Cloud, Aerosol, and Boundary Layer Structure across the Northeast Pacific Stratocumulus–Cumulus Transition as Observed during CSET

CHRISTOPHER S. BRETHERTON, ISABEL L. MCCOY, JOHANNES MOHRMANN, AND ROBERT WOOD
Department of Atmospheric Sciences, University of Washington, Seattle, Washington

VIRENDRA GHATE
Argonne National Laboratory, Argonne, Illinois

ANDREW GETTELMAN AND CHARLES G. BARDEEN
National Center for Atmospheric Research, Boulder, Colorado

BRUCE A. ALBRECHT AND PAQUITA ZUIDEMA
Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida

(Manuscript received 29 November 2018, in final form 4 April 2019)

ABSTRACT

During the Cloud System Evolution in the Trades (CSET) field study, 14 research flights of the National Science Foundation G-V sampled the stratocumulus–cumulus transition between Northern California and Hawaii and its synoptic variability. The G-V made vertically resolved measurements of turbulence, cloud microphysics, aerosol characteristics, and trace gases. It also carried dropsondes and a vertically pointing W-band radar and lidar. This paper summarizes these observations with the goals of fostering novel comparisons with theory, models and reanalyses, and satellite-derived products. A longitude–height binning and compositing strategy mitigates limitations of sparse sampling and spatiotemporal variability. Typically, a 1-km-deep decoupled stratocumulus-capped boundary layer near California evolved into 2-km-deep precipitating cumulus clusters surrounded by patches of thin stratus that dissipated toward Hawaii. Low cloud cover was correlated with estimated inversion strength more than with cloud droplet number, even though the thickest clouds were generally precipitating and ultraclean layers indicative of aerosol–cloud–precipitation interaction were common west of 140°W. Accumulation-mode aerosol concentration correlated well with collocated cloud droplet number concentration and was typically largest near the surface. Aitken mode aerosol concentration was typically larger in the free troposphere. Wildfire smoke produced spikes of aerosol and trace gases on some flights. CSET data are compared with space–time collocated output from MERRA-2 reanalysis and from the CAM6 climate model run with winds and temperature nudged toward this reanalysis. The reanalysis compares better with the observed relative humidity than does nudged CAM6. Both vertically diffuse the stratocumulus cloud layer versus observations. MERRA-2 slightly underestimates in situ carbon monoxide measurements and underestimates ozone depletion within the boundary layer.

1. Introduction

The climatological stratocumulus to cumulus (Sc–Cu) transition over the eastern subtropical oceans has been a long-standing test of our physical understanding and modeling skill. Through a combination of field and satellite observations and detailed process modeling such as large-eddy simulation (LES), the Sc–Cu transition has been explained as due to the deepening and warming of a cloud-topped marine boundary layer under a strong inversion as it advects toward warmer sea surface temperature (SST) in the subtropical trade winds (Krueger et al. 1995a,b; Wyant et al. 1997; Bretherton and Wyant 1997). Strong mean subsidence over cool coastal waters maintains a shallow, fairly well-mixed stratocumulus boundary layer capped by a strong inversion. The boundary layer deepens as it advects away from the coast toward higher SST and weaker mean subsidence.
driving decoupling between surface-driven and cloud-top driven cumulus clouds. Farther downstream, warming of the boundary layer supports deeper and more vigorous cumuli that mix down enough dry free-tropospheric air and may precipitate out enough water to bring down the inversion-base relative humidity and evaporate the capping stratocumulus, completing the transition.

Countless field studies have sampled stratocumulus-topped boundary layers within 1000 km of the California coast. There are far fewer detailed in situ observations farther offshore where the stratocumulus gradually breaks up. The 1992 Atlantic Stratocumulus Transition Experiment (ASTEX), based in the northeast Atlantic, documented the persistently decoupled nature of the cumulus-under-stratocumulus boundary layers in this transitional regime and was among the first field studies to use ground-based millimeter-wavelength radars for studying boundary layer cloud and precipitation processes and their mesoscale organization (Albrecht et al. 1995). ASTEX included two 2-day intensive observing periods in which the Lagrangian evolution of cloudy boundary layers was tracked by aircraft (Bretherton and Pincus 1995; Bretherton et al. 1995). The first of these cases, in which partial breakup of Sc into Cu was observed, was subsequently used for model intercomparisons (Bretherton et al. 1999; van der Dussen et al. 2013).

The GCSS Pacific Cross-Section Intercomparison (GPCI) (Teixeira et al. 2011) created summertime climatologies of a suite of remote sensing observations and reanalysis outputs along a vertical section stretching from the California coast past Hawaii to the central Pacific ITCZ. These were compared with a large group of global numerical weather prediction and climate models to identify biases in the simulated vertical atmospheric structure and cloud distribution, with particular attention to the cloud-topped boundary layer and the Sc–Cu transition.

The GPCI motivated the MAGIC (Marine ARM GPCI Investigation of Clouds) project in 2012–13. MAGIC deployed the Department of Energy Atmospheric Radiation Measurement Mobile Facility (AMF2) on a container ship that traversed between Long Beach, California, and Honolulu, Hawaii. The AMF2 included rawinsondes, vertically pointing cloud radars, a multipulse lidar, aerosol, cloud, radiation, and meteorological measurements. MAGIC sampled through the Sc–Cu transition on numerous cruises across the annual cycle, as documented by Zhou et al. (2016) and Kalmus et al. (2014). MAGIC data have been comprehensively compared with large-eddy simulations of many of the cruises (McGibbon and Bretherton 2017) as well as satellite retrievals (Painemal et al. 2015), global weather forecast model output (Ahlgrimm et al. 2018), and reanalysis (Kalmus et al. 2015).

The focus of this paper is the Cloud System Evolution in the Trades (CSET) campaign in July–August 2015 (Albrecht et al. 2019; hereafter A19). CSET aimed to document and understand cloud processes in the summertime Sc–Cu transition using airborne measurements between California and Hawaii that complement past surface-based and satellite observations.

CSET goals, strategy, and observations were described in detail by A19, as was the instrumentation aboard the National Science Foundation Gulfstream-V (G-V) research aircraft. For this study, we will use the seven flight pairs RF02-03, RF04-05, ..., RF14-15, each of which started with a westbound flight from Sacramento, California, to Kona, Hawaii, followed 2 days later by an eastbound flight. Figure 1 shows these flight tracks, with solid blue lines for westbound flights and dashed blue lines for eastbound flights, superposed on a geostationary satellite-derived estimate of low cloud fraction (detailed in section 2a) for the CSET period of 1 July–15 August 2015. The thickened regions of the tracks indicate low-level sampling, while the thin regions indicate survey legs during which the G-V was flying at approximately 6-km altitude. Figure 1 also shows the nearby GPCI and MAGIC cross sections in red. The solid yellow lines indicate a coordinate system discussed in section 2 that we will use for CSET data analysis.

CSET sampling provided an unprecedentedly rich and extensive suite of observations of the cloud-topped boundary layer between California and Hawaii. The CSET flight plans were based on a Lagrangian low-level sampling strategy for the eastbound flights, as described by A19 and scientifically exploited in the analysis of Mohrmann et al. (2019, manuscript submitted to Mon. Wea. Rev.). Because of this strategy, the low-level sampling on the eastbound flights was typically substantially farther downwind (south and west) than on the westbound flights, as can be seen on Fig. 1. The CSET flights also provided a comprehensive quasi-Eulerian sampling of the Sc–Cu transition.

In this paper, we use this dataset to describe the boundary layer, clouds, and aerosols during each flight, with three main goals. Our first goal is to develop a composite description of the transition suitable for comparison with regional summertime climatology simulated by climate models, with an emphasis on those measurements unique to CSET. Our second goal is to briefly characterize day-to-day variability of the Sc–Cu transition and compare the space–time variation of cloud
properties with our current understanding, including the correlation with "cloud-controlling factors" such as the estimated inversion strength (EIS) (Wood and Bretherton 2006) and aerosol concentration. Our third goal is to demonstrate the potential for comparing data from individual flights with observationally constrained global model simulations sampled along the flight tracks, to evaluate model performance in this challenging regime. This includes reanalysis, as well as climate model simulations with winds and temperature (but not humidity, clouds, and aerosol) nudged toward reanalysis. We attempt to average across the prominent mesoscale organization seen in the cloudy boundary layers of this region, focusing on space scales of 500 km or more for which the CSET flights provide more representative sampling.

In section 2, we briefly summarize the observations used and we describe our binning and compositing technique. We demonstrate that the aircraft-based sampling is sufficiently representative to serve as a rough climatology. Sections 3a and 3b present the basic CSET-composite thermodynamic structure across the Sc–Cu transition and its flight-to-flight variability. With this context, the remainder of the paper emphasizes vertically resolved measurements across this transect that are unique to CSET. Sections 3c and 3d analyze turbulence and trace gas concentrations. Section 4 presents cloud microphysics, in situ aerosol, lidar–radar-derived cloud extent and radar-derived precipitation. Section 5 uses the binned results to test the correlation of low cloud cover with two potential cloud-controlling factors, inversion stability and cloud droplet number concentration. Section 6 compares observations from an illustrative CSET flight with reanalysis and a weather-nudged CSET flight, followed by a summary in section 7.

2. CSET observations and analysis methods

a. Measurements used in this study

The G-V instrumentation used for CSET was described in detail by A19. It included a 94-GHz cloud radar, a high spectral resolution lidar, dropsondes, and in situ probes for basic meteorology, cloud microphysics, aerosol, up- and downwelling radiation, ozone, and carbon monoxide. All G-V data analyzed here were obtained from the National Center for Atmospheric Research Earth Observing Laboratory CSET data archive (https://www.eol.ucar.edu/field_projects/cset). For our study, most in situ measurements are analyzed at 1-Hz frequency. The G-V has faster-response sensors for several of these quantities, such as velocity components. This study will introduce a 1-Hz proxy for dissipation rate of turbulent kinetic energy derived from the 25-Hz vertical velocity measurements.

Some aircraft measurements are compared with products from NOAA’s Geostationary Operational Environmental Satellite-15 (GOES-15), provided to CSET by NASA Langley’s SatCORPS group as described in A19 (and downloaded from https://satcorps.larc.nasa.gov). Our study uses low cloud fraction computed at 15-min frequency based on pixels of 4 km size at nadir using GOES visible and infrared imagery to determine cloud-top phase, temperature, and height (Minnis et al. 2008).

For comparison with flight data, a time-matched subset of GOES data over a 2° × 2° box centered on the aircraft is extracted. The low cloud fraction is defined as the fraction of cloud-filled pixels (based on a reflectance threshold) identified as warm (liquid phase with cloud top temp >273K) and low (cloud top height <4 km). The box size is chosen to encompass a large enough area around the aircraft to represent the cloud statistics in the surrounding region by smoothing across mesoscale cloud variability. Thus the GOES cloud fraction will not be identical to the point cloud fraction measurement made by the aircraft. We also use cloud droplet concentration N_d derived following Painemal and Zuidema (2011) by combining single-pixel cloud optical depth and effective radius retrievals. To mitigate possible biases due to subpixel cloud inhomogeneity and partly cloud-filled pixels, we use the...
75th percentile of the $N_d$s sorted from low to high over all low cloud pixels within the $2^\circ \times 2^\circ$ box.

b. CSET sampling strategy

By design, CSET surveyed the lower atmospheric structure across large distances while flying at varying heights in and above the cloud-topped boundary layer. As a result, the CSET in situ sampling was sparse and inhomogeneous. In addition, the cloud structures had ubiquitous mesoscale variability on scales of tens to hundreds of kilometers, and each flight followed a different track in a different synoptic setting. These challenges shaped our compositing method and affect its representativeness. We will briefly summarize the CSET flight patterns, shown in Fig. 4 of A19.

The CSET flights began and ended with survey legs at about 6 km altitude, typically launching dropsondes every 2° of longitude. The central part of the CSET flights consisted of 2000–2500 km of low-level sampling, repeating “modules” consisting of 10-min level legs at 150-m altitude, within the cloud layer, and approximately 300 m above the cloud layer, and sawtooth legs that sampled across the cloud layer and the capping inversion. The left column of A19’s Fig. 7 shows radar, lidar, and geostationary satellite views of a typical flight RF10 (27 July 2015) through the Sc–Cu transition that we will use as an illustrative example throughout this paper, following A19. This figure illustrates the pronounced small-scale and mesoscale variability in cloud and precipitation sampled during each CSET flight.

Four low-level sampling modules were completed on this flight. The first module went in a near southerly direction, so it is almost collapsed in the longitude–height presentation of A19’s Fig. 7. This motivates our use of a modified longitude coordinate (section 2c) for single-flight analysis and multiflight compositing.

The flight pairs were planned such that the eastbound flights resampled the same boundary layer air mass as sampled on the westbound flights two days earlier, as estimated from isobaric trajectories initialized at 500-m altitude (approximately 960 hPa) from NCEP operational analyses and short-range forecasts (A19). The westbound low-level sampling was weighted toward stratocumulus which was projected to break up as it advected downwind over the following 2 days. The eastbound low-level sampling was substantially farther south and west, as can be seen by comparing the thick dashed lines and the thick solid lines in Fig. 1, and predominantly sampled shallow cumulus cloud.

c. Longitude–height binning

For statistical analysis of day-to-day variability and for developing as representative a summertime “climatology” as possible from the CSET observations, we use longitude–height binning. This approach takes advantage of as much of the G-V data as possible, including ascents and descents as well as level legs. First, we define a simple modified longitude coordinate that we use to measure distance along the transect:

$$\text{lon}' = \text{lon}_0 + 0.8(\text{lon} - \text{lon}_0) + 0.4(\text{lat} - \text{lat}_0),$$

$$\text{lon}_0 = -140^\circ \text{E}, \quad \text{lat}_0 = 30^\circ \text{N}. \quad (1)$$

Figure 1 shows contours of lon’ as thin northwest (NW)–southeast (SE)-oriented solid yellow lines. Along the orthogonal thick solid yellow line extending approximately from Kona to Sacramento, lon’ is equal to the true longitude. Thus, the use of lon’ projects the flight data onto a longitude–height section following that central line. This coordinate has two advantages over our initial choice of using unmodified longitude while retaining a simple mathematical form. First, contours of the climatological cloud fraction and estimated inversion strength trend NW–SE in this region. Hence lon’ is optimized for distilling the systematic variations in cloud and boundary layer properties across the Sc–Cu transition. In particular, use of lon’ mitigates biases due to the Hawaii–California flights being systematically south of the California–Hawaii flights, and therefore tending to sample a lower cloud fraction and higher EIS at a given longitude. Second, with one exception, flight paths never doubled back over the same lon’. That is, for each flight there is a unique correspondence between elapsed time and lon’. In contrast, several flights had nearly north–south sections of low-level sampling that collapse when projected onto regular longitude.

We specify a bin size in lon’ and altitude $z$. For any given field and for each flight, we bin all of the 1-Hz samples into lon’–$z$ bins. For fields measured by the dropsondes, those measurements are also included in this binning. On any given flight, some such bins will be unsampled, some will have a few samples (e.g., where the aircraft was ascending or descending at 7.5 m s$^{-1}$), and some will be sampled during level flight legs and include hundreds or thousands of samples. For each field, we compute a bin mean, except for outlier-prone fields (cloud droplet number concentration, aerosol number concentrations, and chemical tracers), for which we compute a bin-median. We experimented with only retaining bins with a minimum number of 1-Hz samples (e.g., 10), but this did not yield obvious improvements.

A few fields can only be usefully computed with conditional sampling. The cloud droplet concentration $N_d$ is only used in cloud, empirically defined as where 1-Hz cloud liquid water content measured by the Cloud Droplet Probe (CDP) exceeds 0.01 g m$^{-3}$. 

Unauthenticated | Downloaded 12/13/22 08:58 PM UTC
Two measured aerosol number concentrations, UHSAS100 (aerosol particle concentration in the 100–1000 μm diameter range, the “accumulation mode” (Rogers and Yau 1989, p. 93) and CN (concentration of aerosol particles on which gas vapors can condense in the diameter range 10–1000 μm), can be contaminated by shattering of cloud and precipitation droplets. They are only used where the CDP detects less than 0.01 g m\(^{-2}\) of liquid water and the concentration of precipitation-size drops with radii >31 μm detected by the G-V 2D-C optical array probe is less than 1 L\(^{-1}\) (empirically chosen thresholds). Temperature, humidity, and winds are sampled by the dropsondes as well as in situ measurements. For these fields, both data sources are used for the binning; in any flight they mostly sample mutually exclusive lon’–z bins.

Figure 2 shows some examples for a representative flight RF10, using bins of 0.5° in lon’ and 100 m in z that resolve the aircraft flight path.
Figure 2c shows the UHSAS100 (accumulation-mode aerosol) concentration. Small gaps along the flight track show where it has been screened to avoid clouds and precipitation. It shows large horizontal variability both in and above the boundary layer. Figure 2d shows the lon’–z bin-average $N_d$, which is only used inside of clouds. This drastically reduces the sampling of bins during each flight, making it very spotty and less statistically robust.

To get representative spatial averages in the face of sparse sampling, we use coarser averaging bins of 5° in lon’ and 200 m in z for the remainder of this paper. The low-level sampling modules typically span 3°–4° in lon’, so within a 5° lon’ bin, they are sure to sample the full range of altitudes spanned by the module. The 150-m subcloud legs populate the lowest altitude bin and sampling typically extends at least 500 m above the trade inversion. Within the survey legs, dropsondes were typically released every 2° of longitude, so two or more dropsondes are included in each lon’ bin.

d. All-flight compositing

The composite lon’–z cross sections presented later in this paper are averages over the binned cross sections for each flight. By definition, only those flights that sampled each bin contribute to the bin average. Figure 3 shows the representativeness of such averaging. For each bin, we calculate the fraction of flights with no samples in that bin, and color that fraction of the bin in white. The fraction of the flights with over 100 samples in a particular bin is shaded dark blue, with 10–100 samples in medium blue, and with 1–10 or less in light blue.

Figure 3a shows the combined G-V and radiosonde measurements. It shows large horizontal variability both in and above the boundary layer. Small gaps along the flight track show where it has been screened to avoid clouds and precipitation. Figure 3b shows the G-V in situ measurements, and Figure 3c shows the G-V in-cloud measurements only. For each bin in each panel, the fraction of the flights including no, 1–10, 11–100, and 101 + 1-Hz measurements is colored in white, light blue, medium blue, and dark blue, respectively.

Figure 3a shows the UHSAS100 (accumulation-mode aerosol) concentration. Small gaps along the flight track show where it has been screened to avoid clouds and precipitation. It shows large horizontal variability both in and above the boundary layer. Figure 2d shows the lon’–z bin-average $N_d$, which is only used inside of clouds. This drastically reduces the sampling of bins during each flight, making it very spotty and less statistically robust.

To get representative spatial averages in the face of sparse sampling, we use coarser averaging bins of 5° in lon’ and 200 m in z for the remainder of this paper. The low-level sampling modules typically span 3°–4° in lon’, so within a 5° lon’ bin, they are sure to sample the full range of altitudes spanned by the module. The 150-m subcloud legs populate the lowest altitude bin and sampling typically extends at least 500 m above the trade inversion. Within the survey legs, dropsondes were typically released every 2° of longitude, so two or more dropsondes are included in each lon’ bin.

Figure 3 shows the representativeness of bins of 5° in lon’ and 200 m in z for (a) combined G-V and radiosonde measurements, (b) G-V in situ measurements, and (c) G-V in-cloud measurements only. For each bin in each panel, the fraction of the flights including no, 1–10, 11–100, and 101 + 1-Hz measurements is colored in white, light blue, medium blue, and dark blue, respectively.

Figure 2c shows the UHSAS100 (accumulation-mode aerosol) concentration. Small gaps along the flight track show where it has been screened to avoid clouds and precipitation. It shows large horizontal variability both in and above the boundary layer. Figure 2d shows the lon’–z bin-average $N_d$, which is only used inside of clouds. This drastically reduces the sampling of bins during each flight, making it very spotty and less statistically robust.

To get representative spatial averages in the face of sparse sampling, we use coarser averaging bins of 5° in lon’ and 200 m in z for the remainder of this paper. The low-level sampling modules typically span 3°–4° in lon’, so within a 5° lon’ bin, they are sure to sample the full range of altitudes spanned by the module. The 150-m subcloud legs populate the lowest altitude bin and sampling typically extends at least 500 m above the trade inversion. Within the survey legs, dropsondes were typically released every 2° of longitude, so two or more dropsondes are included in each lon’ bin.

d. All-flight compositing

The composite lon’–z cross sections presented later in this paper are averages over the binned cross sections for each flight. By definition, only those flights that sampled each bin contribute to the bin average. Figure 3 shows the representativeness of such averaging. For each bin, we calculate the fraction of flights with no samples in that bin, and color that fraction of the bin in white. The fraction of the flights with over 100 samples in a particular bin is shaded dark blue, with 10–100 samples in medium blue, and with 1–10 or less in light blue.

Figure 3a shows the combined G-V and radiosonde measurements. It shows large horizontal variability both in and above the boundary layer. Small gaps along the flight track show where it has been screened to avoid clouds and precipitation. Figure 2d shows the lon’–z bin-average $N_d$, which is only used inside of clouds. This drastically reduces the sampling of bins during each flight, making it very spotty and less statistically robust.

To get representative spatial averages in the face of sparse sampling, we use coarser averaging bins of 5° in lon’ and 200 m in z for the remainder of this paper. The low-level sampling modules typically span 3°–4° in lon’, so within a 5° lon’ bin, they are sure to sample the full range of altitudes spanned by the module. The 150-m subcloud legs populate the lowest altitude bin and sampling typically extends at least 500 m above the trade inversion. Within the survey legs, dropsondes were typically released every 2° of longitude, so two or more dropsondes are included in each lon’ bin.

Figure 3 shows the representativeness of bins of 5° in lon’ and 200 m in z for (a) combined G-V and radiosonde measurements, (b) G-V in situ measurements, and (c) G-V in-cloud measurements only. For each bin in each panel, the fraction of the flights including no, 1–10, 11–100, and 101 + 1-Hz measurements is colored in white, light blue, medium blue, and dark blue, respectively.

Figure 2c shows the UHSAS100 (accumulation-mode aerosol) concentration. Small gaps along the flight track show where it has been screened to avoid clouds and precipitation. It shows large horizontal variability both in and above the boundary layer. Figure 2d shows the lon’–z bin-average $N_d$, which is only used inside of clouds. This drastically reduces the sampling of bins during each flight, making it very spotty and less statistically robust.

To get representative spatial averages in the face of sparse sampling, we use coarser averaging bins of 5° in lon’ and 200 m in z for the remainder of this paper. The low-level sampling modules typically span 3°–4° in lon’, so within a 5° lon’ bin, they are sure to sample the full range of altitudes spanned by the module. The 150-m subcloud legs populate the lowest altitude bin and sampling typically extends at least 500 m above the trade inversion. Within the survey legs, dropsondes were typically released every 2° of longitude, so two or more dropsondes are included in each lon’ bin.

Figure 3 shows the representativeness of bins of 5° in lon’ and 200 m in z for (a) combined G-V and radiosonde measurements, (b) G-V in situ measurements, and (c) G-V in-cloud measurements only. For each bin in each panel, the fraction of the flights including no, 1–10, 11–100, and 101 + 1-Hz measurements is colored in white, light blue, medium blue, and dark blue, respectively.

Figure 2c shows the UHSAS100 (accumulation-mode aerosol) concentration. Small gaps along the flight track show where it has been screened to avoid clouds and precipitation. It shows large horizontal variability both in and above the boundary layer. Figure 2d shows the lon’–z bin-average $N_d$, which is only used inside of clouds. This drastically reduces the sampling of bins during each flight, making it very spotty and less statistically robust.

To get representative spatial averages in the face of sparse sampling, we use coarser averaging bins of 5° in lon’ and 200 m in z for the remainder of this paper. The low-level sampling modules typically span 3°–4° in lon’, so within a 5° lon’ bin, they are sure to sample the full range of altitudes spanned by the module. The 150-m subcloud legs populate the lowest altitude bin and sampling typically extends at least 500 m above the trade inversion. Within the survey legs, dropsondes were typically released every 2° of longitude, so two or more dropsondes are included in each lon’ bin.

Figure 3 shows the representativeness of bins of 5° in lon’ and 200 m in z for (a) combined G-V and radiosonde measurements, (b) G-V in situ measurements, and (c) G-V in-cloud measurements only. For each bin in each panel, the fraction of the flights including no, 1–10, 11–100, and 101 + 1-Hz measurements is colored in white, light blue, medium blue, and dark blue, respectively.
was often fully attenuated by cloud, but only rarely across a full 5° bin.

The all-flight composite is also subject to various possible sampling biases. For instance, certain synoptic conditions lead to different trajectories that must be sampled with different flight plans, our sampling favored conditions and regions without extensive high cloud, and the low-level sampling on eastbound flights was typically south and west of the westbound low-level sampling. To compare CSET data to a model constrained to have realistic time-varying meteorology for the CSET period, an ideal approach is to sample the model in time and space along the flight tracks, as done in section 6. However, some climate models may not be practical to run and sample in this way. In that case, all-flight composites of observations unique to CSET (i.e., for which a reanalysis or satellite retrieval does not provide trustworthy information, such as the vertical profile of rainwater content or ozone) provide a convenient and worthwhile comparison to the summertime climatology from a climate model, because even a substantial representativeness uncertainty may still be much smaller than the model bias.

3. Boundary layer structure and vertical mixing

a. Composite thermodynamic structure

Figure 4 shows the composite structure of potential temperature $\theta$ and relative humidity RH in the lower troposphere sampled by the 14 CSET cross-sectional flights, including both in situ and dropsonde data. In the CSET study region, there is typically an inversion layer whose base coincides with the top of any boundary layer stratocumulus. For each flight the Heffter (1980) method was used to diagnose the height of the inversion base from all individual aircraft and dropsonde profiles. This inversion base was averaged across 5° lon' bins for inclusion in this and other plots. The solid black line is the mean inversion base across all flights, and the dashed black lines indicate the interquartile range of inversion base between flights. As expected, the mean inversion base slopes upward to the west from 750 m near the coast to nearly 2 km west of 145°W, and there are strong vertical gradients of $\theta$ and RH centered around this height, somewhat sharper nearer the coast. The RH profile has a local maximum at an altitude of 500 m to the west of 140°W, suggestive of the climatological base of a cumulus layer that rises into any overlying stratocumulus.

b. Decoupling

The vast majority of individual CSET profiles show significant vertical humidity gradients below the main inversion, evidence of boundary layer decoupling. Figure 5 shows $\Delta q$, the difference in humidity between samples in the lowest 25% of the boundary layer and the uppermost 25% of the boundary layer, based on G-V profiles and dropsonde soundings from all CSET flights. Jones et al. (2011) characterized $\Delta q > 0.5 \text{ g kg}^{-1}$ as a decoupled boundary layer. Over 95% of the profiles or soundings are decoupled, and the decoupling index increases to the west as the typical inversion base deepens.
All points in Fig. 5 are colored by the GOES-derived low cloud fraction in a $2^\circ \times 2^\circ$ box centered on the sounding location. There is a clear but imperfect correlation between $\Delta q$ and cloud fraction. The well-mixed soundings with $\Delta q < 0.5 \, \text{g kg}^{-1}$ are all in 100% cloudy stratocumulus layers, but there are many more profiles with 80%–100% cloud fraction that have $\Delta q$ as large as 4 g kg$^{-1}$. Many profiles west of 145$^\circ$W have high decoupling index and low cloud fraction, indicative of a trade cumulus regime, but there are numerous exceptions. CSET lidar profiles often had less aerosol backscatter between the lowest cloud bases and the inversion than in the subcloud layer, consistent with decoupling (A19).

c. Turbulence

Approximately half of the CSET low-level sampling consisted of constant-altitude legs, either at 150 m, or within or above the main cloud layer. One could estimate vertical velocity variance or turbulent fluxes from these legs, but very few estimates are obtained for each flight, and the ubiquitous mesoscale variability diminishes the representativeness of those estimates.

Instead, as an example of an alternative interesting measure of turbulence that could usefully be compared with boundary layer parameterizations and large-eddy simulations, we introduce a new aircraft-based estimate of turbulent kinetic energy (TKE) dissipation rate $\epsilon$ per unit volume using high-frequency vertical velocity variability. Figure 7 of Wood et al. (2018) shows how this high-frequency vertical velocity variability clearly discriminates between higher turbulence levels in thicker stratocumulus clouds than in the thin veil clouds that commonly formed in ultraclean layers sampled in CSET. TKE dissipation rate depends on the intensity and continuity of turbulence. Continuous weaker turbulence and intermittent stronger turbulence (e.g., in a cumulus cloud field) may produce comparable dissipation rates. TKE dissipation rate is strongly controlled by the buoyancy (and to a lesser extent, shear) production of turbulence in subtropical marine boundary layers. In the appendix we describe how we estimate $\epsilon$ in lon–$z$ bins for each flight. Our estimate assumes that within each bin, the vertical velocity variance is due to turbulence. Gravity waves in stable stratification, such as in the free troposphere, will alias into this estimate, but they have comparatively little small-scale vertical velocity variance. This estimate may be less representative where the turbulence is highly variable (e.g., in cumulus cloud regimes).

Figure 6a shows our estimate of $\epsilon$ for RF10. Figure 6b shows the all-flight composite. In both cases, the dissipation rate estimates for the free troposphere are small and probably do not reflect truly turbulent flow. Between 500-m altitude and the trade inversion, dissipation is substantial but it has no clear trends in height or longitude that rise above the significant sampling
variability. Below 500 m, in the subcloud layer, we expect that the turbulence is surface-driven and more horizontally homogeneous, so dissipation estimates are thus more consistent across sampling bins, both in height and longitude. The all-flight composite dissipation estimates show an expected tendency to be larger near to the sea surface than within the cloud layer; this is less apparent in RF10 alone. Our estimates are comparable in magnitude and vertical structure to past measurements in stratocumulus (e.g., Lothon et al. 2005, Fig. 8) and large-eddy simulation of shallow cumulus (e.g., Cuijpers and Duynkerke 1993, Fig. 13).

d. Trace species

During CSET, the trace gases ozone (O₃) and carbon monoxide (CO) were measured at 1 Hz using instrumentation described by A19. Figure 7 shows their all-flight composites. These provide additional insight into boundary layer processes and are interesting references for chemical transport models (see section 6 for such a comparison).

Figure 7a shows lower ozone concentrations in the boundary layer than the free troposphere, as expected due to photolysis. At all altitudes, ozone also decreases slowly to the west, where recent advection from sources of continental smoke and pollution is less common.

Figure 7b shows that CO has comparatively little vertical gradient across the trade inversion in this region because it does not have a strong boundary layer sink. It shows the same westward decrease at all altitudes as does O₃, presumably for similar reasons.

4. Clouds and aerosols

a. Cloud cover

Figure 8 shows two all-flight composite estimates of the vertically resolved profile of cloud fractional coverage. The first (Fig. 8a) is derived from the aircraft Cloud Droplet Probe as described in section 2c. Because of the sparsity of the in situ sampling, this composite is noisy and its details are not robust. Nevertheless, some general trends are clear. There is substantial low cloud below the inversion in the east half of the cross section. Examination of individual flights shows this is mostly stratocumulus with underlying scud or cumulus cloud patches with lower bases. Some CSET westbound flights encountered very shallow PBLs with cloud tops less than 400 m at the start of their low-level sampling. Cloud bases could occasionally be less than 200 m there, even though fog was not encountered. Farther to the west, cloud is spread over a wide range of altitudes between 500 m and the mean inversion base. There is still some stratiform cloud (e.g., see Wood et al. 2018; A19), but it is thin and the weak inversion fluctuates in altitude such that a preferred altitude for cloud occurrence is not obvious.

The second composite (Fig. 8b) uses a joint cloud and precipitation occurrence mask created from the vertically pointing cloud radar and lidar (Schwartz et al. 2019). We use this joint “lidar–radar” mask because precipitation cannot be cleanly separated from cloud in the common occurrence that the lidar beam is fully attenuated. Because the lidar and radar sample all altitudes simultaneously in whichever orientation they are both pointed, this composite encompasses a larger
sampling domain and has less sampling variability. Reassuringly, it looks quite similar to the in situ cross section. Because it includes precipitation as well as cloud, the lidar–radar “cloud fraction” remains substantial below the actual cloud base (cf. Fig. 8b with Fig. 8a).

The vertically integrated low (altitude <3 km) cloud cover is particularly important because it is a strong control on the regional albedo. The lidar–radar mask can be used to detect the presence of any low cloud or precipitation above or below the aircraft. We assume that if there is low cloud in the column, the lidar is pointed in the appropriate direction to see it; this may occasionally not hold during sawtooths but overall is a good assumption. We also count a lidar–radar column as cloudy if the G-V Cloud Droplet Probe detects in situ cloud, because the lidar and/or radar may miss cloud that is so close to the aircraft as to lie within their respective dead zones (approximately 150 and 200 m, respectively). Figure 9a shows the lidar–radar-derived low cloud fraction averaged in lon’ bins for each flight (gray curves). It shows strong flight-to-flight variability with irregular cloud cover fluctuations between adjacent longitude bins. This emphasizes the synoptic and mesoscale variability of cloud within the Sc–Cu transition.
Nevertheless, the all-flight mean low cloud cover (black) varies smoothly from 0.8 in the stratocumulus regime between 130° and 135°W to 0.45 in the clustered cumulus regime west of 150°W. Also shown is the all-flight mean of the lon’-binned GOES-derived cloud cover (red), derived for each flight in a 2° × 2° box centered on the flight location and matched within 15 min in time. This shows a very similar trend as the lidar–radar low cloud cover but is higher near the coast, where the lidar–radar data show less low cloud than farther offshore.

Figure 9b shows a scatterplot across all flights and lon’ bins of GOES low cloud fraction versus the lidar–radar low cloud fraction; they are reassuringly consistent given their different sampling and measurement approaches, with a correlation coefficient of 0.73.

Figure 8 of Xiao et al. (2014) shows GOCCP (CALIPSO, day and night average) and MODIS (daytime only) summertime low cloud cover along the GCSS Pacific cross section slightly farther south; our CSET lidar–radar results lie between them, with a similar east–west gradient. We conclude that CSET provides a climatologically representative sample of data across the northeast Pacific Sc–Cu transition.

b. Liquid water

Figure 10a shows lon’–z composites of area-averaged “cloud water” content in droplets of less than 25 μm radius, the size range measured by the CDP and consistent with the 20–30-μm droplet radius threshold for rapid further collision-coalescence growth into a drizzle or rain drop (Rogers and Yau 1989, p. 121). Figure 10b shows “rainwater” content in precipitation-size drops from the 2D-C probe (which sizes drops with radii larger than 31 μm). When rainwater content is substantial, the precipitation mass in the missing 25–31-μm radius range is a small fraction of the total, so we use the sum of the CDP and 2D-C liquid water content as a proxy for total liquid water content.

The composite cloud water content is larger in the Sc region than in the Cu region. Even in the Sc region, the cloud water spans a range of altitudes, mainly due to variations in the boundary layer depth. Individual profiles in this region generally show a single stratocumulus layer less than 500 m thick capped by the trade inversion base, even though there may be substantial underlying cumulus or broken stratocumulus. West of 140°W, the cloud water content is broadly distributed in the vertical between 500 m and the inversion layer, both in this composite and in most individual flights. The rainwater content is larger in the Cu region than in the Sc region, and it is distributed fairly uniformly across the boundary layer depth in both cases.

Figure 10c shows the CSET composite cloud and rainwater path (column-integrated mass) and their sum, the liquid water path (LWP), in each lon’ bin. Within the cumulus regime, rainwater comprises more than half of the LWP, while cloud water comprises about 80% of the LWP in the stratocumulus regime. The scatter in Figs. 10a and 10b suggests substantial sampling uncertainties in these estimates. With this caveat, Fig. 10c also compares the CSET-mean LWP with the MAC-LWP climatology based on satellite microwave radiometer measurements (Elsaesser et al. 2017), sampled in
summertime along the CSET transect. The aircraft composite is smaller but within a factor of two of MAC-LWP across most of the transect. There is a diurnal cycle of boundary layer cloud and drizzle in this region, but at the mean time of in situ sampling (around 0800 local solar time on westbound flights for the Sc region, 130°–135°W, and 0845 local solar time on eastbound flights for the clustered Cu region, 145°–150°W), the in situ LWP is expected to be similar to the diurnal mean.

c. Droplet and aerosol concentration

Figure 11a shows the CDP composite in-cloud bin-median droplet concentration $N_d$. The hatched bins have fewer than 10 contributing in-cloud observations across all flights, but their bin-median $N_d$s appear consistent with the surrounding better-sampled bins. Near the California coast, the composite $N_d$ ranges from 80 to 150 cm$^{-3}$. West of 135°W the composite $N_d$ decreases to 30–75 cm$^{-3}$ below 1 km and tends to decrease with height above that altitude, with values below 10 cm$^{-3}$ near cloud top (darker green and blue shades) in some compositing boxes that are indicative of systematic occurrence of ultraclean layers, as extensively documented for CSET by Wood et al. (2018). O et al. (2018) explain these ultraclean layers as a result of coalescence scavenging of CCN in cumulus updrafts and sedimentation of droplets out of veil clouds formed in the detrainment layers of cumulus clusters.

Figure 11b shows bin-median UHSAS100, the measured concentration of accumulation mode aerosol particles with diameters between 100 and 1000 nm. Figure 11c shows bin-median CN, the concentration of aerosol particles with diameter 10–1000 nm measured by the condensation nucleus counter. CN is usually dominated by the Aitken mode (diameters 10–100 nm) but also includes the accumulation mode. As noted in section 2c, these measurements are screened for cloud
and precipitation to minimize contamination by droplet shattering.

UHSAS100 and CN show similar trends to \( N_d \) but more clearly, because more 1-Hz samples are retained with out-of-cloud sampling. Relatively low UHSAS100 concentrations are found near the inversion base west of 145°W (Fig. 11b). Farther east, free tropospheric UHSAS100 concentrations are comparable to those below the trade inversion at most longitudes. Free tropospheric CN concentrations are mostly 200–300 cm\(^{-3}\), with some higher outliers near the coast. The ratio of small aerosol particles (CN) to larger aerosol particles (UHSAS100) is systematically larger in the free troposphere than in the boundary layer. This suggests that small aerosol particles entrained into the cloud-topped boundary layer grow or coagulate, adding to its reservoir of potential CCN.

There is huge flight-to-flight variability in both \( N_d \) and UHSAS100, presumably reflecting different airmass histories as well as intermittent injections of forest fire smoke into the lower free troposphere from British Columbia, Alaska, and California during CSET (A19). This biomass burning signal is consistent with the high CO concentrations seen in the composite in Fig. 7b.

Figure 12 shows the all-flight mean of the lon’-bin-median CN and UHSAS100 in two representative altitude ranges, 0–200 m (surface, dark shades) and 3000–3200 m (above cloud, light shades). Gray dots connected by colored lines show all-flight means. For each lon’ bin, flight-to-flight variability is shown with a box (interquartile range) and whiskers (minimum–maximum, thin bars). Bars are offset right (left) of lon’ bin center for surface (above cloud) altitude ranges.

Across flights, locations and altitudes, there was a strong correlation between \( N_d \) and UHSAS100. Figure 13 is a scatterplot of the single-flight lon’–z bin-mediants of these fields, for bins where both were simultaneously measured. Both fields span nearly three orders of magnitude. Their correlation coefficient on a log–log plot is 0.71, with a power-law best-fit \( N_d \sim \text{UHSAS100}^{0.8} \). Most of the bins scatter around a 1:1 line, but there is a tail of low-concentration bins in which the UHSAS100 concentration can be double or more the cloud droplet number concentration, suggesting activation of droplets on smaller aerosols. The lowest concentrations (<10 cm\(^{-3}\) in either variable) are indicative of ultraclean layers. They mainly occur in altitude bins above 1 km (yellow, green, and blue shading). In each bin, \( N_d \) concentrations include cumulus cores (which typically have higher \( N_d \) than thin layer clouds), resulting in a lower fraction of bins that have median droplet concentrations below the 10 cm\(^{-3}\) ultraclean layer threshold compared to Wood et al. (2018) and O et al. (2018).

Satellite-derived cloud-top \( N_d \), when reliable, is an invaluable proxy for in situ \( N_d \). For instance, Painemal et al. (2015) compared ship-based UHSAS observations during the MAGIC campaign with collocated...
geostationary \( N_d \) retrievals of boundary layer stratocumulus cloud layers. They found a strong correlation east of 135°W and a weaker correlation farther west, which they interpreted as a consequence of decoupling of the deeper PBL farther to the west. We obtained similar results by correlating the lon\(^{-1}\)-binned UHSAS100 data for the subcloud legs of the CSET flights with GOES retrievals of \( N_d \) in 2° \( \times \) 2° collocated boxes described in section 2a (not shown).

But particularly in the cumulus regime, the assumptions underlying the satellite retrieval are dubious, so it is important to conduct an in situ comparison with that retrieval. In Fig. 14, we compare the in situ \( N_d \) vertically weighted by cloud water content (as a proxy for cloud-top \( N_d \), given the range of cloud heights sampled by the G-V) with GOES retrievals of \( N_d \) made following Painemal et al. (2015). There is a reasonable correlation, but especially to the west of 140°W (open circles in Fig. 14) where the cloud is more cumuliform, the GOES \( N_d \) is often significantly lower than the in situ \( N_d \).

This could be due to a breakdown of the homogeneous plane-parallel assumption for clouds made in the GOES retrievals, to precipitation-size drops that affect the cloud-top effective radius, or to systematic vertical gradients in \( N_d \) that make the cloud-weighted in situ \( N_d \) a poor estimate of the cloud top \( N_d \). More careful study is needed to fully resolve this discrepancy.

d. Precipitation

The airborne cloud radar was the primary tool for sensing precipitation in CSET. Here, we use the CSET flights to examine the precipitation frequency across the summertime Sc–Cu transition. We use the column-maximum reflectivity \( \text{dBZ}_{\text{max}} \) as a precipitation proxy, and consider two thresholds: \( \text{dBZ}_{\text{max}} > -10 \) (light drizzle) and \( \text{dBZ}_{\text{max}} > 0 \) (heavy drizzle). Following Eq. (10) of Comstock et al. (2004), these two thresholds correspond to cloud base precipitation fluxes of 0.3 and 2 mm day\(^{-1}\) respectively. A pragmatic motivation for these thresholds was that the cloud radar was not sensitive enough to reliably detect weaker drizzle down to \(-20 \text{dBZ}\) at the increased range needed on the survey legs.

As shown in Fig. 7 of A19, the radar was typically pointed down (up) when the aircraft was above (below) the cloud layer. We assume that the radar is oriented in the appropriate direction to sample \( \text{dBZ}_{\text{max}} \) throughout. Visual inspection suggests this was generally true, although during sawtooths and level in-cloud legs there may be some underestimate of maximum radar echo.

Figure 15 shows the fraction of sampled columns in each lon\(^{-1}\) bin with reflectivity exceeding the two thresholds. For each threshold, the average over all flights is shown as a solid line. For the sectors west and east of 140°W, the interquartile range between flights is shown as a vertical gray line segment, with a center dot showing the sectoral mean. The probability of \( \text{dBZ}_{\text{max}} \) exceeding –10 (0) dBZ was approximately 15% (5%) at all sampled longitudes. The occurrence fractions from individual flights varied from near zero to three times the mean in each longitude bin, with strong spatial variability associated with mesoscale regions of suppressed and enhanced precipitation. Figure 4 of a CloudSat study by Smalley and L’Ecuyer (2015) shows a map of June–July–August (JJA) mean rain occurrence frequency corresponding to a threshold of about –7.5 dBZ. Along the CSET cross section, values
ranged from 8% to 12%. The qualitative consistency between CSET radar observations and a comparable satellite climatology gives us confidence that CSET sampled a seasonally representative population of pre-cipitating boundary layer clouds.

On top of this variability, there is a systematic difference between longitude-binned precipitation occurrence frequencies averaged over the westbound (dotted) and the eastbound (dashed) flights. East of 140°W, the pre-cipitation frequency is higher on the westbound flights. We interpret this as a manifestation of the diurnal cycle of stratocumulus precipitation, with higher precipitation frequency in the early morning than in early afternoon.

On westbound (eastbound) flights, the G-V typically crossed 132.5°W around 0800 (1300) LST and 147.5°W around 1045 (0845) LST. Thus, the lower precipitation frequencies correspond to times significantly later in the day. Colored dots show all-flight mean precipitation fraction for the west and east sectors, separated by the thin dashed line at lon′ = −140. Thick gray lines (offset for the two reflectivity thresholds for clarity) indicate the interquartile range of the precipitation fraction across individual flights in these two sectors.

5. Role of cloud-controlling factors (EIS and \( N_d \))

CSET aimed to contribute to a better understanding of the Sc–Cu transition. In this section, we correlate the low cloud cover observed during CSET to two known controlling factors, Estimated Inversion Strength (EIS; Wood and Bretherton 2006) and cloud droplet number concentration \( N_d \). Here \( N_d \) is regarded as a marker of the interaction of aerosol with the cloud. CSET has extensive in situ \( N_d \) and cloud fraction measurements in a regime that challenges satellite \( N_d \) retrievals (as seen in Fig. 14), so it is well suited for disentangling cloud-controlling effects of \( N_d \) from those of EIS. We calculate lon′-binned EIS from ERA5 reanalysis, since the G-V profiles during low-level sampling modules often did not reach the 3 km elevation needed to compute it. In support of this approach, we compared EIS calculated from G-V dropsonde profiles and in situ profiles, which spanned 150–3000-m altitude, with EIS based on collocated ERA5 output. We found the ERA5 EIS to be unbiased, with a mean absolute deviation slightly larger than 1 K compared to the EIS derived from the G-V observations.

Figure 16 shows a scatterplot of the lon′-binned lidar–radar low cloud fraction introduced in section 4a versus EIS, colored by in situ cloud water-content weighted \( N_d \). This shows an expected positive correlation (Klein and Hartmann 1993; Wood and Bretherton 2006) with a fit line (black) consistent with the gray seasonal mean line from Wood and Bretherton (2006). There is much more scatter than on seasonal time scales, reflecting the strong internal variability of the instantaneous cloud field. There is usually a decrease of EIS toward the west, where ocean temperatures are warmer; this results in the open circles, which come from the west sector, mainly lying to the left of the closed circles, which come from the east sector.

For a given EIS, if cloud fraction was also substantially correlated with \( N_d \), Fig. 16 would show a vertical stratification of dot colors, but we do not see...
this. Stepwise linear regression corroborates that conclusion; after accounting for EIS, using in situ \( N_d \) as a second predictor does not add significant skill in predicting cloud fraction.

6. Comparison with models

A goal of CSET was to produce a dataset useful for testing global model representations of the Sc–Cu transition. Despite sampling and representativeness uncertainties, the composites above should prove useful comparisons for summertime climatologies from global climate models.

A more powerful model test is to sample global weather simulations of the flight days and sample them along the position and time of the flight track to compare with the observations. Since clouds and aerosols respond rapidly to synoptic variations, this allows every flight to be used as a quasi-independent test of the model.

We compare two simple approaches. The first is to sample a global reanalysis. We use NASA’s Modern-Era Retrospective Analysis for Research and Applications, version 2, or MERRA-2 (Gelaro et al. 2017, data used here obtained from https://disc.gsfc.nasa.gov/datasets). MERRA-2 has 72 vertical levels and approximately 60 km horizontal grid spacing. It includes a selection of chemical species and aerosols. Neither the CSET in situ airborne observations nor dropsondes were transmitted to forecast centers, so they are independent checks on this reanalysis. The second approach is to use a global climate model. We use version 6 of the Community Atmosphere Model (CAM6), with the wind, temperature and surface pressure fields (but not humidity, clouds, or aerosols) nudged at every grid point toward a time/space interpolated version of MERRA-2 using a 24-h relaxation time scale. The CAM6 data used here can be requested from A. Gettelman (andrew@ucar.edu). Wu et al. (2017) describe an earlier application of this approach to CAM5 for the HIAPER Pole-to-Pole Observations (HIPPO) campaign in 2009–11, with a focus on ice cloud microphysics.

CAM6 has 32 vertical levels and approximately 1° horizontal grid spacing. The purpose of this second approach is to test the skill of CAM6 in representing clouds, aerosols and boundary layer structure when constrained to approximately follow the observed weather. Because humidity and clouds are freely evolving in this framework, they can develop errors that can point to needed model improvements. We now show an illustrative example of this methodology; its systematic exploitation for CSET (particularly for aerosols and warm cloud microphysics) is left for a sequel.

Figure 17 shows a comparison of the RF10 measurements of relative humidity and cloud water content with MERRA-2 and nudged CAM6, each sampled along the latitude, longitude, and time of the flight path. Since CAM6 is nudged toward MERRA-2 meteorology, differences in humidity and cloud fields in the two models likely stem mainly from their different parameterized cloud and boundary layer physics.

The top of the “moist layer” of high RH is a marker of boundary layer height. MERRA-2 captures most of the longitudinal variability in RH (Fig. 17a) and places the boundary layer stratocumulus top at approximately the right height (Fig. 17c). The steady increase of RH between the surface and 1 km altitude indicates a well-mixed boundary layer in MERRA-2 east of lon’ = 140°W, while the observed RH already shows a weaker increase with height at lon’ = 135°W and even in the profile just east of lon’ = 130°W, indicative of weak decoupling. The nudged CAM6 simulates a shallower moist layer than observed across most of the flight track (Fig. 17b). In the east sector of the cross section, both models, especially MERRA-2, simulate stratocumulus layers that are vertically diffuse compared to the observed cloud water (Figs. 17c,d), an expected consequence of comparing point samples with gridbox averages. MERRA-2 simulates more cloud water than CAM6 in the shallow cumulus regime to the west of the in situ sampling. These trends are also evident in all-flight averages of these comparisons (not shown).

MERRA-2 also provides \( O_3 \) and CO concentration estimates. Figure 18 compares these with the in situ data for RF10. MERRA-2 overestimates the distinctly lower in situ \( O_3 \) concentrations below the inversion for this and other CSET flights, suggesting that ozone loss processes in the boundary layer are underestimated. The MERRA-2 CO is somewhat lower than the in situ observations.

7. Summary

CSET was the first airborne field study to gather extensive in situ observations across the entire summertime northeast Pacific Sc–Cu transition. CSET was designed for Lagrangian sampling of this transition, as discussed by A19 and a companion paper by Mohrmann et al. (2019, manuscript submitted to Mon. Wea. Rev.). In the process, the 14 CSET flights also gathered a superb quasi-Eulerian dataset representative of a line between Northern California and Hawaii, slightly north of the nearby GPCI and MAGIC cross sections. We have presented an Eulerian view of the Sc–Cu transition and its synoptic variability based on these flights, including unique new in situ vertically resolved measurements of
FIG. 17. Comparison of RF10 measurements (colored channels) of relative humidity and cloud water content with MERRA-2 and nudged CAM6, each sampled along the latitude, longitude, and time of the flight path and indicated by the background shading.
turbulence, cloud microphysics, aerosol characteristics, trace gases, and vertically pointing W-band radar and lidar returns. We developed a longitude–height binning and compositing strategy to mitigate limitations of sparse sampling and internal spatial and temporal variability.

Overall, the observed characteristics of the Sc–Cu transition matched expectations, with a Sc-topped boundary layer becoming systematically more decoupled farther to the southwest, followed by a cloudiness transition at 135°–145°W into precipitating cumulus clusters organized on 100-km scales. These clusters were surrounded by patches of thin Sc ("veil clouds") and often ultraclean layers near the trade inversion base, as analyzed by Wood et al. (2018).

The fractional coverage of the Sc patches diminished toward Hawaii, where the inversion base was typically 2 km. Multiscale cloud organization was apparent throughout the flights.

Aircraft lidar-measured low cloud cover, averaged during each flight over 5° longitude bins, was positively correlated with estimated inversion strength (EIS) with a regression slope close to past climatology. After controlling for EIS, the low cloud cover was not well correlated with cloud droplet concentration, which suggests that the Sc–Cu transition is not strongly controlled by aerosol processes. This contrasts with some recent idealized large-eddy simulation studies (e.g., Yamaguchi et al. 2017), but is consistent with another such study using data-constrained forcings and boundary conditions (McGibbon and Bretherton 2017).

The CSET lidar–radar cloud fraction compared well with GOES satellite retrievals of cloud fraction, but the GOES retrievals of droplet concentration for CSET were often biased low, especially in the cumulus regime. This issue merits further investigation.

In almost all flights and regimes, the thickest clouds were precipitating, but the areal coverage of column-maximum radar echoes exceeding 0 dBZ was only about 5% in both Cu and Sc regimes, with more extensive precipitation earlier in the morning. The ultraclean layers observed in accumulation mode aerosol and cloud droplet concentration are evidence of active
aerosol–cloud–precipitation feedbacks within the cumulus clusters.

The flights measured the partitioning of liquid water between cloud (≤25 μm droplet radius) and drizzle/rain (larger drops). This was highly variable, but in general, in the shallow boundary layers typically seen farther east, most of the liquid water is cloud, while farther west, where Cu occasionally reached 4 km deep, the average column mass of drizzle and rain was larger than that of cloud. A multiflight composite of the column-integrated liquid water is within a factor of 2 of a satellite climatology of liquid water path (MAC-LWP) during summertime at most longitudes across this transect.

Above the boundary layer, Aitken mode aerosol concentration was typically 150–200 cm$^{-3}$, about twice as much as near the surface. Accumulation-mode aerosol concentration near the surface was typically 50–100 cm$^{-3}$, and often higher near the coast, with somewhat smaller values at 3 km. Within the boundary layer, cloud droplet number correlated quite well with accumulation mode aerosol number concentration from the same height. Smoke from some large forest fires advected over the region, producing spikes of aerosol, CO and O3 on some flights. The CO mixed down into the boundary layer with little systematic gradient at the inversion, but the O3 was half or less as large within the boundary layer as above, particularly in the deeper Cu boundary layers at the west end of the trajectory.

Measured profiles of turbulent kinetic energy dissipation rate across the transition are presented as a possible comparison for LES and those GCM parameterizations of subtropical cloud-topped boundary layers that compute dissipation.

Last, we showed a single-flight comparison of the CSET dataset with space–time collocated output from the MERRA-2 global reanalysis and the CAM6 climate model nudged toward MERRA-2 winds and temperatures. MERRA-2 compares well with the aircraft-measured trade inversion base; CAM6 has a slightly lower inversion base. Both models vertically diffuse the observed stratocumulus cloud layers. In a future paper, we intend to use such comparisons more systematically as a tool for GCM diagnosis and improvement of parameterized cloud and aerosol processes. The breadth of CSET observations analyzed in this paper provides a unique and rigorous test of GCM performance across a subtropical Sc–Cu transition.

Acknowledgments. The success of CSET is primarily due to the dedicated staff of NCAR’s Research Aviation Facility who operated the G-V and its core instrumentation, provided invaluable advice about analyzing the data, and made CSET a pleasure for its scientific participants. The authors wish to acknowledge support from NSF grants AGS-1445813, AGS-1445832, and AGS-1445831. I. McCoy was supported by the National Science Foundation Graduate Research Fellowship Program under grant DGE-1762114. V. Ghate acknowledges support from the U.S. Department of Energy’s (DOE) Atmospheric System Research (ASR) under Contract DE-AC02-06CH11357 awarded to Argonne National Laboratory, and computing resources provided on Blues, a high-performance computing cluster operated by the Laboratory Computing Resource Center (LCRC) at Argonne National Laboratory.

APPENDIX

Estimate of TKE Dissipation from Airborne Measurements Using High-Frequency Vertical Velocity Variability

In homogeneous isotropic turbulence with TKE dissipation rate $\varepsilon$ per unit volume, the power spectrum of the vertical velocity has the form

$$P_{ww}(k) = 0.8 \varepsilon^{2/3} k^{-5/3}, \quad (A1)$$

where $k$ is wavenumber (e.g., Garratt 1992, p. 71). Our goal is to use this formula to derive a spatially localized estimate for $\varepsilon$. Our approach is broadly based on Cornish et al. (2006). Like that study we infer $\varepsilon$ from high-frequency vertical velocity fluctuations, but our approach is based on time-windowed statistics rather than their more sophisticated but also more complex wavelet filtering method.

For a plane obtaining a time series of vertical velocity fluctuation $\omega(t)$ by flying horizontally through the turbulence at speed $U_0$, the wavenumber $k$ is equivalent to a frequency $f = kU_0/2\pi$ and the temporal power spectrum is

$$P_{ww}(f) = 0.8 \left( \frac{U_0}{2\pi} \right)^{2/3} \varepsilon^{2/3} f^{-5/3}. \quad (A2)$$

During CSET, the velocity components were all measured using pressure transducers and provided at 25 Hz. On level subcloud legs, in which the turbulence usually appears fairly homogeneous, the power spectrum of vertical velocity fluctuations $\omega$ follows the expected Kolmogorov $k^{-5/3}$ spectrum between 0.5 Hz and the Nyquist frequency of 12.5 Hz. On the other hand, the power spectra of horizontal velocity fluctuations hit white noise floors at 1 Hz in the direction of aircraft motion and 5 Hz in the transverse direction, suggesting that they have larger measurement errors. Thus we
estimate \( \varepsilon \) from \( w' \), filtered to only retain frequencies within the inertial range.

This can be done in many ways. We chose to chunk \( w' \) into windows of length \( \tau = 1 \) s, each corresponding to a second of aircraft sampling with \( N = 25 \) samples. Within each window, we choose \( w_{hi} \) to be the 25-Hz vertical velocity fluctuation about the 1-s mean and define \( w_{h'} \) as the “high-pass” sample variance of \( w_{hi} \) within that window. The tilde indicates that this is a sample variance calculated every 1 s. If the turbulence is assumed to be statistically stationary over some interval, we can define \( w_{h'} w_{h'} \) to be the expected high-pass variance averaged over many samples, denoted by an overline. This has the nice property that it can be added to the 1-Hz vertical velocity variance to get the full 25-Hz vertical velocity variance; it mainly responds to frequencies of 1–2 Hz.

We chose a 1-Hz turbulence measure because it is convenient for correlating with other aircraft-measured quantities such as liquid water content (see Wood et al. 2018 for an application to clouds in ultra-clean layers). However, the 1-Hz estimate is quite variable, even in homogeneous turbulence, because TKE dissipation is distributed highly nonuniformly within a turbulent fluid. We find that 20 s or longer averages of \( w_{hi} w_{hi} \) smooth out much of this intrinsic variability to give a more robust estimate of \( w_{h'} w_{h'} \). The binning used for our single-flight composites effectively does this. When the plane is sampling along a level leg, a 5° lon’–200-m height bin will naturally include at least 10 min of sampling. During G-V climbs and descents, the plane maintained a vertical velocity of approximately 7.5 m s\(^{-1}\), traversing a 200-m vertical bin in approximately 27 s. Thus, for each flight we use the bin-mean of \( w_{hi} w_{hi} \), which we denote by angle brackets, to locally estimate

\[
\langle w_{hi} w_{hi} \rangle \approx w_{h'} w_{h'}.
\]  

(A3)

From \( w_{hi} w_{hi} \) we can estimate the dissipation rate \( \varepsilon \). Within each 1 s window, we can Fourier analyze \( w' \) into the frequencies \( f_j \), \( \Delta f = 1/\tau \), \( j = 0, 1, \ldots, (N - 1)/2; \)

\( w_{h'} \) is obtained by removing the zero-frequency component from \( w' \). The expected value of the remaining variance is

\[
\langle w_{hi} w_{hi} \rangle = \sum_{j=1}^{(N-1)/2} \Delta f \sum_{j=1}^{(N-1)/2} f_j^{-5/3} \Delta f = 0.8a \left( \frac{U}{2\pi} \right)^{2/3} \varepsilon^{2/3} \cos \left( \frac{U \tau}{2\pi} \right)^{2/3},
\]  

(A4)

where

\[
a = \sum_{j=1}^{(N-1)/2} j^{-5/3} = 1.85.
\]

Approximately half of the high-pass variance comes from the 1-Hz frequency, and the rest comes from the higher harmonics. This can be solved for the estimated TKE dissipation rate:

\[
\varepsilon \approx \left( \frac{C}{(0.8a)} \right)^{3/2} \tau^{3/2} = 2\pi/(0.8a)^{3/2} = 3.5, \quad \tau = 1 \text{s}.
\]  

(A5)

REFERENCES


