrated in Fig. 5-5. Cold pools suppress convection in their cold interior, but promote convection on their edge associated with a gust front (e.g., Rochetin et al., 2021). We find, however, that only a single sounding of the largest-5% vertical gradients is identified as a cold pool sounding using the method in Touzé-Peiffer et al. (2021). The influence of a cold pool could, however, be temporally delayed, such that the cold pool passage both enhances surface sensible heat fluxes (larger temperature difference between the surface and mixed layer and and stronger wind speed at the gust front edge) and suppresses convection due to the colder mixed layer value, giving rise to the canonical dry boundary layer structure. With the dropsonde data, it not possible, however, to establish whether a clear-sky patch was due to the lingering influence of a cold pool, or simply restoration to a background environmental state, such as of large-scale subsidence, independent of the cold pool's passage.

5.4.2 Two populations of clouds

A population of clouds is identified that we hypothesize is responsible for changing vertical gradients in the transition layer relative to cloud-free conditions. Fig. 5-7 shows the distribution of WALES cloud top heights. This distribution is bimodal, with peaks around 850 m and 1900 m. There is an apparent scale separation around 1300 m. The first peak is associated with shallow, likely non-precipitating clouds, and the second peak is associated with deeper, potentially precipitating clouds and stratiform clouds resulting from detrainment near the trade-wind inversion around 2–3 km. Fig. 5-7 also shows that the first cloud peak can be well-described by an exponential distribution. An exponential distribution would be consistent with cloud formation following a Poisson process, or forming continuously, independently, and at a constant rate. The exponential fit cannot, however, explain the second cloud top height distribution, and it appears to overestimate the fraction of cloud top heights near the scale break around 1300 m, suggesting that deeper cloud formation violates

the Poisson process assumptions, for instance that clouds do not form independently, as numerous updrafts may aggregate to form a deeper cloud.

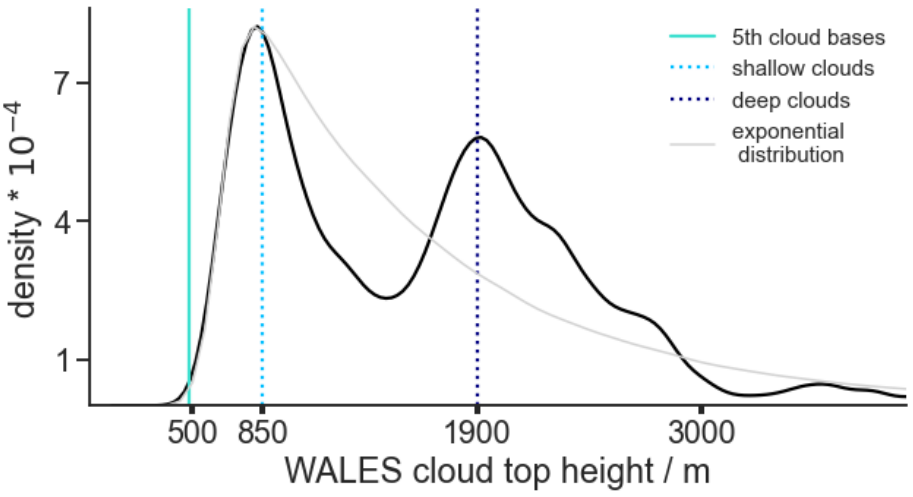


Figure 5-7: Distribution of cloud top heights estimated from WALES lidar data as described in Sec. 5.2. The distribution is bimodal, with peaks around 850 m and 1900 m, corresponding to shallow and deeper clouds. The grey line plots the best-fit exponential distribution fit to clouds with tops below 1900 m. The solid turquoise line marks the 5$^{\text{th}}$ cloud base heights from R/V Meteor and BCO ceilometers, giving a lower bound of cloud bases.

The bimodality in cloud top heights appears to be a robust feature of trade-wind cumuli (e.g., Genkova et al., 2007; Leahy et al., 2012b; Mieslinger et al., 2019). Genkova et al. (2007), for instance, use various remote sensing retrievals and find peaks with maxima at 650 m and 1500 m. These peaks are similar to ours, although their estimate had a vertical uncertainty of 250–500 m, highlighting that the WALES data with its 10 m vertical resolution is more amenable to such a fine-scale analyses. Malkus (1958) already implicitly represents this bimodality in cloud top heights (see her Fig. 1 and Fig. 2, reproduced in Fig. 5-8), though she did not explicitly discuss a bimodality in cloud top heights given the lack of data to analyze this question at the time. The two populations could, moreover, reflect the long-established notion of 'active' clouds, associated with thermals and cloud cores, and 'passive' clouds, sometimes referred to as forced convection, or non-buoyant thermals that overshoot their lifting condensation level. Radar measurements often miss the smaller, optically-thin passive clouds and only capture larger, active clouds, whereas lidar measurements

like WALES better sample both active and passive clouds (e.g., Jacob et al., 2020; Klingebiel et al., 2021).

A back-of-the-envelope scaling provides further evidence that the transition layer is populated by shallow clouds. In the WALES data, shallow clouds are more frequent than deeper clouds, and they appear to vary relatively little in height, with a standard deviation of 184 m for cloud tops below 1.3 km (Fig. 5-7). As a heuristic example, consider the 5^{th}-percentile of cloud bases around 500 m, which give a lower bound of cloud bases and roughly correspond to the mixed layer top from individual dropsondes. The transition layer extends about 150–200 m above this mixed layer top. The mean shallow cloud top height is 850 m, or about 350 m above the mixed layer top. Shallow clouds therefore appear to be rooted in the transition layer and grow only a few hundred meters above this layer. Viewed differently, the distributions in Fig. 5-7 show that 14% of clouds have their tops below 800 m, and 29% of all clouds have their tops below 1000 m. Selecting only the shallow clouds below 1.3 km (first peak in Fig. 5-7), 35% of cloud tops are below 800 m and 72% are below 1000 m.

5.4.3 Shallow clouds appear responsible for transition layer structure

To more formally test how shallow clouds shape the transition layer structure, we revisit the vertical gradient distribution using WALES cloud top estimates. The goal is to select large clear-sky swaths that are relatively free from cloud influences. Using the WALES cloud flag and cloud top heights, measurements are separated into three categories: large clear-sky areas, cloudy areas with cloud tops below 1.3 km (shallow clouds, e.g. Fig. 5-7), and all areas that are not large clear-sky areas, including cloudy and smaller clear-sky areas. This separation allows us to test whether the presence of shallow clouds is sufficient to change the transition layer structure from the dry boundary layer case, independent of the influence of deeper clouds.

Large clear-sky areas are selected by first identifying all clear-sky segments using the cloud flag, calculating the 95^{th}-percentile of segment lengths, and then considering

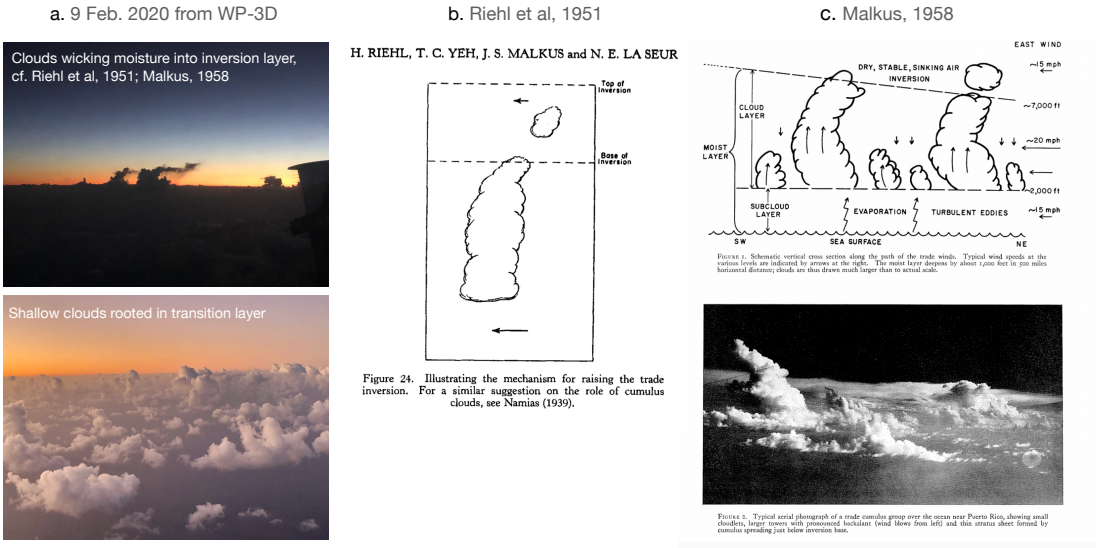


Figure 5-8: In column a., photos I took during a EUREC^{4}A flight in the WP-3D aircraft that illustrate shallow clouds rooted in the transition layer (bottom), and deeper clouds wicking or injecting moisture to maintain the inversion layer (top). Riehl et al. (1951) propose that the evaporation of deeper clouds maintains the cloud layer is by Riehl et al. (1951), as illustrated by the schematic in column b. Column c. reproduces figures from Malkus (1958) showing two populations of shallow and deeper clouds both in schematic and photograph form.

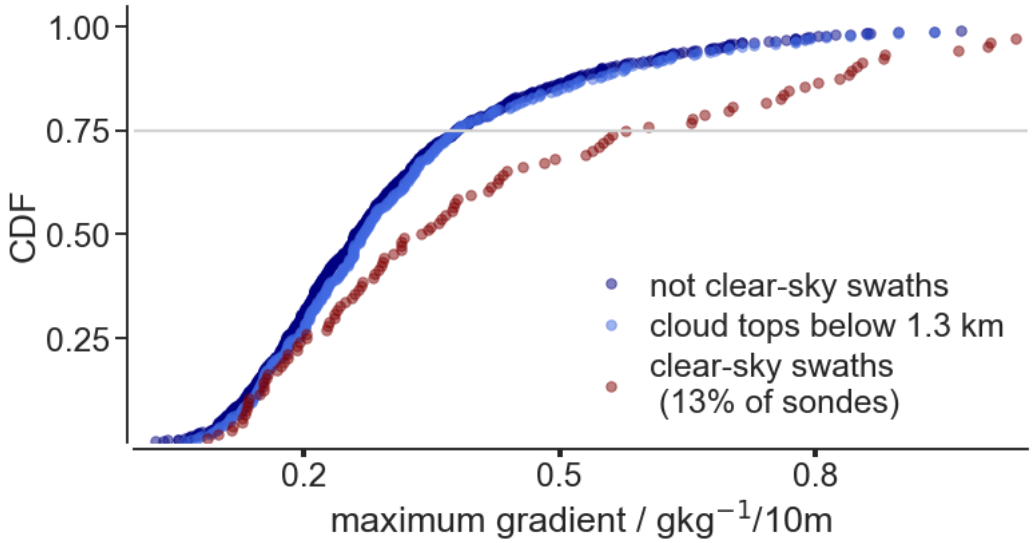


Figure 5-9: Similar to Fig. 5-6, except with distributions conditioned on large clear-sky scenes of at least 15 minutes without encountering a cloud (red), shallow clouds with cloud top heights estimated from WALES lidar data below 1.3 km (light blue), and all areas outside of large clear-sky scenes (dark blue).

segments that are greater than this 95$^{\text{th}}$-percentile. These large clear-sky patches correspond to 15 minutes of flight time without encountering a cloud, or about 180 km at a typical flight speed of 200 m/s. The associated dropsondes are then chosen for these clear-sky patches, corresponding to 13% of all dropsondes. For non-clear-sky areas, the disjoint set of the large clear-sky swaths is chosen. Scenes with shallow clouds below 1.3 km do not have a minimum time or length requirement, given that the influence of clouds is expected to extend outside the immediate cloudy vicinity.

We then repeat the analysis from Fig. 5-6 for these three categories. Fig. 5-9 shows that gradients are stronger in large clear-sky areas than other scenes. A second result is that the distribution of gradients estimated for shallow clouds is nearly identical with the distribution of gradients for all conditions (including deeper clouds), suggesting that the presence of shallow clouds is sufficient to cause mixing that smooths vertical gradients relative to cloud-free conditions. Fig. 5-10 also shows that large clear-sky areas exhibit stronger vertical gradients over a shorter distance than sondes launched outside such clear-sky areas.

From this conditional sampling, a physical picture emerges that the life cycle of shallow clouds forming and dissipating in the transition layer changes vertical gradients relative to cloud-free conditions. Similarly to Malkus (1958), we find that the transition layer structure differs in cloudy and clear-sky conditions. Yet, we find an opposite result that the transition layer occurs over a smaller distance (sharper gradient) in clear-sky than cloudy regions, though Malkus (1958) also noted an ambiguity in how the depth of the transition layer was defined for cloudy profiles in her study.

5.5 Implications and interpretation

The general picture of the role of shallow clouds in maintaining the transition layer structure is that when these clouds form, they warm and dry the surrounding air, and when they evaporate, they cool and moisten, such that the ambient air takes on properties that more closely resemble mixed layer air. Does this different transition layer structure and the role of shallow clouds in maintaining it have any implications

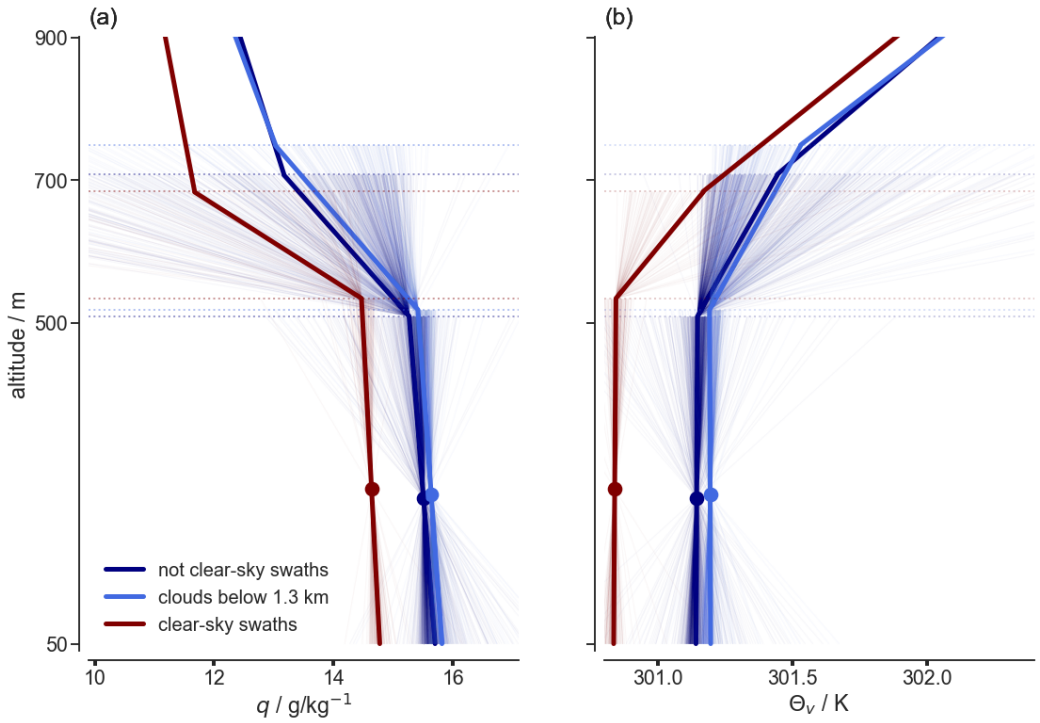


Figure 5-10: Similar composite profiles of (a.) specific humidity, q, and (b.) virtual potential temperature, θ_v, as in Fig. 5-4, but for the conditional sampling as described in Sec. 5.4.3. The profiles for each category are constructed from the mean mixed layer value, mean gradients over the mixed layer, transition layer, and lower cloud layer, and mean heights of the mixed layer and transition layer top. Colors correspond to at least 15 minutes of sampling (about 180 km) without encountering a cloud (red), shallow clouds with cloud top heights below 1.3 km (light blue), and all areas outside of large clear-sky scenes (dark blue).

for how entrainment is modeled? As discussed in Sec. 5.1, common entrainment closure schemes are based on the cloud-free boundary layer structure. To begin to answer this question, we return to mixed layer theory, as introduced in the previous chapter, albeit from a different angle. The height of air incorporated into the mixed layer has the potential to be informative about the conceptualization of entrainment mixing, in particular whether it is a local process in the transition layer or rather extends deeper into the cloud and potentially inversion layers.

As discussed in Chapt. 4, the subcloud layer bulk budget of a scalar, ϑ, is expressed as the vertical divergence of the turbulent fluxes balancing a source term, Q_ϑ, which comprises material derivatives (e.g. advection and time-rate of change or storage terms) and diabatic tendencies (e.g. radiation, evaporation, and precipitation),

$$hQ_\vartheta = \overline{w'\vartheta'}\big|_0 - \overline{w'\vartheta'}\big|_1 . \tag{5.1}$$

Here w denotes vertical velocity, and the prime symbol denotes fluctuations from the mean value. Subscript 0 denotes values at the lower interface (e.g. the ocean surface), subscript 1 denotes values at the upper interface, and $h = z_1 - z_0$, is the layer thickness, which can either refer to the mixed layer thickness, or the subcloud layer thickness as in Chapt. 4. Note that in this chapter, h refers to the mixed layer top given our focus on processes in the transition layer above the mixed layer top.

In Chapt. 4, two closure assumptions were discussed for this equation. First, the fluxes at an interface were modeled by a mean exchange velocity, V_i, and a 'jump' at the interface,

$$\overline{w'\vartheta'}\big|_i = -V_i \Delta_i \vartheta, \tag{5.2}$$

where $\Delta_i \vartheta$ defines the difference in ϑ across the interface, i, from top to bottom. This assumption allows us to rewrite Eq. (5.1) as,

$$-V_0 \Delta_0 \vartheta + V_1 \Delta_1 \vartheta = Q_\vartheta h. \tag{5.3}$$

Second, it is assumed that the surface θ_v flux (proportional to the surface buoyancy

flux) at the upper interface is energetically constrained by its source of surface θ_v fluxes, such that,

$$V_1 \Delta_1 \theta_v = -A V_0 \Delta_0 \theta_v, \tag{5.4}$$

with A defining the entrainment efficiency and assumed to be constant. In the previous chapter, a mean value $A = 0.43$ was found from a Bayesian inversion of uncertain entrainment parameters in subcloud layer thermodynamic budgets closed with EUREC4^{4}A data (Albright et al., 2022).

With this second closure assumption in Eq. (5.4) and taking $\vartheta = \theta_v$, Eq. (5.3) becomes,

$$-V_0(1 + A)\Delta_0 \theta_v = Q_{\theta_v} h. \tag{5.5}$$

Rearranging Eq. (5.5) to solve for the θ_v-jump at the surface yields

$$\Delta_0 \theta_v = -\frac{Q_{\theta_v} h}{V_0(1 + A)}. \tag{5.6}$$

With Eq. (5.6), Eq. (5.4) can be manipulated to solve for the velocity at the upper interface, V_1,

$$V_1 = A^\star V_0 \quad \text{where} \quad A^\star = -\frac{A \Delta_0 \theta_v}{\Delta_1 \theta_v}. \tag{5.7}$$

The modified constant $A^\star$ is the velocity scale analogue to the entrainment buoyancy flux ratio, A. That is, whereas A relates the buoyancy flux at the upper and lower interfaces, $A^\star$ relates the *exchange velocities* at the upper and lower interfaces. $A^\star$ is, however, not energetically constrained by the surface buoyancy source (e.g., that the buoyancy flux used at the upper interface is generated at the surface, such that A cannot exceed one). As a result, $A^\star$ is not expected to be fixed across conditions.

Using the above formulations, including $A^\star$, the generalized budget equation in Eq. (5.3) can be solved for $\overline{\vartheta}$,

$$\overline{\vartheta} = \frac{\vartheta_0 + A^\star \vartheta_1 + Q_\vartheta \tau}{(1 + A^\star)}, \quad \text{where} \quad \tau = \frac{h}{V_0}. \tag{5.8}$$

This equation gives an expression for the layer-mean value of any scalar based on

external boundary conditions, such as values at the lower and upper interfaces (ϑ_0, ϑ_1) and the surface exchange velocity, V_0. Other components could be considered emergent properties of the system, such as its height, h, and diabatic tendencies and material derivatives in Q_ϑ. If $Q_\vartheta = 0$, then $\overline{\vartheta}$ is simply a weighted average between its lower value, ϑ_0, and upper value, ϑ_1, with a larger $A^\star$ denoting that $\overline{\vartheta}$ is more influenced by the upper than the lower interface. If $Q_\vartheta \neq 0$ term, the $\tau = \frac{h}{V_0}$, can be thought of as an adjustment time scale that changes values from pure two-point mixing between the upper and lower interfaces. That is, this adjustment time scale weights the material derivative and diabatic tendencies.

5.5.1 Mixing diagrams

A first way of visualizing these equations is through 'mixing diagrams', also known as Paluch diagrams (Paluch, 1979). Fig. 5-11 plots a mixing diagram with observed pairs of campaign-mean q and θ values (every 10 m). This visualization shows that q-θ pairs fall along lines, often referred to as mixing lines. A linear structure is what is expected from two-point mixing, or mixtures of air derived from two sources (e.g., Paluch, 1979; Betts and Albrecht, 1987; Heus et al., 2008; Böing et al., 2014). Fig. 5-11 shows that there appears to be a mixing line that includes the mixed and transition layers. This structure suggests that air in the transition layer is incorporated into the mixed layer, and air at various heights in the transition layer cannot be distinguished from the point of view of mixing. The observed mixing line does, however, appear to change slope around 900 m, suggesting that air above this height is not directly incorporated into the mixed layer, and that different processes are responsible for its structure.

A theoretical mixing line is calculated from Eq. (5.8) for q and θ. The best-fit mixing line to observations incorporates air from 10–150 m above the mixed layer top into the mixed layer. The theoretical mixing line increasingly diverges from observations as air is incorporated from higher levels (not shown). That is, theoretical and observed mixing lines agree best when air is entrained into the mixed layer from the transition layer, rather than from deeper in the cloud layer. This analysis

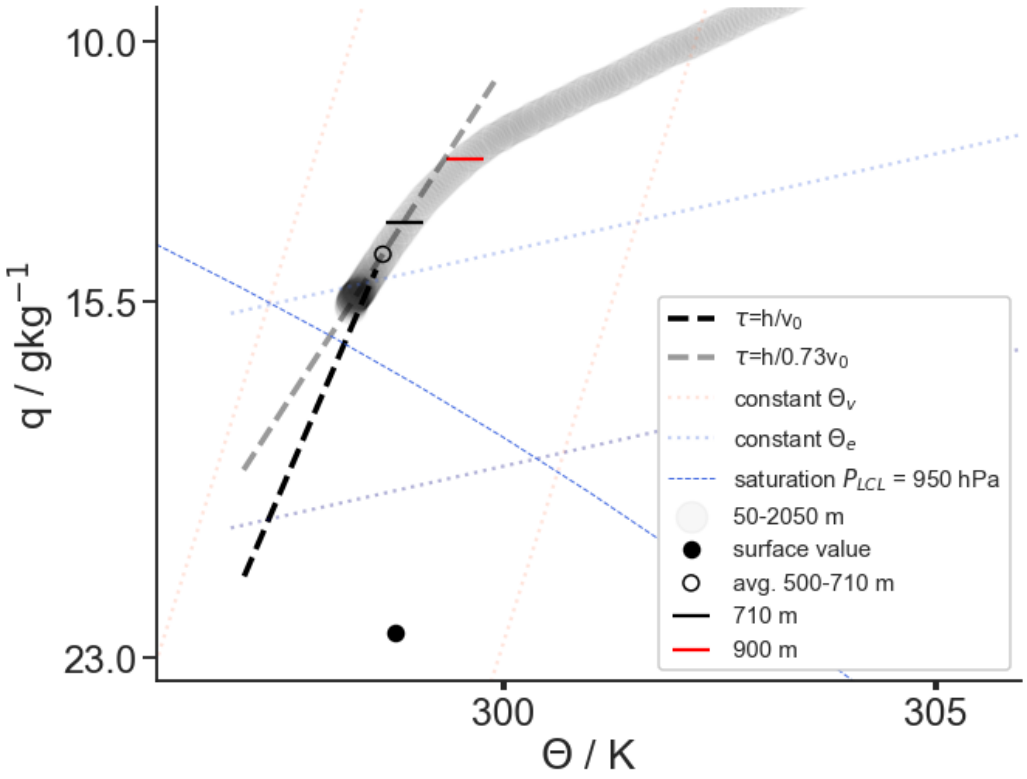


Figure 5-11: Scatter plot of observed, campaign-mean pairs of q and θ up to 1.6 km (grey points). Two theoretical mixing lines are shown with q and θ values calculated from Eq. (5.8). The black dashed line is the best-fit theoretical mixing line, taking all values at their campaign-mean, varying $A^\star$, and incorporating air from 10–150 m above the mixed layer layer. The grey dashed line multiples the surface wind speed by 0.73, thereby increasing τ. Additional curves and markers are plotted for reference: values for a constant lifting condensation level at 950 hPa (blue dotted line), constant θ_v (orange dotted lines), constant θ_e (darker blue dotted lines), the surface values (black filled circle), mean values from 540–690 m (black open circle), and the heights 720 m (black horizontal marker) and 900 m (red horizontal marker).

suggests that entrainment is a local mixing across the transition layer, rather than one wherein eddies bring down air from higher aloft. This inference is also consistent with the kink in the observed mixing line around 900 m (Fig. 5-11). The agreement of this theoretical mixing line, incorporating air from within the transition layer, with the observed values can be further improved by increasing $\tau = \frac{h}{V_0}$, such as by multiplying the wind speed by a factor 0.73 in the V_0 term, or by increasing the source terms, Q_ϑ (Fig. 5-11, grey dashed line). These adjustments suggest potential limitations in how the surface exchange velocity is represented or missing source terms, such as cloudy radiative heating. Also shown for reference are lines of constant θ_v (constant buoyancy), constant θ_e lines, which illustrate the slope that a saturated, cloudy updraft would follow, and the curve indicating a constant lifting condensation level at a pressure of 950 hPa. Of note, the observed and best-fit theoretical mixing line are perpendicular to the constant LCL curve. Whether this perpendicularity is a coincidence, or whether mixing aligns to maximize variance in the lifting condensation level merits further study.

5.5.2 Further support for a shallow mixing layer

A complementary approach is to set Eq. (5.8) equal for mixed layer mean specific humidity, $\overline{q}$, potential temperature, $\overline{\theta}$, and then solve for the two unknowns, $A^\star$ and τ. All other terms can be calculated from EUREC^{4}A circling-mean data and are set to their campaign-mean value. Again h refers to the mixed layer top, not the subcloud layer top, as in the previous chapter. Solving these two equations for the two unknowns yields an equation for $A^\star$,

$$A^\star = \frac{-\overline{\theta}\frac{Q_q}{Q_\theta} + \overline{q} + \theta_1\frac{Q_q}{Q_\theta} - q_1}{\overline{\theta}\frac{Q_q}{Q_\theta} - \overline{q} - \theta_2\frac{Q_q}{Q_\theta} - q_2}. \tag{5.9}$$

Physically, $A^\star = \frac{V_1}{V_0}$ can be thought of weighting air with properties of the upper and low interfaces. A large $A^\star$ denotes that air at the upper interface has more influence

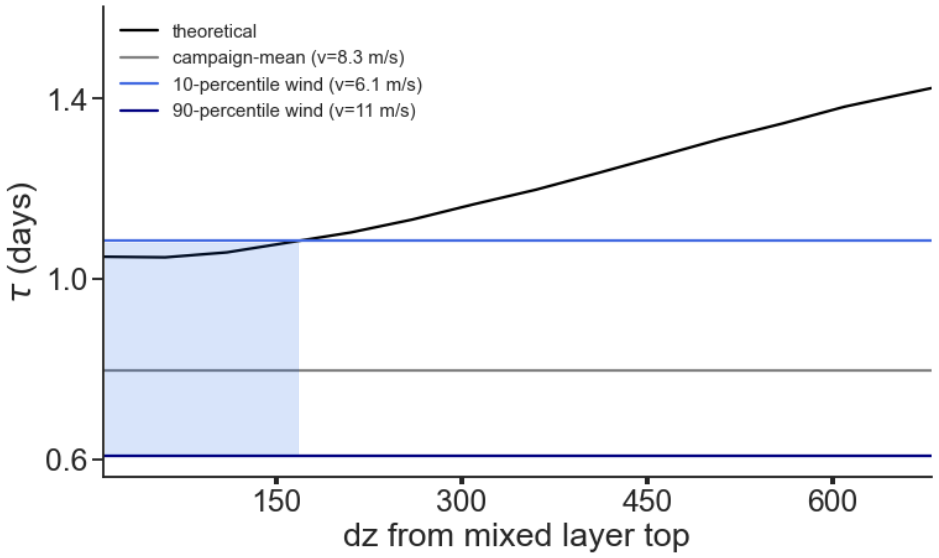


Figure 5-12: Theoretical values of τ (black curve) calculated from Eq. (5.10) for different depths, dz, above the mixed layer top, h. The campaign-mean τ (grey horizontal line), calculated from the campaign-mean mixed layer height, h, and campaign-mean 10 m wind speed, is always less than the predicted τ. Also shown are τ values when using the 10^{th}-percentile wind speed from the campaign (medium blue horizontal line, top) and 90^{th}-percentile of wind speed (dark blue horizontal line, bottom).

on the mixed layer value, and vice versa. Solving for τ as a function of $A^\star$ yields,

$$\tau = \frac{(1 + A^\star)\overline{\theta} - \theta_1 - A^\star\theta_2}{Q_\theta}. \tag{5.10}$$

The parameter τ has units of time and can be thought of as a time scale that adjusts how quickly the mixed layer responds to other influences than simply mixing between the upper and lower interface – that is, influences from the material derivative and diabatic processes such as radiation.

Calculating this adjustment time scale, τ, for a range of heights, dz, about the mixed layer top, and comparing these values to observed values of τ provides further information about the depth of air that is entrained into the mixed layer. The campaign-mean τ calculated from the campaign-mean h and surface exchange velocity, V_0, is smaller than the theoretical values (Fig. 5-12), suggesting that the the-

ory predicts a slower adjustment time scale. Also plotted is $\tau = \frac{h}{V_0}$ for the 10^{th}-percentile and 90^{th}-percentile observed 10 m wind speeds. These 10^{th}-percentile and 90^{th}-percentile values approximately correspond to the wind speed in the first and second half of the campaign (George et al., 2021; Stephan et al., 2020; Bony et al., 2022). For expository purposes, variations in the near-surface wind speed (which drives variations in V_0) are plotted, keeping h fixed at its campaign-mean value.

Fig. 5-12 allows for two inferences. First, the predicted τ and the observed τ for weak wind speeds converge when air is taken from within 150 m depth above the mixed layer top, corresponding to the approximate transition layer depth. If air is sourced from above the transition layer, a less physical τ must increasingly be adopted. Second, Fig. 5-12 generally indicates a τ value that is larger than the observed campaign-mean value. This overestimation of τ could suggest that the surface mixing velocity, V_0, and hence the surface fluxes are overestimated, there is a flaw in how this surface mixing is conceptualized, or source terms are missing in Q_ϑ. The first inference suggests that air is incorporated into the mixed layer from the transition layer, rather than from deeper eddies extending higher into the cloud layer. The second inference could be consistent with the greater skill of mixed layer theory for closing subcloud layer moisture than temperature budgets in Chapt. 4, perhaps suggesting that it is important to include cloudy, rather than simply clear-sky radiative heating rates in Q_ϑ. It could also potentially suggest that the mixed layer framework is more skillful when wind speeds are weak (e.g., associated with the 'sugar' pattern of cloud organization (Bony et al., 2020)).

5.6 Discussion and conclusions

This ongoing analysis suggests that the transition layer structure is determined by cloudy, not dry processes. An offset exists between mixed and subcloud layer tops in individual soundings, as well as more aggregated vertical profiles. We associate this offset with a transition layer. Strong vertical gradients at the mixed layer top are only found rarely and when they occur, they tend to occur in scenes that are cloud-free over

large areas. In areas influenced by clouds, vertical gradients are instead smoother, extending over a larger depth. The transition layer is populated by small clouds that have their bases starting around 500 m and tend to only grow a few hundred meters above these bases.

Based on these findings, we propose a new conceptual picture that the formation and dissipation of shallow clouds maintains the transition layer, in analogy with the maintenance of the trade-wind inversion by deeper clouds, as proposed by Riehl et al. (1951) and developed by Stevens (2007). This conceptual model suggests that small clouds beget larger clouds. Small clouds precondition the large-scale environment and decrease the resistance to convection through weaker vertical gradients in the transition layer, making it easier for deeper clouds to form.

From this analysis emerges the potential for an alternate view of entrainment mixing, which is based on the ability to detrain condensate into the overlying stable layer and thus induce gentle sinking motion through negative buoyancy. Clouds forming and evaporating by mixing with ambient air in the transition layer 'precondition' the environment, both decreasing the barrier to convection and causing gentle sinking motion into the mixed layer. This view contrasts with the classical view established from dry boundary layers. The classical view of entrainment posits that overshooting plumes must do work on the overlying stratified fluid in order to bring down tendrils or filaments of overlying, more stable fluid and mix it into the turbulent boundary layer below (e.g., Bretherton, 1997). The preconditioning picture is one of 'subtle persuasion' rather than 'brute force' in dry layers (Bjorn Stevens, personal communication). An injection of moisture into the transition layer, moreover, leads to additional radiative cooling from water vapor and further reduces the work required to incorporate this air into the mixed layer.

Our findings also suggest a symmetry between shallow and deep clouds, with both populations growing their own layer. The shallow clouds maintain the transition layer, whereas deeper clouds maintain the trade-wind inversion layer. This idea is illustrated schematically in Fig. 5-13. This idea forms in analogy with the maintenance of the trade-wind inversion layer as described by Riehl et al. (1951):

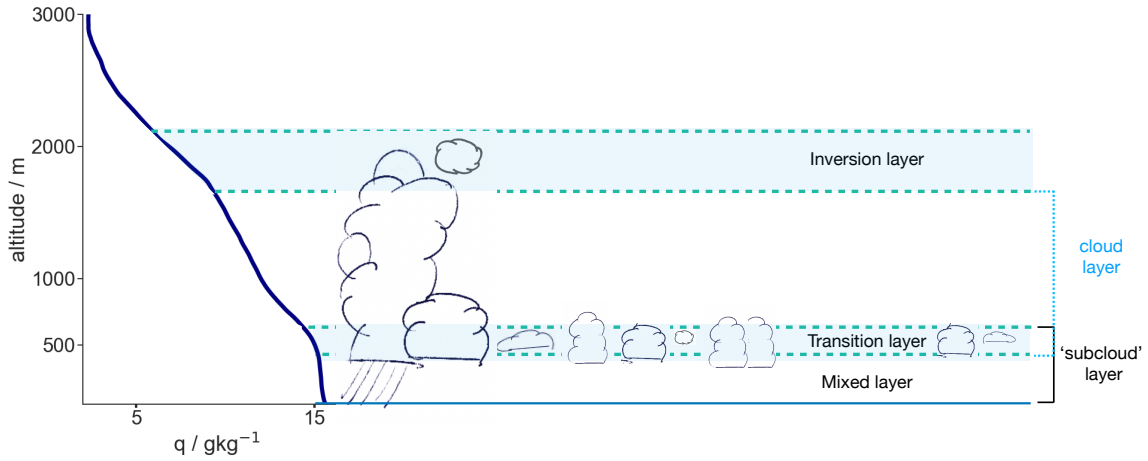


Figure 5-13: Illustration of a new conceptual picture that suggests a symmetry between deeper clouds growing and maintaining the trade-wind inversion layer following Riehl et al. (1951) and Stevens (2007), and shallower clouds growing the transition layer. The formation and dissipation of shallow clouds in the transition layer moistens and cools (denoted by the transparent blue area) the transition layer, rendering gradients smoother and weaker compared to the dry boundary layer case.

.

> "It is well known that the bases of the cumuli have a nearly uniform height, but that the tops are very irregular. Some are found within the cloud layer, many near the inversion base, and some within the inversion layer as active clouds penetrate the base. As shown by visual observation and many photographs, the tops of these clouds break off and evaporate quickly. In this way moisture is introduced into the lower portions of the inversion layer, and the air there situated gradually takes on the characteristics of the cloud layer."

In this view, overshooting convective plumes collapse or break off and, in the process, inject condensate into the trade-wind inversion layer. The evaporation of this condensate maintains the depth of the cloud layer against compensating subsidence, which evacuates mass and shrinks the layer. These ideas are also illustrated by Fig. 5-8b reproduced from Riehl et al. (1951), as well as Fig. 5-8c from Malkus (1958). In a similar way, clouds evaporating in the transition layer grow and maintain the transition layer compared to the abrupt discontinuity observed in the dry

boundary layer case. The transition layer then more closely resembles the properties of the mixed layer. Ongoing work further investigates such cloud dissipation in the transition layer with high-frequency turbulence data from the ATR-42 aircraft. These data allow for tracking, with high-frequency, the cooling and moistening achieved by clouds dissipating in the transition layer. Other ongoing work partitions whether the observed thermodynamic structure outside of large clear-sky swaths can be explained by radiative cooling, subsidence, or cooling from cloud evaporation.

Building upon findings in Riehl et al. (1951), Stevens (2007), moreover, show that the cloud layer growth could be represented using concepts based on cloud-free, or dry mixed layer theory. He finds that a simple model that is formulated in terms of an 'effective' dry buoyancy flux is skillful at predicting the cloud layer growth in large eddy simulations. These findings suggest that cloudy processes can still be represented using cloud-free mixed layer theory if effective parameters are introduced. This result is similar to the findings of Chapt. 4 wherein an entrainment formulation based on mixed layer theory and dry boundary layers was used, but effective parameters were introduced to account for ambiguities resulting from the finite-thickness transition layer. We find that mixed layer theory, with slight modifications, still closes subcloud layer moisture and energy budgets. These considerations highlight that extensive observations from the field, such as from the EUREC^{4}A campaign, open the door to revisiting old concepts and testing their applicability when confronted with new data. This confrontation, in turn, allows for improved understanding of the origins of the characteristic vertical structure of trade-wind air.

Chapter 6

Uncertainty in trade cumulus feedbacks still contributes to uncertainty in global cloud feedbacks in CMIP6

6.1 Outline

In the second part of this thesis, physical understanding developed in chapters 2–5 is applied to the evaluation of general circulation models (GCMs). In previous CMIP ensembles, uncertainties regarding trade cumulus cloud changes were shown to explain large differences in global cloud feedbacks and climate sensitivity (e.g., Bony and Dufresne, 2005; Webb et al., 2006; Vial et al., 2013; Brient and Schneider, 2016). The CMIP6 models have been updated in numerous ways, such as regarding their convection schemes, microphysics schemes, and vertical resolution (e.g., Gettelman et al., 2019; Danabasoglu et al., 2020; Zelinka et al., 2020). It is therefore not *a priori* clear how large a role trade-wind clouds still play in explaining global uncertainties. This chapter uses CMIP6 models to motivate a continued focus on trade cumulus cloud feedbacks. It acts as preparation for Chapt. 7, which presents the first process-based constraint on this trade cumulus feedback. In this chapter and Chapt. 6, Jessica Vial pre-processed the amip and *amip4K* model output and provided valuable guidance throughout this analysis. The structure of this chapter follows the well-known quip from Winston Churchill (1874–1965) of a 'riddle wrapped in a mystery

Analysis in this chapter contributed to Boucher et al. (2020), published in the *Journal of Advances in Modeling Earth Systems* in May 2020.

inside an enigma'. Here the question is whether the riddle of trade-wind clouds, wrapped in the mystery of tropical clouds, still contribute to enigmas surrounding global cloud feedbacks and climate sensitivity.

6.2 Methods

6.2.1 Estimating equilibrium climate sensitivity

As introduced in Chapt. 1, the equilibrium climate sensitivity (ECS) is one of the earliest metrics for quantifying the climate system response to carbon dioxide (CO_2) forcing (Charney et al., 1979; Cess et al., 1989). The top-of-atmosphere (TOA) radiative imbalance, ΔR, can be expressed linearly as the sum of the radiative forcing, F, and the product of a feedback parameter, λ, and the global-mean surface temperature change, ΔT,

$$\Delta R = F + \lambda \Delta T. \tag{6.1}$$

This product, $\lambda \Delta T$, is also referred to as radiative damping.

The ECS is defined as the global-mean surface temperature change, T_{eq}, that restores a TOA radiative balance, ΔR=0, after a carbon dioxide doubling,

$$ECS = \Delta T_{\text{eq}} = -\frac{F_{\text{2xCO2}}}{\lambda}. \tag{6.2}$$

The ECS depends both on the radiative forcing associated with a CO_2 doubling, F_{2xCO2} and the strength of the radiative feedback parameter, λ. A climate system that more effectively reflects shortwave radiation to space, such as through greater low cloud fraction, or more effectively radiates longwave radiation away to space, such through fewer upper-tropospheric clouds, requires less warming, ΔT, to restore a planetary energy balance (e.g., Zelinka et al., 2020; Meehl et al., 2020). A key assumption is that the radiative damping, $\lambda \Delta T$, in Eq. **??** can be expressed as a product of a single, time-invariant feedback parameter, λ, and ΔT. A large body of research has highlighted that the constant λ assumption is imperfect, such as due to

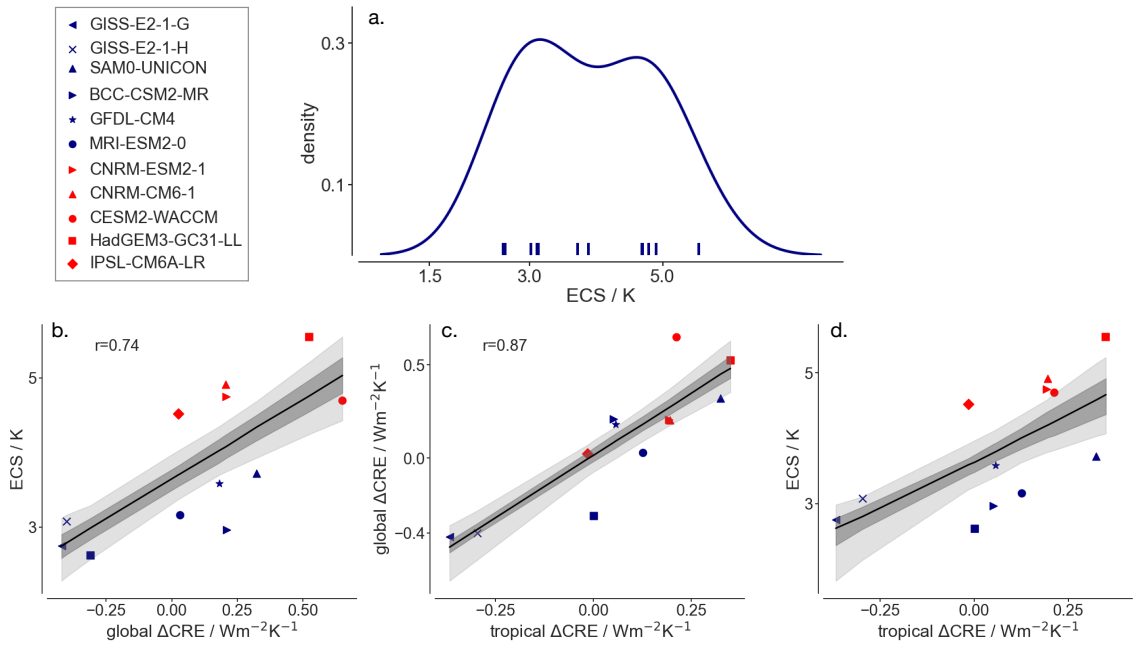


Figure 6-1: (a) Distribution of ECS estimated for the models listed in the legend, and scatter plots of (b) global ΔCRE and ECS (c), tropical ΔCRE and global ΔCRE, and (d) tropical ΔCRE and ECS. High climate sensitivity models, defined as ECS > 4°C are red, and low climate sensitivity models (ECS < 4°C) are blue. 5–95% uncertainty on the linear regression (grey shading) is defined from bootstrapping.

different radiative feedbacks actualized on different response timescales (e.g., Armour et al., 2013; Proistosescu and Huybers, 2017) and the state-dependence of feedbacks (e.g., Bloch-Johnson et al., 2015).

Here ECS is estimated following Gregory et al. (2004), similar to Andrews et al. (2012) and Zelinka et al. (2020). Anomalies are calculated as the difference between the first 150 years of global-annual output from *abrupt-4xCO2* simulations (coupled GCM simulations wherein atmospheric CO_2 concentrations are abruptly quadrupled from their preindustrial baseline and then held fixed) and a preindustrial control simulation, *piControl*. Positive radiative fluxes are defined downwards. Note that ECS using 150 years of model output is sometimes referred to as the 'effective' climate sensitivity instead of the equilibrium climate sensitivity since the model has not reached equilibrium above 150 years (e.g., Rugenstein et al., 2020; Meehl et al., 2020). Here the two terms are used interchangeably, given the focus on understanding relative differences among models.

Using a subset of 12 CMIP6 models, Fig. 6-1a shows that there is a bimodal distribution of ECS, with a mean of 4°C and two cluster of models around 3°C and 5°C, consistent with Zelinka et al. (2020) and Flynn and Mauritsen (2020). In the following analysis, low and high ECS are defined relative to the CMIP6 mean of 4°C. The second distribution around 5°C was not present in CMIP5, which had a range of 2.1–4.7°C (Taylor et al., 2012; Andrews et al., 2012), and it is above the canonical Charney range (Sec. 1.1). As discussed in Sec. 1.2.1, the shift towards higher ECS in the CMIP6 ensemble is thought to result from correcting too-negative extratropical low cloud feedbacks, compared with satellite constraints. This correction then unmasked the consistently coupled too-positive trade cumulus feedbacks, driving high ECS values (Myers et al., 2021).

A different way of quantifying climate sensitivity is also proposed, which examines the influence of the nonlinear relationship between forcing and response. Methodologically, the approach bins annual-mean data and then repeatedly samples from each bin to perform a linear regression, rather than performing a linear regression on all data directly. Binning the data better illustrates a potential nonlinearity in the forcing-response relationship (Fig. 6-2a vs. Fig. 6-2b). Repeating this bootstrapping procedure, for instance, 1000 times, yields a distribution, such as for climate sensitivity in Fig. 6-2c, or the climate feedback parameter. The mean ECS from this method is similar to results from the Gregory et al. (2004) regression. Using 300 years of data, ECS from the Gregory et al. (2004) method is 4.8°C, equal to the mean from the bootstrapping approach. The 5-95% uncertainty from the bootstrapping approach (4.2–5.7 °C) is, however, larger than the 5-95% uncertainty on the regression (4.3–5.2 °C). This result suggests that influences from the nonlinear forcing-response relationship tend to be small but influence the upper bound, at least for the IPSL-CM6A-LR model as shown in Fig. 6-2. In the following analysis, we return to the Gregory et al. (2004) method for consistency with other studies.

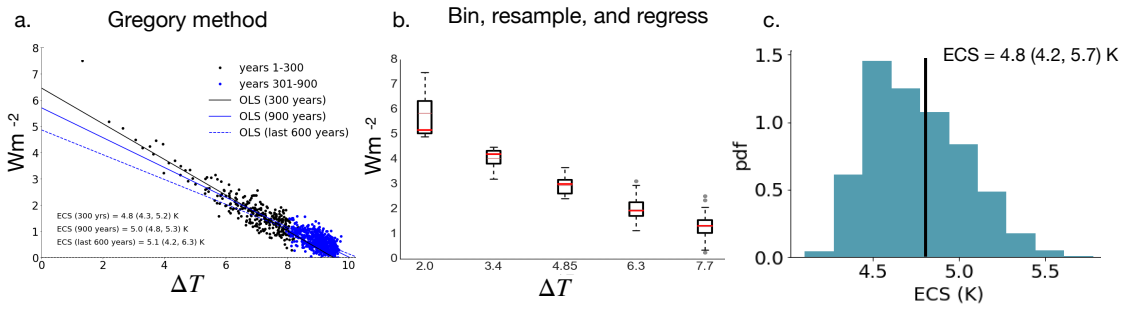


Figure 6-2: (a) Canonical Gregory ordinary least-squares (OLS) regression using different lengths of global-annual data as given in the legend; (b) an example of binning 300 years of model output from IPSL-CM6A-LR to better see the concavity in forcing and response; (c) a quantile regression performed for different percentiles in 300 years of model output; (d) the ECS distribution that results from binning data as in (b) and repeatedly sampling (1000 times) global-annual mean data from each bin to perform the least-squares regression. The vertical black line denotes the mean

6.2.2 Quantifying cloud radiative responses

In previous ensembles, the spread in cloud feedbacks was the primary driver of the spread in climate sensitivity (e.g., Bony and Dufresne, 2005; Vial et al., 2013; Zelinka et al., 2020). Here we quantify the spread in cloud radiative responses among a subset of models using a simplified metric, the change in the cloud radiative effect (ΔCRE), in order to quantify how much differences in this ΔCRE, globally, tropically, and in the trades, contribute to uncertainty in ECS.

The cloud feedback is a component of the radiative feedback, λ, in Eq. (6.1) and Eq. (6.2). A cloud feedback is a change in the top-of-atmosphere radiative flux resulting from a change in cloud albedo (e.g., due to changes in droplet size, water content), fraction, or altitude with warming. A positive feedback denotes that clouds change radiatively in a way that amplifies the initial warming, whereas a negative feedback denotes that clouds change in a way that damps the initial warming. A closely related quantity, the change in the cloud radiative effect (ΔCRE), was shown to be a skillful proxy for intermodel differences in the cloud feedback (e.g., Soden et al., 2008; Vial et al., 2013). ΔCRE is defined as the difference between all-sky (all, with clouds) and clear-sky (clr, clouds artificially removed) net downward radiative

fluxes,

$$CRE = R_{\text{all}} - R_{\text{clr}} = (LW_{\text{clr}} - LW_{\text{all}}) + (SW_{\text{all}} - SW_{\text{clr}}) = CRE_{\text{LW}} + CRE_{\text{SW}}, \quad (6.3)$$

with positive radiative fluxes defined downward. In present-day climate, net CRE is negative (about -20 Wm^{-2}) due to $\text{CRE}_{\text{LW}} \approx 30$ Wm^{-2}, reflecting the longwave warming effect of high clouds, and $\text{CRE}_{\text{SW}} \approx -50$ Wm^{-2}, due to clouds enhancing the planetary albedo (e.g., Ramanathan et al., 1989). A small change in the cooling effect of clouds (ΔCRE on the order of a few Wm^{-2}) due to global-mean warming, ΔT, could induce a strong feedback, motivating the large, long-standing focus on better constraining the cloud feedback, proportional to $\Delta CRE/\Delta T$ (e.g., Ramanathan et al., 1989; Bony et al., 2006).

Here ΔCRE is defined as the difference in CRE between a perturbed and control climate simulation. Cloud feedbacks are equal to the ΔCRE with an offset, typically 0.3 Wm^{-2} (e.g., Soden et al., 2008), accounting for the masking effects of clouds on clear-sky fluxes. For instance, removing a high cloud in a dry atmosphere would have a larger influence on the outgoing longwave radiation than would removing such a cloud in an already moist and opaque atmosphere. Whereas the ΔCRE is not a reliable metric of the sign or absolute magnitude of the cloud feedback, it is a skillful proxy for intermodel differences in the cloud feedback as calculated from other methods (e.g., Vial et al., 2013; Zelinka et al., 2020). The global ΔCRE calculated for this subset of models has a Pearson correlation coefficient r=0.99 with global cloud feedback estimates from Zelinka et al. (2020) using the kernel method.

6.3 Uncertainty in ECS still driven by spread in global and tropical cloud feedbacks

Fig. 6-1b shows that, in this representative subset of models, the spread in the global ΔCRE can explain 55% of the variance in ECS (r=0.74). Variability in climate sensitivity is driven more by variance in the shortwave ΔCRE (r=0.61) than by

variance in the longwave ΔCRE (r=-0.25) (not shown). The explanatory power of the global ΔCRE for ECS in this sample is similar to that in the previous CMIP5 ensemble, r=0.73 (e.g., Ceppi et al., 2017). The global ΔCRE can result from multiple physical processes, such as rising high clouds (e.g., due to fixed anvil temperature hypothesis), the iris effect (e.g., decreasing anvil cloud fraction with warming), cloud fraction changes in shallow clouds over tropical oceans, and cloud fraction and phase changes in extratropical clouds (e.g., Bony et al., 2006; Ceppi et al., 2017). It is thus not *a priori* clear how large a role the tropics, or cloud fraction changes in the trades, in particular, play in explaining global ΔCRE and climate sensitivity.

Only the *tropical* ΔCRE can, however, still explain 44% of the variance in climate sensitivity (r=0.66, Fig. 6-1c) and 76% of the variance in the global cloud feedback (r=0.87, Fig. 6-1d). The tropics are defined equatorward of 30°. The extratropical ΔCRE, by contrast, explains only 18% of the variance in ECS (r=0.42). While changes in extratropical cloud feedback can explain the shift towards higher mean values from CMIP5 to CMIP6, variance in the tropical ΔCRE can still explain more variance *within* ensembles. The origins of the spread in tropical ΔCRE are therefore examined in greater detail in the following sections.

6.4 Trade-wind cloud responses differ between high and low climate sensitivity models

6.4.1 Conceptualizing the tropical circulation

The tropical cloud responses can further be decomposed using the framework from Bony et al. (2004). The tropical radiation budget depends on the distribution of cloud types, which in turn is controlled, to first-order, by the large-scale atmospheric circulation (Bony et al., 2004). Large-scale atmospheric subsidence favors the formation of shallow clouds, such as trade-wind cumuli and stratocumulus clouds, whereas large-scale ascending motion is associated with deeper convective clouds (Fig. 6-3a). This framework uses the large-scale vertical velocity at 500 hPa, ω_{500} (expressed in

hPa/day), as a proxy of the large-scale atmospheric circulation. ω_{500} is the first baroclinic mode of the large-scale circulation in the deep tropics and correlates well with the total diabatic heating of the column, and hence with the precipitation. The tropical circulation can be discretized as a series of dynamical regimes, wherein the ascending branches of large-scale circulations with warm sea surface temperatures correspond to negative values of ω_{500}, and regions of colder sea surface temperatures and large-scale subsiding motions correspond to positive values of ω_{500}. The trades are often identified between 10 to 30 hPa/day (e.g., Bony et al., 2004; Brient et al., 2016).

The tropical circulation can also be conceptualizing using the large-scale vertical motion at 700 hPa, ω_{700}, and estimated inversion strength, EIS, more precisely discriminate among low-cloud regimes, such as between the trades and stratocumulus regimes. This approach is similar to Medeiros and Stevens (2011) and Medeiros et al. (2015) using lower tropospheric stability and ω_{500}, and is analogous to the approach in Myers et al. (2021). Following Myers et al. (2021), we identify the trades as areas with climatological annual-mean EIS below 1 K and ω_{700} between 0 and 15 hPa/day.

In the following analysis, ten *amip* models from CMIP5 and CMIP6 are used, which are a different sample of models presented above. The *amip* simulations are global simulations of the atmosphere and land surface forced by observed sea surface temperatures (rather than letting these freely evolve as in coupled *piControl* and *abrupt-4xCO2* experiments), sea ice cover, and greenhouse gas concentrations (Gates et al., 1999). Using observed sea surface temperatures (SST) avoids known biases in coupled models. In coupled models, Liu et al. (2013) show that SSTs are too cold and not variable enough in the Atlantic warm pool compared to observations, and Zhou et al. (2016) show that SSTs are too warm in the eastern equatorial Pacific compared with observed SSTs.

ΔCRE values are calculated as the difference between CRE in *amip4K* and *amip* simulations and then normalized by the 4 K temperature difference in these simulations to give a proxy for the cloud feedback. Qin et al. (2021) find good agreement between cloud feedbacks across geographic locations calculated using cou-

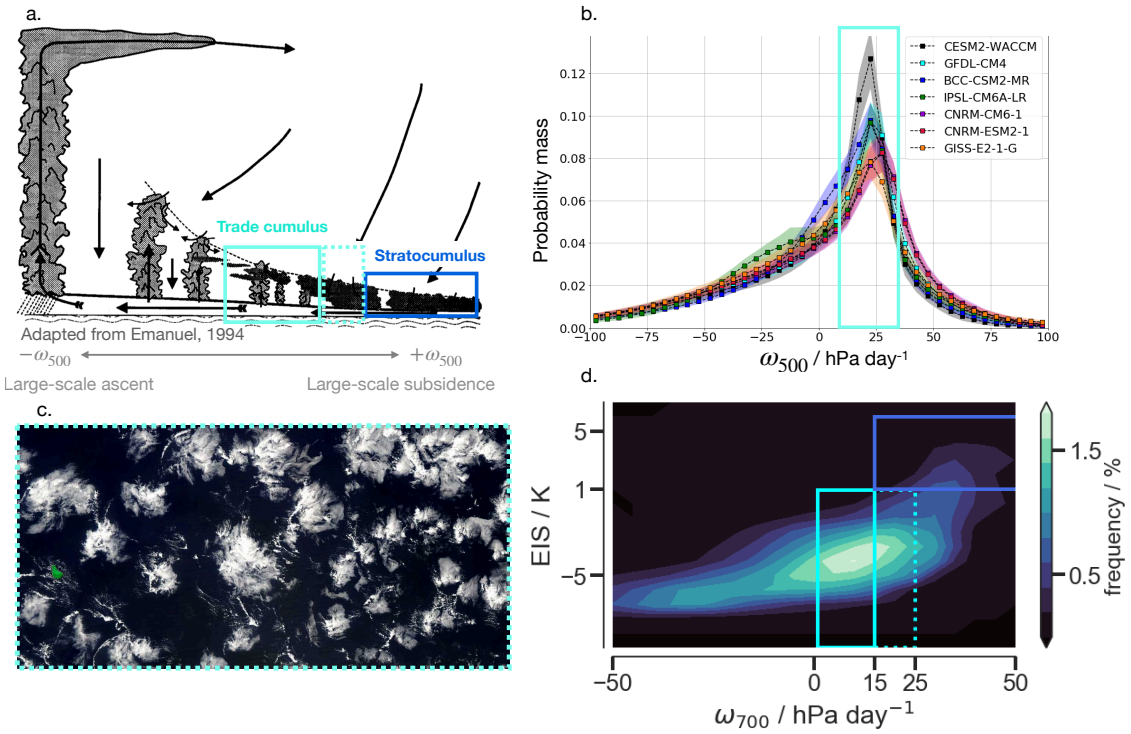


Figure 6-3: (a) An idealization of the tropical circulation as given in Emanuel et al. (1994), ranging from regions of large-scale ascent and deep convection to regions of moderate large-scale subsidence and trade cumulus clouds (turquoise box) and strong large-scale subsidence and stratocumulus clouds (dark blue box), as well as transition zones between trade cumuli and stratocumuli (dashed turquoise box). (b) The probability density function of ω_{500} in seven CMIP6 *piControl* simulations, with the colored shading denoting one standard deviation of monthly values around 20-year climatology. (c) A satellite view of 'flower' clouds that could represent clouds found in the cumulus-stratocumulus transition zone, perhaps corresponding to the dashed turquoise box in other panels. (d) The mean frequency distribution of tropical dynamical regimes averaged over ten *amip* simulations in EIS and ω_{700} space. Frequency of discretized dynamical regimes is defined as the area covered by these regions, normalized by the total area of the tropics. Following Myers et al. (2021), the trades are defined as areas with climatological annual mean EIS below 1 K and ω_{700} between 0 and 15 hPa/day. A potential intermediate zone between cumulus and stratocumulus clouds is extended from 15–25 hPa/day (dashed turquoise box).

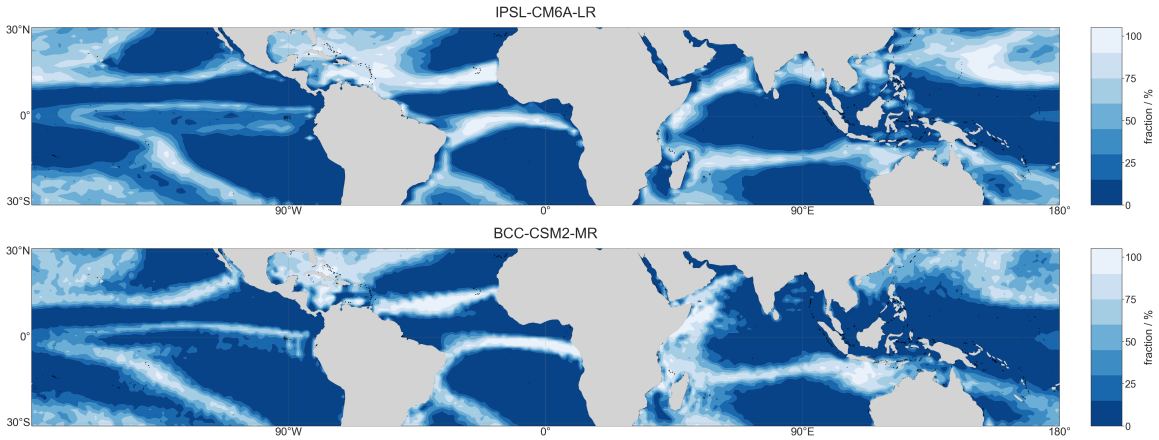


Figure 6-4: Frequency of occurrence of trade-wind regions in geographic space in the IPSL-CM6A-LR model (high ECS) and BCC-CSM2-MR (low ECS) using 30 years of annual model output from *amip* simulations. Trade-wind regions are identified as having EIS<1 and ω_{700} between 0–15 hPa/day, as introduced previously and following Myers et al. (2021).

pled (e.g., *abrupt-4xCO2* and *piControl*) and uncoupled, atmosphere-only simulations (e.g., *amip4K* and *amip*), justifying our use of atmosphere-only simulations to analyze the spread of cloud feedbacks. For these *amip* experiments, for heuristic purposes, we take ECS values from the corresponding coupled models (e.g., as calculated in Zelinka et al. (2020) and Sec. 6.2.1), with low and high ECS again defined relative to the CMIP6 mean of 4°C.

Fig. 6-3b,d shows how the trades are the most common cloud regime in the tropics in these two coordinate systems. In the ensemble-mean frequency distribution (Fig. 6-3d), the trade-wind area covers 32% of the tropics, compared with 4.4% for stratocumulus, using the respective EIS and ω_{700} definitions from Myers et al. (2021) and 39% of the tropics when defining the trades between 10–30 in hPa/day in ω_{500}. The frequency distribution structure is similar for individual models. Perhaps surprisingly, there is also substantial weight in between what Myers et al. (2021) define as either trade-wind or stratocumulus based on the EIS-ω_{700} decomposition. The area from 16–25 hPa/day and EIS $<$ 1 K covers 14% of the tropics and could correspond to transition zones between stratocumuli and trade cumuli, such as when stratocumulus decks recede towards continents. These clouds could perhaps also correspond to

clouds named 'flowers' by Stevens et al. (2020b) and Bony et al. (2020), which tend to be associated with stronger subsidence (Bony et al., 2020; Schulz et al., 2021).

In geographic space, Fig. 6-4 shows the frequency of occurrence of trade-wind regions in two representative climate models, one having high ECS (IPSL-CM6A-LR) and another with low ECS (BCC-CSM2-MR). Fig. 6-4 shows that trade-wind regions cover wide swaths of tropical oceans. The EUREC^{4}A measurement area east of Barbados is nearly always classified as a trade-wind region across the ten models.

6.4.2 Different thermodynamic and dynamic cloud radiative responses for high and low ECS models

Clouds are sensitive to changes in both temperature and circulation. The Bony et al. (2004) framework separates these two influences, decomposing a cloud variable, such as the ΔCRE, into the sum of a *thermodynamic* component that is related to temperature changes, a *dynamic* component that is related to circulation changes, and a *covariance* term that tends to be negligibly small,

$$\overline{\delta C_\omega} = \int_{-\infty}^{+\infty} C_\omega \delta P_\omega + \int_{-\infty}^{+\infty} P_\omega \delta C_\omega \int_{-\infty}^{+\infty} \delta P_\omega \delta C_\omega. \tag{6.4}$$

Discretizing this equation gives,

$$\overline{\delta C_\omega} = \sum_\omega C_\omega \Delta P_\omega + \sum_\omega P_\omega \Delta C_\omega + \sum_\omega \Delta P_\omega \Delta C_\omega \tag{6.5}$$

The dynamic component, $C_\omega \Delta P_\omega$, represents the effect of changing circulation on the cloud variable, without letting the clouds adjust in response to temperature. The thermodynamic component, $P_\omega \Delta C_\omega$, by contrast, quantifies the effect of changing temperature structure the cloud response, holding circulation fixed.

To illustrate the association between clouds and circulation, Fig. 6-5 projects the tropical *thermodynamic* cloud radiative response, $P_\omega \Delta CRE_\omega$, into the two circulation spaces. A difference emerges between high and low ECS models in the trades. Trade-wind clouds exhibit a moderate radiative sensitivity to temperature change,

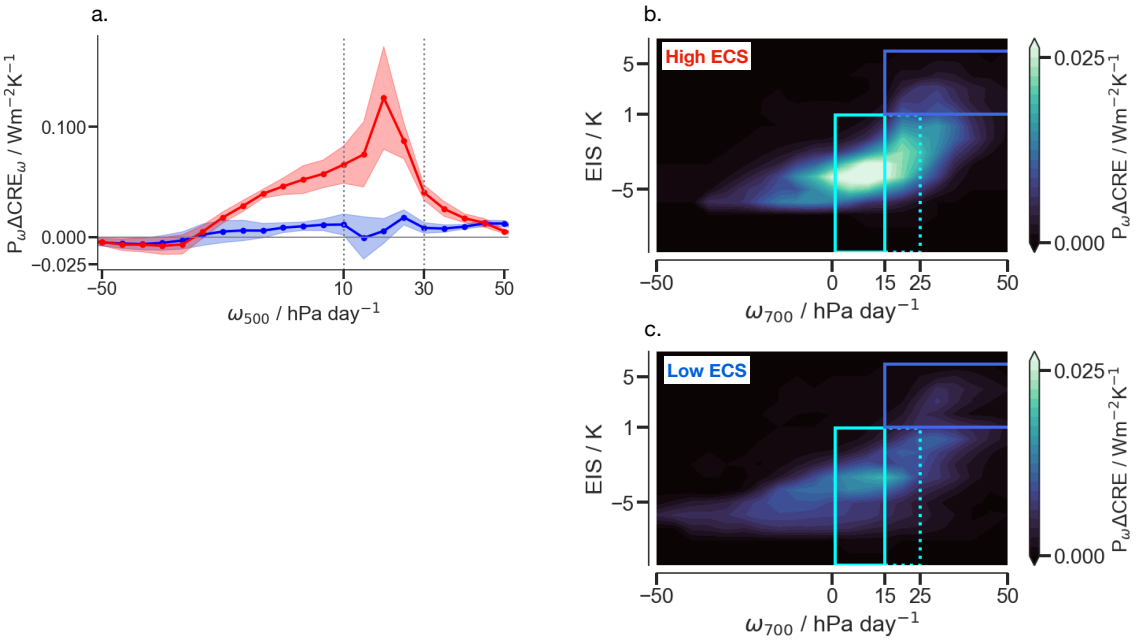


Figure 6-5: (a) The thermodynamic component, $P_\omega\Delta CRE$, in ω_{500} space for 4 high (red) and 6 low (blue) ECS models, with the mean (solid line) and standard error (shading). Trade-wind regions are demarcated by the dashed lines from 10–30 hPa/day. (b) and (c) show the thermodynamic component in EIS-ω_{700} space with the boxes defined in the same way as in Fig. 6-3.

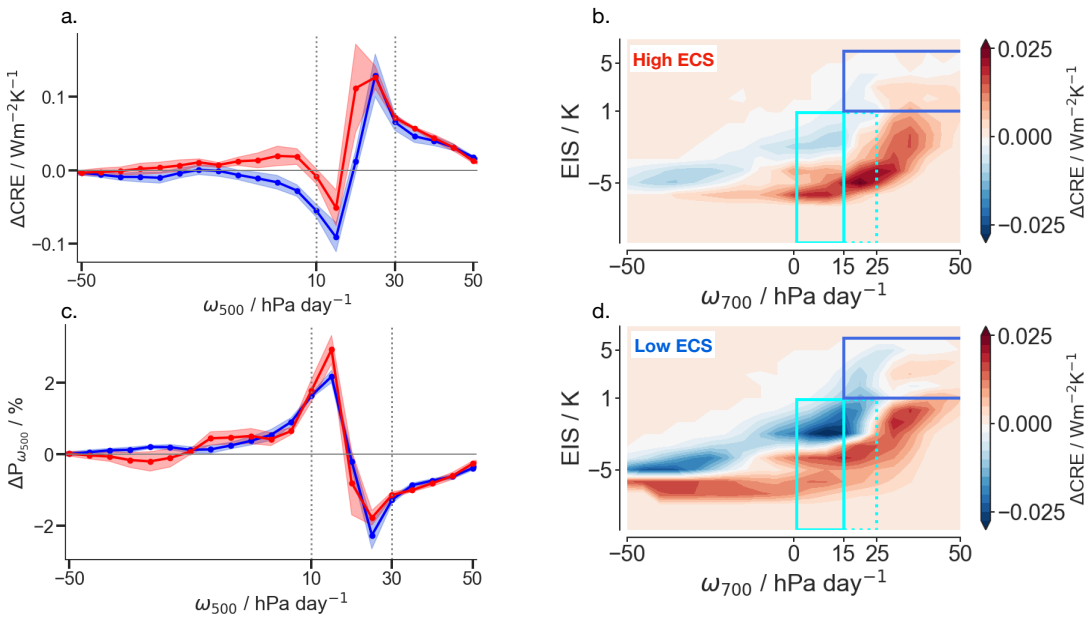


Figure 6-6: (a) The net ΔCRE in ω_{500} space for 4 high (red) and 6 low (blue) ECS models, with the mean and standard error in shading. Trade-wind regions are demarcated by the dashed lines from 10–30 hPa/day. (b) and (d) show the net ΔCRE in EIS-ω_{700} space with the boxes defined in the same way as in Fig. 6-3. (c) shows the change in regime frequency, $\delta P_{\omega_{500}}$ between *amip4K* and *amip* simulations.

yet through their ubiquity, or large statistical weight in P_ω, they have a large net influence on the tropical radiation budget. If defining the trades from 10–30 hPa/day in ω_{500}, the fractional contribution of the trades to the tropical ΔCRE is 20% for low ECS models and 46% for high ECS models. In general, the contribution to the spread of a variable, such as ΔCRE, can be quantified as,

$$\frac{\sum_{\omega_{\text{trades}}}(\Delta CRE_\omega^{\text{h}} - \Delta CRE_\omega^{\text{l}})}{\overline{\Delta CRE}^{\text{h}} - \overline{\Delta CRE}^{\text{l}}}, \tag{6.6}$$

where $\overline{\Delta CRE}$ is the tropical-mean ΔCRE for high (h) or low (l) ECS models. As one example, applying Eq. (6.6) to the thermodynamic ΔCRE and defining the trades from 10–30 hPa/day in ω_{500}, trade-wind regimes contribute to 65% of the spread in the tropical thermodynamic ΔCRE. This contribution is similar for different definitions of the trades: 57% if defining the trades from 10–25 hPa/day, or 78% if defining from 0–30h hPa/day.

Fig. 6-6 shows the tropical net ΔCRE, including both the dynamic and thermodynamic component, and the small covariance term. The correlation between the tropical thermodynamic and net ΔCRE values is r=0.66. The change in regime frequency, ΔP_ω, (Fig. 6-6c) illustrates the well-established finding that the large-scale tropical circulation weakens with warming. In trade cumulus regimes, weakly-subsiding regions become more frequent at the expense of more strongly-subsiding regimes, with ΔP_ω values crossing zero around 15 hPa/day. The structure and magnitude of the ΔCRE is driven more by the dynamic response to changes in the large-scale circulation than the thermodynamic response. In particular, the dipole structure in the net ΔCRE reflects the influence of the dynamic component of ΔCRE, driven by the dipole structure in ΔP_ω (Fig. 6-6c). Low ECS models have more negative, stabilizing ΔCRE in weakly-subsiding regimes than do high ECS models (Fig. 6-6a,d). In low ECS models, the more stabilizing ΔCRE likely arises from a positive dynamic change in cloud fraction near cloud base and the trade-wind inversion (Fig. 6-8b) and a less-negative thermodynamic component of cloud fraction changes (Fig. 6-8a). These results may suggest that low ECS models simulate stronger estimated inversion strength (EIS) (Wood and Bretherton, 2006) in these weakly-subsiding regimes, whereas high ECS models are associated with weaker EIS, as larger EIS is associated with increased cloudiness (e.g., Wood and Bretherton, 2006; Myers and Norris, 2013).

6.4.3 Global influence of spread in trade cumulus feedbacks

Fig. 6-7, moreover, shows that these differences in the trade-wind ΔCRE can explain differences in the global ΔCRE and ECS. Here the trades are selected using the EIS and ω_{700} criteria, but results are qualitatively similar if using the ω_{500} criterion. Differences in the net trade-wind ΔCRE explain 69% of the variance in the global cloud feedback (Fig. 6-7a) and 25% of the variance in ECS (Fig. 6-7b). Defining the trades using ω_{700} values up to 25 hPa/day instead of 15 hPa/day, as in Myers et al. (2021), following discussions in Sec. 6.4.1, increases the mean trade-wind ΔCRE from a mean of 0.31 to 0.46 $\mathrm{Wm^{-2}K^{-1}}$. The correlation between the different ΔCRE values is, however, one-to-one, r=0.996, showing that the choice of upper bound on ω_{700} does

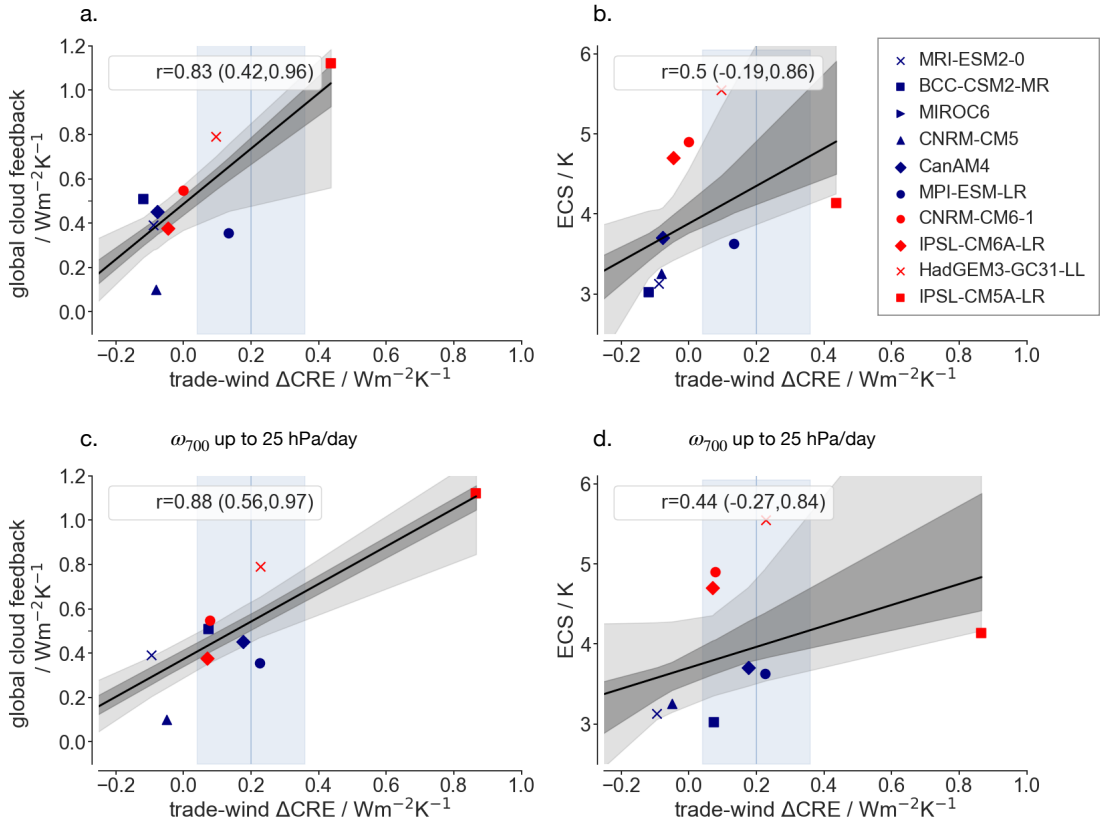


Figure 6-7: Scatter plots between the (a) trade-wind net ΔCRE and global ΔCRE (b) and trade-wind net ΔCRE and ECS. Red is ECS > 4°C, and blue is < 4°C. Uncertainty is 5–95% uncertainty on the regression defined from bootstrapping. Trade-wind regions are selected from the EIS-ω_{700} criterion. Also shown is the mean (blue vertical line) and standard deviation (blue shading) of the marine shallow cloud feedback given in IPCC AR6.

not strongly influence the relative differences among models (cf. Fig. 6-7c,d). Also illustrated is the mean and standard deviation of the marine shallow cloud feedback given in the International Panel of Climate Change's Sixth Assessment Report (IPCC AR6), which is assessed to be 0.2±0.16 Wm^{-2}K^{-1} from multiple lines of evidence (Arias et al., 2021). The broad standard deviation reflects uncertainties associated with different approaches of constraining trade cumulus feedbacks (Arias et al., 2021), as discussed in Chapt. 1.

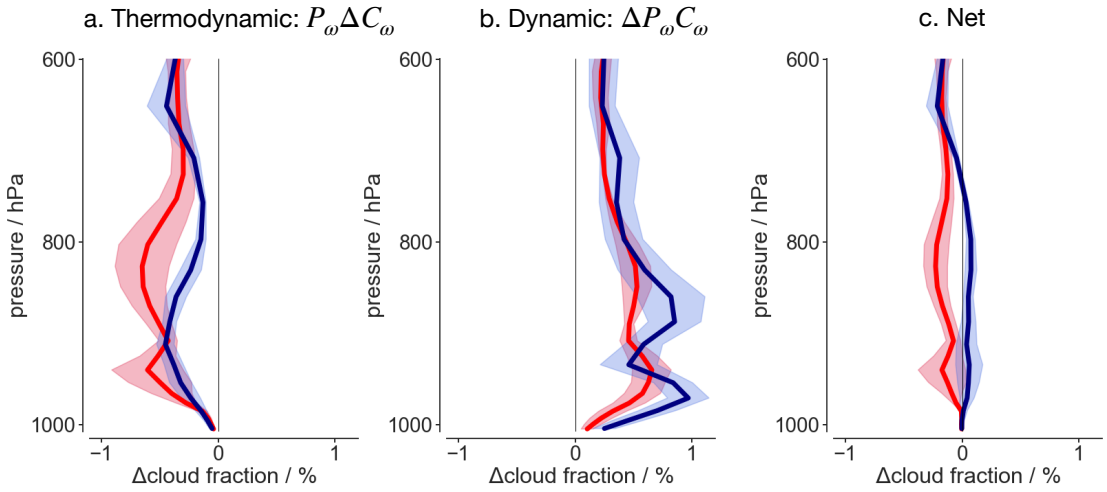


Figure 6-8: Contribution of cloud fraction changes in trades to the tropical cloud fraction change between *amip4K* and *amip* simulations for high (4 models, red) and low ECS models (6 models, blue). Shown are the (a) thermodynamic change, $P_\omega \Delta C_\omega$, (b) dynamic change, ${}_\omega C_\omega$, and (c) total change, $\overline{\delta C_\omega}$, (right), where C_ω refers to cloud fraction. The shading is the standard error. Trade-wind regions are selected from the ω_{700} and EIS criteria as described in the text.

6.4.4 Vertical cloud fraction changes in high and low ECS models

The change in the vertical cloud fraction in the trades also differs between high and low ECS models, similar to previous ensembles (e.g., Vial et al., 2013; Brient et al., 2016). In previous ensembles, the trade cumulus feedback was shown to largely be governed by changes in cloud fraction near cloud-base, which is, in turn, sensitive to the representation of turbulence, convection, and radiation in models (e.g., Brient et al., 2016; Vial et al., 2016).

Fig. 6-8 shows the change in trade-wind cloud fraction, expressed as the contribution to the net tropical cloud fraction change, which is on the order of one- or two-percent across models. Trade-wind regions are selected from the ω_{700} and EIS criteria, but results are similar when selecting on the basis of ω_{500}. The three panels correspond to the thermodynamic change (Fig. 6-8a), dynamic change (Fig. 6-8b), and the total change (Fig. 6-8c). Cloud changes between high and low sensitivity

models tend to be most pronounced at two levels – around cloud base at about 930 hPa ($\sim$ 700 m) and the inversion base at about 800 hPa ($\sim$ 2 km). The next chapter will focus on cloud fraction changes at cloud-base.

6.5 Analysis of climate sensitivity in the IPSL-CM6A-LR model

Before concluding, a more in-depth examination of the causes of higher ECS in the IPSL-CM6A-LR model relative to its predecessor, IPSL-CM5A-LR, is discussed. This analysis contributed to Boucher et al. (2020) (its Sec. 6 on climate sensitivity) and is described below.

When calculated with the Gregory et al. (2004) method, the effective ECS between IPSL-CM5A-LR and IPSL-CM6A-LR increases from 4.0 to 4.5 K using 150 years of data, or from 4.1 to 4.8 K using 300 years of data. The relative contributions to ECS are calculated following Dufresne and Bony (2008) and Vial et al. (2013), and illustrated in Fig. 6-9a. This method decomposes the contributions to ECS into (i) rapid tropospheric and stratospheric adjustments to carbon dioxide and (ii) temperature-mediated feedbacks operating on longer time scales. More specifically, the rapid tropospheric adjustment includes the climate response associated with all tropospheric adjustments (temperature, water vapor, and clouds), surface albedo change, and the small land surface warming due to the CO2 forcing (Vial et al., 2013). The method also quantifies the relative contributions of the water vapor and temperature lapse rate, surface albedo, and cloud feedbacks. Individual feedbacks are calculated by the radiative kernel method (Bony et al., 2006; Soden et al., 2008; Shell et al., 2008). A radiative kernel acts as a partial derivative, representing the sensitivity of the radiative flux, R, to changes in a climate variable, X, such as water vapor, temperature, and surface albedo. To find the kernel, the radiative code of a climate model is run offline with a standard perturbation, such as 1 K warming at each vertical level, moistening that would occur from warming by 1 K at constant

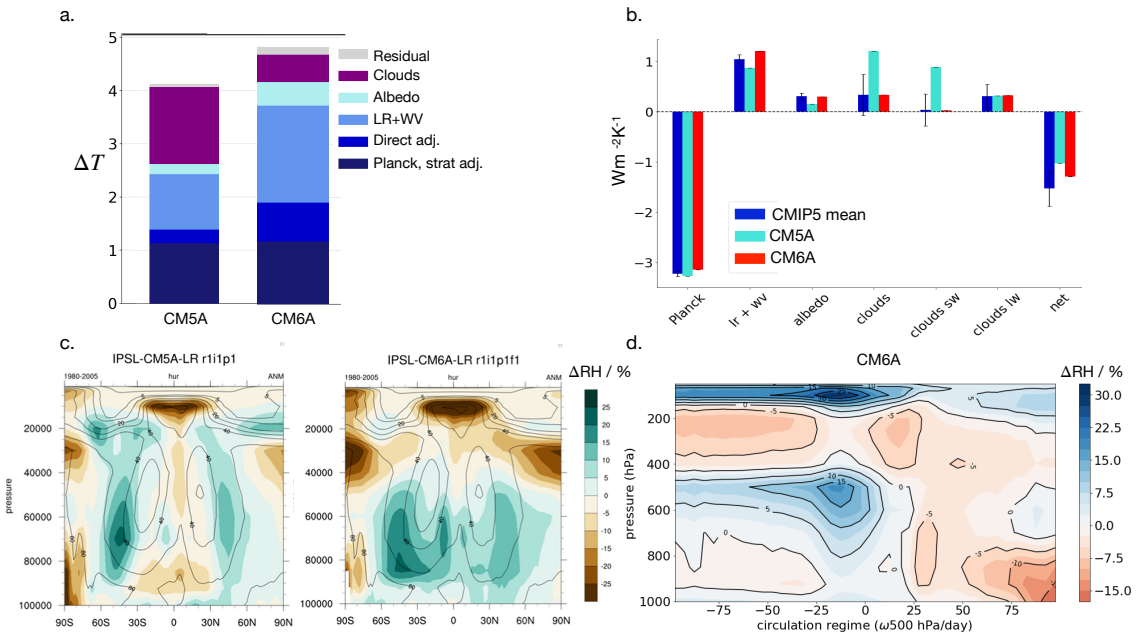


Figure 6-9: (a) Changes in radiative feedbacks calculated with the kernel method following Vial et al. (2013) between IPSL-CM5A-LR and IPSL-CM6A-LR, with each model labelled as either 'CM5A' or 'CM6A' (b) also compared with the CMIP5 mean feedback values given in Ceppi et al. (2017). Panel (c) shows that a relative humidity bias in the CMIP5 model version (left) increased further in IPSL-CM6A-LR (right) relative to ERA Interim reanalysis. (d) Differences between relative humidity (%) after 150 years of the *abrupt-4xCO2* experiment and values in the *piControl* experiment for the IPSL-CM6A-LR model

.

relative humidity, and changing surface albedo by 1%. The radiative kernel, $\frac{\partial R}{\partial X}$, is multiplied by the change in the climate variable of interest diagnosed from a model simulation and then normalized by the global-mean surface temperature change to yield the feedback value,

$$\lambda_x = \frac{\partial R}{\partial X}\frac{\partial X}{d\overline{T}}. \tag{6.7}$$

We employ the same kernels as in Shell et al. (2008) for water vapor, temperature, and surface albedo. The cloud feedback is calculated as a corrected residual term, correcting for a cloud-masking term following Vial et al. (2013), which adds a consistent offset to net cloud feedback value estimated from the cloud radiative effect method (e.g., Andrews et al., 2012). A small residual term reflects nonlinearities in the relationship between radiative perturbation and the temperature response.

The main drivers of this larger ECS in IPSL-CM6A-LR are, somewhat unexpectedly, more positive rapid tropospheric adjustment to CO2, and a stronger combined lapse rate and water vapor feedback (Fig. 6-9a), rather than a stronger cloud feedback. We diagnose the strong tropospheric adjustment from *aqua-4xCO2* and *amip-4xCO2* simulations, as well as the *abrupt-4xCO2* simulations, and find that the stronger adjustments come from clear-sky regimes (not shown). The stronger water vapor feedback primarily results from strong moistening tendencies in weakly-ascending regimes around 500 hPa (Fig. 6-9d). We diagnose this moistening tendency in weak ascent regimes by projecting the relative humidity anomalies, defined as the difference between relative humidity after 150 years of the *abrupt-4xCO2* simulation and the *piControl*, into a circulation regime basis, based on ω_{500} (Bony et al., 2004). Relative humidity anomalies reach up to 15% in these weak ascent regimes. The IPSL-CM6A-LR model is, however, also too moist in the historical tropical atmosphere compared with ERA-Interim data (Fig. 6-9c), which suggests the moistening under warming might be exaggerated as well. The net cloud feedback is less positive than in previous model versions (Fig. 6-9a). The IPSL-CM6A-LR model is somewhat of an outlier that predicts a high ECS without a strong global cloud feedback (e.g., Fig. 6-1a.,c).

An anticorrelation between the strength of the cloud and combined lapse rate and water vapor feedback was noted by Huybers (2010) for the CMIP3 ensemble. The two IPSL model versions appear to trade-off in the strength of the combined global lapse rate and water vapor feedback and cloud feedback. In the subset of 12 CMIP6 models analyzed here, such an anticorrelation is also evident (r=-0.68), driven more by the water vapor (r=-0.62) rather than lapse rate feedback (r=0.25). One physical hypothesis is that a mid-tropospheric increase in relative humidity and cloudiness in weakly-ascending regimes could reflect a trade-off between shallow and deep convection schemes. If thermals are not strong enough to trigger deep convection, water vapor and cloudiness accumulate in the mid-troposphere after being transported there by shallow convection, yielding a positive water vapor feedback and negative cloud feedback. This mechanism is, however, confined to a single tropical regime, whereas the feedbacks are global. It would be useful to more systematically examine

potential physical origins of such compensations, such as whether they are tied to a single regime or physical mechanism across models.

6.6 Initial conclusions

In a representative sample of CMIP6 models, differences in trade-wind cloud responses can still discriminate between high and low sensitivity models. The trade-wind thermodynamic and net change in the cloud radiative effect (ΔCRE), as well as vertical changes in cloud fraction, differ between high and low sensitivity models. The trade-wind net ΔCRE can, moreover, explain about 70% of the variance in the *global* cloud feedback. These analyses show that trade-wind cloud feedbacks are still a large source of uncertainty in global cloud feedbacks, even if other regions contribute to the spread in ECS in CMIP6 more so than in previous CMIP ensembles (e.g., Zelinka et al., 2020). These findings are similar to those shown for CMIP5, such as in Vial et al. (2013) and Brient et al. (2016), albeit for a larger number of models.

Three main conclusions are drawn from this chapter that motivate analyses in the following chapter. First, large differences remain in CMIP6 among modeled trade cumulus cloud responses to warming. Second, these differences in trade cumulus cloud responses to warming differ between high and low ECS models. Third, the inability to assess which group of modeled responses are more physically credible highlights a fundamental gap in our understanding of the environmental controls on trade-wind cloudiness. The next chapter tries to improve this understanding with EUREC^{4}A observations in order to constrain trade cumulus feedbacks.

Chapter 7

Constraining trade cumulus feedbacks with EUREC^{4}A

7.1 Introduction

The skill of mixed layer theory for obsessionally closing subcloud layer moisture and energy budgets (Chapt. 4) gives confidence that it can also be applied to the subcloud layer mass budget. In the following chapter, the motivation and key results of Vogel et al. are summarized, with a focus on my contributions. Vogel et al. use mixed layer theory and novel sampling strategies to perform the first observational test of the 'mixing-desiccation' hypothesis.

I am the second author on the Vogel et al. study and contributed to the analysis in two main ways. My first contribution was constraining the entrainment rate from the subcloud layer thermodynamic budgets and estimating the subcloud layer depth, as discussed in Chapt. 4. Together with Jessica Vial, my second contribution is analyzing how GCMs represent the mixing-desiccation hypothesis in present-day (Sec. 7.4) and relating these representations to trade-wind cloud radiative responses to warming (Sec. 7.5).

Analysis in this chapter contributes to Vogel et al., which is in preparation for submission to *Science*.

7.2 Mixing-desiccation hypothesis

EUREC^{4}A was originally conceived to observationally constrain trade-wind cloud feedbacks that were shown to explain large differences in climate sensitivity (Bony et al., 2017; Stevens et al., 2021), as introduced in Chapt. 1 and further discussed in Chapt. 6. A substantial chain of research has formulated a specific physical idea about what influences changes in cloudiness with warming. This idea can be expressed as the *mixing-desiccation hypothesis*, or the *cloud fraction vs. mass flux dilemma*. An increased convective mass flux evacuates more mass from the subcloud layer, deepening the cloud layer. This deepening through cloud formation causes mixing at the cloud top, which brings down dry air from the free troposphere. As a result, the moist lower troposphere becomes relatively drier, leading to the evaporation, or desiccation, of clouds near their base. As a result, the mixing-desiccation hypothesis predicts an inverse relationship between convective mixing and cloud-base cloud fraction (e.g., Gettelman et al., 2012; Rieck et al., 2012; Zhang et al., 2013; Sherwood et al., 2014; Tomassini et al., 2014; Brient et al., 2016; Vial et al., 2016; Bony et al., 2017). In a warmer climate, the lower tropospheric humidity gradient increases, following the nonlinear Clausius-Clapeyron relationship. All else equal, vertical mixing even more efficiently dries the environment at cloud-base. The decrease in cloud-base cloudiness is expected to lead to a more positive cloud feedback (e.g., Bony et al., 2017).

Early support for this conceptual picture comes by analogy with the process of stratocumulus breaking up into cumulus. Bretherton and Wyant (1997) use a mixed layer model and find that increasing surface fluxes, relative to cloud-top radiative cooling, drive the decoupling of the stratocumulus layer. That is, increased surface fluxes, and thus convective mixing (e.g., Tiedtke, 1989), appear to break up stratocumulus into cumulus and reduce cloud fraction. Studies with GCM-ensembles and single GCMs also support the idea that increased convective mixing decreases cloudiness at cloud-base. Sherwood et al. (2014) use 43 GCMs (in CMIP5) to show that differences in the simulated strength of convective mixing between the lower and mid-troposphere in the tropics can explain about 50% of the spread in climate sensitivity.

Stevens et al. (2016), Tomassini et al. (2014), and Vial et al. (2016) use single climate models run in configurations that differ in their convective mixing parameters and also find evidence that increased convective mixing decreases cloud-base cloudiness.

Other approaches, however, do not find a strong reduction in cloud-base cloudiness with convective mixing. As discussed in Chapt. 1, process-based studies using large-eddy simulations suggest that low clouds are more resilient to changing environmental conditions than many GCM studies suggest (e.g., Blossey et al., 2013; Zhang et al., 2013; Bretherton, 2015; Vogel et al., 2016; Radtke et al., 2021), supporting a neutral or only slightly positive trade cumulus cloud feedback. Using the MIROC model, Kamae et al. (2016) find that differences in lower-tropospheric mixing can explain the spread in low-cloud feedbacks in only half of a perturbed physics ensemble. Myers et al. (2021) and Cesana and Del Genio (2021) use satellite measurements and also find evidence for a near-zero trade-wind cloud feedback. The physical mechanisms behind an apparent robustness of trade-wind clouds to warming are, however, not yet understood.

An alternate null hypothesis is that increased convective mixing *increases* cloud-base cloudiness by moistening the large-scale environment at cloud-base. Such a hypothesis is not the mixing-desiccation hypothesis, but could instead be called the mixing-moistening hypothesis. There are multiple reasons why the mixing-desiccation hypothesis could be wrong. Observed mesoscale vertical velocity is, for instance, larger than the longer timescale-mean values that are typically used in models of varying complexity (Bony and Stevens, 2019; George et al., 2021). This larger variability in vertical velocity could introduce dynamic controls on cloudiness that might be missing in models. It is also not *a priori* clear whether an increased mass flux and entrainment necessarily decrease relative humidity, or could instead increase relative humidity by deepening the subcloud layer (Vogel et al., 2020).

7.3 First observational test of this mechanism

7.3.1 Measurements and methods

To observationally test this hypothesis requires jointly measuring the cloud fraction (C), mass flux (M), and relative humidity (RH) near cloud base. Vogel et al., in prep estimate all of these quantities for the first time at the process level using EUREC^{4}A observations. Here the key points related to the measurements and methods are summarized. Further information is found in Vogel et al., in prep.

To measure the cloud fraction at cloud base, C, the ATR-42 aircraft had horizontally-staring backscatter lidar operating at 355 nm ('ALIAS', *Airborne Lidar for Atmospheric Studies*) (Chazette et al., 2020) and horizontally-staring Doppler cloud radar ('BASTA', *Bistatic rAdar SysTem for Atmospheric studies*) (Bony et al., 2022). A radar-lidar synergy product of cloud fraction near cloud-base product agrees well with independent, coincident estimates (see Fig. 17 of Bony et al. (2022)). The estimate includes drizzle, making it an upper bound on cloud-base cloud fraction.

Vogel et al., in prep consider the mass flux, M, as a proxy for the lower-tropospheric convective mixing. It can be estimated as a residual from the subcloud layer mass budget,

$$\frac{Dh}{dt} = E - M + W. \tag{7.1}$$

The mass balance of the subcloud layer is based on mixed layer theory, as also evaluated in Chapt. 4 for subcloud layer moisture and energy budgets. In the mass balance, the subcloud layer, having a depth h, is controlled by the entrainment rate, E (a mass source), the mass flux, M (a mass sink), and the large-scale vertical velocity, W (either a mass source or sink), as, for instance, described in Stevens (2006). As preparation for EUREC^{4}A, Vogel et al. (2020) also evaluated the skill of this mass budget framework using large-eddy simulations. These terms have units of height per time, such as mm/s. Note that depth, h, multiplied by area and density corresponds to mass.

If assuming stationarity and homogeneity, $\frac{Dh}{dt}$=0, M can be solved for as the

residual of the subcloud layer mass budget. That it, the mass flux is the sum of E and W,

$$M = E + W. \tag{7.2}$$

Vogel et al., in prep finds that the storage and advection terms are small and unbiased, such that these assumptions are justified. Results are also qualitatively similar when including the total derivative term.

Whereas C is measured from the ATR-42, other terms, such as those to calculate M, are estimated using over 800 dropsondes launched from the coincident HALO aircraft. The subcloud layer height, h, is estimated using the θ_v-gradient method as described in Chapt. 2. The method for calculating the entrainment rate, E, is also described in Chapt. 2. Chapt. 4 constrain uncertain entrainment parameters using a Bayesian inversion of subcloud layer moisture and energy budgets. This approach assesses the skill of mixed layer theory and provides an independent constraint on E for the mass budget in Vogel et al., in prep. The vertical velocity, W, is historically challenging to measure. During two smaller field campaigns in preparation for EUREC^{4}A, novel sampling strategies were developed and tested to measure W (Bony and Stevens, 2019). With EUREC^{4}A data, George et al. (2021) compute vertical profiles of W by vertically integrating the divergence of the horizontal wind field measured by dropsondes from the surface up to the flight level. These measurements allow for capturing variations in the strength of mesoscale circulations (Bony and Stevens, 2019; George et al., 2021). In this analysis, W is taken at the subcloud layer top, h. Relative humidity, RH, is also estimated from HALO dropsonde measurements at h. Results are similar when considering RH measured by the ATR-42. All terms are computed at the one-hour circle-scale and then aggregated to the three-hourly circling-scale (see terminology in Chapt. 4, Table 1).

7.3.2 Observed relationships

To test the mixing-desiccation hypothesis, Vogel et al., in prep perform a multiple linear regression between cloudiness, C, and mass flux, M, and relative humidity,

RH,

$$C = a_{\mathrm{M}} M + a_{\mathrm{RH}} RH + C_0. \tag{7.3}$$

M and RH are not correlated, with a Pearson correlation coefficient r=-0.075.

Contrary to the anticorrelation between C and M outlined by the mixing-desiccation hypothesis, the observed correlation between C and M near cloud-base is strong and positive, with r=0.72. That is, increased convective mixing is associated with increased, rather than decreased cloudiness. Also including RH further tightens the relationship, giving a correlation between predicted and observed C of r=0.83, though the relationship remains driven mostly by M variations. Adding relative humidity presumably improves the correlation because the large-scale moisture conditions influence the persistence of clouds and because not all clouds are active clouds and associated with a mass flux (e.g., Stull, 1985). The physical mechanism will be further discussed in Sec. 7.5, and more thoroughly presented in Vogel et al., in prep.

7.4 Comparison with GCMs

As a point of comparison, we examine how GCMs represent the couplings among the same terms, C, M, and RH. Hourly modeling output at so-called *cfSites*-locations is produced as part of the *Cloud Feedback Model Intercomparison Project* (CFMIP). These 120 locations are sometimes coincident with measurement stations, such as the Barbados Cloud Observatory, and are generally located in areas exhibiting a large spread in intermodel cloud feedbacks (e.g. see Fig. 2 in Webb et al. (2017)). Despite the general notion that GCMs do not produce output in the 'space of observables', these point-wise model outputs are more amenable to comparison with *in situ* observations than GCM outputs on larger grid scales.

C, M, and RH are calculated for ten CMIP models (four from CMIP5 and six from CMIP6) using the *cfSites* output. As in Vial et al., in prep, we use the *amip* configuration from 1979–2008, selecting data in January and February to correspond to EUREC^{4}A measurements. For each model, between 2–6 *cfSites* locations are available in the north Atlantic trades between 59–44 °W and 11–16 °N. All profiles with

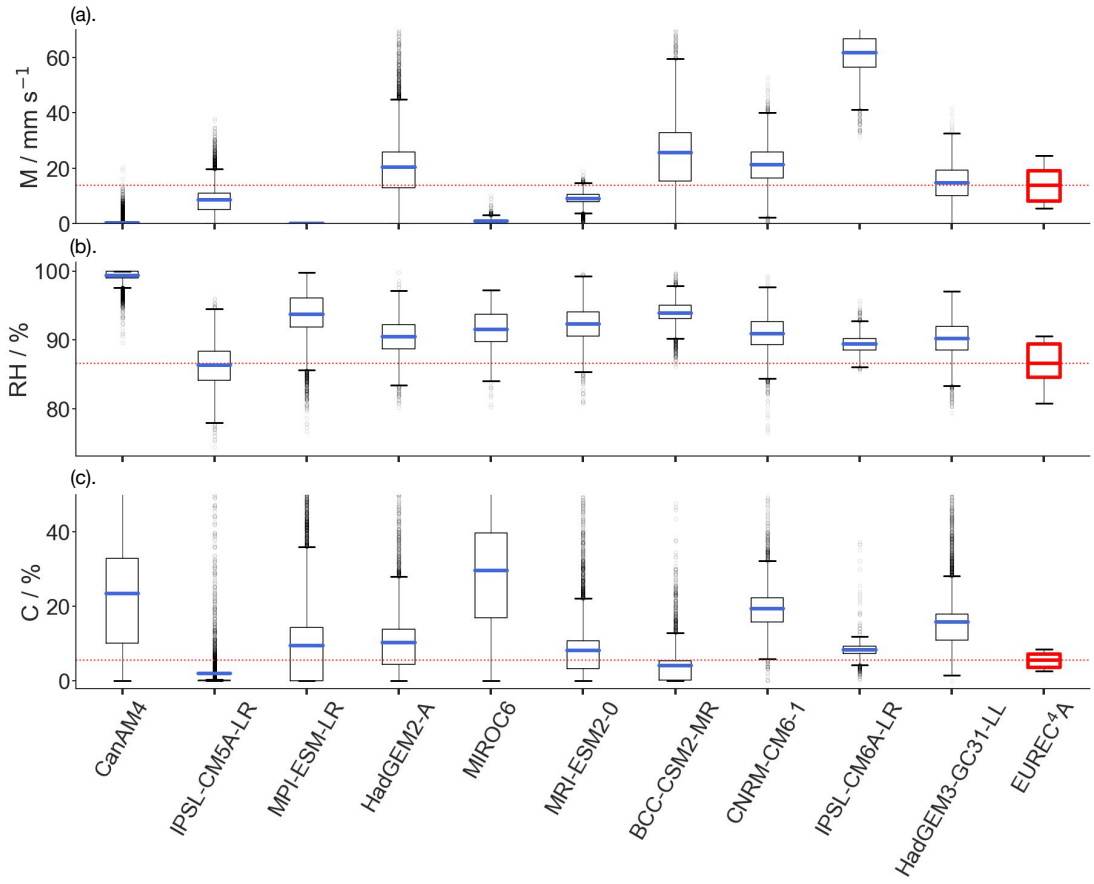


Figure 7-1: Box plots of *cfSite* output from ten *amip* model simulations for (a) mass flux, M, (b), relative humidity, RH, and (c), cloud fraction, C. The box plot corresponds to the interquartile range (from the first quartile to the third quartile), with mean values shown in red (medians not shown). Whiskers extend from the box by 1.5 times the interquartile range. Flier points are those extending past the whiskers. The interquartile range of observed EUREC^{4}A values are outlined in red, and the mean observed value is given by the dotted red line.

clouds above 600 hPa (about 4.2 km) are dropped to ensure a focus on shallow convection. Values near cloud-base are selected as the first maximum below 850 hPa (about 1.5 km). These vertical levels can differ for cloudiness, mass flux, and relative humidity, but the height differences are, on average, zero. Results are qualitatively similar when constraining all values to occur at the same level. After calculating the near cloud-base values for all sites, we average across the available sites for each model. These spatially-averaged, hourly outputs are then aggregated to the three-hourly timescale, which corresponds to the circling-mean timescale of EUREC^{4}A data (see description in Chapt. 4). Hourly outputs are also aggregated to monthly-means for studying longer timescale variability.

Fig. 7-1 shows that these ten models simulate a diversity of cloud-base mass flux, relative humidity, and cloud fraction values for three-hourly data. Across models, the mean M ranges enormously from $1.25*10^{-11}$–61.6 mm/s, the mean RH ranges from 86.4–99.4%, and the mean C ranges from 2.06–29.7%. The observed M values is 15.2±6.54 mm/s, with values denoting the mean and standard deviation. Observed RH values are 86.6±3.06%, and observed C values are 5.38±1.94%. Except for one model (IPSL-CM5A-LR), the mean modeled RH value is higher than observed mean RH. Modeled M values vary widely, including by many orders of magnitude. Eight out of ten models predict C values greater than the observed value, and four out of ten models predict C values that are more than twice the observed value.

Among models, the relationships among these variables also differ strongly, both at the three-hourly and monthly timescale. Fig. 7-2a,b plot relationships among C and M and between C and RH using non-standardized data (e.g., not divided by the standard deviation) to show the diversity of simulated relationships. Given the sometimes large differences among modeled values, for instance, in the magnitude of M, values are, however, standardized in the following analysis. As a result of this standardization, the slope and correlation values are equal. In GCMs, the correlation between standardized C and M ranges from -0.12 to 0.55 in three-hourly data and -0.42 to 0.61 in monthly data. Three of ten GCMs have a negative correlation at the three-hourly timescale and seven GCMs have a weakly positive correlation of

r<0.3. The observed correlation between C and M is, by contrast, 0.72, with a 50% confidence interval from 0.65 to 0.82, outside the values spanned by these ten GCMs.

The relationships between C and RH in models also show large variability (Fig. 7-2b), yet these relationships overlap with the observed correlation. In the GCMs, correlations range from 0.097–0.84 in three-hourly model output and 0.21–0.84 in monthly model output. The correlation in three-hourly observed data is 0.36, towards the center of the GCM range, with a 50% confidence interval from 0.16–0.55.

7.5 Can present-day variability constrain future cloud changes?

One condition for the relevance of EUREC4A measurements for evaluating cloud feedbacks, as discussed in Chapt. 2, is that variability expressed on shorter timescales is informative of variability expressed on longer timescales. In the ten GCMs, a strong association emerges between correlations estimated at the three-hourly and monthly timescales, both for the C-M relationship (r=0.80, 0.33–0.95, 5-95% c.i.) and C-RH relationship (r=0.80, 0.21–0.96, 5-95% c.i.), as shown in Fig. 7-3. These correlations provide important support that inferences from EUREC^{4}A data are informative about climate scale behaviors. More generally, these correlations suggest that the physical processes underlying the relationships among C, M, and RH are relatively timescale-invariant.

A common approach to narrowing uncertainty in climate variables involves so-called 'emergent constraints' (e.g., Eyring et al., 2019; Hall et al., 2019). Emergent constraints relate an observable quantity with an uncertain, non-observable climate response parameter, such as cloud feedbacks or climate sensitivity. If a relationship (e.g., linear) emerges between these two variables, estimating the observable quantity provides a 'shortcut' to constrain the non-observable variable. Early emergent constraints were proposed for the hydrological cycle (Allen and Ingram, 2002) and snow-albedo feedback and (Hall and Qu, 2006). Regarding Northern Hemisphere

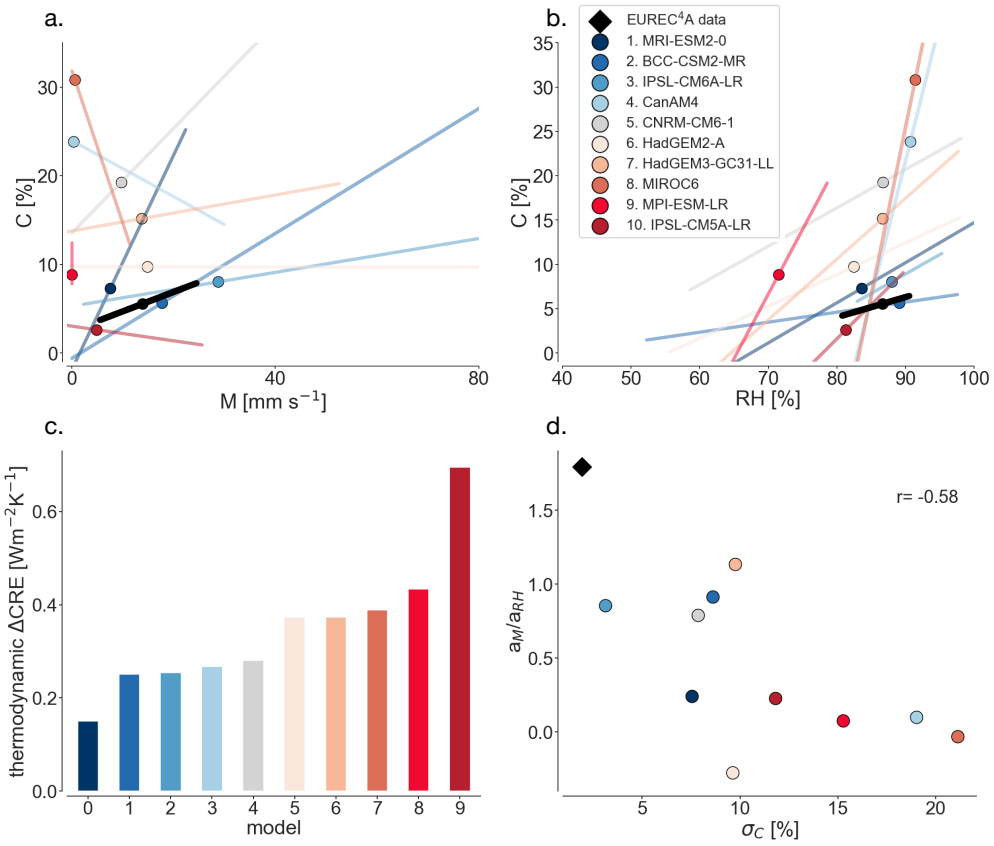


Figure 7-2: Diversity of relationships between (a) C and M and (b) between C and RH. Note that in panels a. and b., data are not standardized (divided by standard deviation) in order to show the variability among models, but the mass flux for the IPSL-CM6A-LR model is divided by three so that its magnitude is comparable to the other models. The circle refers the mean value, and the line is the linear regression fit to model output or observations. Colors are consistent throughout panels and correspond to values of the *thermodynamic* trade-wind ΔCRE, divided by the 4 K change in global-mean surface temperature, with blue to red corresponding to increasing values. Panel c. shows a bar plot of these thermodynamic ΔCRE values. Panel d. plots a_M/a_{RH} versus σ_C as described in Sec. 7.5.

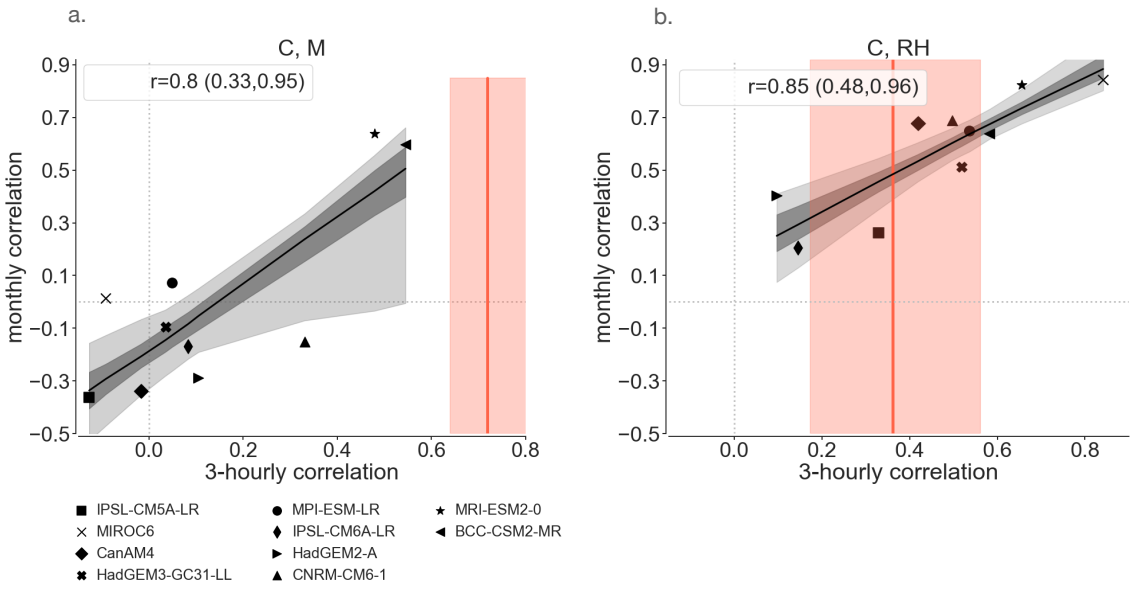


Figure 7-3: Relationship between three-hourly correlation and monthly correlation between cloud fraction and mass flux (a) and between cloud fraction and relative humidity (b). Symbols correspond to models as listed in the legend, and the black line is the ordinary least squares regression. Dark and light grey shading correspond to 50% and 90% confidence interval on the regression, as estimated by bootstrapping, or repeatedly sampling with replacement and performing the regression 1000-times. The (three-hourly) observed correlation is given by the orange vertical line, and shading corresponds to 50% uncertainty on the Pearson correlation as estimated by bootstrapping (rather than the Fisher transformation).

snow cover, Hall and Qu (2006) demonstrated that, in 17 CMIP3 models, there is a tight correlation between the amplitude of the seasonal cycle and decrease in snow cover per degree of local warming. This constraint is compelling because it is physically plausible that the same mechanism controls snow cover changes, whether due to seasonal changes or CO_2-mediated radiative forcing. Indeed, this constraint has persisted across multiple climate model ensembles (Qu et al., 2014; Thackeray et al., 2018).

Many other emergent constraints are less robust, casting doubt on the utility of this framework. Caldwell et al. (2018) found that only four of 19 published emergent constraints on climate sensitivity remained plausible when applied to out-of-sample data. Similarly, Schlund et al. (2020b) recalculated previous emergent constraints for the CMIP6 ensemble and found much lower correlations of these metrics with climate sensitivity than with the ensembles for which these constraints were calculated. More generally, the limited sample size of models and near-infinite number of observable quantities implies that some spurious emergent relationships can arise solely due to chance (e.g., Caldwell et al., 2014). The framework is therefore criticized as being the result of 'data mining' whose inferences should be treated with caution (e.g., Caldwell et al., 2014, 2018; Lutsko et al., 2021). As GCMs become increasingly sophisticated, multiple factors can explain the spread in global quantities, such as climate sensitivity, (e.g., Zelinka et al., 2020), suggesting that a single metric will no longer have global explanatory power. Emergent constraints should therefore instead target specific physical processes or regions (e.g., Klein and Hall, 2015; Brient and Schneider, 2016), which we adopt as the approach in this work. Such improved physical understanding can then be used to assess the credibility of model representations, or even rule out certain models, whether or not a linear relationship emerges between present-day and future variables.

7.5.1 Observational constraints

We consider three metrics, which can be applied to both three-hourly observations and model output and then related to future cloud changes: (1) the correlation between C

and M, as shown in Fig. 7-2a; (2) a_M/a_{RH}, the ratio of the multiple linear regression coefficients of Eq. (7.3); and (3) σ_C, the standard deviation of C.

The first metric quantifies how increasing convective mixing changes the cloud-base cloud fraction. Fig. 7-4 shows that a linear relationship emerges between the correlation between C and M and the change in the trade-wind thermodynamic ΔCRE (r=-0.71). There are indications that models simulating a more positive correlation between C and M, in better agreement with observations, predict less-positive thermodynamic ΔCRE values, and vice versa. With the other two metrics, a linear relationship does not emerge with the thermodynamic or net ΔCRE. The correlation with the thermodynamic ΔCRE is -0.19 with a_M/a_{RH} and 0.30 with σ_C.

The two other metrics nonetheless allow for assessing models relative to observational constraints, as modeled metrics both have a large spread and differ from observational metrics. The ratio a_M/a_{RH} quantifies the relative dependence of cloudiness on M, a dynamic control, versus RH, a more thermodynamic control. A greater dependence on RH is, moreover, expected if increasing the mass flux decreases relative humidity, which then decreases cloudiness, in line with the mixing-desiccation hypothesis. Fig. 7-2d shows that all models underestimate a_M/a_{RH} relative to its observed value. This finding suggests that models have a larger dependence on RH (larger a_{RH} regression coefficient) and a weaker dependence on M than observations. Expressed differently, models whose cloudiness depends more on M variations, rather than RH variations are in better agreement with observations. These models, moreover, tend to predict weaker thermodynamic ΔCRE values, though a clear linear relationship does not emerge. These findings suggest that observed clouds are more dynamically controlled, by M, and less thermodynamically controlled, by RH, than clouds in GCMs.

Fig. 7-2d also shows that all models overestimate σ_C, the variability in cloud-base cloud fraction. Nine of ten models have σ_C values more than three times the observed value. The standard deviation in the IPSL-CM6A-LR (3.2%) is closest to the observed value (1.9%). Both the mean cloud fraction (Fig. 7-1c) and its variability through σ_C are overestimated. Moreover, models that produce a larger mean C also

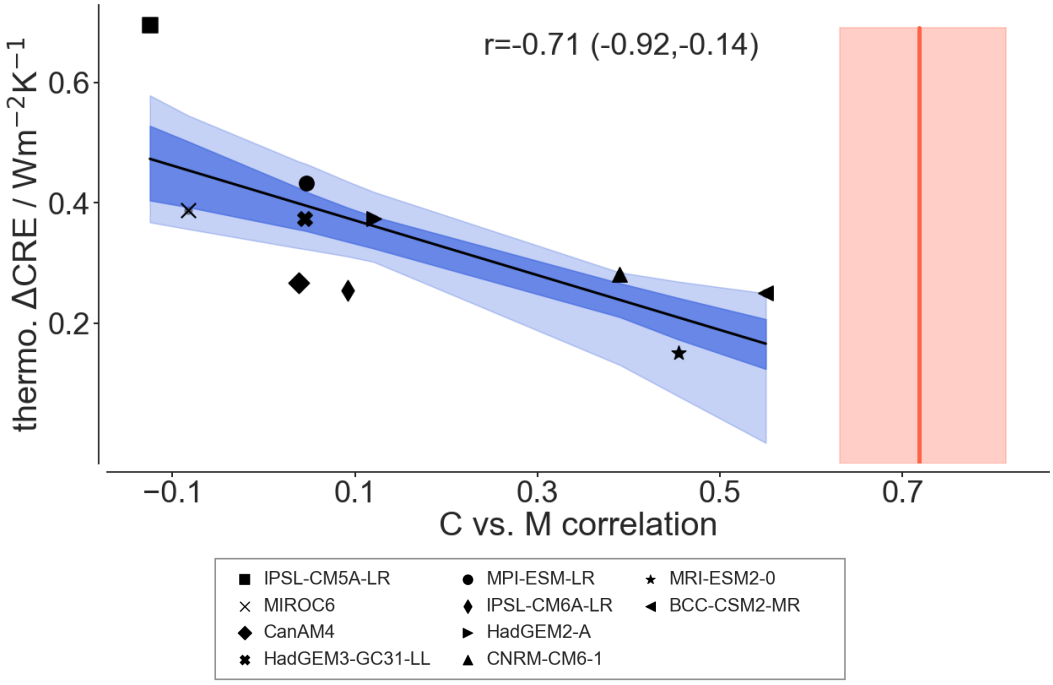


Figure 7-4: Correlation between cloud fraction and mass flux versus trade-wind thermodynamic ΔCRE. As in Fig. 7-3, symbols again correspond to models as listed in the legend, and the black line is the ordinary least squares regression. Dark and light blue shading correspond to 50% and 90% confidence interval on the regression, as estimated by bootstrapping, or repeatedly sampling with replacement and performing the regression 1000-times. The (three-hourly) observed correlation is given in the orange vertical line with shading corresponds to 50% uncertainty on the Pearson correlation as estimated by bootstrapping, rather than the Fisher transformation.

simulate more variation about this mean value, with a correlation r=0.62 between the mean C and σ_C. One speculation is that some GCMs represent trade-cumulus clouds more similarly to stratocumulus clouds. The larger mean cloud fraction and larger variability could perhaps be more analogous to stratocumulus decks forming and breaking up than the more scattered trade cumulus fields with lower mean cloud fraction and lower variability.

In terms of obtaining a probabilistic estimate of the trade cumulus feedback, the three metrics co-vary and therefore cannot be treated as fully independent constraints. The correlation between the C and M correlation is 0.49 with a_M/a_{RH} and -0.53 with σ_C. The correlation between a_M/a_{RH} and σ_C is -0.58. These considerations motivate

the development of a framework to consider constraints jointly (Sec. 7.6 and Chapt. 8).

7.5.2 Comparison with other cloud feedback estimates

In Fig. 7-5, one metric, the C vs. M correlation is related to the trade cumulus feedbacks estimated from Myers et al. (2021) as a check of consistency. The Myers et al. (2021) trade-wind cloud feedbacks are calculated from different model simulations (*abrupt-4xCO2* rather than *amip4K* simulations), but we use the same regime partitioning based on ω_{700} and EIS as in Myers et al. (2021) to identify the trades as described in Chapt. 6. There are seven models that overlap between our analyses. Fig. 7-5 relates the Myers et al. (2021) trade cumulus feedbacks to the C vs. M correlation for GCMs and observations. The MIROC6 model emerges as an outlier whose trade cumulus cloud feedback is much smaller in the coupled (*abrupt-4xCO2*) than uncoupled (*amip4K*) simulations (Fig. 7-5b). Including the MIROC6 model, the correlation between the Myers et al. (2021) trade cumulus feedback and C vs. M correlation is -0.37. If excluding MIROC6, the correlation is much stronger, r=-0.92. The best linear regression fit to GCMs is also extrapolated to the observed correlation value. The two lines including and excluding MIROC6 span the 90% confidence interval from Myers et al. (2021), as shown in Fig. 7-5, which provides additional support for a weak trade cumulus feedback.

7.6 Discussion and initial conclusions

This chapter described key results of Vogel et al., with a focus on my contributions. Vogel et al. present novel measurements of the convective mass flux, cloud fraction and relative humidity at cloud base from the recent EUREC^{4}A field campaign. These measurements allow for performing the first observational test of the mixing-desiccation hypothesis, which is thought to explain large differences in trade cumulus feedbacks and thus equilibrium climate sensitivity.

Instead of the negative relationship between cloudiness and the convective mass flux outlined by the mixing-desiccation hypothesis, observations show a strong, posi-

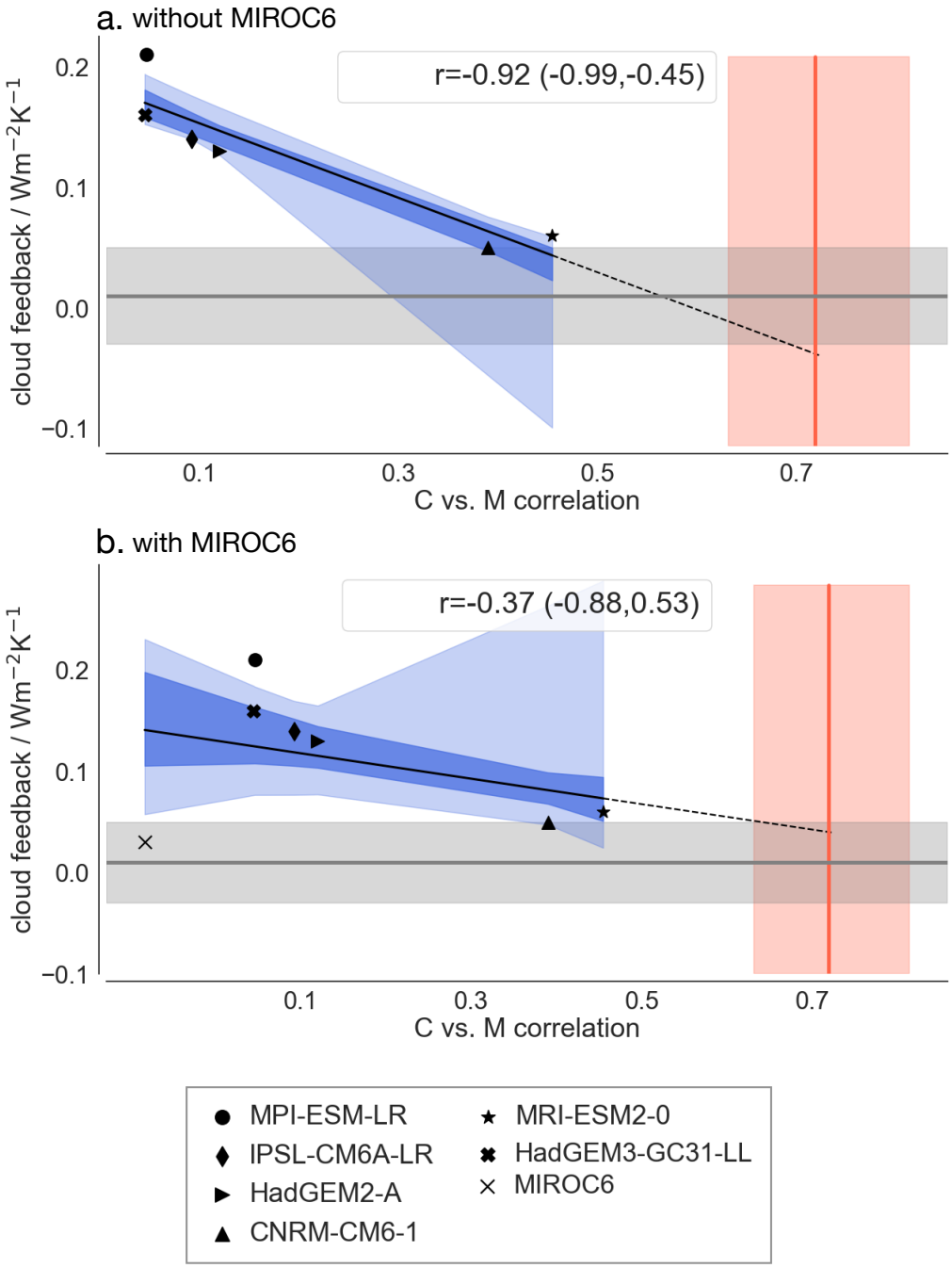


Figure 7-5: As in Fig. 7-4, but with trade-wind cloud feedbacks estimated from Myers et al. (2021) for the six models that overlap with our analysis. Also shown is the best estimate of the trade cumulus feedback (grey horizontal line) and the 90% confidence interval (grey horizontal shading), both from Myers et al. (2021). Also shown is an extrapolation of the best-fit regression line from modeled correlations between the cloud fraction and mass flux to the observed correlation (dashed black line).

tive relationship between cloudiness and the mass flux. This positive relationship is strengthened by also including relative humidity at cloud base. Observations therefore suggest that cloudiness is principally controlled dynamically through the mass flux, whereas an analysis of ten GCMs shows that cloudiness in these models is more controlled thermodynamically by relative humidity variations. GCMs also tend to overestimate both the mean and variability in cloud fraction compared with observations, which could potentially result from a more stratocumulus- than trade cumulus-like representation in GCMs.

We propose three metrics related to these couplings among cloudiness, mass flux, and relative humidity, which can be used to observationally constrain the trade cumulus feedback. Models that simulate a negative coupling between cloudiness and the mass flux, as opposed to the positive observed relationship, tend to produce a stronger thermodynamic cloud feedback (thermodynamic ΔCRE) with warming. Different approaches exist for combining multiple physical constraints in a statistical framework to yield a probabilistic estimate of the trade cumulus feedback, which are discussed in Chapt. 8 as ongoing work.

The findings described in this chapter and more-fully in Vogel et al. provide the first constraint on the trade cumulus feedback at the process-level. The results support a weak trade cumulus feedback, consistent with large-eddy simulations and satellite studies, as discussed in Chapt. 1 and Sec. 7.2. The present work distinguishes itself, however, from these foregoing studies because it also elucidates a physical mechanism behind the robustness of trade cumuli to changes in their environment. Whereas clouds in GCMs are more sensitive to thermodynamic variations, clouds in nature are more robust to thermodynamic variations and instead more controlled by dynamic variations that are not well-represented in GCMs.

Chapter 8

Conclusions and perspectives

Clouds may first strike an observer by their poetry, which sparks the imagination. As introduced in Chapt. 1, clouds have long fascinated human imagination, and this fascination finds manifold expressions in art. When something is striking, there is often a greater desire to understand it. That leads to deeper analysis and deeper appreciation of what we see. The British painter John Constable (1776–1837) contended that "we see nothing truly until we understand it" (Hamblyn, 2002). In this spirit, this thesis can also be interpreted as an attempt to better understand, and therefore to better see the visible and invisible motions that surround us.

This chapter reviews key results from this thesis, organized in two parts (Sec. 8.1), and then discusses new questions that follow from these results (Sec. 8.2). The first part (Chapters 2–5) used EUREC^{4}A observations to improve understanding of the characteristic vertical structure of trade-wind air and the processes that determine this structure. Better understanding how the trades vary in the present-day is a necessary condition for evaluating how they will change in the future. In the second part (Chapters 6–7), this improved physical understanding was applied to the evaluation of general circulation models (GCMs) to yield the first process-based constraint on the trade cumulus feedback.

To return to the challenges outlined in Chapt. 1, estimates of future warming, as quantified by the equilibrium climate sensitivity (ECS), disagree by several degrees. One of the largest sources of disagreement is uncertainties about how trade-wind clouds over the oceans will respond to warming (e.g., Bony and Dufresne, 2005; Webb et al., 2006; Vial et al., 2013; Myers et al., 2021). During EUREC^{4}A, we collected the necessary data to test whether models predicting large trade cumulus feedbacks

are physically plausible (Bony et al., 2017; Stevens et al., 2021). A single set of measurements cannot disprove a hypothesis, as individual measurements may be biased or non-reproducible. EUREC^{4}A, however, brings together a wealth of instruments and approaches. This comprehensive view informed by many coincident observations allows us to propose new conceptual models of the structure of the trade-wind boundary layer and the role of clouds in determining this structure and to conclude that there is little evidence for a strong trade cumulus feedback to warming.

8.1 Summary of key results

8.1.1 Clear-sky radiation in the trades

In the uncertain couplings among clouds, convection, and circulation in the trades, radiative heating plays an important role, yet it is difficult to observe from space. Passive remote sensing cannot capture the sharp vertical moisture gradients, especially in the lower troposphere, that are essential for calculating atmospheric radiative heating profiles (e.g., Maddy and Barnet, 2008; Chazette et al., 2014; Stevens et al., 2017). To fill this gap, we calculate clear-sky radiative profiles from 2580 *in situ* soundings launched during EUREC^{4}A, which form the largest radiative profiles data set for the trades (Chapt. 3). An updated radiative transfer code RRTMGP (Rapid Radiative Transfer Model for GCMs, Parallel) from Pincus et al. (2019) is used to calculate the radiative profiles. Variability in radiative heating is evident at multiple scales, such as related to the diurnal cycle, synoptic variability, and mesoscale cloud organization. An uncertainty assessment, moreover, shows that errors resulting from uncertainties in observed sounding profiles and ERA5 reanalysis employed as upper and lower boundary conditions are small. In the context of this thesis, the clear-sky radiative profiles are a necessary component of observationally closing subcloud layer moisture and energy budgets, which is the subject of Chapt. 4.

8.1.2 Observed subcloud layer moisture and energy budgets

The trade-wind subcloud layer is an important structural component of the tropical atmosphere (e.g., Riehl, 1954; Malkus, 1958; Tiedtke, 1989; LeMone et al., 2018), which has long been characterized using simple, bulk frameworks (Betts, 1973; Arakawa and Schubert, 1974; Deardorff, 1972; Stevens, 2006), of which mixed layer models are the simplest kind. The adequacy of the mixed layer description of the subcloud layer has, however, only been assessed from relatively few observations and large-eddy simulations often performed for small domains and idealized conditions (e.g., Stevens, 2006). These limitations render *in situ* observations especially important for testing the skill of mixed layer theory.

Chapt. 4 shows that mixed layer theory, evaluated with EUREC^{4}A observations and with uncertain parameters constrained in a Bayesian approach, provides a closed description of subcloud layer moisture and energy budgets. Campaign-residuals are 3.6 Wm^{-2} for moisture and 2.9 Wm^{-2} for energy, and residuals for synoptic variability are small and unbiased. In defining the subcloud layer height using observed thermodynamic profiles, we find evidence for a transition layer that separates the well-mixed part of the subcloud layer from the subcloud layer top. The presence of a transition layer and its vertical gradients introduce ambiguity in the application of mixed layer theory, which are addressed through the introduction of effective parameters estimated through a Bayesian methodology. We find that constrained entrainment parameters reflect mixing over a finitely-thick transition layer. These entrainment parameters, notably a mean effective entrainment efficiency, $A_\mathrm{e} = 0.43$, that is greater than 0.2 as often assumed, are consistent with expectations both from theory and direct numerical simulations. The small residuals, moreover, suggest that closing moisture and energy budgets does not require knowledge of additional processes, such as phase changes associated with evaporating precipitation in downdrafts.

Regarding large-scale external influences, we find that the net influence of the surface wind speed on mean subcloud layer moisture is weak due to its compensating influences through surface moistening and entrainment drying. Instead, knowing

moisture and temperature values above the subcloud layer has the most skill for predicting subcloud layer mean moisture and energy, presumably because these are the properties of the air mixed into the subcloud layer by entrainment.

8.1.3 A new conceptual picture of the transition layer

The presence of a finitely-thick transition layer, as discussed in Chapt. 4, contrasts with foregoing theory based on cloud-free convective boundary layers, which have an abrupt discontinuity at their top (e.g., Lilly, 1968). This discrepancy is investigated further in Chapt. 5. The transition layer in the trades has long been observed and simulated, but its origins remain little investigated. This layer is often associated with an 150–200 m stable layer that separates dry turbulent processes in the well-mixed part of the subcloud layer from moist convection in the overlying cloud layer (e.g, Malkus, 1958; Augstein et al., 1974; Yin and Albrecht, 2000). Extensive observations from EUREC^{4}A indicate that the majority of clouds are already rooted in the transition layer, and cloud-mediated mixing causes its vertical structure. Strong jumps at the layer top, as expected from the theory of cloud-free convective boundary layers, are only found rarely and when they occur, they tend to occur in large cloud-free areas. A population of small clouds with their bases around 600–700 m is shown to be responsible for smoothing vertical gradients over the transition layer.

These findings lead to a new conceptual picture that the formation and dissipation of shallow clouds maintains the transition layer, in analogy with the maintenance of the trade-wind inversion by deeper clouds as proposed by Riehl et al. (1951) and elaborated by Stevens (2007). Small clouds precondition the large-scale environment and decrease the resistance to convection through weaker vertical gradients in the transition layer, making it easier for deeper clouds to form. This conceptual model suggests that small clouds beget larger clouds. From this analysis also emerges the potential for an alternate view of entrainment mixing, which is based on the ability to detrain condensate into the overlying stable layer and thus induce gentle sinking motion through negative buoyancy. Inferences from mixed layer theory and Paluch mixing diagrams (e.g., Paluch, 1979) are also used to support inferences of air being

sourced from the transition layer for entrainment, rather than being mixed directly from deeper in the cloud layer.

8.1.4 Towards the first process-based constraint on trade cumulus feedbacks

In the second part of this thesis, physical understanding developed in the previous chapters 2–5 is applied to constrain trade-wind cloud feedbacks simulated by GCMs. In CMIP3 and CMIP5, uncertainties in trade-wind cloud responses were the largest source of uncertainty in climate sensitivity (e.g., Bony and Dufresne, 2005; Webb et al., 2006; Vial et al., 2013; Brient and Schneider, 2016).

Chapt. 6 finds that, in a representative sample of CMIP6 models, trade-wind cloud responses still differ between high and low sensitivity models, in the terms of the trade-wind thermodynamic and net ΔCRE, as well as vertical changes in cloud fraction. The trade-wind net ΔCRE can, moreover, explain about 70% of the variance in the global cloud feedback. These analyses show that trade-wind cloud feedbacks are still a large source of uncertainty in global cloud feedbacks, even if other regions also contribute to the spread in ECS in CMIP6 (e.g., Zelinka et al., 2020; Flynn and Mauritsen, 2020; Myers et al., 2021). Given the multiple influences on ECS, efforts should focus on constraining physical processes or regional cloud responses rather than ECS (e.g., Klein and Hall, 2015; Brient and Schneider, 2016; Lutsko et al., 2021), which is the focus of Chapt. 7.

Chapt. 7 presents analysis supporting the first process-based constraint on the trade cumulus feedback. This chapter describes results of Vogel et al., with a focus on my contributions. Vogel et al. present novel measurements from EUREC^{4}A of the cloud-base cloud fraction, convective mass flux, and relative humidity. These measurements allow for testing the mixing-desiccation hypothesis, which predicts a negative relationship between convective mixing and cloudiness and is thought to explain large differences among trade cumulus feedbacks.

Contrary to this hypothesis, observations show a positive relationship between the

convective mass flux and cloud fraction (r=0.72), which is strengthened by the inclusion of relative humidity (r=0.83). Observations indicate that cloudiness is principally controlled dynamically by variations in the mass flux, which is, in turn, influenced by variability in the entrainment rate and mesoscale vertical velocity. Both the entrainment rate and mesoscale vertical velocity have, historically, been challenging to estimate with observations. The entrainment rate is independently constrained in Chapt. 4. Novel sampling strategies during EUREC^{4}A provided the most extensive data set of the mesoscale vertical velocity (Bony and Stevens, 2019; George et al., 2021), revealing its influence on cloudiness.

An analysis of ten GCMs, by contrast, indicates that cloudiness in GCMs is more controlled thermodynamically, by relative humidity variations, than dynamically, by vertical velocity variations, as in the observations. We speculate that this dynamical control makes clouds in nature more resilient to relative humidity decreases due to increased convective mixing, which has tended to dissipate clouds in GCMs.

To obtain a probabilistic estimate of the trade cumulus feedback, we propose three metrics related to the couplings among cloudiness, mass flux, and relative humidity, which can be applied to both observations and GCMs. In the ten GCMs, a relationship, for instance, emerges between the correlation between the cloudiness and convective mass flux and a proxy for the trade cumulus cloud feedback, the thermodynamic ΔCRE. Models predicting the most negative coupling between convective mixing and cloudiness, most inconsistent with observational constraints, are associated with the strongest thermodynamic ΔCRE. This constraint suggests a weak trade cumulus feedback, consistent with large-eddy simulations and satellite studies as discussed in Chapt. 1 and Sec. 7.2. Calculating a probabilistic estimate of the trade cumulus feedback using the three proposed metrics is the subject of ongoing work and described in more detail in Chapt. 8. The present work, moreover, goes a step further than previous constraints because its aim is to elucidate the physical mecha-

This mass flux is also estimated using mixed layer theory. Closing subcloud layer moisture and energy budgets in Chapt. 4 provided an observational assessment of this framework, as well as an independent estimate of the entrainment rate, E, which is used to calculate M in Eq. (7.2).

nism behind the resilience of trade-wind cumuli to changes in their environment, with observations indicating that dynamic, rather thermodynamic controls make clouds in nature more resilient than those in GCMs to changing environment conditions.

8.2 Perspectives

8.2.1 Short-term perspectives

Several projects emerge as relatively logical extensions to ideas proposed in this thesis. Given that mixed layer theory offers a closed description for observations (Chapt. 4), it offers an appealing framework for evaluating larger-scale models that must parameterize the processes regulating this important component of the tropical atmosphere. The mixed layer framework would ensure a consistent definition of different terms and processes and allow for more like-for-like comparisons among terms across models, including GCMs, storm-resolving models, and large-eddy simulations. Such a project would allow the quantification of the relative magnitudes of different processes in these budgets, test how well the budgets close in different models, and examine how individual terms vary according to large-scale environmental conditions. Observations from the EUREC^{4}A field campaign, as analyzed in Chapt. 4, act as a novel benchmark for these modeled representations.

Another extension is calculating moist static energy (MSE) budgets for the total atmospheric column. It is not immediately clear whether the budgets will close on time- and spatial-scales relevant for EUREC^{4}A, such as the three-hourly circling and monthly campaign-mean timescale and about 220 km-diameter EUREC^{4}A circle spatial scale. Previous studies used remote sensing observations, but on much larger monthly and annual time scales and continental-wide spatial scales, to observationally constrain water (Rodell et al., 2015) and energy budgets (L'Ecuyer et al., 2015), still with large uncertainties. Inoue and Back (2015) close MSE budgets using TOGA COARE data and Lanczos filters to separate variability with different timescales (about 2-, 5-, 10-day, and MJO timescales). Preliminary analysis of EUREC^{4}A sound-

ings shows that horizontal and vertical advection terms become increasingly large and variable when approaching the three-hourly circling timescale.

If the total-column MSE budget can, however, be closed to within small residuals using EUREC^{4}A observations, this framework could be used to study the interplay of clouds and their environment through an energetic lens. The atmospheric cloud radiative effect (ACRE) emerges as a residual from the total-column MSE budget, yielding an energetic estimate of clouds. It would be useful to compare these estimates of the ACRE with coincident satellite and geometric (i.e. cloud fraction) estimates of cloudiness, as well as examine how the ACRE varies on different scales and relates to large-scale environmental conditions. In a different approach, Brient and Bony (2012) and Brient and Bony (2013) estimate the ACRE from total-column MSE budgets calculated for the IPSL-CM5A-LR model run in a hierarchy of configurations, diagnose the contributions to the ACRE, and conclude that the impact of an external perturbation on low-cloud cover depends on how the perturbation influences the vertical gradient of moist static energy within the boundary layer.

A third extension, especially of Chapt. 7, is to combine different observational constraints on the trade cumulus feedback in a physical, statistical framework. Different approaches exist for combining physical constraints, as introduced in Sec. 7.6. Stevens et al. (2016) proposed a quantitative 'storyline' approach, which was expanded into the Bayesian statistical framework in Stevens et al. (2016) to constrain climate sensitivity. Bretherton and Caldwell (2020) discuss an approach that uses multiple physical constraints and an ensemble of GCMs to yield a multivariate Gaussian distribution of an uncertain climate sensitivity proxy. Their approach accounts for observational uncertainties in the different metrics, sampling uncertainties, and covariance among the metrics. One drawback to this approach is an assumption of normality of the underlying variables, which could give overconfidence about the tails of the constrained distribution. Other approaches include the Bayesian approaches of Renoult et al. (2020) and Bowman et al. (2018), and machine learning approach of Schlund et al. (2020a). It is ongoing work to develop a statistical framework to combine the multiple constraints discussed in Chapt. 7 into a single probabilistic estimate

of the trade cumulus feedback.

8.2.2 Bias correction using atmospheric radiative profiles from *in situ* measurements and machine learning approaches

Chapt. 3 presented over 2500 radiative profiles derived from *in situ* dropsondes and radiosondes launched during the EUREC^{4}A campaign. The soundings more accurately represent fine-scale vertical moisture features, which are critical for calculating atmospheric radiative cooling profiles but are often smoothed in passive remote sensing retrievals and reanalysis data, as for example shown by Prange et al. (2021) and discussed in Chapt. 3. The EUREC^{4}A soundings, and therefore the radiative profiles, also have much higher temporal and spatial sampling than the soundings that are typically launched worldwide. The greater accuracy and high sampling intensity of the soundings and radiative profiles provide an opportunity to improve satellite and reanlysis products.

Machine learning is an increasingly common approach for studying clouds and radiation, such as by training on cloud-resolving model output for use in coarser climate models (Rasp et al., 2018; Gentine et al., 2018; O'Gorman and Dwyer, 2018; Beucler et al., 2021, e.g.,). Bias correction is a common problem to which machine learning is applied (e.g., Lary et al., 2016; Karpatne et al., 2018; Mathieu and Aubrecht, 2018). A successful application of machine learning depends primarily on two factors: the machine learning algorithm and a comprehensive training data set (e.g., Mathieu and Aubrecht, 2018; Rolnick et al., 2019; Beucler et al., 2021). The radiative profiles presented in Chapt. 3, and its underlying moisture profiles, would be a useful training set in the application of machine learning, in particular using deep neural networks. The soundings and radiative profiles introduce information about short-term, small-scale features that is missing from passive remote sensing and reanalysis products and therefore has the potential to improve biases in these global products.

8.2.3 How do tropical forests modulate atmospheric moisture transports?

As discussed in Sec. 1.2.1, the trades can be pictured as an expansive river in the sky, advecting moisture into the Inter-tropical Convergence Zone. A longer-term, broad question is, how and why does the path of water change when it moves over land? How do atmospheric moisture transports work over land where the sources are more variable and the flows more influenced by topography? In other words, how does the broad, moist wave in the trade-wind atmosphere evolve as it snakes and meanders across tropical continents?

One way that land influences atmospheric water flows is through moisture recycling by rainforests, such as the vast rainforests surrounding the Amazon and Congo Rivers. In many tropical regions, moisture appears to be transported in a sort of 'atmospheric river', traveling across borders, channeled narrowly in the vertical, but spread out widely in the horizontal, compared to atmospheric rivers in the extratropics, which are narrow in the horizontal and often associated with extreme precipitation (e.g., Gimeno et al., 2014; Rutz et al., 2019). There are many open questions regarding these expansive moisture transports in the tropics and how they differ from atmospheric rivers in the midlatitudes, which are better characterized (e.g., Espinoza et al., 2018). The sources of water, for instance, remain unclear. Jasechko et al. (2013) use isotope measurements to show that transpiration is the largest source of water over continents, representing 80–90% of terrestrial evapotranspiration, though other studies contend that such partitioning is subject to large uncertainties (e.g., Schlesinger and Jasechko, 2014; Ellison et al., 2017). Beyond the origins of the moisture, how do these tropical atmospheric 'rivers' meander across borders, how variable are their flows, and how do they vary according to large-scale conditions?

These transcontinental moisture flows produce rainfall that feeds lakes and rivers, yet too often such bodies of water are considered as national entities. Political conflicts often find root in river systems, such as Egypt, Sudan, and Ethiopia's conflicts over the Nile (e.g., Swain, 1997; Carlson, 2013). There is an impression that 'Egypt is

the Nile' when examining a map of its population density (e.g., as shown in Fig. 1 of Haars et al. (2016)), illustrating the preeminence of these freshwater sources for human welfare and livelihood. Yet the main source of the Nile is the Ethiopian upper Blue Nile Basin (e.g., Mellander et al., 2013), and rainfall in Ethiopia has its origins, in part, from moisture recycling from rainforests in the Congo.

How would deforestation affect these extensive tropical atmospheric rivers? Attributing precipitation changes to such human influences must be treated with caution. Ellison (2018), however, suggest that deforestation in West Africa may have decreased Nile flows from Ethiopia in the last decades of the 20^{th} century. In other African regions, Keys et al. (2016) estimate that up to 40% of sub-Saharan rainfall is presently due to moisture recycled by vegetation. Is the ongoing famine in Madagascar, for instance, in any way connected to ongoing deforestation in the Congo River basin? Similar questions about the influence of deforestation of atmospheric circulation and precipitation for the Amazon region (e.g., Coe et al., 2017; Staal et al., 2018), and southeast Asian regions (e.g., Paul et al., 2016).

Tropical atmospheric moisture flows and their interactions with land are not yet fully understood, and they have never been governed, as would be a river that crosses international borders on land. Storm-resolving models, in concert with remote sensing and *in situ* observations, are a new tool and potentially have advantages for studying these atmospheric moisture flows and precipitation compared to GCMs. There are indications that storm-resolving models may better represent precipitation than GCMs with certain well-documented deficiencies, such as raining too early in the day and too frequently compared to observations (e.g., Palmer and Stevens, 2019; Stevens et al., 2020a). The influences of deforestation on atmospheric water flows could potentially be studied by selectively changing the evaporative capacity over land to reflect how forests modulate water flows compared to deforested areas. Increasing physical uncertainty about these atmospheric moisture flows and precipitation would provide valuable information for scientists, stakeholders, and policymakers. It therefore seems reasonable to approach these questions with a diversity of tools, and new storm-resolving models complement existing studies using *in situ* observations, satel-

lite products, and GCMs. Idealized bulk models, similar to those used in this thesis, have provided insights to land-atmosphere couplings (e.g., Betts, 2000; Cronin, 2013) and could potentially be used in an exploration of the influence of forests on atmospheric moisture variability.

8.2.4 Final thoughts

Behind the apparent steadiness of the trades, as evoked by Pierre Loti, with its "same regular breath, warm, exquisite to breathe; and the same transparent sea, and the same small white clouds", a more dynamic picture emerges. Ideas in this thesis allow for proposing new conceptual models of the lower trade-wind atmosphere and a more active role for cloud mixing processes in determining its vertical structure, and falsifying a strong trade-wind cloud response to warming, based on clouds being dynamically, rather than thermodynamically controlled. I hope that this thesis was at times enjoyable to read. Participating in EUREC^{4}A was a tremendous opportunity to be immersed in the natural environment that I study and approach these questions with first-hand experience, having collected data, watched clouds form and dissipate, and imagined the circulations that fuel these clouds. There remains much to be understood about the trades and how they will evolve with warming. I can only hope that some of the ideas in this thesis can help improve our understanding of the surprises still hidden in the clouds.

Chapter 9

Additional materials

A list of publications as first or co-author is included at the end of this manuscript.

As a final note, another part of my research, developed with Peter Huybers, involves studying aerosols and their multifaceted influences, on radiation and the climate system, as well as on paintings. Regarding the former, our recent paper Albright et al. (2021b) is included after the list of publications, at the end of this thesis. This work uses a Bayesian model of aerosol forcing and Earth's multi-time-scale temperature response to radiative forcing to understand the origins of different lower bounds on aerosol forcing. We first demonstrate the ability of a simple aerosol model to emulate aerosol radiative forcing simulated by 10 general circulation models. A joint inference of climate sensitivity and effective aerosol forcing from historical surface temperatures is then made over 1850–2019. We obtain a maximum likelihood estimate of aerosol radiative forcing of -0.85 Wm^{-2} (5–95% credible interval from -1.3 to -0.50 Wm^{-2}) for 2010–19 relative to 1750. A relatively tight bound on aerosol forcing is obtained from the structure of temperature and aerosol precursor emissions and, particularly, from the rapid growth in emissions between 1950 and 1980. Obtaining a 5^{th} percentile lower bound on aerosol forcing around -2.0 Wm^{-2} requires prescribing internal climate variance that is a factor of 5 larger than the CMIP6 mean and assuming large, correlated errors in global temperature observations. Ocean heat uptake observations may further constrain aerosol radiative forcing but require a better understanding of the relationship between time-variable radiative feedbacks and radiative forcing.

Regarding the latter, after studying aerosol forcing over the instrumental period, we look further back in time to study the influence of 19^{th} century air pollution in

London and Paris. We present evidence that trends in paintings by J.M.W. Turner and Claude Monet depict trends in historical air pollution. The abstract of this paper, which is in preparation, is reproduced here.

Paintings by Turner and Monet depict trends in 19$^{\text{th}}$ century air pollution

Abstract: Anthropogenic aerosol emissions increased greatly during the 19$^{\text{th}}$ century as a result of industrialization (Fouquet, 2011), particularly in Western European cities, leading to an optical environment having less contrast and more intensity (Corton, 2015; Horvath, 1971; Kim and Kim, 2005). Here we argue that the stylistic trends from more figurative to impressionistic styles in J.M.W. Turner and Claude Monet's paintings in London and Paris over the 19$^{\text{th}}$ century are, at least in part, a response to changes in their local optical environment. In particular, we show that changes in local SO_2 emissions are a highly statistically-significant explanatory variable for trends in the contrast and coloration of Turner and Monet's works, including after controlling for time trends and subject matter. Trends in contrast and coloration across the 19$^{\text{th}}$ century are quantitatively similar to differences between pairs of modern clear-sky and polluted photographs. It has previously been shown that paintings by artists including Vincent van Gogh and Edvard Munch render specific atmospheric phenomena (Neuberger, 1970; Olson et al., 2003; Baker and Thornes, 2006; Zerefos et al., 2007, 2014; Fikke et al., 2017). Our results indicate that longer-term stylistic trends are also informed by physical atmospheric phenomena.

Bibliography

Bruce Albrecht, Virendra Ghate, Johannes Mohrmann, Robert Wood, Paquita Zuidema, Christopher Bretherton, Christian Schwartz, Edwin Eloranta, Susanne Glienke, Shaunna Donaher, et al. Cloud system evolution in the trades (cset): Following the evolution of boundary layer cloud systems with the nsf–ncar gv. *Bulletin of the American Meteorological Society*, 100(1):93–121, 2019.

Bruce A Albrecht, Alan K Betts, Wayne H Schubert, and Stephen K Cox. Model of the thermodynamic structure of the trade-wind boundary layer: Part i. theoretical formulation and sensitivity tests. *Journal of Atmospheric Sciences*, 36(1):73–89, 1979.

Anna Lea Albright, Benjamin Fildier, Ludovic Touzé-Peiffer, Robert Pincus, Jessica Vial, and Caroline Muller. Atmospheric radiative profiles during eurec^{4}a. *Earth System Science Data*, 13(2):617–630, 2021a.

Anna Lea Albright, Cristian Proistosescu, and Peter Huybers. Origins of a relatively tight lower bound on anthropogenic aerosol radiative forcing from bayesian analysis of historical observations. *Journal of Climate*, 34(21):8777–8792, 2021b.

Anna Lea Albright, Sandrine Bony, Bjorn Stevens, and Raphaela Vogel. Observed subcloud layer moisture and energy budgets in the trades. *Journal of the Atmospheric Sciences*, 2022.

Myles R Allen and William J Ingram. Constraints on future changes in climate and the hydrologic cycle. *Nature*, 419(6903):228–232, 2002.

Timothy Andrews, Jonathan M Gregory, Mark J Webb, and Karl E Taylor. Forcing, feedbacks and climate sensitivity in cmip5 coupled atmosphere-ocean climate models. *Geophysical research letters*, 39(9), 2012.

Akio Arakawa and Wayne Howard Schubert. Interaction of a cumulus cloud ensemble with the large-scale environment, part i. *Journal of the Atmospheric Sciences*, 31 (3):674–701, 1974.

Paola Arias, Nicolas Bellouin, Erika Coppola, Richard Jones, Gerhard Krinner, Jochem Marotzke, Vaishali Naik, Matthew Palmer, G-K Plattner, Joeri Rogelj, et al. Climate change 2021: The physical science basis. contribution of working group14 i to the sixth assessment report of the intergovernmental panel on climate change; technical summary. 2021.

Kyle C Armour, Cecilia M Bitz, and Gerard H Roe. Time-varying climate sensitivity from regional feedbacks. *Journal of Climate*, 26(13):4518–4534, 2013.

Svante Arrhenius. *Worlds in the making: the evolution of the universe*. Harper & brothers, 1908.

Ernst Augstein, Herbert Riehl, Feodor Ostapoff, and Volker Wagner. Mass and energy transports in an undisturbed atlantic trade-wind flow. *Monthly Weather Review*, 101(2):101–111, 1973.

Ernst Augstein, Heiner Schmidt, and Feodor Ostapoff. The vertical structure of the atmospheric planetary boundary layer in undisturbed trade winds over the atlantic ocean. *Boundary-Layer Meteorology*, 6(1-2):129–150, 1974.

Jacob Baker and John E Thornes. Solar position within monet's houses of parliament. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 462(2076):3775–3788, 2006.

François Bellec. *Marchands au long cours*. Ed du Chêne, 2003.

Gilles Bellon and Bjorn Stevens. Using the sensitivity of large-eddy simulations to evaluate atmospheric boundary layer models. *Journal of the Atmospheric Sciences*, 69(5):1582–1601, 2012.

Gilles Bellon and Bjorn Stevens. Time scales of the trade wind boundary layer adjustment. *Journal of the atmospheric sciences*, 70(4):1071–1083, 2013.

AK Betts. Non-precipitating cumulus convection and its parameterization. *Quarterly Journal of the Royal Meteorological Society*, 99(419):178–196, 1973.

Alan Betts. Modeling subcloud layer structure and interaction with shallow cumulus layer. *Journal of the Atmospheric Sciences*, 33:2363–2382, 1976.

Alan K Betts. Thermodynamic classification of tropical convective soundings. *Monthly Weather Review*, 102(11):760–764, 1974.

Alan K Betts. Idealized model for equilibrium boundary layer over land. *Journal of Hydrometeorology*, 1(6):507–523, 2000.

Alan K Betts and Bruce A Albrecht. Conserved variable analysis of the convective boundary layer thermodynamic structure over the tropical oceans. *Journal of Atmospheric Sciences*, 44(1):83–99, 1987.

Alan K Betts and W Ridgway. Climatic equilibrium of the atmospheric convective boundary layer over a tropical ocean. *Journal of the Atmospheric Sciences*, 46(17): 2621–2641, 1989.

Tom Beucler, Imme Ebert-Uphoff, Stephan Rasp, Michael Pritchard, and Pierre Gentine. Machine learning for clouds and climate (invited chapter for the agu geophysical monograph series" clouds and climate"). 2021.

Frank Beyrich and Jens-Peter Leps. An operational mixing height data set from routine radiosoundings at lindenberg: Methodology. *Meteorologische Zeitschrift*, 21(4):337–348, 2012.

Jonah Bloch-Johnson, Raymond T Pierrehumbert, and Dorian S Abbot. Feedback temperature dependence determines the risk of high warming. *Geophysical Research Letters*, 42(12):4973–4980, 2015.

Peter N Blossey, Christopher S Bretherton, Minghua Zhang, Anning Cheng, Satoshi Endo, Thijs Heus, Yangang Liu, Adrian P Lock, Stephan R de Roode, and Kuan-Man Xu. Marine low cloud sensitivity to an idealized climate change: The CGILS LES intercomparison. *Journal of Advances in Modeling Earth Systems*, 5(2):234–258, 2013.

Olivier Bock, Pierre Bosser, Cyrille Flamant, Erik Doerflinger, Friedhelm Jansen, Romain Fages, Sandrine Bony, and Sabrina Schnitt. Integrated water vapour observations in the caribbean arc from a network of ground-based gnss receivers during EUREC^{4}A. *Earth System Science Data*, 13(5):2407–2436, 2021.

Steven J Böing, Harm JJ Jonker, Witek A Nawara, and A Pier Siebesma. On the deceiving aspects of mixing diagrams of deep cumulus convection. *Journal of the Atmospheric Sciences*, 71(1):56–68, 2014.

Sandrine Bony and Jean-Louis Dufresne. Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophysical Research Letters*, 32(20), 2005. doi: 10.1029/2005GL023851.

Sandrine Bony and Bjorn Stevens. Measuring area-averaged vertical motions with dropsondes. *Journal of the Atmospheric Sciences*, 76(3):767–783, 2019. doi: 10.1175/JAS-D-18-0141.1.

Sandrine Bony, Jean-Louis Dufresne, Herve Le Treut, Jean-Jacques Morcrette, and Catherine Senior. On dynamic and thermodynamic components of cloud changes. *Climate Dynamics*, 22(2-3):71–86, 2004.

Sandrine Bony, Robert Colman, Vladimir M Kattsov, Richard P Allan, Christopher S Bretherton, Jean-Louis Dufresne, Alex Hall, Stephane Hallegatte, Marika M Holland, William Ingram, et al. How well do we understand and evaluate climate change feedback processes? *Journal of Climate*, 19(15):3445–3482, 2006.

Sandrine Bony, Gilles Bellon, Daniel Klocke, Steven Sherwood, Solange Fermepin, and Sébastien Denvil. Robust direct effect of carbon dioxide on tropical circulation and regional precipitation. *Nature Geoscience*, 6(6):447–451, 2013a.

Sandrine Bony, Bjorn Stevens, Isaac H Held, John F Mitchell, Jean-Louis Dufresne, Kerry A Emanuel, Pierre Friedlingstein, Stephen Griffies, and Catherine Senior. Carbon dioxide and climate: Perspectives on a scientific assessment. In *Climate science for serving society*, pages 391–413. Springer, 2013b.

Sandrine Bony, Bjorn Stevens, Félix Ament, Sebastien Bigorre, Patrick Chazette, Susanne Crewell, Julien Delanoë, Kerry Emanuel, David Farrell, Cyrille Flamant, Silke Gross, Lutz Hirsh, Johannes Karstensen, Bernhard Mayer, Louise Nuijens, James H. Ruppert, Irina Sandu, Pier Siebesma, Sabrina Speich, Frédéric Szczap, Julien Totems, Raphaela Vogel, Manfred Wendisch, and Martin Wirth. EUREC^{4}A: A field campaign to elucidate the couplings between clouds, convection and circulation. *Surveys in Geophysics*, 38(6):1529–1568, 2017. doi: 10.1007/s10712-017-9428-0.

Sandrine Bony, Hauke Schulz, Jessica Vial, and Bjorn Stevens. Sugar, gravel, fish, and flowers: Dependence of mesoscale patterns of trade-wind clouds on environmental conditions. *Geophysical Research Letters*, 47(7):e2019GL085988, 2020.

Sandrine Bony et al. EUREC^{4}A observations from the SAFIRE ATR42 aircraft. *Earth System Science Data Discussions*, 2022.

Olivier Boucher, Jérôme Servonnat, Anna Lea Albright, Olivier Aumont, Yves Balkanski, Vladislav Bastrikov, Slimane Bekki, Rémy Bonnet, Sandrine Bony, Laurent Bopp, et al. Presentation and evaluation of the ipsl-cm6a-lr climate model. *Journal of Advances in Modeling Earth Systems*, 12(7):e2019MS002010, 2020.

Kevin W Bowman, Noel Cressie, Xin Qu, and Alex Hall. A hierarchical statistical framework for emergent constraints: Application to snow-albedo feedback. *Geophysical Research Letters*, 45(23):13–050, 2018.

Kamau Brathwaite. *Born to Slow Horses*. Wesleyan University Press, 2005.

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

Christopher S Bretherton. Insights into low-latitude cloud feedbacks from high-resolution models. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373(2054):20140415, 2015.

Christopher S Bretherton and Peter M Caldwell. Combining emergent constraints for climate sensitivity. *Journal of Climate*, 33(17):7413–7430, 2020.

Christopher S Bretherton and Matthew C Wyant. Moisture transport, lower-tropospheric stability, and decoupling of cloud-topped boundary layers. *Journal of Atmospheric Sciences*, 54(1):148–167, 1997.

Christopher S Bretherton, Malcolm K MacVean, Peter Bechtold, Andreas Chlond, William R Cotton, Joan Cuxart, Hans Cuijpers, M Mhairoutdinov, Bronko Kosovic, Dave Lewellen, et al. An intercomparison of radiatively driven entrainment and turbulence in a smoke cloud, as simulated by different numerical models. *Quarterly Journal of the Royal Meteorological Society*, 125(554):391–423, 1999.

Christopher S Bretherton, Taneil Uttal, Christopher W Fairall, Sandra E Yuter, Robert A Weller, Darrel Baumgardner, Kimberly Comstock, Robert Wood, and Graciela B Raga. The epic 2001 stratocumulus study. *Bulletin of the American Meteorological Society*, 85(7):967–978, 2004.

Christopher S. Bretherton, Peter N. Blossey, Marat Khairoutdinov, Christopher S. Bretherton, Peter N. Blossey, and Marat Khairoutdinov. An Energy-Balance Analysis of Deep Convective Self-Aggregation above Uniform SST. *Journal of the Atmospheric Sciences*, 62(12):4273–4292, dec 2005. ISSN 0022-4928. doi: 10.1175/JAS3614.1.

CS Bretherton. Entrainment, detrainment and mixing in atmospheric convection. In *The Physics and Parameterization of Moist Atmospheric Convection*, pages 211–230. Springer, 1997.

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

F Brient and S Bony. How may low-cloud radiative properties simulated in the current climate influence low-cloud feedbacks under global warming? *Geophysical Research Letters*, 39(20), 2012.

Florent Brient and Sandrine Bony. Interpretation of the positive low-cloud feedback predicted by a climate model under global warming. *Climate Dynamics*, 40(9): 2415–2431, 2013.

Florent Brient and Tapio Schneider. Constraints on climate sensitivity from space-based measurements of low-cloud reflection. *Journal of Climate*, 29(16):5821–5835, 2016.

Florent Brient, Tapio Schneider, Zhihong Tan, Sandrine Bony, Xin Qu, and Alex Hall. Shallowness of tropical low clouds as a predictor of climate models' response to warming. *Climate Dynamics*, 47(1-2):433–449, 2016.

Peter M Caldwell, Christopher S Bretherton, Mark D Zelinka, Stephen A Klein, Benjamin D Santer, and Benjamin M Sanderson. Statistical significance of climate sensitivity predictors obtained by data mining. *Geophysical Research Letters*, 41 (5):1803–1808, 2014.

Peter M Caldwell, Mark D Zelinka, and Stephen A Klein. Evaluating emergent constraints on equilibrium climate sensitivity. *Journal of Climate*, 31(10):3921–3942, 2018.

Guylaine Canut, Fleur Couvreux, Marie Lothon, David Pino, and Frédérique Saïd. Observations and large-eddy simulations of entrainment in the sheared sahelian boundary layer. *Boundary-layer meteorology*, 142(1):79–101, 2012.

Guangxia Cao, Thomas W Giambelluca, Duane E Stevens, and Thomas A Schroeder. Inversion variability in the hawaiian trade wind regime. *Journal of Climate*, 20(7): 1145–1160, 2007. doi: 10.1175/JCLI4033.1.

Andrew Carlson. Who owns the nile? egypt, sudan, and ethiopia's history-changing dam. *Origins*, 6(6), 2013.

K. S. Carslaw, L. A. Lee, L. A. Regayre, and J. S. Johnson. Climate models are uncertain, but we can do something about it. *EOS*, 99, 2018.

Paulo Ceppi, Florent Brient, Mark D Zelinka, and Dennis L Hartmann. Cloud feedback mechanisms and their representation in global climate models. *Wiley Interdisciplinary Reviews: Climate Change*, 8(4):e465, 2017.

Grégory V Cesana and Anthony D Del Genio. Observational constraint on cloud feedbacks suggests moderate climate sensitivity. *Nature Climate Change*, 11(3): 213–218, 2021.

Robert D Cess, Gerald L Potter, JP Blanchet, GJ Boer, SJ Ghan, JT Kiehl, H Le Treut, Z-X Li, X-Z Liang, JFB Mitchell, et al. Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science*, 245 (4917):513–516, 1989.

Jule G Charney, Akio Arakawa, D James Baker, Bert Bolin, Robert E Dickinson, Richard M Goody, Cecil E Leith, Henry M Stommel, and Carl I Wunsch. Carbon dioxide and climate: a scientific assessment, 1979.

P. Chazette, F. Marnas, J. Totems, and X. Shang. Comparison of IASI water vapor retrieval with H2O-Raman lidar in the framework of the Mediterranean HyMeX and ChArMEx programs. *Atmospheric Chemistry and Physics*, 14(18):9583–9596, sep 2014. ISSN 1680-7324. doi: 10.5194/acp-14-9583-2014.

Patrick Chazette, Julien Totems, Alexandre Baron, Cyrille Flamant, and Sandrine Bony. Trade-wind clouds and aerosols characterized by airborne horizontal lidar measurements during the EUREC^{4}A field campaign. *Earth System Science Data*, 12(4):2919–2936, 2020.

H Chepfer, S Bony, D Winker, G Cesana, JL Dufresne, P Minnis, CJ Stubenrauch, and S Zeng. The gcm-oriented calipso cloud product (calipso-goccp). *Journal of Geophysical Research: Atmospheres*, 115(D4), 2010.

Michael T Coe, Paulo M Brando, Linda A Deegan, Marcia N Macedo, Christopher Neill, and Divino V Silverio. The forests of the amazon and cerrado moderate regional climate and are the key to the future. *Tropical Conservation Science*, 10: 1940082917720671, 2017.

Christine L Corton. *London fog*. Harvard University Press, 2015.

Timothy W Cronin. A sensitivity theory for the equilibrium boundary layer over land. *Journal of Advances in Modeling Earth Systems*, 5(4):764–784, 2013.

C Dai, Q Wang, JA Kalogiros, DH Lenschow, Z Gao, and M Zhou. Determining boundary-layer height from aircraft measurements. *Boundary-layer meteorology*, 152(3):277–302, 2014.

Gokhan Danabasoglu, J-F Lamarque, J Bacmeister, DA Bailey, AK DuVivier, Jim Edwards, LK Emmons, John Fasullo, R Garcia, Andrew Gettelman, et al. The community earth system model version 2 (cesm2). *Journal of Advances in Modeling Earth Systems*, 12(2), 2020.

James W Deardorff. Parameterization of the planetary boundary layer for use in general circulation models. *Monthly Weather Review*, 100(2):93–106, 1972.

Thomas G Dopplick. Radiative heating of the global atmosphere. *Journal of the Atmospheric Sciences*, 29(7):1278–1294, 1972.

Jean-Louis Dufresne and Sandrine Bony. An assessment of the primary sources of spread of global warming estimates from coupled atmosphere–ocean models. *Journal of Climate*, 21(19):5135–5144, 2008.

Christoph Dyroff, Andreas Zahn, Emanuel Christner, Richard Forbes, Adrian M Tompkins, and Peter FJ van Velthoven. Comparison of ecmwf analysis and forecast humidity data with caribic upper troposphere and lower stratosphere observations. *Quarterly Journal of the Royal Meteorological Society*, 141(688):833–844, 2015. doi: https://doi.org/10.1002/qj.2400.

David Ellison. Forests and water. *Background Analytical Study*, 2, 2018.

David Ellison, Cindy E Morris, Bruno Locatelli, Douglas Sheil, Jane Cohen, Daniel Murdiyarso, Victoria Gutierrez, Meine Van Noordwijk, Irena F Creed, Jan Pokorny, et al. Trees, forests and water: Cool insights for a hot world. *Global Environmental Change*, 43:51–61, 2017.

Kerry Emanuel. The effect of convective response time on wishe modes. *Journal of Atmospheric Sciences*, 50(12):1763–1776, 1993.

Kerry Emanuel, Allison A Wing, and Emmanuel M Vincent. Radiative-convective instability. *Journal of Advances in Modeling Earth Systems*, pages 75–90, 2014. doi: 10.1002/2013MS000270.Received.

Kerry A Emanuel. An air-sea interaction theory for tropical cyclones. part i: Steady-state maintenance. *Journal of Atmospheric Sciences*, 43(6):585–605, 1986.

Kerry A Emanuel et al. *Atmospheric convection*. Oxford University Press on Demand, 1994.

Vicky Espinoza, Duane E Waliser, Bin Guan, David A Lavers, and F Martin Ralph. Global analysis of climate change projection effects on atmospheric rivers. *Geophysical Research Letters*, 45(9):4299–4308, 2018.

European Centre for Medium-Range Weather Forecasts. Era5 reanalysis, 2017.

Veronika Eyring, Peter M Cox, Gregory M Flato, Peter J Gleckler, Gab Abramowitz, Peter Caldwell, William D Collins, Bettina K Gier, Alex D Hall, Forrest M Hoffman, et al. Taking climate model evaluation to the next level. *Nature Climate Change*, 9(2):102–110, 2019.

Chris W Fairall, Edward F Bradley, JE Hare, Andrey A Grachev, and James B Edson. Bulk parameterization of air–sea fluxes: Updates and verification for the coare algorithm. *Journal of climate*, 16(4):571–591, 2003.

Svein M Fikke, Jón Egill Kristjánsson, and Øyvind Nordli. Screaming clouds. *Weather*, 72(5):115–121, 2017.

James R Fleming. Joseph fourier, the 'greenhouse effect', and the quest for a universal theory of terrestrial temperatures. *Endeavour*, 23(2):72–75, 1999.

Clare Marie Flynn and Thorsten Mauritsen. On the climate sensitivity and historical warming evolution in recent coupled model ensembles. *Atmospheric Chemistry and Physics*, 20(13):7829–7842, 2020.

Roger Fouquet. Long run trends in energy-related external costs. *Ecological Economics*, 70(12):2380–2389, 2011.

Ž Fuchs-Stone, David J Raymond, and S Sentić. Otrec2019: Convection over the east pacific and southwest caribbean. *Geophysical Research Letters*, 47(11): e2020GL087564, 2020.

L Garand, D S Turner, M Larocque, J Bates, S Boukabara, P Brunel, F Chevallier, G Deblonde, R Engelen, M Hollingshead, D Jackson, G Jedlovec, J Joiner, T Kleespies, D S McKague, L McMillin, J L Moncet, J R Pardo, P J Rayer, E Salathe, R Saunders, N. A. Scott, P Van Delst, and H Woolf. Radiance and Jacobian intercomparison of radiative transfer models applied to HIRS and AMSU channels. *J. Geophys. Res.*, 106(D20):24017–24031, October 2001. doi: 10.1029/2000JD000184.

Jade Rachele Garcia and Juan Pedro Mellado. The two-layer structure of the entrainment zone in the convective boundary layer. *Journal of the Atmospheric Sciences*, 71(6):1935–1955, 2014.

John Roy Garratt. The atmospheric boundary layer. *Earth-Science Reviews*, 37(1-2): 89–134, 1994.

Michael Garstang, Edward Zipser, Robert Ellingson, Kenneth Warsh, Peter Grose, Stanley Ulanski, Ronald Holle, Ward Seguin, David Fitzjarrald, Steven Greco, et al. Three early tropical field experiments. *Bulletin of the American Meteorological Society*, 100(11):2243–2258, 2019.

W Lawrence Gates, James S Boyle, Curt Covey, Clyde G Dease, Charles M Doutriaux, Robert S Drach, Michael Fiorino, Peter J Gleckler, Justin J Hnilo, Susan M Marlais, et al. An overview of the results of the atmospheric model intercomparison project (amip i). *Bulletin of the American Meteorological Society*, 80(1):29–56, 1999.

Iliana Genkova, Gabriela Seiz, Paquita Zuidema, Guangyu Zhao, and Larry Di Girolamo. Cloud top height comparisons from aster, misr, and modis for trade wind cumuli. *Remote sensing of environment*, 107(1-2):211–222, 2007.

Pierre Gentine, Mike Pritchard, Stephan Rasp, Gael Reinaudi, and Galen Yacalis. Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters*, 45(11):5742–5751, 2018.

Geet George, Bjorn Stevens, Sandrine Bony, Robert Pincus, Chris Fairall, Hauke Schulz, Tobias Kölling, Quinn T Kalen, Marcus Klingebiel, Heike Konow, et al. Joanne: Joint dropsonde observations of the atmosphere in tropical north atlantic meso-scale environments. *Earth System Science Data Discussions*, pages 1–33, 2021.

A Gettelman, JE Kay, and KM Shell. The evolution of climate sensitivity and climate feedbacks in the community atmosphere model. *Journal of Climate*, 25(5):1453–1469, 2012.

Andrew Gettelman, Cecile Hannay, Julio T Bacmeister, Richard B Neale, AG Pendergrass, G Danabasoglu, J-F Lamarque, JT Fasullo, DA Bailey, DM Lawrence, et al. High climate sensitivity in the community earth system model version 2 (cesm2). *Geophysical Research Letters*, 46(14):8329–8337, 2019.

Luis Gimeno, Raquel Nieto, Marta Vázquez, and David A Lavers. Atmospheric rivers: A mini-review. *Frontiers in Earth Science*, 2:2, 2014.

William M Gray and Robert W Jacobson Jr. Diurnal variation of deep cumulus convection. *Monthly Weather Review*, 105(9):1171–1188, 1977. doi: 10.1175/1520-0493(1977)105<1171:DVODCC>2.0.CO;2.

W.M. Gray and R.W. Jacobson. Diurnal variation of deep cumulus convection. *Monthly Weather Review*, 105:1171–1187, 1977. doi: 10.1175/1520-0493(1977)105<1171:DVODCC>2.0.CO;2.

JM1789797 Gregory, WJ Ingram, MA Palmer, GS Jones, PA Stott, RB Thorpe, JA Lowe, TC Johns, and KD Williams. A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical research letters*, 31(3), 2004.

F. Guichard, D. Parsons, and E. Miller. Thermodynamic and radiative impact of the correction of sounding humidity bias in the tropics. *Journal of Climate*, 13(20): 3611–3624, 2000.

Manuel Gutleben, Silke Martha Groß, and Martin Wirth. Cloud macro-physical properties in saharan-dust-laden and dust-free north atlantic trade wind regimes: a lidar case study. *Atmospheric Chemistry and Physics (ACP)*, 19(16):10659–10673, 2019. doi: 10.5194/acp-19-10659-2019.

Manuel Gutleben, Silke Groß, Martin Wirth, and Bernhard Mayer. Radiative effects of long-range-transported saharan air layers as determined from airborne lidar measurements. *Atmospheric Chemistry and Physics*, 20(20):12313–12327, 2020.

Christian Haars, EM Lönsjö, Bianca Mogos, and Bram Winkelaar. The uncertain future of the nile delta. *NASA/GSFC: Greenbelt, MD, USA*, 2016.

Alex Hall and Xin Qu. Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophysical Research Letters*, 33(3), 2006.

Alex Hall, Peter Cox, Chris Huntingford, and Stephen Klein. Progressing emergent constraints on future climate change. *Nature Climate Change*, 9(4):269–278, 2019.

Richard Hamblyn. *The invention of clouds: How an amateur meteorologist forged the language of the skies*. Macmillan, 2002.

Dennis L Hartmann, Maureen E Ockert-Bell, and Marc L Michelsen. The effect of cloud type on earth's energy balance: Global analysis. *Journal of Climate*, 5(11): 1281–1304, 1992.

W.K. Hastings. Monte Carlo sampling methods using Markov chains and their application. *Biometrika*, 57(1):97—-109, 1970.

William A Heckley. Systematic errors of the ecmwf operational forecasting model in tropical regions. *Quarterly Journal of the Royal Meteorological Society*, 111(469): 709–738, 1985.

Jerome L Heffter. Air resources laboratories atmospheric transport and dispersion model (arl-atad). Technical report, 1980.

Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

Thijs Heus, Gertjan Van Dijk, Harm JJ Jonker, and Harry EA Van den Akker. Mixing in shallow cumulus clouds studied by lagrangian particle tracking. *Journal of the Atmospheric Sciences*, 65(8):2581–2597, 2008.

Joshua Z Holland. Preliminary report on the bomex sea-air interaction program. *Bulletin of the American Meteorological Society*, 51(9):809–821, 1970.

Joshua Z Holland and Eugene M Rasmusson. Measurements of the atmospheric mass, energy, and momentum budgets over a 500-kilometer square of tropical ocean. *Monthly Weather Review*, 101(1):44–55, 1973.

George C Holzworth. Estimates of mean maximum mixing depths in the contiguous united states. *Mon. Weather Rev*, 92(5):235–242, 1964.

Helmuth Horvath. On the applicability of the koschmieder visibility formula. *Atmospheric Environment (1967)*, 5(3):177–184, 1971.

Peter Huybers. Compensation between model feedbacks and curtailment of climate sensitivity. *Journal of Climate*, 23(11):3009–3018, 2010.

AJ Illingworth, RJ Hogan, EJ O'connor, Dominique Bouniol, ME Brooks, Julien Delanoë, DP Donovan, JD Eastment, N Gaussiat, JWF Goddard, et al. Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations. *Bulletin of the American Meteorological Society*, 88(6):883–898, 2007.

Kuniaki Inoue and Larissa Back. Column-integrated moist static energy budget analysis on various time scales during toga coare. *Journal of the Atmospheric Sciences*, 72(5):1856–1871, 2015.

Marek Jacob, Pavlos Kollias, Felix Ament, Vera Schemann, and Susanne Crewell. Multilayer cloud conditions in trade wind shallow cumulus–confronting two icon model derivatives with airborne observations. *Geoscientific Model Development*, 13 (11):5757–5777, 2020.

Scott Jasechko, Zachary D Sharp, John J Gibson, S Jean Birks, Yi Yi, and Peter J Fawcett. Terrestrial water fluxes dominated by transpiration. *Nature*, 496(7445): 347–350, 2013.

Nadir Jeevanjee and Stephan Fueglistaler. Simple spectral models for atmospheric radiative cooling. *Journal of the Atmospheric Sciences*, 77(2):479–497, 2020. doi: https://doi.org/10.1175/JAS-D-18-0347.1.

Richard H Johnson, Thomas M Rickenbach, Steven A Rutledge, Paul E Ciesielski, and Wayne H Schubert. Trimodal characteristics of tropical convection. *Journal of climate*, 12(8):2397–2418, 1999.

Peter Kalmus, Matthew Lebsock, and João Teixeira. Observational boundary layer energy and water budgets of the stratocumulus-to-cumulus transition. *Journal of Climate*, 27(24):9155–9170, 2014.

Youichi Kamae, Hideo Shiogama, Masahiro Watanabe, Tomoo Ogura, Tokuta Yokohata, and Masahide Kimoto. Lower-tropospheric mixing as a constraint on cloud feedback in a multiparameter multiphysics ensemble. *Journal of Climate*, 29(17): 6259–6275, 2016.

Anuj Karpatne, Imme Ebert-Uphoff, Sai Ravela, Hassan Ali Babaie, and Vipin Kumar. Machine learning for the geosciences: Challenges and opportunities. *IEEE Transactions on Knowledge and Data Engineering*, 31(8):1544–1554, 2018.

Seiji Kato, Thomas P Ackerman, Eugene E Clothiaux, James H Mather, Gerald G Mace, Marvin L Wesely, Frank Murcray, and Joseph Michalsky. Uncertainties in modeled and measured clear-sky surface shortwave irradiances. *Journal of Geophysical Research: Atmospheres*, 102(D22):25881–25898, 1997. doi: 10.1029/97JD01841.

SR Kawa and R Pearson Jr. An observational study of stratocumulus entrainment and thermodynamics. *Journal of the atmospheric sciences*, 46(17):2649–2661, 1989.

Patrick W Keys, Lan Wang-Erlandsson, and Line J Gordon. Revealing invisible water: moisture recycling as an ecosystem service. *PloS one*, 11(3):e0151993, 2016.

Kyung W Kim and Young J Kim. Perceived visibility measurement using the hsi color difference method. *Journal of the Korean Physical Society*, 46(5):1243, 2005.

Stephen A Klein and Alex Hall. Emergent constraints for cloud feedbacks. *Current climate change reports*, 1(4):276–287, 2015.

Marcus Klingebiel, Heike Konow, and Bjorn Stevens. Measuring shallow convective mass flux profiles in the trade wind region. *Journal of the Atmospheric Sciences*, 78(10):3205–3214, 2021.

Reto Knutti, Maria AA Rugenstein, and Gabriele C Hegerl. Beyond equilibrium climate sensitivity. *Nature Geoscience*, 10(10):727–736, 2017.

Heike Konow, Marek Jacob, Felix Ament, Susanne Crewell, Florian Ewald, Martin Hagen, Lutz Hirsch, Friedhelm Jansen, Mario Mech, and Bjorn Stevens. A unified data set of airborne cloud remote sensing using the halo microwave package (hamp). *Earth System Science Data*, 11(2):921–934, 2019.

Heike Konow, Florian Ewald, Geet George, Marek Jacob, Marcus Klingebiel, Tobias Kölling, Anna E Luebke, Theresa Mieslinger, Veronika Pörtge, Jule Radtke, et al. EUREC^{4}A's HALO. *Earth System Science Data Discussions*, pages 1–26, 2021.

David J Lary, Amir H Alavi, Amir H Gandomi, and Annette L Walker. Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7(1):3–10, 2016.

LV Leahy, R Wood, RJ Charlson, CA Hostetler, RR Rogers, MA Vaughan, and DM Winker. On the nature and extent of optically thin marine low clouds. *Journal of Geophysical Research: Atmospheres*, 117(D22), 2012a. doi: 10.1029/2012JD017929.

LV Leahy, R Wood, RJ Charlson, CA Hostetler, RR Rogers, MA Vaughan, and DM Winker. On the nature and extent of optically thin marine low clouds. *Journal of Geophysical Research: Atmospheres*, 117(D22), 2012b.

Tristan S. L'Ecuyer, Norman B. Wood, Taryn Haladay, Graeme L. Stephens, and Paul W. Stackhouse Jr. Impact of clouds on atmospheric heating based on the r04 cloudsat fluxes and heating rates data set. *Journal of Geophysical Research: Atmospheres*, 113(D8), 2008. doi: 10.1029/2008JD009951. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD009951`.

Margaret A LeMone and William T Pennell. The relationship of trade wind cumulus distribution to subcloud layer fluxes and structure. *Monthly Weather Review*, 104 (5):524–539, 1976.

Margaret A LeMone, Wayne M Angevine, Christopher S Bretherton, Fei Chen, Jimy Dudhia, Evgeni Fedorovich, Kristina B Katsaros, Donald H Lenschow, Larry Mahrt, Edward G Patton, et al. 100 years of progress in boundary layer meteorology. *Meteorological Monographs*, 59:9–1, 2018.

Donald H Lenschow, Paul B Krummel, and Steven T Siems. Measuring entrainment, divergence, and vorticity on the mesoscale from aircraft. *Journal of atmospheric and oceanic technology*, 16(10):1384–1400, 1999.

Douglas K Lilly. Models of cloud-topped mixed layers under a strong inversion. *Quarterly Journal of the Royal Meteorological Society*, 94(401):292–309, 1968.

Richard S Lindzen and Arthur V Hou. Hadley circulations for zonally averaged heating centered off the equator. *Journal of Atmospheric Sciences*, 45(17):2416–2427, 1988.

Richard S Lindzen and Sumant Nigam. On the role of sea surface temperature gradients in forcing low-level winds and convergence in the tropics. *Journal of Atmospheric Sciences*, 44(17):2418–2436, 1987.

Hailong Liu, Chunzai Wang, Sang-Ki Lee, and David Enfield. Atlantic warm pool variability in the cmip5 simulations. *Journal of climate*, 26(15):5315–5336, 2013.

Shuyan Liu and Xin-Zhong Liang. Observed Diurnal Cycle Climatology of Planetary Boundary Layer Height. *Journal of Climate*, 23(21):5790–5809, 11 2010. doi: 10.1175/2010JCLI3552.1.

Norman G Loeb, Bruce A Wielicki, David R Doelling, G Louis Smith, Dennis F Keyes, Seiji Kato, Natividad Manalo-Smith, and Takmeng Wong. Toward optimal closure of the earth's top-of-atmosphere radiation budget. *Journal of Climate*, 22 (3):748–766, 2009.

Charles N Long, Sally A McFarlane, A Del Genio, Patrick Minnis, Thomas P Ackerman, J Mather, J Comstock, Gerald G Mace, Michael Jensen, and Christian Jakob.

Arm research in the equatorial western pacific: A decade and counting. *Bulletin of the American Meteorological Society*, 94(5):695–708, 2013.

Katrin Lonitz, Bjorn Stevens, Louise Nuijens, Vivek Sant, Lutz Hirsch, and Axel Seifert. The signature of aerosols and meteorology in long-term cloud radar observations of trade wind cumuli. *Journal of the Atmospheric Sciences*, 72(12): 4643–4659, 2015.

Edward N Lorenz. *The nature and theory of the general circulation of the atmosphere*, volume 218. World Meteorological Organization Geneva, 1967.

Bingkun Luo and Peter Minnett. Evaluation of the era5 sea surface skin temperature with remotely-sensed shipborne marine-atmospheric emitted radiance interferometer data. *Remote Sensing*, 12(11):1873, 2020. doi: https://doi.org/10.3390/rs12111873.

Nicholas J Lutsko, Max Popp, Robert H Nazarian, and Anna Lea Albright. Emergent constraints on regional cloud feedbacks. *Geophysical Research Letters*, page e2021GL092934, 2021.

Tristan S L'Ecuyer, Hiroko K Beaudoing, Matthew Rodell, W Olson, B Lin, S Kato, CA Clayson, E Wood, J Sheffield, R Adler, et al. The observed state of the energy budget in the early twenty-first century. *Journal of Climate*, 28(21):8319–8346, 2015.

Eric S. Maddy and Christopher D. Barnet. Vertical Resolution Estimates in Version 5 of AIRS Operational Retrievals. *IEEE Trans. Geosci. Remote Sensing*, 46(8): 2375–2384, August 2008. ISSN 0196-2892. doi: 10.1109/TGRS.2008.917498.

René Magritte, Sarah Whitfield, Michael Raeburn, and Lynette Cawthra. *René Magritte: catalogue raisonné*. Menil Foundation, 1992.

L Mahrt. Mixed layer moisture structure. *Monthly Weather Review*, 104(11):1403–1407, 1976.

Joanne Starr Malkus. *On the structure of the trade wind moist layer*. Massachusetts Institute of Technology, 1958.

Syukuro Manabe and Robert F Strickler. Thermal equilibrium of the atmosphere with a convective adjustment. *Journal of the Atmospheric Sciences*, 21(4):361–385, 1964.

Syukuro Manabe and Richard T Wetherald. Thermal equilibrium of the atmosphere with a given distribution of relative humidity. 1967.

B. E. Mapes. Water's two height scales: The moist adiabat and the radiative troposphere. *Quart. J. Roy. Meteor. Soc.*, 127(577):2353–2366, 2001. doi: 10.1175/1520-0469(1998)055<1354:ETGMSO>2.0.CO;2.

Frank J Marsik, Kenneth W Fischer, Tracey D McDonald, and Perry J Samson. Comparison of methods for estimating mixing height used during the 1992 atlanta field intensive. *Journal of Applied Meteorology*, 34(8):1802–1814, 1995.

Pierre-Philippe Mathieu and Christoph Aubrecht. *Earth observation open science and innovation*. Springer Nature, 2018.

Brian Medeiros and Louise Nuijens. 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, 2016a. doi: 10.1073/pnas.1521494113.

Brian Medeiros and Louise Nuijens. 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, 2016b.

Brian Medeiros and Bjorn Stevens. Revealing differences in gcm representations of low clouds. *Climate dynamics*, 36(1-2):385–399, 2011.

Brian Medeiros, Bjorn Stevens, Isaac M Held, Ming Zhao, David L Williamson, Jerry G Olson, and Christopher S Bretherton. Aquaplanets, climate sensitivity, and low clouds. *Journal of Climate*, 21(19):4974–4991, 2008.

Brian Medeiros, Bjorn Stevens, and Sandrine Bony. Using aquaplanets to understand the robust responses of comprehensive climate models to forcing. *Climate Dynamics*, 44(7):1957–1977, 2015.

Gerald A Meehl, Catherine A Senior, Veronika Eyring, Gregory Flato, Jean-Francois Lamarque, Ronald J Stouffer, Karl E Taylor, and Manuel Schlund. Context for interpreting equilibrium climate sensitivity and transient climate response from the cmip6 earth system models. *Science Advances*, 6(26):eaba1981, 2020.

Per-Erik Mellander, Solomon G Gebrehiwot, Annemieke I Gärdenäs, Woldeamlak Bewket, and Kevin Bishop. Summer rains and dry seasons in the upper blue nile basin: the predictability of half a century of past and future spatiotemporal patterns. *PloS one*, 8(7):e68461, 2013.

W. Paul Menzel, Timothy J. Schmit, Peng Zhang, and Jun Li. Satellite-Based Atmospheric Infrared Sounder Development and Applications. *Bulletin of the American Meteorological Society*, 99(3):583–603, March 2018. ISSN 0003-0007, 1520-0477. doi: 10.1175/BAMS-D-16-0293.1.

N Metropolis, A Rosenbluth, M Rosenbluth, A Teller, and E Teller. Equation of state calculations by fast computing machines. *Journal of Chemistry and Physics*, 21: 1087–1092, 1953.

Theresa Mieslinger, Ákos Horváth, Stefan A Buehler, and Mirjana Sakradzija. The dependence of shallow cumulus macrophysical properties on large-scale meteorology

as observed in aster imagery. *Journal of Geophysical Research: Atmospheres*, 124 (21):11477–11505, 2019.

Theresa Mieslinger, Bjorn Stevens, Tobias Kölling, Manfred Brath, Martin Wirth, and Stefan A Buehler. Optically thin clouds in the trades. *Atmospheric Chemistry and Physics Discussions*, pages 1–33, 2021.

EJ Mlawer, SA Clough, and S Kato. Shortwave clear-sky model measurement intercomparison using rrtm. In *Proceedings of the Eighth ARM Science Team Meeting*, pages 23–27. Citeseer, 1998.

Eli J. Mlawer, Vivienne H. Payne, Jean Luc Moncet, Jennifer S. Delamere, Matthew J. Alvarado, and David C. Tobin. Development and recent evaluation of the MT-CKD model of continuum absorption. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1968):2520–2556, jun 2012. ISSN 1364503X. doi: 10.1098/rsta.2011.0295.

Chin-Hoh Moeng, Peter P Sullivan, and Bjorn Stevens. Including radiative effects in an entrainment rate formula for buoyancy-driven pbls. *Journal of the atmospheric sciences*, 56(8):1031–1049, 1999.

Kenneth P Moran, Brooks E Martner, MJ Post, Robert A Kropfli, David C Welsh, and Kevin B Widener. An unattended cloud-profiling radar for use in climate research. *Bulletin of the American Meteorological Society*, 79(3):443–456, 1998.

C. J. Muller and I. M. Held. Detailed Investigation of the Self-Aggregation of Convection in Cloud-Resolving Simulations. *Journal of the Atmospheric Sciences*, 69 (8):2551–2565, 2012. ISSN 0022-4928. doi: 10.1175/JAS-D-11-0257.1.

Timothy A Myers and Joel R Norris. Observational evidence that enhanced subsidence reduces subtropical marine boundary layer cloudiness. *Journal of Climate*, 26(19):7507–7524, 2013.

Timothy A Myers, Ryan C Scott, Mark D Zelinka, Stephen A Klein, Joel R Norris, and Peter M Caldwell. Observational constraints on low cloud feedback reduce uncertainty of climate sensitivity. *Nature Climate Change*, 11(6):501–507, 2021.

Badrinath Nagarajan and Anantha R Aiyyer. Performance of the ecmwf operational analyses over the tropical indian ocean. *Monthly weather review*, 132(9):2275–2282, 2004. doi: https://doi.org/10.1175/1520-0493(2004)132<2275:POTEOA>2.0.CO;2.

Ann Kristin Naumann, Bjorn Stevens, Cathy Hohenegger, and Juan Pedro Mellado. A conceptual model of a shallow circulation induced by prescribed low-level radiative cooling. *Journal of the Atmospheric Sciences*, 74(10):3129–3144, 2017.

Ann Kristin Naumann, Bjorn Stevens, and Cathy Hohenegger. A moist conceptual model for the boundary layer structure and radiatively driven shallow circulations in the trades. *Journal of the Atmospheric Sciences*, 76(5):1289–1306, 2019.

J David Neelin, Isaac M Held, and Kerry H Cook. Evaporation-wind feedback and low-frequency variability in the tropical atmosphere. *Journal of Atmospheric Sciences*, 44(16):2341–2348, 1987.

Roel Neggers, Bjorn Stevens, and David Neelin. A simple equilibrium model for shallow-cumulus-topped mixed layers. *Theoretical and Computational Fluid Dynamics*, 20(5):305–322, 2006.

Hans Neuberger. Climate in art. *Weather*, 25(2):46–56, 1970.

William Nordhaus. Projections and uncertainties about climate change in an era of minimal climate policies. *American Economic Journal: Economic Policy*, 10(3): 333–60, 2018.

Joel R Norris. Low cloud type over the ocean from surface observations. part ii: Geographical and seasonal variations. *Journal of climate*, 11(3):383–403, 1998.

Louise Nuijens and Bjorn Stevens. The influence of wind speed on shallow marine cumulus convection. *Journal of the atmospheric sciences*, 69(1):168–184, 2012.

Louise Nuijens, Ilya Serikov, Lutz Hirsch, Katrin Lonitz, and Bjorn Stevens. The distribution and variability of low-level cloud in the north atlantic trades. *Quarterly Journal of the Royal Meteorological Society*, 140(684):2364–2374, 2014.

Louise Nuijens, Brian Medeiros, Irina Sandu, and Maike Ahlgrimm. The behavior of trade-wind cloudiness in observations and models: The major cloud components and their variability. *Journal of Advances in Modeling Earth Systems*, 7(2):600–616, 2015a. doi: 10.1002/2014MS000390.

Louise Nuijens, Brian Medeiros, Irina Sandu, and Maike Ahlgrimm. The behavior of trade-wind cloudiness in observations and models: The major cloud components and their variability. *Journal of Advances in Modeling Earth Systems*, 7(2):600–616, 2015b.

Louise Nuijens, Brian Medeiros, Irina Sandu, and Maike Ahlgrimm. Observed and modeled patterns of covariability between low-level cloudiness and the structure of the trade-wind layer. *Journal of Advances in Modeling Earth Systems*, 7(4): 1741–1764, 2015c.

Kuan-Ting O, Robert Wood, and Hsiu-Hui Tseng. Deeper, precipitating pbls associated with optically thin veil clouds in the sc-cu transition. *Geophysical Research Letters*, 45(10):5177–5184, 2018. doi: 10.1029/2018GL077084. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL077084`.

Paul A O'Gorman and John G Dwyer. Using machine learning to parameterize moist convection: Potential for modeling of climate, climate change, and extreme events. *Journal of Advances in Modeling Earth Systems*, 10(10):2548–2563, 2018.

Donald W Olson, Russell L Doescher, and Marilynn S Olson. Dating van gogh's" moonrise". 2003.

Katsuyuki Ooyama. A theory on parameterization of cumulus convection. *Journal of the Meteorological Society of Japan. Ser. II*, 49:744–756, 1971.

Isidoro Orlanski. A rational subdivision of scales for atmospheric processes. *Bulletin of the American Meteorological Society*, pages 527–530, 1975.

Boutheina Oueslati and Gilles Bellon. Tropical precipitation regimes and mechanisms of regime transitions: Contrasting two aquaplanet general circulation models. *Climate dynamics*, 40(9):2345–2358, 2013.

Tim Palmer and Bjorn Stevens. The scientific challenge of understanding and estimating climate change. *Proceedings of the National Academy of Sciences*, 116(49): 24390–24395, 2019.

Ilga R Paluch. The entrainment mechanism in colorado cumuli. *Journal of Atmospheric Sciences*, 36(12):2467–2478, 1979.

Supantha Paul, Subimal Ghosh, Robert Oglesby, Amey Pathak, Anita Chandrasekharan, and RAAJ Ramsankaran. Weakening of indian summer monsoon rainfall due to changes in land use land cover. *Scientific reports*, 6(1):1–10, 2016.

R Thermostates Pierrehumbert. Thermostats, radiator fins, and the local runaway greenhouse. *Journal of Atmospheric Sciences*, 52(10):1784–1806, 1995.

Robert Pincus, Anton Beljaars, Stefan A. Buehler, Gottfried Kirchengast, Florian Ladstaedter, and Jeffrey S Whitaker. The Representation of Tropospheric Water Vapor Over Low-Latitude Oceans in (Re-)analysis: Errors, Impacts, and the Ability to Exploit Current and Prospective Observations. *Surveys in Geophysics*, 38(6): 1399–1423, October 2017. doi: 10.1007/s10712-017-9437-z.

Robert Pincus, Eli J. Mlawer, and Jennifer S. Delamere. Balancing accuracy, efficiency, and flexibility in radiation calculations for dynamical models. *Journal of Advances in Modeling Earth Systems*, 11(10):3074–3089, 2019. doi: 10.1029/2019MS001621.

Robert Pincus, Chris W Fairall, Adriana Bailey, Haonan Chen, Patrick Y Chuang, Gijs de Boer, Graham Feingold, Dean Henze, Quinn T Kalen, Jan Kazil, et al. Observations from the noaa p-3 aircraft during atomic. *Earth System Science Data*, 13(7):3281–3296, 2021.

Max Planck. *Scientific autobiography: And other papers*. Open Road Media, 2014.

Max Popp and Levi G Silvers. Double and single itczs with and without clouds. *Journal of Climate*, 30(22):9147–9166, 2017.

Marc Prange, Manfred Brath, and Stefan Alexander Buehler. Are elevated moist layers a blind spot for hyperspectral infrared sounders?–a model study. *Atmospheric Measurement Techniques Discussions*, pages 1–29, 2021.

Cristian Proistosescu and Peter J Huybers. Slow climate mode reconciles historical and model-based estimates of climate sensitivity. *Science advances*, 3(7):e1602821, 2017.

Yi Qin, Mark D Zelinka, and Stephen A Klein. On the correspondence between atmosphere-only and coupled simulations for radiative feedbacks and forcing from co2. 2021.

Xin Qu, Alex Hall, Stephen A Klein, and Peter M Caldwell. On the spread of changes in marine low cloud cover in climate model simulations of the 21st century. *Climate dynamics*, 42(9-10):2603–2626, 2014.

Jule Radtke, Thorsten Mauritsen, and Cathy Hohenegger. Shallow cumulus cloud feedback in large eddy simulations–bridging the gap to storm-resolving models. *Atmospheric Chemistry and Physics*, 21(5):3275–3288, 2021.

VLRD Ramanathan, RD Cess, EF Harrison, P Minnis, BR Barkstrom, E Ahmad, and D Hartmann. Cloud-radiative forcing and climate: Results from the earth radiation budget experiment. *Science*, 243(4887):57–63, 1989.

David A Randall and Stephen Tjemkes. Clouds, the earth's radiation budget, and the hydrologic cycle. *Global and Planetary Change*, 4(1-3):3–9, 1991. doi: 10.1016/0921-8181(91)90063-3.

Stephan Rasp, Michael S Pritchard, and Pierre Gentine. Deep learning to represent subgrid processes in climate models. *Proceedings of the National Academy of Sciences*, 115(39):9684–9689, 2018.

Stephan Rasp, Hauke Schulz, Sandrine Bony, and Bjorn Stevens. Combining crowdsourcing and deep learning to explore the mesoscale organization of shallow convection. *Bulletin of the American Meteorological Society*, 101(11):E1980–E1995, 2020.

Robert M Rauber, Bjorn Stevens, Harry T Ochs III, Charles Knight, Bruce A Albrecht, AM Blyth, CW Fairall, JB Jensen, SG Lasher-Trapp, OL Mayol-Bracero, et al. Rain in shallow cumulus over the ocean: The rico campaign. *Bulletin of the American Meteorological Society*, 88(12):1912–1928, 2007.

David J Raymond, GB Raga, Christopher S Bretherton, John Molinari, Carlos López-Carrillo, and Željka Fuchs. Convective forcing in the intertropical convergence zone of the eastern pacific. *Journal of the atmospheric sciences*, 60(17):2064–2082, 2003.

Martin Renoult, James Douglas Annan, Julia Catherine Hargreaves, Navjit Sagoo, Clare Flynn, Marie-Luise Kapsch, Qiang Li, Gerrit Lohmann, Uwe Mikolajewicz,

Rumi Ohgaito, et al. A bayesian framework for emergent constraints: case studies of climate sensitivity with pmip. *Climate of the Past*, 16(5):1715–1735, 2020.

Malte Rieck, Louise Nuijens, and Bjorn Stevens. Marine boundary layer cloud feedbacks in a constant relative humidity atmosphere. *Journal of the Atmospheric Sciences*, 69(8):2538–2550, 2012.

Herbert Riehl. Variations of energy exchange between sea and air in the trades. *Weather*, 9(11):335–340, 1954.

Herbert Riehl, TC Yeh, Joanne S Malkus, and Noel E La Seur. The north-east trade of the pacific ocean. *Quarterly Journal of the Royal Meteorological Society*, 77(334): 598–626, 1951.

Camille Risi, Caroline Muller, and Peter Blossey. What controls the water vapor isotopic composition near the surface of tropical oceans? results from an analytical model constrained by large-eddy simulations. *Journal of advances in modeling earth systems*, 12(8):e2020MS002106, 2020.

Nicolas Rochetin, Cathy Hohenegger, Ludovic Touzé-Peiffer, and Najda Villefranque. A physically based definition of convectively generated density currents: Detection and characterization in convection-permitting simulations. *Journal of Advances in Modeling Earth Systems*, 13(7):e2020MS002402, 2021.

Matthew Rodell, Hiroko Kato Beaudoing, TS L'Ecuyer, William S Olson, James Stephen Famiglietti, Paul Raymond Houser, Robert Adler, Michael G Bosilovich, Carol Anne Clayson, D Chambers, et al. The observed state of the water cycle in the early twenty-first century. *Journal of Climate*, 28(21):8289–8318, 2015.

MJ Rodwell and TN Palmer. Using numerical weather prediction to assess climate models. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 133(622):129–146, 2007.

David Rolnick, Priya L Donti, Lynn H Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, Nikola Milojevic-Dupont, Natasha Jaques, Anna Waldman-Brown, et al. Tackling climate change with machine learning. *arXiv preprint arXiv:1906.05433*, 2019.

Maria Rugenstein, Jonah Bloch-Johnson, Jonathan Gregory, Timothy Andrews, Thorsten Mauritsen, Chao Li, Thomas L Frölicher, David Paynter, Gokhan Danabasoglu, Shuting Yang, et al. Equilibrium climate sensitivity estimated by equilibrating climate models. *Geophysical Research Letters*, 47(4):e2019GL083898, 2020.

James H Ruppert and Richard H Johnson. On the cumulus diurnal cycle over the tropical warm pool. *Journal of Advances in Modeling Earth Systems*, 8(2):669–690, 2016. doi: 10.1002/2015MS000610.

James H Ruppert Jr and Morgan E O'Neill. Diurnal cloud and circulation changes in simulated tropical cyclones. *Geophysical Research Letters*, 46(1):502–511, 2019. doi: 10.1029/2018GL081302.

Lynn M Russell, Armin Sorooshian, John H Seinfeld, Bruce A Albrecht, Athanasios Nenes, Lars Ahlm, Yi-Chun Chen, Matthew Coggon, Jill S Craven, Richard C Flagan, et al. Eastern pacific emitted aerosol cloud experiment. *Bulletin of the American Meteorological Society*, 94(5):709–729, 2013.

Jonathan J Rutz, Christine A Shields, Juan M Lora, Ashley E Payne, Bin Guan, Paul Ullrich, Travis O'brien, L Ruby Leung, F Martin Ralph, Michael Wehner, et al. The atmospheric river tracking method intercomparison project (artmip): quantifying uncertainties in atmospheric river climatology. *Journal of Geophysical Research: Atmospheres*, 124(24):13777–13802, 2019.

Jerôme Schalkwijk, Harmen JJ Jonker, and A Pier Siebesma. Simple solutions to steady-state cumulus regimes in the convective boundary layer. *Journal of the atmospheric sciences*, 70(11):3656–3672, 2013.

William H Schlesinger and Scott Jasechko. Transpiration in the global water cycle. *Agricultural and Forest Meteorology*, 189:115–117, 2014.

Manuel Schlund, Veronika Eyring, Gustau Camps-Valls, Pierre Friedlingstein, Pierre Gentine, and Markus Reichstein. Constraining uncertainty in projected gross primary production with machine learning. *Journal of Geophysical Research: Biogeosciences*, 125(11):e2019JG005619, 2020a.

Manuel Schlund, Axel Lauer, Pierre Gentine, Steven C Sherwood, and Veronika Eyring. Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6? *Earth System Dynamics*, 11(4):1233–1258, 2020b.

Timothy J. Schmit, Jun Li, Steven A. Ackerman, and James J. Gurka. High-Spectral- and High-Temporal-Resolution Infrared Measurements from Geostationary Orbit. *Journal of Atmospheric and Oceanic Technology*, 26(11):2273–2292, November 2009. ISSN 1520-0426, 0739-0572. doi: 10.1175/2009JTECHA1248.1.

Hauke Schulz, Ryan Eastman, and Bjorn Stevens. Characterization and evolution of organized shallow convection in the downstream north atlantic trades. *Journal of Geophysical Research: Atmospheres*, 126(17):e2021JD034575, 2021.

Petra Seibert, Frank Beyrich, Sven-Erik Gryning, Sylvain Joffre, Alix Rasmussen, and Philippe Tercier. Review and intercomparison of operational methods for the determination of the mixing height. *Atmospheric environment*, 34(7):1001–1027, 2000.

Dian J Seidel, Yehui Zhang, Anton Beljaars, Jean-Christophe Golaz, Andrew R Jacobson, and Brian Medeiros. Climatology of the planetary boundary layer over the continental united states and europe. *Journal of Geophysical Research: Atmospheres*, 117(D17), 2012.

Axel Seifert and Thijs Heus. Large-eddy simulation of organized precipitating trade wind cumulus clouds. *Atmospheric Chemistry and Physics*, 13:5631–5645, 2013. doi: 10.5194/acpd-13-1855-2013.

Axel Seifert, Thijs Heus, Robert Pincus, and Bjorn Stevens. Large-eddy simulation of the transient and near-equilibrium behavior of precipitating shallow convection. *Journal of Advances in Modeling Earth Systems*, 7(4):1918–1937, 2015. doi: 10.1002/2015MS000489.

Karen M Shell, Jeffrey T Kiehl, and Christine A Shields. Using the radiative kernel technique to calculate climate feedbacks in ncar's community atmospheric model. *Journal of Climate*, 21(10):2269–2282, 2008.

SC Sherwood, Mark J Webb, James D Annan, KC Armour, Piers M Forster, Julia C Hargreaves, Gabriele Hegerl, Stephen A Klein, Kate D Marvel, Eelco J Rohling, et al. An assessment of earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, 58(4):e2019RG000678, 2020.

Steven C Sherwood, Sandrine Bony, and Jean-Louis Dufresne. Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, 505(7481):37–42, 2014.

A Pier Siebesma, Christopher S Bretherton, Andrew Brown, Andreas Chlond, Joan Cuxart, Peter G Duynkerke, Hongli Jiang, Marat Khairoutdinov, David Lewellen, Chin-Hoh Moeng, et al. A large eddy simulation intercomparison study of shallow cumulus convection. *Journal of the Atmospheric Sciences*, 60(10):1201–1219, 2003.

A Pier Siebesma, Sandrine Bony, Christian Jakob, and Bjorn Stevens. *Clouds and Climate: Climate Science's Greatest Challenge*. Cambridge University Press, 2020.

Joseph Smagorinsky. General circulation experiments with the primitive equations: I. the basic experiment. *Monthly weather review*, 91(3):99–164, 1963.

Brian J Soden. Tracking upper tropospheric water vapor radiances: A satellite perspective. *Journal of Geophysical Research: Atmospheres*, 103(D14):17069–17081, 1998.

Brian J Soden, Isaac M Held, Robert Colman, Karen M Shell, Jeffrey T Kiehl, and Christine A Shields. Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, 21(14):3504–3520, 2008.

Roy W Spencer and William D Braswell. How dry is the tropical free troposphere? implications for global warming theory. *Bulletin of the American Meteorological Society*, 78(6):1097–1106, 1997.

Arie Staal, Obbe A Tuinenburg, Joyce HC Bosmans, Milena Holmgren, Egbert H van Nes, Marten Scheffer, Delphine Clara Zemp, and Stefan C Dekker. Forest-rainfall cascades buffer against drought across the amazon. *Nature Climate Change*, 8(6):539–543, 2018.

Claudia Christine Stephan, Sabrina Schnitt, Hauke Schulz, Hugo Bellenger, Simon P de Szoeke, Claudia Acquistapace, Katharina Baier, Thibaut Dauhut, Rémi Laxenaire, Yanmichel Morfa-Avalos, et al. Ship-and island-based atmospheric soundings from the 2020 EUREC4A field campaign. *Earth System Science Data Discussions*, pages 1–35, 2020. doi: 10.5194/essd-2020-174.

Graeme L. Stephens, Juilin Li, Martin Wild, Carol Anne Clayson, Norman Loeb, Seiji Kato, Tristan L'Ecuyer, Paul W. Stackhouse, Matthew Lebsock, and Timothy Andrews. An update on earth's energy balance in light of the latest global observations. *Nature Geoscience*, 5(10):691–696, 2012. doi: 10.1038/ngeo1580.

Nicholas Stern. Stern review: The economics of climate change. 2006.

Bjorn Stevens. Bulk boundary-layer concepts for simplified models of tropical dynamics. *Theoretical and Computational Fluid Dynamics*, 20(5):279–304, 2006.

Bjorn Stevens. On the growth of layers of nonprecipitating cumulus convection. *Journal of the atmospheric sciences*, 64(8):2916–2931, 2007.

Bjorn Stevens and Sandrine Bony. What are climate models missing? *Science*, 340 (6136):1053–1054, 2013.

Bjorn Stevens, Andrew S Ackerman, Bruce A Albrecht, Andrew R Brown, Andreas Chlond, Joan Cuxart, Peter G Duynkerke, David C Lewellen, Malcolm K Macvean, Roel AJ Neggers, et al. Simulations of trade wind cumuli under a strong inversion. *Journal of the atmospheric sciences*, 58(14):1870–1891, 2001.

Bjorn Stevens, Donald H Lenschow, Gabor Vali, Hermann Gerber, A Bandy, B Blomquist, J-L Brenguier, CS Bretherton, F Burnet, Teresa Campos, et al. Dynamics and chemistry of marine stratocumulus—dycoms-ii. *Bulletin of the American Meteorological Society*, 84(5):579–594, 2003.

Bjorn Stevens, David Farrell, Lutz Hirsch, Friedhelm Jansen, Louise Nuijens, Ilya Serikov, Björn Brügmann, Marvin Forde, Holger Linne, Katrin Lonitz, and Joseph M. Prospero. The Barbados Cloud Observatory: Anchoring investigations of clouds and circulation on the edge of the itcz. *Bulletin of the American Meteorological Society*, 97(5):787–801, 2016. doi: 10.1175/BAMS-D-14-00247.1.

Bjorn Stevens, Hélène Brogniez, Christoph Kiemle, Jean-Lionel Lacour, Cyril Crevoisier, and Johannes Kiliani. Structure and dynamical influence of water vapor in the lower tropical troposphere. *Surveys in Geophysics*, 38(6):1371–1397, 2017. doi: 10.1007/s10712-017-9420-8.

Bjorn Stevens, Claudia Acquistapace, Akio Hansen, Rieke Heinze, Carolin Klinger, Daniel Klocke, Harald Rybka, Wiebke Schubotz, Julia Windmiller, Panagiotis Adamidis, et al. The added value of large-eddy and storm-resolving models for simulating clouds and precipitation. *Journal of the Meteorological Society of Japan. Ser. II*, 2020a.

Bjorn Stevens, Sandrine Bony, Hélène Brogniez, Laureline Hentgen, Cathy Hohenegger, Christoph Kiemle, Tristan S L'Ecuyer, Ann Kristin Naumann, Hauke Schulz, Pier A Siebesma, et al. Sugar, gravel, fish and flowers: Mesoscale cloud patterns in the trade winds. *Quarterly Journal of the Royal Meteorological Society*, 146(726): 141–152, 2020b.

Bjorn Stevens, Sandrine Bony, D Farrell, and 293 co authors. EUREC^{4}A. *Earth System Science Data*, 2021.

Claudia J Stubenrauch, William B Rossow, Stefan Kinne, S Ackerman, G Cesana, H Chepfer, L Di Girolamo, B Getzewich, A Guignard, A Heidinger, et al. Assessment of global cloud datasets from satellites: Project and database initiated by the gewex radiation panel. *Bulletin of the American Meteorological Society*, 94(7): 1031–1049, 2013.

Roland B Stull. The energetics of entrainment across a density interface. *Journal of the atmospheric sciences*, 33(7):1260–1267, 1976.

Roland B Stull. A fair-weather cumulus cloud classification scheme for mixed-layer studies. *Journal of Applied Meteorology and Climatology*, 24(1):49–56, 1985.

Roland B Stull. *An introduction to boundary layer meteorology*, volume 13. Springer Science & Business Media, 2012.

Ashok Swain. Ethiopia, the sudan, and egypt: the nile river dispute. *The Journal of Modern African Studies*, 35(4):675–694, 1997.

Karl E Taylor, Ronald J Stouffer, and Gerald A Meehl. An overview of cmip5 and the experiment design. *Bulletin of the American meteorological Society*, 93(4):485–498, 2012.

Hendrik Tennekes. A model for the dynamics of the inversion above a convective boundary layer. *Journal of Atmospheric sciences*, 30(4):558–567, 1973.

Chad W Thackeray, Xin Qu, and Alex Hall. Why do models produce spread in snow albedo feedback? *Geophysical Research Letters*, 45(12):6223–6231, 2018.

K Thayer-Calder and David Randall. A numerical investigation of boundary layer quasi-equilibrium. *Geophysical Research Letters*, 42(2):550–556, 2015.

David W. J. Thompson, Sandrine Bony, and Ying Li. Thermodynamic constraint on the depth of the global tropospheric circulation. *Proceedings of the National Academy of Sciences*, 114(31):8181–8186, 2017. ISSN 0027-8424. doi: 10.1073/pnas.1620493114.

MICHAEL Tiedtke. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly weather review*, 117(8):1779–1800, 1989.

Richard SJ Tol. The economic effects of climate change. *Journal of economic perspectives*, 23(2):29–51, 2009.

Lorenzo Tomassini, Aiko Voigt, and Bjorn Stevens. On the connection between tropical circulation, convective mixing, and climate sensitivity. *Quarterly Journal of the Royal Meteorological Society*, 141(689):1404–1416, 2014.

Ludovic Touzé-Peiffer, Raphaela Vogel, and Nicolas Rochetin. Detecting cold pools from soundings during EUREC^{4}A. *arXiv preprint arXiv:2104.09146*, 2021.

Vaisala. Rd41 technical data. Technical report, University Corporation for Atmospheric Research, 2018.

Margreet C Vanzanten, Peter G Duynkerke, and Joannes WM Cuijpers. Entrainment parameterization in convective boundary layers. *Journal of the atmospheric sciences*, 56(6):813–828, 1999.

Jessica Vial, Jean-Louis Dufresne, and Sandrine Bony. On the interpretation of intermodel spread in cmip5 climate sensitivity estimates. *Climate Dynamics*, 41(11-12): 3339–3362, 2013.

Jessica Vial, Sandrine Bony, Jean-Louis Dufresne, and Romain Roehrig. 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):1892–1911, 2016.

Jessica Vial, Sandrine Bony, Bjorn Stevens, and Raphaela Vogel. Mechanisms and model diversity of trade-wind shallow cumulus cloud feedbacks: a review. *Shallow clouds, water vapor, circulation, and climate sensitivity*, pages 159–181, 2017.

Jessica Vial, Raphaela Vogel, Sandrine Bony, Bjorn Stevens, David M Winker, Xia Cai, Cathy Hohenegger, Ann Kristin Naumann, and Hélène Brogniez. A new look at the daily cycle of trade wind cumuli. *Journal of advances in modeling earth systems*, 11(10):3148–3166, 2019. doi: 10.1029/2019MS001746.

Raphaela Vogel, Louise Nuijens, and Bjorn Stevens. 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):843–862, 2016.

Raphaela Vogel, Sandrine Bony, and Bjorn Stevens. Estimating the shallow convective mass flux from the subcloud-layer mass budget. *Journal of Atmospheric Sciences*, 77(5):1559–1574, 2020.

Gernot Wagner and Martin L Weitzman. *Climate shock*. Princeton University Press, 2016.

Mark J Webb, CA Senior, DMH Sexton, WJ Ingram, KD Williams, MA Ringer, BJ McAvaney, R Colman, Brian J Soden, R Gudgel, et al. On the contribution of local feedback mechanisms to the range of climate sensitivity in two gcm ensembles. *Climate Dynamics*, 27(1):17–38, 2006.

Mark J Webb, Timothy Andrews, Alejandro Bodas-Salcedo, Sandrine Bony, Christopher S Bretherton, Robin Chadwick, Hélène Chepfer, Hervé Douville, Peter Good, Jennifer E Kay, et al. The cloud feedback model intercomparison project (cfmip) contribution to CMIP6. *Geoscientific Model Development*, 10(1):359–384, 2017.

Martin L Weitzman. Fat-tailed uncertainty in the economics of catastrophic climate change. *Review of Environmental Economics and Policy*, 5(2):275–292, 2011.

Martin L Weitzman. Fat tails and the social cost of carbon. *American Economic Review*, 104(5):544–46, 2014.

Keith D Williams, Alejandro Bodas-Salcedo, Michel Déqué, Solange Fermepin, B Medeiros, M Watanabe, Christian Jakob, Stephen A Klein, Catherine A Senior, and David L Williamson. The transpose-amip ii experiment and its application to the understanding of southern ocean cloud biases in climate models. *Journal of Climate*, 26(10):3258–3274, 2013.

Martin Wirth, Andreas Fix, Peter Mahnke, Horst Schwarzer, Friedrich Schrandt, and Gerhard Ehret. The airborne multi-wavelength water vapor differential absorption lidar wales: system design and performance. *Applied Physics B*, 96(1):201–213, 2009.

R Wood, CS Bretherton, D Leon, AD Clarke, P Zuidema, G Allen, and H Coe. An aircraft case study of the spatial transition from closed to open mesoscale cellular convection over the southeast pacific. *Atmospheric Chemistry and Physics*, 11(5): 2341–2370, 2011.

Robert Wood. Stratocumulus Clouds. *Monthly Weather Review*, 140(8):2373–2423, 2012. ISSN 0027-0644. doi: 10.1175/MWR-D-11-00121.1.

Robert Wood and Christopher S Bretherton. On the relationship between stratiform low cloud cover and lower-tropospheric stability. *Journal of climate*, 19(24):6425–6432, 2006.

Robert Wood, Kuan-Ting O, Christopher S Bretherton, Johannes Mohrmann, Bruce A Albrecht, Paquita Zuidema, Virendra Ghate, Chris Schwartz, Ed Eloranta, Susanne Glienke, et al. Ultraclean layers and optically thin clouds in the stratocumulus-to-cumulus transition. part i: Observations. *Journal of the Atmospheric Sciences*, 75(5):1631–1652, 2018a. doi: 10.1175/JAS-D-17-0213.1.

Robert Wood, Hsiu-Hui Tseng, et al. Deeper, precipitating pbls associated with optically thin veil clouds in the sc-cu transition. *GeoRL*, 45(10):5177–5184, 2018b. doi: 10.1029/2018GL077084.

Bingfan Yin and Bruce A Albrecht. Spatial variability of atmospheric boundary layer structure over the eastern equatorial pacific. *Journal of climate*, 13(9):1574–1592, 2000.

Mark D Zelinka, Timothy A Myers, Daniel T McCoy, Stephen Po-Chedley, Peter M Caldwell, Paulo Ceppi, Stephen A Klein, and Karl E Taylor. Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, 47(1): e2019GL085782, 2020.

Christos S Zerefos, VT Gerogiannis, Dimitris Balis, SC Zerefos, and Andreas Kazantzidis. Atmospheric effects of volcanic eruptions as seen by famous artists and depicted in their paintings. *Atmospheric Chemistry and Physics*, 7(15):4027–4042, 2007.

CS Zerefos, P Tetsis, A Kazantzidis, V Amiridis, SC Zerefos, J Luterbacher, K Eleftheratos, E Gerasopoulos, S Kazadzis, and A Papayannis. Further evidence of important environmental information content in red-to-green ratios as depicted in paintings by great masters. *Atmospheric Chemistry and Physics*, 14(6):2987–3015, 2014.

Minghua Zhang, Christopher S Bretherton, Peter N Blossey, Phillip H Austin, Julio T Bacmeister, Sandrine Bony, Florent Brient, Suvarchal K Cheedela, Anning Cheng, Anthony D Del Genio, et al. Cgils: Results from the first phase of an international project to understand the physical mechanisms of low cloud feedbacks in single column models. *Journal of Advances in Modeling Earth Systems*, 5(4):826–842, 2013.

Y Zhang, Z Gao, D Li, Y Li, N Zhang, X Zhao, and J Chen. On the computation of planetary boundary-layer height using the bulk richardson number method. *Geoscientific Model Development*, 7(6):2599–2611, 2014.

Youtong Zheng. Theoretical understanding of the linear relationship between convective updrafts and cloud-base height for shallow cumulus clouds. part i: Maritime conditions. *Journal of the Atmospheric Sciences*, 76(8):2539–2558, 2019.

Chen Zhou, Mark D Zelinka, Andrew E Dessler, and Stephen A Klein. The relationship between interannual and long-term cloud feedbacks. *Geophysical Research Letters*, 42(23):10–463, 2015.

Chen Zhou, Mark D Zelinka, and Stephen A Klein. Impact of decadal cloud variations on the earth's energy budget. *Nature Geoscience*, 9(12):871–874, 2016.

Sergej Zilitinkevich and Alexander Baklanov. Calculation of the height of the stable boundary layer in practical applications. *Boundary-Layer Meteorology*, 105(3):389–409, 2002.

List of publications

1. **Albright, A.L.**, Fildier, B., Touzé-Peiffer, L., Pincus, R., Vial, J. and Muller, C., 2021. Atmospheric radiative profiles during EUREC4A. *Earth System Science Data*, *13*(2), pp.617-630.

2. **Albright, A.L.**, Proistosescu, C. and Huybers, P., 2021. Origins of a Relatively Tight Lower Bound on Anthropogenic Aerosol Radiative Forcing from Bayesian Analysis of Historical Observations. *Journal of Climate*, *34*(21), pp.8777-8792.

3. **Albright, A.L.**, Bony, S., Stevens, B., and Vogel, R, 2022. Observed subcloud layer moisture and energy budgets in the trades. Submitted to *Journal of the Atmospheric Sciences*

4. **Albright, A.L.**, Bony, S., Stevens, B., and Vogel, R, 2022. A new conceptual picture of the trade-wind transition layer. In preparation for *Journal of the Atmospheric Sciences.*

5. **Albright, A.L.** and Huybers, P, 2022. Paintings by Turner and Monet depict trends in 19th century air pollution. In preparation.

6. Boucher, O., Servonnat, J., **Albright, A.L.**, Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L. and Braconnot, P. et al, 2020. Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling Earth Systems*, *12*(7), p.e2019MS002010.

7. Stevens, B., Bony, S., and the EUREC4A team (including **Albright, A.L**), 2021. EUREC4A. *Earth System Science Data*, *13*(8), pp.4067-4119.

8. Lutsko, N.J., Popp, M., Nazarian, R.H. and **Albright, A.L.**, 2021. Emergent constraints on regional cloud feedbacks. *Geophysical Research Letters*, p.e2021GL092934.

9. Bony, S. and the ATR team (including **Albright, A.L.**), 2022. EUREC4A observations from the SAFIRE ATR-42 aircraft. Submitted to *Earth System Science Data*.

10. Vogel, **Albright, A.L.**, Vial, J., George, G., Stevens, S., and Bony, S. No observational evidence for a strong trade cumulus cloud feedback. In preparation for *Science.*

11. Vial, J., **Albright, A.L.**, Vogel, R., Musat, I. and Bony, S. The daily cycle of trade-wind cumuli: a process-based constraint for cloud feedback. In preparation for *PNAS.*

Origins of a Relatively Tight Lower Bound on Anthropogenic Aerosol Radiative Forcing from Bayesian Analysis of Historical Observations

ANNA LEA ALBRIGHT,[a] CRISTIAN PROISTOSESCU,[b] AND PETER HUYBERS[c]

[a] *Laboratoire de Météorologique Dynamique, Sorbonne Université, École Normale Supérieure, École Polytechnique, Paris*
[b] *Department of Atmospheric Sciences and Department of Geology, University of Illinois at Urbana–Champaign, Champaign, Illinois*
[c] *Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts*

(Manuscript received 26 February 2021, in final form 4 August 2021)

ABSTRACT: A variety of empirical estimates have been published for the lower bounds on aerosol radiative forcing, clustered around -1.0 or $-2.0\,\mathrm{W\,m^{-2}}$. The reasons for obtaining such different constraints are not well understood. In this study, we explore bounds on aerosol radiative forcing using a Bayesian model of aerosol forcing and Earth's multi-time-scale temperature response to radiative forcing. We first demonstrate the ability of a simple aerosol model to emulate aerosol radiative forcing simulated by 10 general circulation models. A joint inference of climate sensitivity and effective aerosol forcing from historical surface temperatures is then made over 1850–2019. We obtain a maximum likelihood estimate of aerosol radiative forcing of $-0.85\,\mathrm{W\,m^{-2}}$ (5%–95% credible interval from -1.3 to $-0.50\,\mathrm{W\,m^{-2}}$) for 2010–19 relative to 1750 and an equilibrium climate sensitivity of 3.4°C (5%–95% credible interval from 1.8° to 6.1°C). The wide range of climate sensitivity reflects difficulty in empirically constraining long-term responses using historical temperatures, as noted elsewhere. A relatively tight bound on aerosol forcing is nonetheless obtained from the structure of temperature and aerosol precursor emissions and, particularly, from the rapid growth in emissions between 1950 and 1980. Obtaining a 5th percentile lower bound on aerosol forcing around $-2.0\,\mathrm{W\,m^{-2}}$ requires prescribing internal climate variance that is a factor of 5 larger than the CMIP6 mean and assuming large, correlated errors in global temperature observations. Ocean heat uptake observations may further constrain aerosol radiative forcing but require a better understanding of the relationship between time-variable radiative feedbacks and radiative forcing.

KEYWORDS: Aerosol radiative effect; Forcing; Bayesian methods; Idealized models; Internal variability

1. Introduction

Historical aerosol forcing is a major source of uncertainty in Earth's energy budget, with attendant consequences for observation-based estimates of the transient climate response (TCR) and equilibrium climate sensitivity (ECS) (Andreae et al. 2005; Otto et al. 2013; Forster 2016; Knutti et al. 2017). The most recent comprehensive review of anthropogenic aerosol radiative forcing F_{aer} indicates a 5%–95% confidence interval from -2.0 to $-0.40\,\mathrm{W\,m^{-2}}$ relative to a preindustrial baseline (Bellouin et al. 2020, sections 10 and 11 therein). This range of F_{aer} is broadly consistent with the previous range from -1.9 to $-0.1\,\mathrm{W\,m^{-2}}$ of the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR5; Myhre et al. 2013). The magnitude of F_{aer} is inferred from both process-based constraints and empirical constraints over the historical climate record. Process-based constraints alone, however, yield a broad 5%–95% range from -3.5 to $-0.40\,\mathrm{W\,m^{-2}}$ (Bellouin et al. 2020), and our focus is understanding empirical constraints on the lower bound of F_{aer}.

Empirical lower bounds on F_{aer} are based on the inference that more negative values would imply temperature trends over parts of the twentieth century that are inconsistent with observed warming (Bellouin et al. 2020). Specific lower bounds, however, vary across studies. In one set of studies, lower bounds cluster in the range from -1.8 to $-1.7\,\mathrm{W\,m^{-2}}$ (Aldrin et al. 2012; Skeie et al. 2014, 2018), and, in another, in the range from -1.3 to $-0.70\,\mathrm{W\,m^{-2}}$ (Andronova and Schlesinger 2001; Forest et al. 2006, 2008; Murphy et al. 2009; Libardoni and Forest 2011; Stevens 2015). The origins of these discrepant estimates of F_{aer} are not entirely clear, particularly given that the frameworks make similar assumptions regarding simplified aerosol forcing and climate response models and employ similar sampling techniques. Differences, however, may stem from multiple sources, including the formulation of the forward model, sensitivity to prior distributions, choice of global or regional historical observations, interval of analysis, representation of observational error and internal climate variability, uncertainty associated with forcing efficacy, and assumptions regarding how ocean heat content observations constrain estimates (Annan 2015; Bodman and Jones 2016; Forest 2018).

In this study we examine lower bounds on F_{aer} using estimates of global-annual historical radiative forcing and surface temperature since 1850 and simple models of aerosol forcing and of the temperature response to forcing. We first present a flexible Bayesian model formulation in section 2. In section 3 we describe the inference of a $-1.3\,\mathrm{W\,m^{-2}}$ lower bound on 2010–19 F_{aer}, and in section 4 we explore the origin of these constraints, along with prospects for further constraints based upon ocean heat uptake. Given its focus on drawing physical insight from simple models, our study can be considered to occupy a rung on the hierarchical approach to climate science (Hoskins 1983; Held 2005; Polvani et al. 2017), complementing

Corresponding author: Anna Lea Albright, anna-lea.albright@lmd.jussieu.fr

DOI: 10.1175/JCLI-D-21-0167.1

analyses from more comprehensive models (Pincus et al. 2016; Kretzschmar et al. 2017).

2. Bayesian model

We use a model that is an extension of the linear response framework of Proistosescu and Huybers (2017). This simple framework uses temporal Green's functions to represent the response of global-mean surface temperature and net heat uptake to an imposed radiative forcing. The model is able to emulate the joint evolution of global-average surface temperature and net heat uptake across an ensemble of general circulation models (GCMs) runs from phase 5 of the Coupled Model Intercomparison Project (CMIP5). Here we extend the model to represent the radiative forcing associated with anthropogenic aerosol and volcanic emissions in greater detail.

Some previous studies have used regionally resolved simple models to bound climate sensitivity and F_{aer} (e.g., Andronova and Schlesinger 2001; Forest et al. 2002; Skeie et al. 2014). Although using regional temperature and heat uptake trends makes use of more information, it also engenders considerable complexity that is not easily accounted for. For example, hemispheric energy balance constraints ignore variations in cross-equatorial heat transport (Skeie et al. 2014; Bellouin et al. 2020), whereas models with constant hemispheric- or zonal-mean feedbacks are unlikely to reproduce the time dependence of the net radiative feedback that is thought to be primarily controlled by east–west equatorial temperature gradients (Dong et al. 2019). Furthermore, inferences of F_{aer} and climate sensitivity were found to be sensitive to regional differences in surface temperature observational estimates (Libardoni and Forest 2011, 2013), and regional uncertainties in surface temperatures since 1850 are not yet fully quantified (Chan and Huybers 2019; Davis et al. 2019).

a. Modeled and observed temperatures

Over the historical record, the temperature response to radiative forcing is well captured by linear models (e.g., Hasselmann et al. 1993; Held et al. 2010; Geoffroy et al. 2013; Caldeira and Myhrvold 2013; Haustein et al. 2019). Proistosescu and Huybers (2017) generalized the response of these linear models as a superposition of three independent linear modes, representing the dominant eigenmodes of the climate system:

$$T_{\text{forced}}(t) = \sum_{n=1}^{3} \left(\frac{\alpha_n T_{2\times}}{\tau_n F_{2\times}} \right) \exp(-t/\tau_n) \bigstar F_{\text{net}}(t). \tag{1}$$

The term $T_{2\times}$, or ECS, is the equilibrium climate response to a radiative forcing per doubling of atmospheric CO_2, $F_{2\times}$. Each mode has a response time scale τ_n and contributes a fraction α_n of the equilibrium response. The forced temperature response T_{forced} is obtained by convolving each eigenmode with the net radiative forcing F_{net} and summing across all modes. Convolution in Eq. (1) is indicated by the star, and τ_n^{-1} is a normalizing factor that makes the integral $\int_{t=o}^{t=\infty} \tau_n^{-1} \exp(-t/\tau_n)$ equal one.

Following Proistosescu and Huybers (2017), we employ three eigenmodes because fewer modes result in systematic residuals in fits to GCM simulations during initial model adjustment to increased greenhouse gas forcing, whereas additional eigenmodes do not improve the fit, as judged by a Bayesian information criterion. These three modes, moreover, correspond to annual, decadal, and centennial response times to forcing, consistent with foregoing work (e.g., Caldeira and Myhrvold 2013; Tsutsui 2017; Cummins et al. 2020; Leach et al. 2021).

We assume that the difference between observed global-annual historical temperatures T_{obs} and the forced temperature response T_{forced} is due to a combination of observational uncertainty and unforced internal variability. The difference between observations and the forced response is modeled as a multivariate Gaussian,

$$P(T_{\text{forced}} - T_{\text{obs}}) \sim N(0, \boldsymbol{\Sigma}^2), \tag{2}$$

with the covariance matrix $\boldsymbol{\Sigma}$ representing both internal variability and observational error:

$$\boldsymbol{\Sigma}^2 = \rho^{|i-j|}\sigma_{\text{int}}^2 + \delta_{i,j}\sigma_{\text{obs}}^2. \tag{3}$$

Internal variability, the first term on the right-hand side of Eq. (3), is represented as a first-order autoregressive, or AR1, process with autocorrelation coefficient ρ in keeping with earlier studies (e.g., Frankignoul and Hasselmann 1977; Feldstein 2000; Aldrin et al. 2012; Johansson et al. 2015). Observational uncertainty is represented as uncorrelated error by the second term on the right-hand side of Eq. (3), with $\delta_{i,j}$ equaling 1 when i equals j and 0 otherwise.

We estimate the distribution of internal climate variability using 39 CMIP6 preindustrial control simulations of detrended global-mean temperature, wherein each time series is the last 400 years of the control run (Parsons et al. 2020). Internal variability is estimated as $\sigma_{\text{int}}^2 = 0.019 \pm 0.011°\text{C}^2$, with values equaling the mean and standard deviation across these 39 simulations. Following Schneider and Neumaier (2001), we estimate an AR1 coefficient, $\rho = 0.57 \pm 0.16$. By way of comparison, 41 CMIP5 control simulations (Taylor et al. 2011) yield $\sigma_{\text{int}}^2 = 0.017 \pm 0.014°\text{C}^2$ and $\rho = 0.55 \pm 0.18$. Similarly, for the 40-member Community Earth System Model version 1 Large Ensemble (CESM1 LE) (Kay et al. 2015), we obtain $\sigma_{\text{int}}^2 = 0.010 \pm 0.0019°\text{C}^2$ and $\rho = 0.46 \pm 0.07$. Because the CESM1 LE includes only forced runs, we use the ensemble mean of global-annual surface temperatures as the forced component and attribute ensemble spread to internal variability. We specify ρ as a free parameter in the inversion and use these results as a point of comparison because the appropriate value is not obvious (Bodman and Jones 2016). In our inversion, we set the variance associated with internal variability equal to the CMIP6 ensemble mean, $\sigma_{\text{int}}^2 = 0.019°\text{C}^2$. Allowing for larger internal variability by employing the CMIP6 mean, rather than the CMIP5 or CESM1 LE mean, provides a more conservative estimate of the constraints placed on aerosol forcing by historical temperatures.

For historical temperature observations we use a version of the HadCRUT4 dataset developed by Cowtan and Way (2014) that they refer to as version 2.0. Temperature data span 1850–2019 and are referenced to a 1960–90 baseline. We use time-variable values for σ_{obs}^2, squaring the standard deviation estimated by

Cowtan and Way (2014), which decreases from 0.07°C in 1850 to 0.03°C in 2019. Using NOAA temperature observations and error estimates (Zhang et al. 2019) yields qualitatively similar results, which is unsurprising given the agreement of different temperature products in the global average. Accounting for correlated error structures within observational estimates (Morice et al. 2012; Karl et al. 2015; Cowtan et al. 2015; Chan et al. 2019; Davis et al. 2019) would increase the influence of observational error, and we discuss this possibility further in section 4c.

b. Anthropogenic aerosol forcing model

We follow Stevens (2015) in representing aerosol radiative forcing as a log-linear function of global-annual anthropogenic aerosol precursor emissions. We, however, extend the Stevens (2015) model to include not only SO_2 emissions, but also contributions from black carbon, B, and organic carbon, O, in units of Tg of each per year:

$$F_{\text{aer}} = \alpha_{SO_2} SO_2(t) + \alpha_B B(t) + \alpha_O O(t) + \beta \ln\left[1 + \frac{SO_2(t)}{S_o} + \frac{B(t) + O(t)}{B_o + O_o}\right]. \quad (4)$$

The first three terms on the right-hand side represent the direct effects of aerosol–radiation interactions, wherein sulfur dioxide and organic carbon are expected to have a cooling effect, whereas black carbon could have a warming effect (Penner et al. 1994; Chylek and Wong 1995).

The last term represents aerosol–cloud interactions, defined here as the first indirect, or Twomey, effect (Twomey 1977). This logarithmic term follows Charlson et al. (1992), yet is a highly simplified term given the complex processes underlying aerosol–cloud interactions (Boucher et al. 2013). Whereas the aerosol direct effect is proportional to atmospheric aerosol burden, assumed to follow linearly from SO_2, B, and O, the indirect effect of cloud brightening involves the ratios of SO_2 to the natural aerosol background, S_o, and the sum of B and O to their natural aerosol background, $B_o + O_o$, and is represented as becoming saturated (Twomey 1977; Wigley and Raper 1992; Boucher and Pham 2002). A further rationale for the functional form of Eq. (4) is provided in Stevens (2015, appendices A and B therein). For global-annual SO_2, B, and O emissions, we employ the most recent estimate from O'Rourke et al. (2021), although using estimates from Hoesly et al. (2018) or Smith et al. (2011) gives qualitatively similar results for the period of overlap.

Equation (4) permits for more detail than empirical studies that have rescaled IPCC AR4 or AR5 aerosol effective radiative forcing (e.g., Padilla et al. 2011; Aldrin et al. 2012; Schwartz 2018), yet still involves potentially consequential simplifications. For example, we neglect nonsulfate and noncarbonaceous precursor emissions, such as nitrate emissions, which could impact aerosol radiative forcing (e.g., Bellouin et al. 2011; Hauglustaine et al. 2014). Further, rapid tropospheric adjustments, such as the cloud lifetime effect, are potentially important (Albrecht 1989), though uncertain (Stevens and Feingold 2009; Gettelman 2015; Malavelle et al. 2017). Finally, aerosol forcing is found to be sensitive to emission location in some studies (Zhang et al. 2016; Gettelman and Sherwood 2016; Persad and Caldeira 2019) but insensitive in others (e.g., Fiedler et al. 2019).

To check whether Eq. (4) adequately represents historic variations in aerosol radiative forcing, we fit it to the effective aerosol radiative forcing from 10 CMIP6 models participating in the Radiative Forcing Model Intercomparison Project (RFMIP) and Aerosols and Chemistry Model Intercomparison Project (AerChemMIP) (Pincus et al. 2016). Figure 1 shows the least squares fit of Eq. (4) to these RFMIP and AerChemMIP simulations. After applying a 5-yr smoothing to the effective aerosol radiative forcing, the standard deviation across models is 0.43 W m^{-2} over 1850–2014. Over this interval, the root-mean-square residual between each model and the least squares fit of Eq. (4) has a mean value of 0.09 W m^{-2}, ranging from 0.07 to 0.14 W m^{-2}. If using solely SO_2, the root-mean-square residual is slightly larger, with a mean of 0.13 W m^{-2}, demonstrating that the inclusion of black and organic carbon improves the simple model fit. This small residual suggests that Eq. (4) affords sufficient flexibility to emulate the magnitude and temporal evolution of aerosol radiative forcing generated by more sophisticated models. That is, contributions to global-scale aerosol radiative forcing from nonsulfate and noncarbonaceous aerosols, rapid tropospheric adjustments, and spatial variability in aerosol emissions, as simulated in sophisticated models, can largely be absorbed by appropriate fitting of Eq. (4).

c. Net radiative forcing

To obtain net radiative forcing F_{net}, we sum time series of F_{aer} with nonaerosol effective radiative forcing from Smith et al. (2021), which overlap with observed temperatures from 1850 to 2019. These nonaerosol effective radiative forcing time series are updated AR5 time series that are extended to 2019 using the SSP2.45 scenario (Smith et al. 2021). Nonaerosol forcing agents include long-lived greenhouse gases, tropospheric and stratospheric ozone, land-use change, black carbon on snow, contrails, solar variability, and stratospheric water vapor. These terms together contribute a net positive forcing over the historical period. We address uncertainties in nonaerosol forcing by including a scaling parameter γ_p. We note that greater forcing uncertainty was estimated in AR5 for methane (e.g., Etminan et al. 2016; Hoesly et al. 2018) and solar forcing (e.g., Dudok de Wit et al. 2017). Persisting uncertainty in these nonaerosol forcing terms in the updated forcing data, insomuch as they allow for positive trends, is expected to shift the inferred lower bound on F_{aer} to more negative values (Stevens 2018).

Volcanic forcing is also taken from Smith et al. (2021), but it is important to account for the fact that the temperature response to volcanic forcing appears to be damped relative to the temperature response to other forcing. That is, volcanic forcing appears to have a forcing efficacy less than one (Meinshausen et al. 2011; Marvel et al. 2016; Ceppi and Gregory 2019). We admit for uncertainty in the efficacy of volcanic forcing by introducing a scaling parameter γ_v. Although our focus is on anthropogenic aerosol forcing, we account for uncertainties in volcanic aerosol forcing because of the possibility that it trades off against F_{aer}. In another approach, Stevens (2015) focused

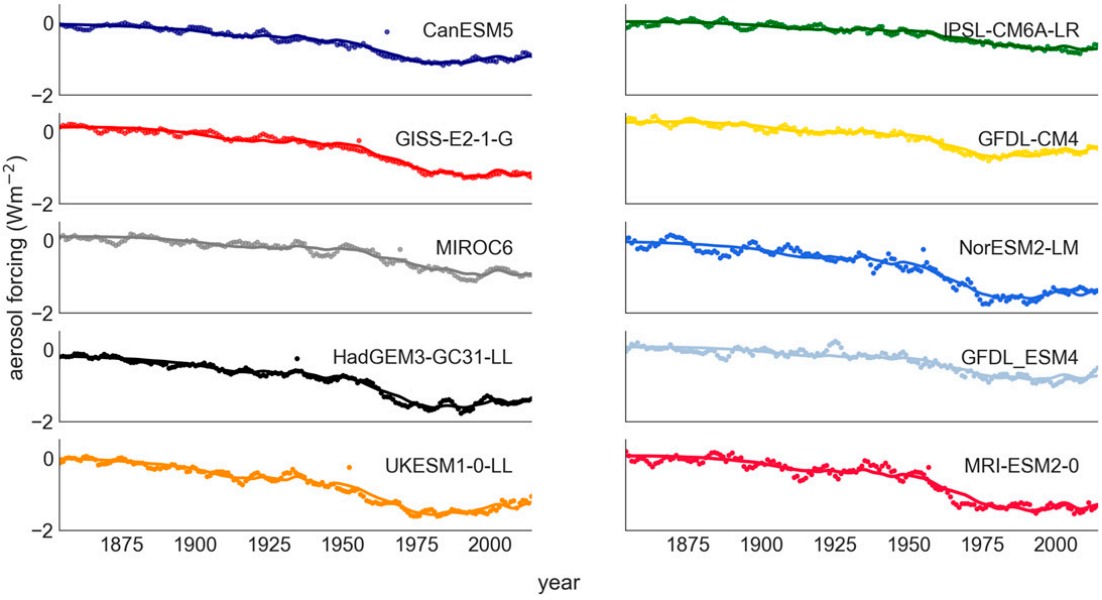


FIG. 1. Fits of Eq. (4) to time histories of aerosol forcing from 10 CMIP6 models participating in the Radiative Forcing Model Intercomparison Project (RFMIP) and Aerosols and Chemistry Model Intercomparison Project (AerChemMIP). Points are 5-yr smoothed values of aerosol forcing anomalies. For RFMIP simulations we calculate aerosol forcing anomalies as the difference of the piClim-Histaer and piClim-control experiments, and for AerChemMIP simulations we calculate anomalies using the histSST experiment relative to the histSST-piAer control experiment. The solid line is the least squares fit of Eq. (4) to each CMIP6 model simulation. Best fits suggest that the highly simplified aerosol model we use can approximate the evolution of historical effective aerosol radiative forcing from more sophisticated models. Furthermore, the parametric uncertainty represented in our Bayesian framework reflects model uncertainty associated with Eq. (4).

on the early record before 1950 as a period of relatively quiescent volcanism in order to separate F_{aer} from volcanic forcing. Such a time-limited approach is discussed in section 4a.

d. Bayesian priors and inversion

Our Bayesian model comprises 16 parameters: 6 associated with aerosol radiative forcing in Eq. (4), 3 associated with nonaerosol radiative forcing variance and the autocorrelation of internal variability, and 7 further temperature response parameters from Eq. (1). Note that the three temperature eigenmode weights α_n are constrained to sum to one, such that it is only necessary to invert for two eigenmode weights. We collectively refer to these 16 parameters as θ.

Table 1 summarizes the prior distribution specified for each parameter. Volcanic forcing efficacy γ_v is assigned a uniform prior from 0 to 1, consistent with findings of lower efficacy than for CO_2 (e.g., Meinshausen et al. 2011; Marvel et al. 2016; Gregory et al. 2016). For nonaerosol, nonvolcanic forcing efficacy γ_p, we assign a normal prior with mean of 1 and standard deviation 0.2. Equilibrium climate sensitivity $T_{2\times}$ is assigned a broad uniform prior from 1° to 10°C (Sherwood et al. 2020). The AR1 coefficient for internal variability, ρ, is specified as uniform from 0 to 1. Priors for τ_n and α_n are normal distributions with mean and standard deviation taken from Proistosescu and Huybers (2017). For adjusted forcing to a doubling of carbon dioxide $F_{2\times}$, we follow the Sherwood et al. (2020) normal prior with a mean of 4.0 W m^{-2} and standard deviation of 0.3 W m^{-2}.

There exist large and long-standing uncertainties associated with aerosol–radiation interactions (e.g., Bellouin et al. 2013; Su et al. 2013; Samset et al. 2014) and aerosol–cloud interactions (e.g., Carslaw et al. 2013; Gettelman 2015; Seinfeld et al. 2016; Feingold et al. 2016). The natural aerosol background parameters, S_o and $B_o + O_o$, are also uncertain (e.g., Carslaw et al. 2010; McCoy et al. 2015), despite their being fixed in previous analyses (e.g., Andronova and Schlesinger 2001; Aldrin et al. 2012; Skeie et al. 2014). Furthermore, to avoid circular reasoning (Rodhe et al. 2000; Anderson et al. 2003), the aerosol forcing prior should not be informed by temperature trends. Accordingly, we assume broad and uniform priors for the six aerosol parameters in Eq. (4) (see Table 1) that are only weakly informative in assuming a uniform distribution from −12 to 0 W m^{-2}. The prior constraint that F_{aer} is negative comes from process-based models (Bellouin et al. 2020). In comparison, Stevens (2015) bounded Eq. (4) such that present-day aerosol forcing varied between −1.5 and 0 W m^{-2}. Other analyses have also been found to be sensitive to the choice of prior (Frame et al. 2005; Hegerl et al. 2006; Annan 2015),

Following Bayes's rule, we invert for the joint posterior distribution of our 16 parameters θ,

$$P(\theta|T_{\text{obs}}) \propto P(T_{\text{obs}}|\theta)P(\theta), \tag{5}$$

where the likelihood $P(T_{\text{obs}}|\theta)$ is formulated based on the multivariate Gaussian distribution of the difference between observed temperature and the forced response of our model:

TABLE 1. Model specification for the baseline inversion using temperature and forcing data from 1850 to 2019. Listed are posterior maximum likelihood estimates (MLE), the associated 5%–95% credible intervals (c.i.), and prior distributions for aerosol forcing parameters (α_{SO_2},α_B, α_O, β, S_o, and $B_o + O_o$), volcanic efficacy (γ_v), nonaerosol forcing efficacy (γ_p), the AR1 coefficient of internal variability (ρ), and temperature response parameters (T_{2x}, τ_1, τ_2, τ_3, α_1, α_1, and F_{2x}). We also report the net aerosol forcing (F_{aer}), aerosol direct (F_{aer_d}), and aerosol indirect (F_{aer_i}) effects that are calculated from Eq. (4); the joint posterior distributions of aerosol parameters; and SO_2, B, and O time series, as described in section 2b. Aerosol forcing values are for the 2010–19 interval, with values for the 2000–10 interval given in parentheses. The parameter α_{SO_2} has units of W m^{-2} (Tg SO_2 yr^{-1})$^{-1}$ and is multiplied by 10^3 in the table. Similarly, parameters α_B and α_O have units of W m^{-2} (Tg B yr^{-1})$^{-1}$ and W m^{-2} (Tg O yr^{-1})$^{-1}$, respectively, and are multiplied by 10^3. Columns μ and σ are the mean and standard deviation of normal prior distributions.

	Units	MLE	5% c.i.	95% c.i.	Prior distribution	Range	μ	σ
α_{SO_2}	See caption	−2.8	−5.5	−0.37	Uniform	−50 to 0	—	—
α_B	See caption	7	0.39	21	Uniform	−50 to 50	—	—
α_O	See caption	−16	−36	−8.5	Uniform	−50 to 50	—	—
β	W m^{-2}	−0.45	−0.80	−0.039	Uniform	−3 to 0	—	—
S_o	Tg SO_2	97	35	144	Uniform	10–200	—	—
$B_o + O_o$	Tg B + O	31	5.4	57	Uniform	10–60	—	—
ρ	—	0.70	0.59	0.77	Uniform	0–1	—	—
$T_{2\times}$	°C	3.4	1.8	6.1	Uniform	1–10	—	—
γ_v	—	0.72	0.59	0.97	Uniform	0–1	—	—
γ_p	—	1.01	0.75	1.3	Normal	—	1	0.2
τ_1	yr	0.67	0.30	1.1	Normal	—	0.7	0.25
τ_2	yr	12	4.9	21	Normal	—	9	7
τ_3	yr	352	221	482	Normal	—	354	80
α_1	—	0.20	0.069	0.38	Normal	—	0.24	0.15
α_2	—	0.38	0.12	0.57	Normal	—	0.32	0.15
$F_{2\times}$	W m^{-2}	4.0	3.5	4.5	Normal	—	4.0	0.3
F_{aer}	W m^{-2}	−0.85 (−0.95)	−1.3 (−1.4)	−0.50 (−0.56)	Uniform	−12 to 0	—	—
F_{aer_d}	W m^{-2}	−0.56	−0.91	−0.16	Uniform	−6 to 0	—	—
F_{aer_i}	W m^{-2}	−0.45	−0.99	−0.04	Uniform	−6 to 0	—	—

$$P(T_{obs}|\theta) \sim N[T_{forced}(\theta), \mathbf{\Sigma}^2(\theta)]. \quad (6)$$

Sampling is performed using the Metropolis–Hastings algorithm (Metropolis et al. 1953; Hastings 1970), and we run three chains of one million samples. The first 5000 burn-in samples are discarded for each chain, and the remainder are thinned by a factor of 5 to reduce serial correlation. Results are consistent among chains, indicating that our model is adequately sampled. Reflecting how observed temperatures are represented, model temperatures between 1960 and 1990 are removed from each realization.

3. Results

Table 1 lists the maximum likelihood estimate (MLE) and 5%–95% posterior credible interval ranges for the 16 parameters, θ, as well as ranges for F_{aer} and aerosol direct and indirect effect forcing. Modeled and observed temperatures fit well, with the posterior 5%–95% posterior credible interval covering 99% of observations (Fig. 2). The credible interval is determined by integrating Eq. (1) using each posterior realization of θ and adding realizations of observational error and internal variability. A mismatch does appear, however, in the form of warmer-than-predicted temperatures observed in the early 1940s. This mismatch, noted in other attempts to reproduce surface temperature from forcing (Folland et al. 2018), likely arises because of systematic errors in the observations, given that such a warm anomaly is absent in nearshore land station temperature records (Cowtan et al. 2018) and is absent from sea surface temperature records after correcting for systematic biases associated with engine-room-intake and bucket estimates (Chan and Huybers 2021).

The posterior distribution of F_{aer} is calculated from the joint posterior distribution of aerosol parameters associated with Eq. (4) (Table 1). The posterior MLE for F_{aer} between 2000 and 2010, relative to 1750, is −0.95 W m^{-2}. The lower bound on F_{aer} is −1.4 W m^{-2} using a 5th percentile value or −1.5 W m^{-2} using a 1st percentile value. Calculating the lower bound over 2010–19, the posterior MLE for F_{aer}, again relative to 1750, is −0.85 W m^{-2}, and its lower bound is −1.3 W m^{-2} using a 5th percentile value (Fig. 2c), or −1.4 W m^{-2} using a 1st percentile value. The mean aerosol direct effect over 2010–19 is −0.56 W m^{-2} (from −0.91 to −0.16 W m^{-2} 5%–95% credible interval), and the mean aerosol indirect effect over 2010–19 is −0.45 W m^{-2} (from −0.99 to −0.04 W m^{-2} 5%–95% credible interval).

a. *Origin of lower-bound constraints on anthropogenic aerosol forcing*

Offsetting covariance among the six aerosol forcing parameters in Eq. (4) better constrains the overall range of F_{aer} than its individual components. The natural sulfate background, S_o, for instance, covaries with values of β with a Pearson correlation coefficient of $r = 0.32$. Thus, whereas a more negative value of F_{aer} can come about either through a greater indirect effect, β, or smaller natural aerosol background, the combination of these parameters is constrained.

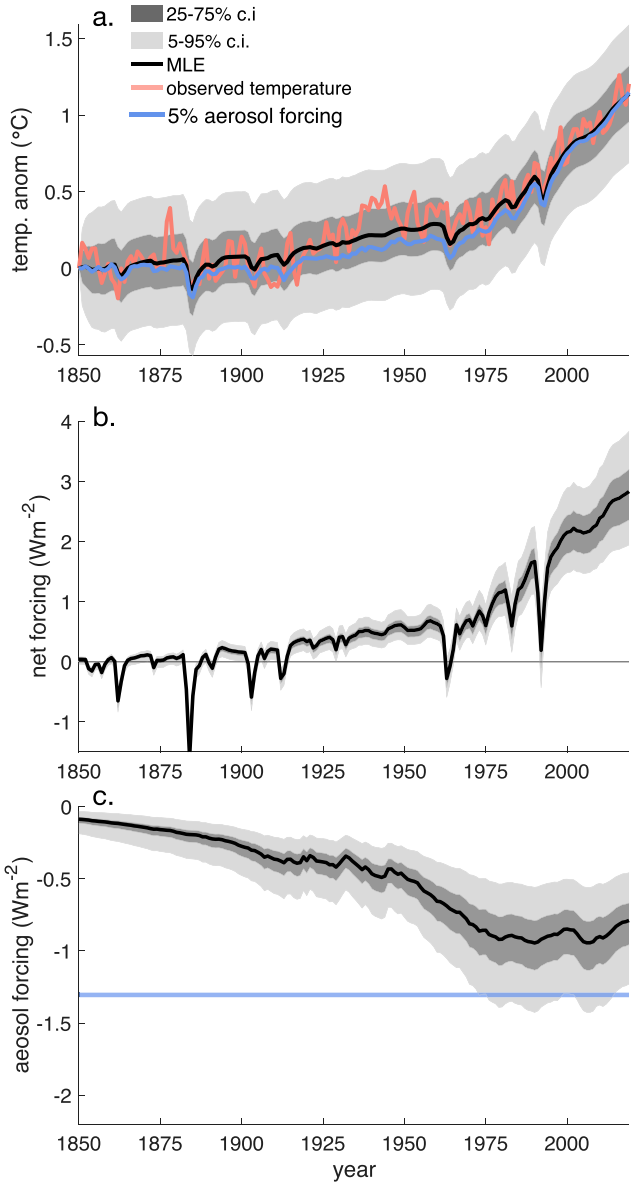


FIG. 2. Posterior estimates of temperature and forcing using data from 1850 to 2019. (a) Observed temperatures from HadCRUT4 (red) and the maximum likelihood estimates (MLE) of temperature (black), where each is referenced to an 1850 baseline. Also shown is temperature if aerosol radiative forcing is conditioned on its 5th percentile lower bound between 2010 and 2019 of $-1.3\,\mathrm{W\,m^{-2}}$ (light blue). (b) Net forcing comprising contributions from greenhouse gases, anthropogenic and volcanic aerosols, land-use change, tropospheric and stratospheric ozone, black carbon on snow, contrails, stratospheric water vapor, and solar variability. (c) Anthropogenic aerosol radiative forcing (F_{aer}) in terms of its MLE (black line) and its 5th percentile lower bound averaged over 2010–19 (blue line). Shading indicates 50% (dark gray) and 90% (light gray) credible intervals for annual averages.

The two direct effect coefficients for carbonaceous aerosols, α_{B} and α_{O}, are anticorrelated, with $r = -0.82$. Aerosol direct and indirect effects also trade off in capturing variability in observed temperature trends. Parameters $\alpha_{\mathrm{SO_2}}$ and β compensate, with $r = -0.54$, as do the aerosol direct and indirect effects, $r = -0.56$, in explaining observed temperature trends, such that the model cannot readily distinguish the aerosol direct from indirect effect. The parameter covariance, even in this highly simplified model, highlights the need for uncertainty quantification in higher-dimensional models, as discussed by Lee et al. (2016) and Carslaw et al. (2018).

We find general agreement between the MLE aerosol forcing parameters and the best-fit parameters across the 10 simulations in Fig. 1. The MLE β of $0.45\,\mathrm{W\,m^{-2}}$ is qualitatively similar to the best-fit β of $0.42\,\mathrm{W\,m^{-2}}$. Our MLE $\alpha_{\mathrm{SO_2}}$ of 2.8 is smaller than the best-fit value of $3.7\,\mathrm{W\,m^{-2}}\,(\mathrm{Tg\,SO_2\,yr^{-1}})^{-1} \times 10^{-3}$, but our MLE S_o of 97 Tg SO_2 is also slightly smaller than the best-fit value across the simulations of 107 Tg SO_2, such that the overall effect is similar. The 10 simulations diverge on the sign and magnitude of the best-fit direct effect parameters for black and organic carbon, α_{B} and α_{O}, consistent with the notion that the direct effect parameters trade-off in capturing observed temperature trends. Note that the Bayesian framework better captures the covariance among these parameters than does the least squares fit used earlier for purposes of initial characterization.

Posterior constraints on the volcanic forcing efficacy parameter γ_v give a MLE value of 0.72 and a 5%–95% posterior credible interval of 0.59–0.97, indicating that the efficacy of volcanic forcing for producing a temperature response is less than that of carbon dioxide forcing. Previous estimates of volcanic efficacy range from 0.5 to 0.7 (e.g., Chylek et al. 2020; Marvel et al. 2016), consistent with our estimate, although these values were for AR5 volcanic forcing data. Performing our inversion with AR5 forcing data produces a MLE volcanic efficacy parameter of 0.67. Lehner et al. (2016), however, suggest that it is premature to infer that volcanic forcing has reduced efficacy because of the coincidence of volcanic eruptions and El Niño events over the twentieth century. A more detailed study of volcanism, El Niño events, and internal climate variability, possibly using this same framework, could be pursued in future work.

Perhaps surprisingly, γ_v does not substantially contribute to uncertainty in F_{aer}. The magnitude of F_{aer} and the volcanic forcing efficacy are weakly correlated ($r = -0.10$). This lack of correlation despite both terms being negative forcing reflects that volcanic forcing tends to decay more rapidly and shows little temporal covariance with trends in F_{aer}. Conditional on the 5% aerosol forcing lower bound, the MLE of γ_v is 0.71, with a similar posterior distribution to the posterior distribution that is not conditioned on 5% aerosol forcing. These results suggest that it is less essential to subselect time periods associated with low volcanic activity in order to isolate F_{aer}, at least when all terms are cointegrated within a Bayesian framework.

Unlike for volcanic forcing, we find that uncertainty in the efficacy of nonaerosol forcing γ_p, relative to values in Smith et al. (2021), is an important contribution to the uncertainty in F_{aer}. The MLE of γ_p is 1.01 with a 5%–95% posterior credible interval of 0.75–1.3. Conditional on the 5th percentile of F_{aer}, the MLE of γ_p is 1.3, giving the expected result that greater nonaerosol forcing allows for more negative values of F_{aer}. Conversely, if γ_p is assigned its MLE of 1.01, the 5th percentile of F_{aer} is $-1.1\,\mathrm{W\,m^{-2}}$.

For the first-order autoregressive coefficient, ρ, we find a MLE value of 0.70 with a 5%–95% credible interval of 0.59–0.77. The MLE value of ρ corresponds to an autocovariance time scale of $\tau = -1/\ln(\rho)$, or 2.5 years, with a 5%–95% credible interval of 1.9–3.8 years. Note that F_{aer} is only weakly sensitive to the time scale of internal variability, as represented using an AR1 process and drawn from the posterior distribution of ρ. If ρ is assigned its MLE value of 0.70, the 5th percentile of the F_{aer} distribution for 2010–19 is unchanged at $-1.3\,W\,m^{-2}$. The narrow distribution of the autocovariance time scale τ may reflect that our representation does not capture low-frequency, interdecadal variability. In section 4c we discuss an extension that better accounts for internal variability at longer time scales.

b. Relationship with climate sensitivity

Our MLE climate sensitivity is 3.4°C and has a broad 5%–95% posterior credible interval of 1.8°–6.1°C. The higher mean and upper bound on ECS than in Sherwood et al. (2020) reflects contributions from the long-response time scale, which is almost unrealized in the historical record (Proistosescu and Huybers 2017). The loading on each eigenmode equals α_n times ECS and has an average that increases from 0.67 (τ_1 = 0.67 yr) to 1.10 (τ_2 = 12 yr) to 1.63°C (τ_3 = 352 yr). The uncertainty on each eigenmode also increases with times scales: standard deviations for the loading on each eigenmode increase from 0.25° to 0.41° to 1.2°C.

Although not the focus of this study, it appears the difference in how informative the transient response is of the equilibrium response helps explain the wide range of foregoing upper bounds on climate sensitivity from empirical studies (e.g., Bodman and Jones 2016). Recent work suggests that as SST patterns evolve from their transient to their equilibrium state, they modulate the net radiative feedback, and, thus, the effective climate sensitivity (Andrews et al. 2015). Representing zonal equatorial temperature gradients appears especially important for capturing changes in radiative feedbacks (Dong et al. 2019). The resulting difference between the transient and long-term feedbacks is called a ''pattern effect'' and needs to be accounted for when inferring ECS from the fast response manifested over the historical record (e.g., Sherwood et al. 2020). Such a pattern effect is generally missing from previous simple models used in inversions of aerosol forcing and climate sensitivity, leading to a strong covariance between the transient and equilibrium feedbacks. In contrast, the pattern effect exhibits a large spread in general circulation models (Andrews et al. 2018) and cannot currently be constrained from the historical climate record (Sherwood et al. 2020).

The Proistosescu and Huybers (2017) model is a general approximation of the leading modes of response that allows for the presence of a pattern effect by allowing for different radiative feedbacks between the fast and slow modes. In this regard, our model appears more flexible than previous box models used for inferring F_{aer}. The three modes can be interpreted as representing time-evolving warming patterns associated with, for example, warming over land on annual time scales, the ocean mixed layer on decadal time scales, and the eventual warming of the deep ocean on centennial time scales. Each mode is assumed to have a different underlying surface temperature pattern and, thus, a different radiative feedback. As the relative loading of each mode changes, so does the SST pattern and the net radiative feedback. The lack of constraints on the long-term feedback and, thus, on the ECS can be understood as a consequence of the fact that this long time scale is almost entirely unrealized in observations.

Previous inverse (e.g., Forest 2018) or coupled model studies (e.g., Andreae et al. 2005; Chylek et al. 2016) found a strong relationship between F_{aer} and climate sensitivity, whereas Smith et al. (2020) do not find such a relationship in the CMIP6 ensemble. We obtain a correlation between the *net* forcing and the fast-mode response, here defined as the summed response of the annual and decadal modes, of $r = -0.76$, or between the net forcing and climate sensitivity of $r = -0.52$. The correlation between F_{aer} and ECS, however, is only $r = 0.01$ on account of uncertainty in the relationship between the fast-mode response and ECS as well as between F_{aer} and F_{net}. That said, this low correlation does not fully capture dependencies between F_{aer} and ECS. Conditioned on the 5th percentile value of the climate sensitivity distribution, the 2010–19 F_{aer} 5th percentile value is $-1.1\,W\,m^{-2}$, whereas when conditioned on 95th percentile climate sensitivity the 5th percentile of F_{aer} becomes $-1.3\,W\,m^{-2}$. Figure 3 illustrates how, despite the lack of a linear relationship, the magnitude of F_{aer} sets the range of other parameters, including the net forcing and the temperature response across different time scales.

4. Discussion

The Bayesian framework allows us to explore how the observational record constrains F_{aer} and climate sensitivity through quantifying how different processes combine to allow fitting to the data within uncertainties, similar to the ''storyline'' approach outlined for constraining climate sensitivity in Stevens et al. (2016). We are particularly interested in reconciling our lower bound with more negative estimates (Bellouin et al. 2020) and, thus, examine several modifications that would alter our 2010–19 5th percentile F_{aer} estimate of $-1.3\,W\,m^{-2}$ to values around $-2.0\,W\,m^{-2}$.

a. Constraining aerosol forcing using pre-1950 data versus the full record

Stevens (2015) used a shorter interval less influenced by volcanic activity and obtained a 2005 F_{aer} lower bound of $-1.3\,W\,m^{-2}$ when employing global energy balance arguments. If only fitting our model using temperature and forcing from 1850 to 1950, we instead obtain a MLE F_{aer} of $-1.4\,W\,m^{-2}$ and a 5th percentile lower bound of $-2.0\,W\,m^{-2}$ for the 2000–10 interval. For 2010–19, we obtain a MLE F_{aer} of $-1.3\,W\,m^{-2}$ and a 5th percentile lower bound of $-1.8\,W\,m^{-2}$.

We illustrate the origin of the more negative bound on F_{aer} than in Stevens (2015) when only fitting to data before 1950 by sequentially modifying three parameters within our model. First, we used broad priors on the aerosol direct effect parameters, α_{SO_2}, α_B, and α_O, thus allowing for a more linear relationship between F_{aer} and precursor emissions, which addresses the particular critique of Stevens (2015) by Kretzschmar et al. (2017)

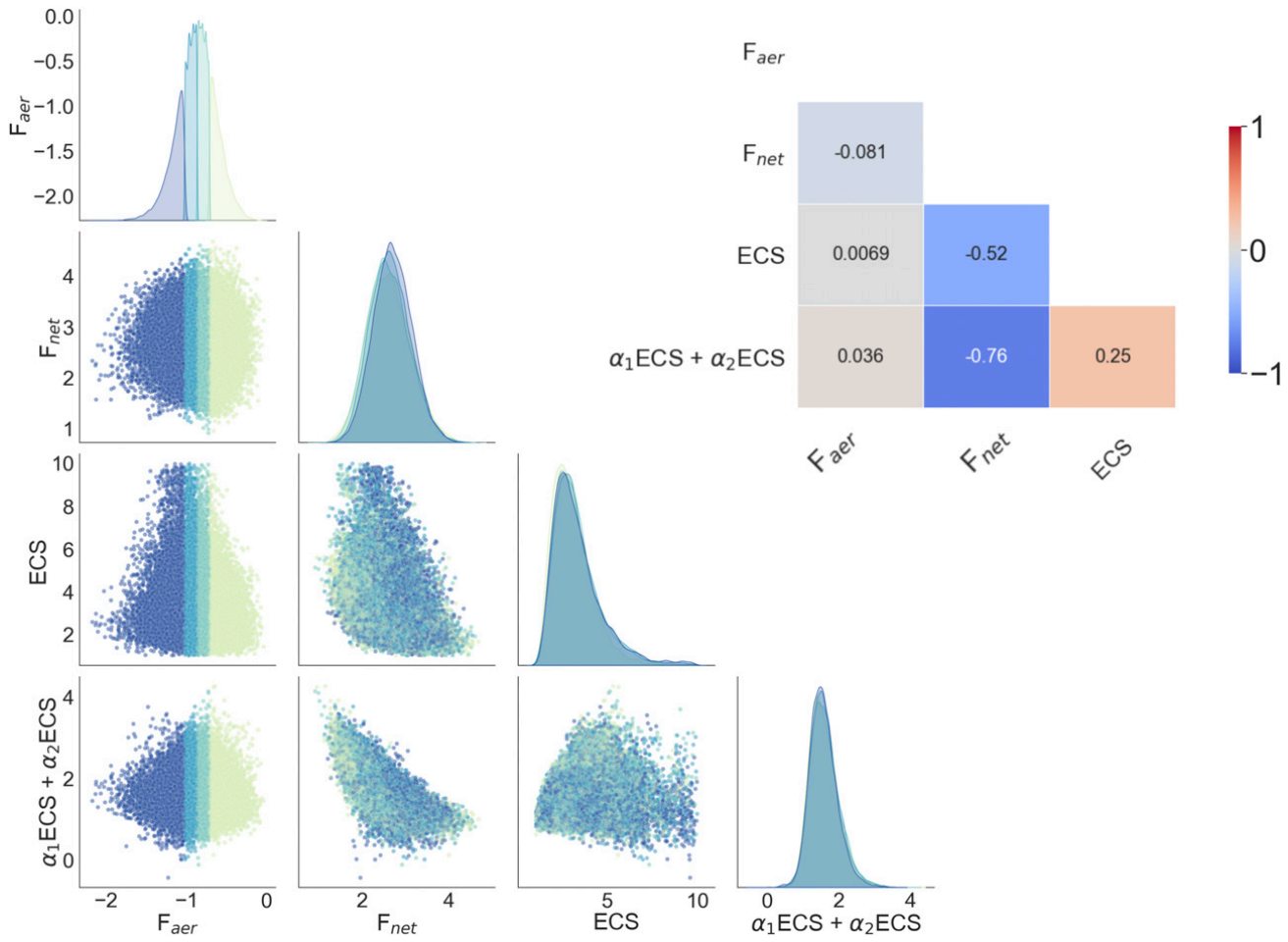


FIG. 3. Parameter covariance among aerosol forcing (F_{aer}), net radiative forcing (F_{net}), climate sensitivity (ECS), and the fast-mode temperature response to forcing (annual contributions from $\alpha_1 \times$ ECS plus decadal from $\alpha_2 \times$ ECS). Binning aerosol forcing by quartiles, shown from the lower quartile (blue) to upper quartile (yellow), illustrates how aerosol forcing influences the permissible range of other terms. The inset shows Pearson correlation coefficients between pairs of terms.

and Booth et al. (2018) that aerosol forcing could be more linear than logarithmic with emissions. If α_{SO_2}, α_B, and α_O are fixed at their MLE values, the 5th percentile lower bound on 2000–10 F_{aer} becomes $-1.7\,W\,m^{-2}$. Second, we incorporated non-aerosol forcing terms beyond solely long-lived greenhouse gas forcing. Excluding all but these greenhouse gas forcing terms gives a 5th percentile of F_{aer} equal to $-1.6\,W\,m^{-2}$. Finally, we allowed for uncertainty in temperature observations and internal variability that, if excluded, yields a 5th percentile bound of $-1.6\,W\,m^{-2}$. Considering these modifications jointly, we can obtain values commensurate with Stevens (2015). Fixing aerosol direct effect parameters at their MLE values, excluding all radiative forcing except that from long-lived greenhouse gases, and excluding uncertainty in temperature observations and internal variability yields a 2000–10 5th percentile aerosol forcing lower bound of $-1.2\,W\,m^{-2}$. This bound is slightly tighter than the Stevens (2015) bound, possibly on account of employing temperature trends rather than comparing net forcing in 1850 and 1950.

Our approach addresses several critiques of the Stevens (2015) framework (Kretzschmar et al. 2017; Booth et al. 2018; Stevens 2018) and indicates that the approach taken by Stevens (2015) gives a larger-magnitude lower bound on aerosol forcing when accounting for additional uncertainties. The fact that our baseline estimate of aerosol radiative forcing is similar to Stevens (2015) is a coincidence whereby greater uncertainties are offset by additional constraints available from considering a longer time series and additional structure in the aerosol emissions and temperature histories.

b. *The role of temporal structure in separating forcing from sensitivity*

It has been suggested that a primary source of uncertainty in constraining aerosol forcing is compensation between radiative forcing and climate sensitivity in explaining historical temperature trends (Andreae et al. 2005; Kiehl 2007; Knutti et al. 2017). This compensation appears when restricting the analysis prior to the 1950s, but less so afterward. Whereas the anticorrelation between ECS and F_{aer} is weak using data from 1850 to 2019 ($r = -0.01$), after restricting to 1850–1950 the anticorrelation increases to -0.30. Using only pre-1950 data, we obtain a highly uncertain ECS distribution

with a 5%–95% credible interval of 2.3°–9.1°C and a MLE value of 5.2°C.

In our framework, limiting model fitting to temperatures from 1850 to 1950 results in a posterior temperature distribution with too much cooling between 1950 and 1980, followed by too much warming (Figs. 4a,b). That is, increased aerosol precursor emissions between 1950 and 1980 cause excessive cooling, and a plateau in emissions after 1980 unmasks the high climate sensitivity and produces excessive warming relative to observations. We infer that the temporal evolution of aerosol precursor emissions and temperature post-1950 results in a reduced capacity for aerosol forcing and sensitivity to trade off against each other in explaining temperature trends.

Previous studies have also pointed to the important role of differences in the post-1950 evolution of aerosol forcing and temperature trends. In the GFDL Global Atmosphere and Land Model AM4.0/LM4.0 (Zhao et al. 2018) and DOE E3SM Coupled Model version 1 (Golaz et al. 2019), strongly negative F_{aer} and high climate sensitivity led to mismatches between modeled and observed temperatures. In particular, these models simulated reduced warming in the 1960s–1980s, followed by an excessive warming trend after aerosol emissions peak around 1980. When using the Held et al. (2010) two-layer energy balance model to emulate their coupled model temperature response, Golaz et al. (2019) found it necessary to reduce F_{aer} from $-1.65\,W\,m^{-2}$ to approximately $-0.8\,W\,m^{-2}$ and ECS from 5.3° to 2.7°C in order to improve the fit of modeled to observed temperatures, qualitatively in keeping with our findings.

c. Increased internal or observational variability

The treatment of low-frequency internal variability was highlighted by Forest (2018) as a key open issue for constraining F_{aer}. Indeed, an important difference in foregoing inverse model studies that obtained a lower bound around $-2.0\,W\,m^{-2}$ appears to be the fact that they allow for more low-frequency internal variability or observational error (e.g., Aldrin et al. 2012; Skeie et al. 2014, 2018). We examine how the lower bound of F_{aer} depends on the specification of the variance of internal variability σ^2_{int} and its autocorrelation ρ. An increase in ρ is motivated by the appearance of greater decadal autocorrelation than would be expected if temperature variations followed an AR1 process (e.g., Vyushin and Kushner 2009; Laepple and Huybers 2014), and that observational products show temperature trends that lie outside the range of temperature variability simulated by coupled models in historical runs (e.g., Zhou et al. 2016; Coats and Karnauskas 2017; Hegerl et al. 2018; Parsons et al. 2020).

One means to obtain an estimate of the autocorrelation of low-frequency internal variability is to calculate the AR1 coefficient ρ from global-mean, detrended temperature time series from 39 CMIP6 coupled model control runs using 10-yr block-averaged temperatures. From block-averaged temperatures, we obtain a mean AR1 coefficient of 0.33 across the 39 simulations. This value implies a mean annual AR1 coefficient of 0.90, where $\rho_{annual} = (\rho_{block})^{1/n}$ and n is the number of years in the block. We obtain a 5th percentile of F_{aer} equal to $-1.7\,W\,m^{-2}$ (MLE $-1.1\,W\,m^{-2}$) for 2000–10, or $-1.6\,W\,m^{-2}$ (MLE $-1.0\,W\,m^{-2}$) for 2010–19 after increasing ρ from 0.5 to 0.90 and doubling innovation variance σ^2_{int} from its CMIP6 ensemble mean value of 0.019° to 0.036°C^2 (Figs. 4c,d). Increasing innovation variance to 0.037°C^2 is within the range simulated by 39 CMIP6 control simulations, which have a maximum value of 0.060°C^2. These modifications increase the variance of the AR1 process fivefold, where variance is $\sigma^2_{int}/(1-\rho^2_{annual})$.

Similar to increased internal variability, representing greater error in observations also permits for a more negative lower bound on F_{aer} (Figs. 4e,f). In particular, low-frequency observational errors in sea surface temperatures could occur at the multidecadal time scales of instrumentation changes (Chan et al. 2019). We add AR1 observational error, similar to the term for internal variability in Eq. (3), with observational error variance estimated as the offset between modeled and observed temperatures between 1936 and 1948, which has a mean of 0.14°C, or 0.02°C^2 (e.g., Fig. 2a). If we assume this variance of 0.02°C^2 and an AR1 coefficient of 0.9, independent of increasing internal variability, we obtain a 5% F_{aer} lower bound of $-1.6\,W\,m^{-2}$ for 2000–10, or $-1.5\,W\,m^{-2}$ for 2010–19. Estimating bounds using both increased internal variability and observational error leads to a MLE of $-1.2\,W\,m^{-2}$ and 5% lower bound on F_{aer} of $-1.9\,W\,m^{-2}$ for 2000–10, or $-1.8\,W\,m^{-2}$ for 2010–19 (Figs. 4e,f).

d. Ocean heat content

Ocean heat content observations could, conceivably, further constrain F_{aer}. The covariance between F_{aer} and ocean heat uptake noted in previous work (e.g., Forest et al. 2002; Forest 2018) argues for inferring both jointly. Skeie et al. (2014) and Johansson et al. (2015) state the importance of ocean heat content observations, yet other studies find that incorporating these observations only weakly constrains climate sensitivity (Knutti et al. 2002; Aldrin et al. 2012) and is influenced by the data source, depth covered, and observational uncertainties (Sokolov et al. 2010).

To explore whether additional constraints on F_{aer} can be obtained from ocean heat content observations, we compare the top-of-atmosphere radiative imbalance predicted by our climate response model with the radiative imbalance estimated from observed ocean heat content estimates. The Proistosescu and Huybers (2017) model that we build upon predicts the global net top-of-atmosphere radiative imbalance $N_{forced}(t)$ as the difference between net forcing F_{net} and the radiative response to forcing,

$$N_{forced}(t) = F_{net}(t) - \sum_{n=1}^{3} \lambda_n T_n(t), \tag{7}$$

where $T_n(t)$ is given by Eq. (1), and λ_n is the radiative feedback associated with each mode of the response. Values of $T_n(t)$ are drawn from our posterior distribution, and values of λ_n are the median radiative feedback coefficients from Proistosescu and Huybers (2017) that were obtained by fitting to CMIP5 simulations.

We compare our results to the top-of-atmosphere radiative imbalance implied by the full-depth ocean heat uptake

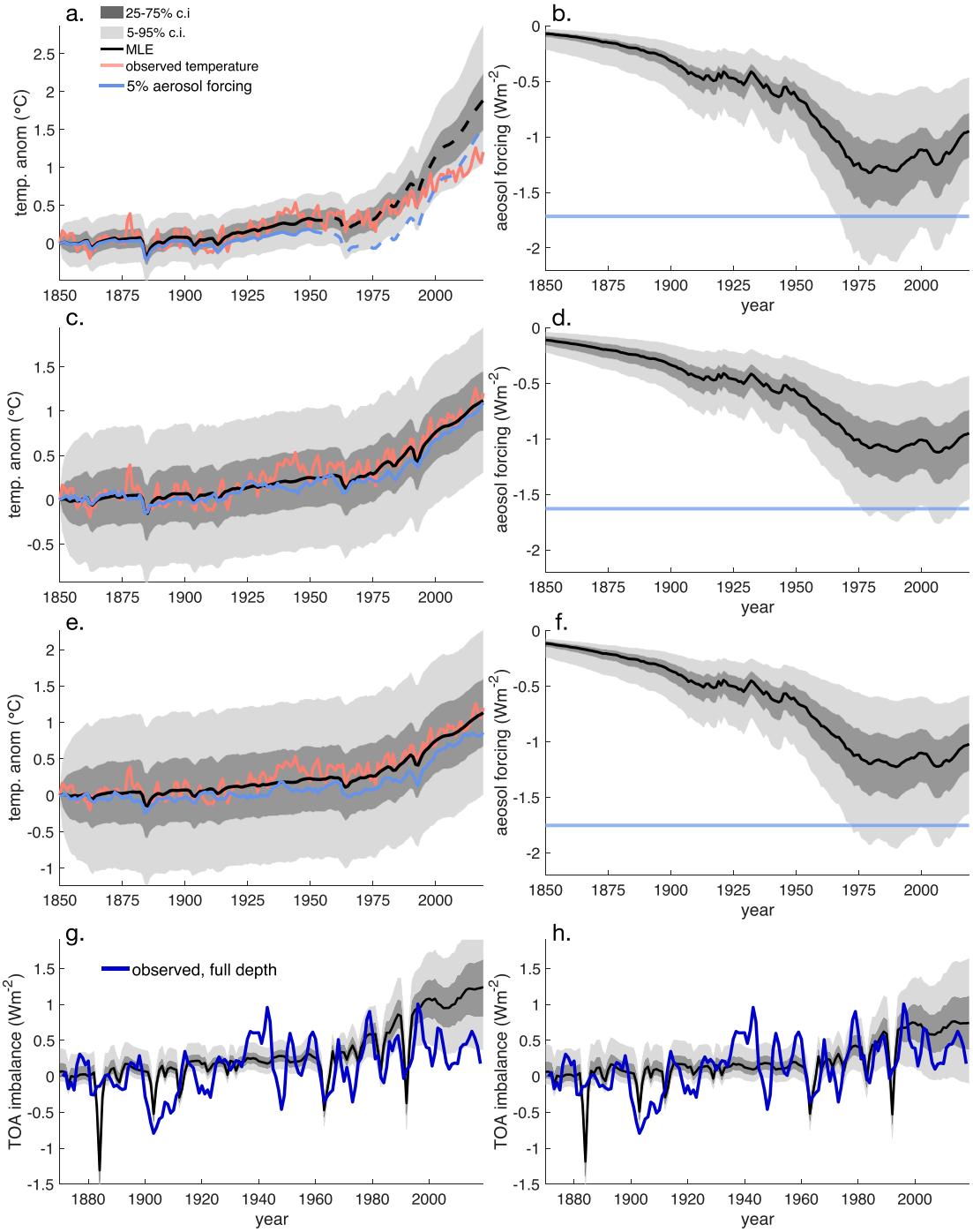


FIG. 4. Posterior estimates of temperature and forcing from various scenarios. (a),(b) Only using data from 1850 to 1950 results in a MLE (solid black line) that predicts warmer temperatures after the 1990s (dashed black line) than observed (red line). Constraining the model to the 5th percentile of aerosol radiative forcing (solid blue line) leads to excessive cooling between 1950 and 1990, followed by stronger warming than observed (dashed blue line). Aerosol radiative forcing is both more negative and more uncertain relative to using the full period (Fig. 2). (c),(d) Increasing the variance of internal variability by a factor of 5, and (e),(f) additionally including similar low-frequency observational error, leads to larger credible intervals. Posterior credible intervals are shown at 50% (dark gray) and 90% (light gray) coverage. Observed and simulated temperatures are referenced to an 1850 baseline. (g) Modeled top-of-atmosphere radiative imbalances (black line) using the median values of CMIP5-inferred radiative damping parameters [λ_n in Eq. (7)]. Radiative imbalances are estimated from ocean heat content estimates Zanna et al. (2019) divided by 0.9. (h) As in (g), but after increasing radiative damping [i.e., λ_n in Eq. (7)].

estimates of Zanna et al. (2019), after dividing ocean heat uptake rates by 0.9, under the assumption that 90% of Earth's energy imbalance is absorbed by the ocean (von Schuckmann et al. 2020). Uncertainties remain, however, regarding how heat uptake is partitioned among the upper and lower ocean, atmosphere, cryosphere, and land system (e.g., Hansen et al. 2005; Cheng et al. 2017).

Comparison of our maximum likelihood estimate of the top-of-atmosphere radiative imbalance against the implied estimate from Zanna et al. (2019) indicates a mismatch in the 1940s and after 1980 (Fig. 4g). Given that the ocean heat uptake estimated by Zanna et al. (2019) is based upon propagating SSTs into the ocean interior, the mismatch in the 1940s may be the result of SST errors during this time period (Cowtan et al. 2018; Chan and Huybers 2021). The fact that our model produces a larger top-of-atmosphere radiative imbalance than estimated by Zanna et al. (2019) after 1980, however, suggests that the recent rate of planetary heat uptake is poorly fit with climate response parameters estimated from the CMIP5 simulations.

For the median CMIP5 λ_n parameters, the MLE of the radiative imbalance from 2006 to 2019, referenced to an 1870 baseline, is 1.16 $\mathrm{W\,m^{-2}}$ (0.45–1.94, 5%–95% credible interval). Sherwood et al. (2020), however, report a MLE of the radiative imbalance of 0.80 $\mathrm{W\,m^{-2}}$ (0.55–1.04, 5%–95% credible interval) for 2006–18. Discrepancies in λ_n are also found in recent work suggesting that coupled GCMs exhibit too little radiative damping relative to GCMs forced with observed SSTs (Zhou et al. 2016; Sherwood et al. 2020). The radiative response to forcing is defined by the second term in Eq. (7), $R = \sum_{n=1}^{3} \lambda_n T_n(t)$, and a too small R leads to an overestimate of the top-of-atmosphere radiative imbalance.

As a heuristic example, we increase λ_n from their CMIP5-fit values of $\lambda_1 = 1.6, \lambda_2 = 1.4$, and $\lambda_3 = 0.8\,\mathrm{W\,m^{-2}\,°C^{-1}}$ to $\lambda_1 = 2.2$, $\lambda_2 = 1.8$, and $\lambda_3 = 0.8\,\mathrm{W\,m^{-2}\,°C^{-1}}$, increasing the net feedback over the historical period from -1.45 to $-1.95\,\mathrm{W\,m^{-2}\,K^{-1}}$, consistent with a $\Delta\lambda = 0.5\,\mathrm{W\,m^{-2}\,K^{-1}}$ mean bias between coupled simulations and simulations prescribed with observed SST patterns (Sherwood et al. 2020). Increased radiative damping decreases the predicted top-of-atmosphere radiative imbalance and increases consistency with ocean heat uptake estimates (Fig. 4h), with MLE radiative imbalance becoming 0.71 $\mathrm{W\,m^{-2}}$ (from -0.02 to 1.54 5%–95% credible interval).

We are thus concerned that biases in simulated radiative damping and planetary heat uptake over recent decades implies that emulators trained on coupled models are similarly biased. Specifically, the discrepancy between the observed planetary heat uptake and net top-of-atmosphere radiative imbalance from coupled models can be thought of as a "ghost forcing." Experiments with prescribed SSTs suggest that this ghost forcing is due to historical SSTs causing a stronger radiative damping and more negative cloud feedback than SSTs in coupled models (Zhou et al. 2016). An inversion with priors informed by coupled models, however, can erroneously attribute this ghost forcing to more negative radiative forcing F_{net}. We speculate that the weakened radiative damping in coupled models leads to inferences of F_{aer} that are biased low when attempting to fit to recent planetary heat uptake. Indeed, a recent paper by Smith et al. (2021) using emulators trained on coupled models finds that the MLE value of F_{aer} becomes more negative by 0.2 $\mathrm{W\,m^{-2}}$ when also fitting planetary heat uptake. In our formulation, the uncertainties in radiative damping are encapsulated in uncertainties in λ_n. Given the lack of independent observational constrains on the true radiative damping and uncertainty about whether recent SST patterns that cause more negative feedbacks are forced or unforced (Andrews et al. 2018; Sherwood et al. 2020), we conclude that ocean heat content data at present do not offer robust additional constraints on F_{aer}. We also note that trends in ocean heat content prior to the 1950s have potentially important uncertainties, particularly in the deep ocean (Gebbie and Huybers 2019).

e. *Reconciling our estimates with foregoing aerosol forcing lower bounds*

Our alternate 2000–10 5% aerosol forcing lower bound of $-1.7\,\mathrm{W\,m^{-2}}$ when including a factor of 5 increase in internal variability is similar to previous finding of a 5% lower bound of $-1.7\,\mathrm{W\,m^{-2}}$ by Aldrin et al. (2012) and Skeie et al. (2014), or a $-1.8\,\mathrm{W\,m^{-2}}$ lower bound by Skeie et al. (2018). These studies reference their lower bounds to a baseline around 2010, such that our 2000–10 is comparable. These foregoing studies used a different model setup than our analysis—for example, including hemispheric temperatures—making it difficult to directly compare the magnitudes of observational error allowed for in foregoing studies, but their specified uncertainties were substantial. Aldrin et al. (2012) specifies AR1 coefficients ranging from 0.5 to 0.9 and hemispheric standard deviations of 0.5°–0.6°C, as well as AR1 model error and a representation of internal variability based on the Southern Oscillation. The posterior distribution of F_{aer} obtained by Aldrin et al. (2012) is nearly identical to their F_{aer} prior distribution from AR5, implying that their observations provide little further constraint on F_{aer}. Skeie et al. (2014, 2018) use a similar approach but also include an autoregressive order three process, i.e., AR3, representing long-term internal variability estimated from long, coupled model control simulations.

A recent study by Smith et al. (2021) uses a similar framework to ours and obtained a 2019 maximum likelihood estimate of $-0.9\,\mathrm{W\,m^{-2}}$ and 5% aerosol forcing lower bound, referenced to 1750, of $-1.6\,\mathrm{W\,m^{-2}}$ (their Table S4), compared to our MLE of $-0.85\,\mathrm{W\,m^{-2}}$ and 5% lower bound of $-1.3\,\mathrm{W\,m^{-2}}$. They use the same aerosol forcing model as in Eq. (4) and a simple emulator for the temperature response to forcing. We hypothesize that three methodological differences account for the approximately 0.3 $\mathrm{W\,m^{-2}}$ discrepancy in the 5% lower bounds in their and our estimate: the use of ocean heat uptake observations, the treatment of observational error, and the treatment of internal climate variability. First, regarding ocean heat uptake, Smith et al. (2021) provide an estimate of aerosol forcing lower bound solely based on observed temperatures, with a MLE of $-0.80\,\mathrm{W\,m^{-2}}$ and a less negative lower bound of 2019 aerosol forcing of $-1.4\,\mathrm{W\,m^{-2}}$. Second, we use time-resolved observational error from HadCRUT4, whereas Smith et al. (2021) use a time-mean value of $0.01°\mathrm{C}^2$. Repeating our inversion with a fixed observational error of $0.01°\mathrm{C}^2$ lowers the 2019 5% lower bound by 0.1 $\mathrm{W\,m^{-2}}$ to $-1.4\,\mathrm{W\,m^{-2}}$. These two

factors already bring our lower bound in line with the Smith et al. (2021) lower bound. Differing treatment of internal variability can further explain differences in the lower bounds. Smith et al. (2021) draw realizations of internal variability from detrended CMIP6 preindustrial control global mean temperature anomalies. We first fix internal variability at the CMIP6 ensemble mean value and in section 4c show that increasing variability by a factor of 5 relative to this CMIP6 mean yields a 2019 5% lower bound of $-1.6\,\mathrm{W\,m^{-2}}$, more negative than the Smith et al. (2020) lower bound when they exclude constraints based on ocean heat uptake.

5. Conclusions

We find that an extension of the Stevens (2015) aerosol forcing model [Eq. (4)], also recently used by Smith et al. (2021), emulates global-annual historical variations in aerosol radiative forcing simulated by 10 state-of-the-art general circulation models. Simulations include contributions from rapid tropospheric adjustments, such as the cloud lifetime effect, and spatial variability in aerosol emissions. The root-mean-square residual between each simulation of global aerosol radiative forcing and a least squares fit using Eq. (4) has a mean value of $0.09\,\mathrm{W\,m^{-2}}$ when the global-annual sulfur dioxide, black carbon, and organic carbon aerosol precursor emissions used in each simulation are prescribed. Furthermore, using a simple aerosol forcing model that emulates the time history of aerosol forcing in CMIP6 models together with parametric uncertainty estimated using a Bayesian framework affords greater flexibility than previous approaches that rely upon rescaling aerosol forcing from GCM simulations (e.g., Padilla et al. 2011; Aldrin et al. 2012; Schwartz 2018).

Our analysis supports a lower bound on F_{aer} that is closer to -1.0 than to $-2.0\,\mathrm{W\,m^{-2}}$. Specifically, we obtain a 95% credible interval on F_{aer} between 2010 and 2019 from -1.3 to $-0.50\,\mathrm{W\,m^{-2}}$, despite ECS being broadly uncertain with a 95% credible interval from 1.8° to 6.1°C. Paleoclimate and process-based constraints elsewhere were argued to yield a tighter upper bound on climate sensitivity than do inferences from the historical record alone (e.g., Sherwood et al. 2020). To ensure that we accurately represent uncertainties in F_{aer}, it is important to jointly infer climate sensitivity and F_{aer}. In practice, however, we find that there is little correlation between climate sensitivity and F_{aer}, apparently because of distinct temporal variability. Similarly, uncertainty in effective volcanic forcing contributes little uncertainty to F_{aer} because their time histories weakly covary.

The origin of the strong F_{aer} constraint is traced to the rapid increase in precursor emissions between 1950 and 1980 and relatively stable emissions thereafter. Support for this emissions structure being critical comes from a restricted analysis using pre-1950 data that yields a 5th percentile 2010–19 lower bound on F_{aer} of $-1.8\,\mathrm{W\,m^{-2}}$ (Figs. 4a,b). In order for our model results to admit a 2010–19 fifth-percentile lower bound of $-1.6\,\mathrm{W\,m^{-2}}$, when employing the historical temperature record from 1850 to 2019, requires increasing the variance of internal variability by a factor of 5 relative to the mean variance inferred from 39 CMIP6 control simulations (Figs. 4c,d). Allowing for comparable variance in global-annual temperature observations further reduces this 2010–19 lower bound to $-1.8\,\mathrm{W\,m^{-2}}$ (Figs. 4e,f). It seems premature to rule out such large values of temperature variance, given differences between coupled and historical atmosphere-only model runs (e.g., Zhou et al. 2016; Parsons et al. 2020), differences between climate model and paleoclimate proxies (Laepple and Huybers 2014), and recent work on systematic SST errors (e.g., Chan and Huybers 2019).

There are several avenues open for improving empirical constraints on F_{aer}. One approach is to test our inference methodology within the context of general circulation model simulations. Temperatures simulated in response to various aerosol emissions and nonaerosol radiative forcing scenarios could be incorporated into our Bayesian methodology to infer F_{aer} and results compared against simulated values. Better quantifying uncertainty associated with low-frequency natural variability and observational error, presently not well characterized in GCMs and SST datasets, would also improve our ability to rule out extreme values of F_{aer}. Ocean heat uptake data cannot, at present, be used as an independent constraint on F_{aer} because radiative damping and net radiative forcing compensate in explaining observed variability in the historical top-of-atmosphere radiative imbalance. Improving the characterization of radiative feedback terms would allow for stronger constraints on F_{aer} from ocean heat uptake data. Further development of a Bayesian methodology to improve estimates of radiative feedbacks, low-frequency internal variability, and observational error should help narrow the stubborn bounds on aerosol radiative forcing.

Acknowledgments. HadCRUT4 historical global-mean surface temperatures from Cowtan and Way (2014), version 2.0, are available at https://www-users.york.ac.uk/~kdc3/papers/coverage2013/series.html. Transient effective aerosol forcing data from the RFMIP and AerChemMIP experiments in CMIP6 were accessed at https://esgf-node.llnl.gov/search/cmip6/. Sulfur dioxide, black carbon, and organic carbon data are from O'Rourke et al. (2021), accessed at https://doi.org/10.5194/gmd-11-369-2018-supplement. Temperature time series from the 40-member Community Earth System Model version 1 Large Ensemble (CESM1 LE) are from Kay et al. (2015), and global-mean temperature time series from CMIP5 control runs were downloaded from https://esgf-node.llnl.gov/projects/cmip5/. We thank Chris Smith for providing updated effective nonaerosol radiative forcing data, Karen McKinnon for assistance in obtaining the CESM1 LE data, Ted Amdur for processing global-mean temperature time series from CMIP5 control runs, and Luke Parsons for processing global-mean temperature time series from CMIP6 control runs. This work was supported by the Harvard University Climate Change Solutions Fund. C.P. acknowledges support from NOAA Grant UWSC12184 and a JISAO postdoctoral fellowship from the University of Washington.

REFERENCES

Albrecht, B. A., 1989: Aerosols, cloud microphysics, and fractional cloudiness. *Science*, **245**, 1227–1230, https://doi.org/10.1126/science.245.4923.1227.

Aldrin, M., M. Holden, P. Guttorp, R. Skeie, G. Myhre, and T. K. Berntsen, 2012: Bayesian estimation of climate sensitivity based

on a simple climate model fitted to observations of hemispheric temperatures and global ocean heat content. *Environmetrics*, **23**, 253–271, https://doi.org/10.1002/env.2140.

Anderson, T. L., R. J. Charlson, S. E. Schwartz, R. Knutti, O. Boucher, H. Rodhe, and J. Heintzenberg, 2003: Climate forcing by aerosols—A hazy picture. *Science*, **300**, 1103–1104, https://doi.org/10.1126/science.1084777.

Andreae, M. O., C. D. Jones, and P. M. Cox, 2005: Strong present-day aerosol cooling implies a hot future. *Nature*, **435**, 1187–1190, https://doi.org/10.1038/nature03671.

Andrews, T., J. M. Gregory, and M. J. Webb, 2015: The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *J. Climate*, **28**, 1630–1648, https://doi.org/10.1175/JCLI-D-14-00545.1.

——, and Coauthors, 2018: Accounting for changing temperature patterns increases historical estimates of climate sensitivity. *Geophys. Res. Lett.*, **45**, 8490–8499, https://doi.org/10.1029/2018GL078887.

Andronova, N., and M. Schlesinger, 2001: Objective estimation of the probability density function for climate sensitivity. *J. Geophys. Res.*, **106**, 22 605–22 611, https://doi.org/10.1029/2000JD000259.

Annan, J., 2015: Recent developments in Bayesian estimation of climate sensitivity. *Curr. Climate Change Rep.*, **1**, 263–267, https://doi.org/10.1007/s40641-015-0023-5.

Bellouin, N., J. Rae, A. Jones, C. Johnson, J. Haywood, and O. Boucher, 2011: Aerosol forcing in the Climate Model Intercomparison Project (CMIP5) simulations by HadGEM2-ES and the role of ammonium nitrate. *J. Geophys. Res.*, **116**, D20206, https://doi.org/10.1029/2011JD016074.

——, G. W. Mann, M. T. Woodhouse, C. Johnson, K. S. Carslaw, and M. Dalvi, 2013: Impact of the modal aerosol scheme GLOMAP-mode on aerosol forcing in the Hadley Centre global environmental model. *Atmos. Chem. Phys.*, **13**, 3027–3044, https://doi.org/10.5194/acp-13-3027-2013.

——, and Coauthors, 2020: Bounding global aerosol radiative forcing of climate change. *Rev. Geophys.*, **58**, e2019RG000660, https://doi.org/10.1029/2019RG000660.

Bodman, R. W., and R. N. Jones, 2016: Bayesian estimation of climate sensitivity using observationally constrained simple climate models. *Wiley Interdiscip. Rev.: Climate Change*, **7**, 461–473, https://doi.org/10.1002/wcc.397.

Booth, B. B., G. R. Harris, A. Jones, L. Wilcox, M. Hawcroft, and K. S. Carslaw, 2018: Comments on ''Rethinking the lower bound on aerosol radiative forcing.'' *J. Climate*, **31**, 9407–9412, https://doi.org/10.1175/JCLI-D-17-0369.1.

Boucher, O., and M. Pham, 2002: History of sulfate aerosol radiative forcings. *Geophys. Res. Lett.*, **29**, 221–224, https://doi.org/10.1029/2001GL014048.

——, and Coauthors, 2013: Clouds and aerosols. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, https://doi.org/10.1017/CBO9781107415324.018..

Caldeira, K., and N. P. Myhrvold, 2013: Projections of the pace of warming following an abrupt increase in atmospheric carbon dioxide concentration. *Environ. Res. Lett.*, **8**, 034039, https://doi.org/10.1088/1748-9326/8/3/034039.

Carslaw, K. S., O. Boucher, D. Spracklen, G. Mann, J. Rae, S. Woodward, and M. Kulmala, 2010: A review of natural aerosol interactions and feedbacks within the Earth system. *Atmos. Chem. Phys.*, **10**, 1701–1737, https://doi.org/10.5194/acp-10-1701-2010.

——, and Coauthors, 2013: Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, **503**, 67–71, https://doi.org/10.1038/nature12674.

——, L. A. Lee, L. A. Regayre, and J. S. Johnson, 2018: Climate models are uncertain, but we can do something about it. *Eos, Trans. Amer. Geophys. Union*, **99**, https://doi.org/10.1029/2018EO093757.

Ceppi, P., and J. M. Gregory, 2019: A refined model for the Earth's global energy balance. *Climate Dyn.*, **53**, 4781–4797, https://doi.org/10.1007/s00382-019-04825-x.

Chan, D., and P. Huybers, 2019: Systematic differences in bucket sea surface temperature measurements among nations identified using a linear-mixed-effect method. *J. Climate*, **32**, 2569–2589, https://doi.org/10.1175/JCLI-D-18-0562.1.

——, and ——, 2021: Correcting observational biases in sea surface temperature observations removes anomalous warmth during World War II. *J. Climate*, **34**, 4585–4602, https://doi.org/10.1175/JCLI-D-20-0907.1.

——, E. C. Kent, D. I. Berry, and P. Huybers, 2019: Correcting datasets leads to more homogeneous early-twentieth-century sea surface warming. *Nature*, **571**, 393–397, https://doi.org/10.1038/s41586-019-1349-2.

Charlson, R., S. Schwartz, J. Hales, R. Cess, J. J. A. Coakley, J. Hansen, and D. Hoffman, 1992: Climate forcing by anthropogenic aerosols. *Science*, **255**, 423–430, https://doi.org/10.1126/science.255.5043.423.

Cheng, L., K. E. Trenberth, J. Fasullo, T. Boyer, J. Abraham, and J. Zhu, 2017: Improved estimates of ocean heat content from 1960 to 2015. *Sci. Adv.*, **3**, e1601545, https://doi.org/10.1126/sciadv.1601545.

Chylek, P., and J. Wong, 1995: Effects of absorbing aerosols on the global radiation budget. *Geophys. Res. Lett.*, **22**, 929–931, https://doi.org/10.1029/95GL00800.

——, T. J. Vogelsang, J. D. Klett, N. Hengartner, D. Higdon, G. Lesins, and M. K. Dubey, 2016: Indirect aerosol effect increases CMIP5 models' projected Arctic warming. *J. Climate*, **29**, 1417–1428, https://doi.org/10.1175/JCLI-D-15-0362.1.

——, C. Folland, J. D. Klett, and M. K. Dubey, 2020: CMIP5 climate models overestimate cooling by volcanic aerosols. *Geophys. Res. Lett.*, **47**, e2020GL087047, https://doi.org/10.1029/2020GL087047.

Coats, S., and K. Karnauskas, 2017: Are simulated and observed twentieth century tropical Pacific sea surface temperature trends significant relative to internal variability? *Geophys. Res. Lett.*, **44**, 9928–9937, https://doi.org/10.1002/2017GL074622.

Cowtan, K., and R. Way, 2014: Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Quart. J. Roy. Meteor. Soc.*, **140**, 1935–1944, https://doi.org/10.1002/qj.2297.

——, and Coauthors, 2015: Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. *Geophys. Res. Lett.*, **42**, 6526–6534, https://doi.org/10.1002/2015GL064888.

——, R. Rohde, and Z. Hausfather, 2018: Evaluating biases in sea surface temperature records using coastal weather stations. *Quart. J. Roy. Meteor. Soc.*, **144**, 670–681, https://doi.org/10.1002/qj.3235.

Cummins, D. P., D. B. Stephenson, and P. A. Stott, 2020: Optimal estimation of stochastic energy balance model parameters. *J. Climate*, **33**, 7909–7926, https://doi.org/10.1175/JCLI-D-19-0589.1.

Davis, L. L., D. W. Thompson, J. J. Kennedy, and E. C. Kent, 2019: The importance of unresolved biases in twentieth-century sea surface temperature observations. *Bull. Amer. Meteor. Soc.*, **100**, 621–629, https://doi.org/10.1175/BAMS-D-18-0104.1.

Dong, Y., C. Proistosescu, K. C. Armour, and D. S. Battisti, 2019: Attributing historical and future evolution of radiative feedbacks

to regional warming patterns using a Green's function approach: The preeminence of the western Pacific. *J. Climate*, **32**, 5471–5491, https://doi.org/10.1175/JCLI-D-18-0843.1.

Dudok de Wit, T., G. Kopp, C. Froehlich, and M. Schoell, 2017: Methodology to create a new total solar irradiance record: Making a composite out of multiple data records. *Geophys. Res. Lett.*, **44**, 1196–1203, https://doi.org/10.1002/2016GL071866.

Etminan, M., G. Myhre, E. J. Highwood, and K. P. Shine, 2016: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing. *Geophys. Res. Lett.*, **43**, 12 614–12 623, https://doi.org/10.1002/2016GL071930.

Feingold, G., A. McComiskey, T. Yamaguchi, J. S. Johnson, K. S. Carslaw, and K. S. Schmidt, 2016: New approaches to quantifying aerosol influence on the cloud radiative effect. *Proc. Natl. Acad. Sci. USA*, **113**, 5812–5819, https://doi.org/10.1073/pnas.1514035112.

Feldstein, S. B., 2000: The timescale, power spectra, and climate noise properties of teleconnection patterns. *J. Climate*, **13**, 4430–4440, https://doi.org/10.1175/1520-0442(2000)013<4430:TTPSAC>2.0.CO;2.

Fiedler, S., and Coauthors, 2019: Anthropogenic aerosol forcing—Insights from multiple estimates from aerosol-climate models with reduced complexity. *Atmos. Chem. Phys.*, **19**, 6821–6841, https://doi.org/10.5194/acp-19-6821-2019.

Folland, C. K., O. Boucher, A. Colman, and D. E. Parker, 2018: Causes of irregularities in trends of global mean surface temperature since the late 19th century. *Sci. Adv.*, **4**, eaao5297, https://doi.org/10.1126/sciadv.aao5297.

Forest, C. E., 2018: Inferred net aerosol forcing based on historical climate changes: A review. *Curr. Climate Change Rep.*, **4**, 11–22, https://doi.org/10.1007/s40641-018-0085-2.

——, P. H. Stone, A. P. Sokolov, M. R. Allen, and M. D. Webster, 2002: Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science*, **295**, 113–117, https://doi.org/10.1126/science.1064419.

——, ——, and ——, 2006: Estimated PDFs of climate system properties including natural and anthropogenic forcings. *Geophys. Res. Lett.*, **33**, L01705, https://doi.org/10.1029/2005GL023977.

——, ——, and ——, 2008: Constraining climate model parameters from observed 20th century changes. *Tellus*, **60A**, 911–920, https://doi.org/10.1111/j.1600-0870.2008.00346.x.

Forster, P. M., 2016: Inference of climate sensitivity from analysis of Earth's energy budget. *Annu. Rev. Earth Planet. Sci.*, **44**, 85–106, https://doi.org/10.1146/annurev-earth-060614-105156.

Frame, D., B. Booth, J. Kettleborough, D. Stainforth, J. Gregory, M. Collins, and M. Allen, 2005: Constraining climate forecasts: The role of prior assumptions. *Geophys. Res. Lett.*, **32**, L09702, https://doi.org/10.1029/2004GL022241.

Frankignoul, C., and K. Hasselmann, 1977: Stochastic climate models, Part II: Application to sea-surface temperature variability and thermocline variability. *Tellus*, **29**, 289–305, https://doi.org/10.3402/tellusa.v29i4.11362.

Gebbie, G., and P. Huybers, 2019: The Little Ice Age and 20th century deep Pacific cooling. *Science*, **363**, 70–74, https://doi.org/10.1126/science.aar8413.

Geoffroy, O., D. Saint-Martin, D. J. L. Olivié, A. Voldoire, G. Bellon, and S. Tyteca, 2013: Transient climate response in a two-layer energy-balance model. Part I: Analytical solution and parameter calibration using CMIP5 AOGCM experiments. *J. Climate*, **26**, 1841–1857, https://doi.org/10.1175/JCLI-D-12-00195.1.

Gettelman, A., 2015: Putting the clouds back in aerosol–cloud interactions. *Atmos. Chem. Phys.*, **15**, 12 397–12 411, https://doi.org/10.5194/acp-15-12397-2015.

——, and S. C. Sherwood, 2016: Processes responsible for cloud feedback. *Curr. Climate Change Rep.*, **2**, 179–189, https://doi.org/10.1007/s40641-016-0052-8.

Golaz, J.-C., and Coauthors, 2019: The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *J. Adv. Model. Earth Syst.*, **11**, 2089–2129, https://doi.org/10.1029/2018MS001603.

Gregory, J. M., T. Andrews, P. Good, T. Mauritsen, and P. Forster, 2016: Small global-mean cooling due to volcanic radiative forcing. *Climate Dyn.*, **47**, 3979–3991, https://doi.org/10.1007/s00382-016-3055-1.

Hansen, J., and Coauthors, 2005: Earth's energy imbalance: Confirmation and implications. *Science*, **308**, 1431–1435, https://doi.org/10.1126/science.1110252.

Hasselmann, K., R. Sausen, E. Maier-Reimer, and R. Voss, 1993: On the cold start problem in transient simulations with coupled atmosphere–ocean models. *Climate Dyn.*, **9**, 53–61, https://doi.org/10.1007/BF00210008.

Hastings, W., 1970: Monte Carlo sampling methods using Markov chains and their application. *Biometrika*, **57**, 97–109, https://doi.org/10.1093/biomet/57.1.97.

Hauglustaine, D. A., Y. Balkanski, and M. Schulz, 2014: A global model simulation of present and future nitrate aerosols and their direct radiative forcing of climate. *Atmos. Chem. Phys.*, **14**, 11 031–11 063, https://doi.org/10.5194/acp-14-11031-2014.

Haustein, K., and Coauthors, 2019: A limited role for unforced internal variability in twentieth-century warming. *J. Climate*, **32**, 4893–4917, https://doi.org/10.1175/JCLI-D-18-0555.1.

Hegerl, G. C., T. J. Crowley, W. T. Hyde, and D. J. Frame, 2006: Climate sensitivity constrained by temperature reconstructions over the past seven centuries. *Nature*, **440**, 1029–1032, https://doi.org/10.1038/nature04679.

——, S. Brönnimann, A. Schurer, and T. Cowan, 2018: The early 20th century warming: Anomalies, causes, and consequences. *Wiley Interdiscip. Rev.: Climate Change*, **9**, e522, https://doi.org/10.1002/wcc.522.

Held, I. M., 2005: The gap between simulation and understanding in climate modeling. *Bull. Amer. Meteor. Soc.*, **86**, 1609–1614, https://doi.org/10.1175/BAMS-86-11-1609.

——, M. Winton, K. Takahashi, T. Delworth, F. Zeng, and G. K. Vallis, 2010: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J. Climate*, **23**, 2418–2427, https://doi.org/10.1175/2009JCLI3466.1.

Hoesly, R. M., and Coauthors, 2018: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). *Geosci. Model Dev.*, **11**, 369–408, https://doi.org/10.5194/gmd-11-369-2018.

Hoskins, B., 1983: Dynamical processes in the atmosphere and the use of models. *Quart. J. Roy. Meteor. Soc.*, **109** (459), 1–21, https://doi.org/10.1002/qj.49710945902.

Johansson, D. J., B. C. O. Neill, C. Tebaldi, and O. Häggström, 2015: Equilibrium climate sensitivity in light of observations over the warming hiatus. *Nat. Climate Change*, **5**, 449–453, https://doi.org/10.1038/nclimate2573.

Karl, T. R., and Coauthors, 2015: Possible artifacts of data biases in the recent global surface warming hiatus. *Science*, **348**, 1469–1472, https://doi.org/10.1126/science.aaa5632.

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal

climate variability. *Bull. Amer. Meteor. Soc.*, **96**, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1.

Kiehl, J. T., 2007: Twentieth century climate model response and climate sensitivity. *Geophys. Res. Lett.*, **34**, L22710, https://doi.org/10.1029/2007GL031383.

Knutti, R., T. F. Stocker, F. Joos, and G.-K. Plattner, 2002: Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature*, **416**, 719–723, https://doi.org/10.1038/416719a.

——, M. A. Rugenstein, and G. C. Hegerl, 2017: Beyond equilibrium climate sensitivity. *Nat. Geosci.*, **10**, 727–736, https://doi.org/10.1038/ngeo3017.

Kretzschmar, J., M. Salzmann, J. Mülmenstädt, O. Boucher, and J. Quaas, 2017: Comment on rethinking the lower bound on aerosol radiative forcing. *J. Climate*, **30**, 6579–6584, https://doi.org/10.1175/JCLI-D-16-0668.1.

Laepple, T., and P. Huybers, 2014: Ocean surface temperature variability: Large model-data differences at decadal and longer periods. *Proc. Natl. Acad. Sci. USA*, **111**, 16 682–16 687, https://doi.org/10.1073/pnas.1412077111.

Leach, N. J., and Coauthors, 2021: FaIRv2.0.0: A generalized impulse response model for climate uncertainty and future scenario exploration. *Geosci. Model Dev.*, **14**, 3007–3036, https://doi.org/10.5194/gmd-14-3007-2021.

Lee, L., C. Reddington, and K. Carslaw, 2016: On the relationship between aerosol model uncertainty and radiative forcing uncertainty. *Proc. Natl. Acad. Sci. USA*, **113**, 5820–5827, https://doi.org/10.1073/pnas.1507050113.

Lehner, F., A. P. Schurer, G. C. Hegerl, C. Deser, and T. L. Frölicher, 2016: The importance of ENSO phase during volcanic eruptions for detection and attribution. *Geophys. Res. Lett.*, **43**, 2851–2858, https://doi.org/10.1002/2016GL067935.

Libardoni, A. G., and C. E. Forest, 2011: Sensitivity of distributions of climate system properties to the surface temperature dataset. *Geophys. Res. Lett.*, **38**, L22705, https://doi.org/10.1029/2011GL049431.

——, and ——, 2013: Correction to sensitivity of distributions of climate system properties to the surface temperature data set. *Geophys. Res. Lett.*, **40**, 2309–2311, https://doi.org/10.1002/grl.50480.

Malavelle, F. F., and Coauthors, 2017: Strong constraints on aerosol–cloud interactions from volcanic eruptions. *Nature*, **546**, 485–491, https://doi.org/10.1038/nature22974.

Marvel, K., G. A. Schmidt, R. L. Miller, and L. S. Nazarenko, 2016: Implications for climate sensitivity from the response to individual forcings. *Nat. Climate Change*, **6**, 386–389, https://doi.org/10.1038/nclimate2888.

McCoy, D. T., S. M. Burrows, R. Wood, D. P. Grosvenor, S. M. Elliott, P. L. Ma, P. J. Rasch, and D. L. Hartmann, 2015: Natural aerosols explain seasonal and spatial patterns of Southern Ocean cloud albedo. *Sci. Adv.*, **1**, e1500157, https://doi.org/10.1126/sciadv.1500157.

Meinshausen, M., and Coauthors, 2011: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213–241, https://doi.org/10.1007/s10584-011-0156-z.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, 1953: Equation of state calculations by fast computing machines. *J. Chem. Phys.*, **21**, 1087–1092, https://doi.org/10.1063/1.1699114.

Morice, C. P., J. J. Kennedy, N. A. Rayner, and P. D. Jones, 2012: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *J. Geophys. Res.*, **117**, D08101, https://doi.org/10.1029/2011JD017187.

Murphy, D., S. Solomon, R. Portmann, K. Rosenlof, P. Forster, and T. Wong, 2009: An observationally based energy balance for the Earth since 1950. *J. Geophys. Res.*, **114**, D17107, https://doi.org/10.1029/2009JD012105.

Myhre, G., and Coauthors, 2013: Anthropogenic and natural radiative forcing. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 659–740, https://doi.org/10.1017/CBO9781107415324.018.

O'Rourke, P. R., and Coauthors, 2021: CEDS v_2021_04_21 release emission data. Zenodo, accessed 1 June 2021, https://doi.org/10.5281/zenodo.4741285.

Otto, A., and Coauthors, 2013: Energy budget constraints on climate response. *Nat. Geosci.*, **6**, 415–416, https://doi.org/10.1038/ngeo1836.

Padilla, L. E., G. K. Vallis, and C. W. Rowley, 2011: Probabilistic estimates of transient climate sensitivity subject to uncertainty in forcing and natural variability. *J. Climate*, **24**, 5521–5537, https://doi.org/10.1175/2011JCLI3989.1.

Parsons, L. A., M. K. Brennan, R. C. Wills, and C. Proistosescu, 2020: Magnitudes and spatial patterns of interdecadal temperature variability in CMIP6. *Geophys. Res. Lett.*, **47**, e2019GL086588, https://doi.org/10.1029/2019GL086588.

Penner, J., and Coauthors, 1994: Quantifying and minimizing uncertainty of climate forcing by anthropogenic aerosols. *Bull. Amer. Meteor. Soc.*, **75**, 375–400, https://doi.org/10.1175/1520-0477(1994)075<0375:QAMUOC>2.0.CO;2.

Persad, G. G., and K. Caldeira, 2019: Divergent global-scale temperature effects from identical aerosols emitted in different regions. *Nat. Commun.*, **9**, 3289, https://doi.org/10.1038/s41467-018-05838-6.

Pincus, R., P. M. Forster, and B. Stevens, 2016: The Radiative Forcing Model Intercomparison Project (RFMIP): Experimental protocol for CMIP6. *Geosci. Model Dev.*, **9**, 3447–3460, https://doi.org/10.5194/gmd-9-3447-2016.

Polvani, L., A. Clement, B. Medeiros, J. Benedict, and I. Simpson, 2017: When less is more: Opening the door to simpler climate models. *Eos, Trans. Amer. Geophys. Union*, **99**, 15–16, https://doi.org/10.1029/2017EO079417.

Proistosescu, C., and P. J. Huybers, 2017: Slow climate mode reconciles historical and model-based estimates of climate sensitivity. *Sci. Adv.*, **3**, e1602821, https://doi.org/10.1126/sciadv.1602821.

Rodhe, H., R. Charlson, and T. Anderson, 2000: Avoiding circular logic in climate modeling. *Climatic Change*, **44**, 419–422, https://doi.org/10.1023/A:1005536902789.

Samset, B. H., G. Myhre, and M. Schulz, 2014: Upward adjustment needed for aerosol radiative forcing uncertainty. *Nat. Climate Change*, **4**, 230–232, https://doi.org/10.1038/nclimate2170.

Schneider, T., and A. Neumaier, 2001: Algorithm 808: ARfit—A Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans. Math. Software*, **27**, 58–65, https://doi.org/10.1145/382043.382316.

Schwartz, S., 2018: Unrealized global temperature increase: Implications of current uncertainties. *J. Geophys. Res. Atmos.*, **123**, 3462–3482, https://doi.org/10.1002/2017JD028121.

Seinfeld, J. H., and Coauthors, 2016: Improving our fundamental understanding of the role of aerosol–cloud interactions in the climate system. *Proc. Natl. Acad. Sci. USA*, **113**, 5781–5790, https://doi.org/10.1073/pnas.1514043113.

Sherwood, S., and Coauthors, 2020: An assessment of Earth's climate sensitivity using multiple lines of evidence. *Rev. Geophys.*, **58**, e2019RG000678, https://doi.org/10.1029/2019RG000678.

Skeie, R., T. Berntsen, M. Aldrin, M. Holden, and G. Myhre, 2014: A lower and more constrained estimate of climate sensitivity using updated observations and detailed radiative forcing time series. *Earth Syst. Dyn.*, **5**, 139–175, https://doi.org/10.5194/esd-5-139-2014.

——, ——, ——, ——, and ——, 2018: Climate sensitivity estimates—Sensitivity to radiative forcing time series and observational data. *Earth Syst. Dyn.*, **9**, 879–894, https://doi.org/10.5194/esd-9-879-2018.

Smith, C. J., and Coauthors, 2020: Effective radiative forcing and adjustments in CMIP6 models. *Atmos. Chem. Phys.*, **20**, 9591–9618, https://doi.org/10.5194/acp-20-9591-2020.

——, and Coauthors, 2021: Energy budget constraints on the time history of aerosol forcing and climate sensitivity. *J. Geophys. Res. Atmos.*, **126**, e2020JD033622, https://doi.org/10.1029/2020JD033622.

Smith, S., J. van Aardenne, Z. Klimont, R. Andres, A. Volke, and S. Delgado Arias, 2011: Anthropogenic sulfur dioxide emissions: 1850–2005. *Atmos. Chem. Phys.*, **11**, 1101–1116, https://doi.org/10.5194/acp-11-1101-2011.

Sokolov, A. P., C. E. Forest, and P. H. Stone, 2010: Sensitivity of climate change projections to uncertainties in the estimates of observed changes in deep-ocean heat content. *Climate Dyn.*, **34**, 735–745, https://doi.org/10.1007/s00382-009-0556-1.

Stevens, B., 2015: Rethinking the lower bound on aerosol radiative forcing. *J. Climate*, **28**, 4794–4819, https://doi.org/10.1175/JCLI-D-14-00656.1.

——, 2018: Reply to ''Comments on 'Rethinking the lower bound on aerosol radiative forcing.''' *J. Climate*, **31**, 9413–9416, https://doi.org/10.1175/JCLI-D-18-0185.1.

——, and G. Feingold, 2009: Untangling aerosol effects on clouds and precipitation in a buffered system. *Nature*, **461**, 607–613, https://doi.org/10.1038/nature08281.

——, S. C. Sherwood, S. Bony, and M. J. Webb, 2016: Prospects for narrowing bounds on Earth's equilibrium climate sensitivity. *Earth's Future*, **4**, 512–522, https://doi.org/10.1002/2016EF000376.

Su, W., N. Loeb, G. Schuster, M. Chin, and F. Rose, 2013: Global all-sky shortwave direct radiative forcing of anthropogenic aerosols from combined satellite observations and GOCART simulations. *J. Geophys. Res. Atmos.*, **118**, 655–669, https://doi.org/10.1029/2012JD018294.

Taylor, K. E., V. Balaji, S. Hankin, M. Juckes, B. Lawrence, and S. Pascoe, 2011: CMIP5 data reference syntax (DRS) and controlled vocabularies. PCMDI, 14 pp., https://pcmdi.llnl.gov/mips/cmip5/docs/cmip5_data_reference_syntax_v1-02_marked.pdf.

Tsutsui, J., 2017: Quantification of temperature response to CO_2 forcing in atmosphere–ocean general circulation models. *Climatic Change*, **140**, 287–305, https://doi.org/10.1007/s10584-016-1832-9.

Twomey, S., 1977: Influence of pollution on the short-wave albedo of clouds. *J. Atmos. Sci.*, **34**, 1149–1152, https://doi.org/10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2.

von Schuckmann, K., and Coauthors, 2020: Heat stored in the Earth system: Where does the energy go? *Earth Syst. Sci. Data*, **12**, 2013–2041, https://doi.org/10.5194/essd-12-2013-2020.

Vyushin, D., and P. Kushner, 2009: Power-law and long-memory characteristics of the atmospheric general circulation. *J. Climate*, **22**, 2890–2904, https://doi.org/10.1175/2008JCLI2528.1.

Wigley, T., and S. Raper, 1992: Implications for climate and sea level of revised IPCC emissions scenarios. *Nature*, **357**, 293–300, https://doi.org/10.1038/357293a0.

Zanna, L., S. Khatiwala, J. M. Gregory, J. Ison, and P. Heimbach, 2019: Global reconstruction of historical ocean heat storage and transport. *Proc. Natl. Acad. Sci. USA*, **116**, 1126–1131, https://doi.org/10.1073/pnas.1808838115.

Zhang, H.-M., and Coauthors, 2019: Release of NOAA global surface temperature (noaaglobaltemp) dataset version 5. *2019 Fall Meeting*, San Francisco, CA, Amer. Geophys. Union, Abstract GC51O-1097, https://agu.confex.com/agu/fm19/meetingapp.cgi/Paper/519861.

Zhang, S., and Coauthors, 2016: On the characteristics of aerosol indirect effect based on dynamic regimes in global climate models. *Atmos. Chem. Phys.*, **16**, 2765–2783, https://doi.org/10.5194/acp-16-2765-2016.

Zhao, M., and Coauthors, 2018: The GFDL global atmosphere and land model AM4.0/LM4.0: 1. Simulation characteristics with prescribed SSTs. *J. Adv. Model. Earth Syst.*, **10**, 691–734, https://doi.org/10.1002/2017MS001208.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2016: Impact of decadal cloud variations on the Earth's energy budget. *Nat. Geosci.*, **9**, 871–874, https://doi.org/10.1038/ngeo2828.