

## RESEARCH ARTICLE

10.1029/2021JD034666

## Wavelet Analysis of Properties of Marine Boundary Layer Mesoscale Cells Observed From AMSR-E

## Key Points:

- Two-dimensional wavelet analysis is applied for the first time to a full year of passive microwave retrievals from Advanced Microwave Scanning Radiometer for Earth Observing System
- Cell-scale cloud water path perturbations are good predictors of the water vapor path and rain rate perturbations, similar between open and closed cells
- Sea surface temperature, estimated inversion strength, and 10-m wind speed do not play a major role in determining the scale of the mesoscale cellular convection

## Supporting Information:

Supporting Information may be found in the online version of this article.

## Correspondence to:

X. Zhou,  
xiaoli.zhou@noaa.gov

## Citation:

Zhou, X., Bretherton, C. S., Eastman, R., McCoy, I. L., & Wood, R. (2021). Wavelet analysis of properties of marine boundary layer mesoscale cells observed from AMSR-E. *Journal of Geophysical Research: Atmospheres*, 126, e2021JD034666. <https://doi.org/10.1029/2021JD034666>

Received 26 JAN 2021

Accepted 22 JUN 2021

Xiaoli Zhou<sup>1,2,3</sup> , Christopher S. Bretherton<sup>1</sup> , Ryan Eastman<sup>1</sup> , Isabel L. McCoy<sup>1,4</sup> , and Robert Wood<sup>1</sup> 

<sup>1</sup>Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA, <sup>2</sup>Now at NOAA Chemical Sciences Laboratory (CSL), Boulder, CO, USA, <sup>3</sup>Now at Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder, CO, USA, <sup>4</sup>Now at Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, FL, USA

**Abstract** Marine boundary layer clouds tend to organize into closed or open mesoscale cellular convection (MCC). Here, two-dimensional wavelet analysis is applied for the first time to passive microwave retrievals of cloud water path (CWP), water vapor path (WVP), and rain rate (RR) from Advanced Microwave Scanning Radiometer for Earth Observing System in 2008 over the Northeast and Southeast Pacific, and the Southeast Atlantic subtropical stratocumulus to cumulus transition regions. The (co-)variability between CWP, WVP, and RR in  $160 \times 160 \text{ km}^2$  analysis boxes is partitioned between four mesoscale wavelength octaves (20, 40, 80, and 160 km). The cell scale is identified as the wavelength of the peak CWP variance. Together with a machine-learning classification of cell type, this allows the statistical characteristics of open and closed MCC of various scales, and its relation to WVP, RR, and potential environmental controlling factors to be analyzed across a very large set of cases. The results show that the cell wavelength is most commonly 40–80 km. Cell-scale CWP perturbations are good predictors of the WVP and RR perturbations. For cells larger than 20 km, there is no obvious dependence of cell scale on the environmental controlling factors tested, suggesting that the cell scale may depend more on its historical evolution than the current environmental conditions.

## 1. Introduction

Marine boundary layer clouds over the colder regions of the ocean often organize into closed or open mesoscale cellular convection (MCC) with cell sizes between 10 and 100 km, modulating cloud water path (CWP), precipitation, and albedo (Agee et al., 1973). MCC is associated with significant mesoscale variations of moisture (~10% relative humidity perturbation), temperature, and winds (Rothermel & Agee, 1980). MCC-like patterns can be simulated in large-eddy simulations, weather, and climate models with horizontal grid resolutions of O (10 km) or less (e.g., Boutle & Abel, 2012). To evaluate their skill requires good documentation and understanding of MCC cloud morphology and scale, of co-variability between observable quantities within closed and open cells, and of the sensitivity of closed and open MCC to potential environmental controlling factors across different boundary-layer cloud regions.

There is a 50-year history of MCC observations from in situ and satellite measurements that has advanced our understanding and provided local data that has been used for model comparisons. MCC was first observed by the first weather satellites in the early 1960s (Agee, 1984). MCC covers extensive regions over the eastern subtropical oceans (Muhlbauer et al., 2014; Wood & Hartmann, 2006), with closed cells forming near the coast and open cells occurring toward the warm oceans (Atkinson & Zhang, 1996; Muhlbauer et al., 2014). Closed and open cells can be considered as different stages of the stratocumulus to cumulus transition (Wood, 2012), which is often associated with the advection of clouds over a warmer ocean surface (Bretherton & Wyant, 1997; Sandu & Stevens, 2011), or with the passage of cyclones and cold air outbreaks (Abel et al., 2017; Field et al., 2014; Fletcher et al., 2016; McCoy et al., 2017; Yamaguchi & Feingold, 2015). Eastman et al. (2021) finds that strong wind speed and heavier drizzle favors the transition from closed to open cells. Previous idealized model and observational studies have suggested that precipitation differs significantly between closed and open cells and can be a microphysical trigger of transition between the two (Abel et al., 2020; Berner et al., 2013; Goren & Rosenfeld, 2012; Savic-Jovicic & Stevens, 2008; Sharon et al., 2006; Stevens et al., 2005; Wang & Feingold, 2009; Xue et al., 2008).

Global statistical analyses of MCC useful for comparing with models and theories are scarce. Agee et al. (1973) published the first global map of open and closed cell occurrences estimated from positions of warm and cold ocean currents, respectively. With the advent of neural network algorithms and their easy application to satellite data, this qualitative MCC climatology has been quantified and statistically connected to environmental controls (McCoy et al., 2017; Muhlbauer et al., 2014; Wood & Hartmann, 2006). Additionally, Wood and Hartmann (2006) found that typical MCC cell sizes increase with boundary layer depth, maintaining an approximate aspect ratio of 40:1.

One underexploited tool to probe MCC is passive microwave remote sensing of CWP, water vapor path (WVP) and precipitation. Zhou and Bretherton (2019b) used ground-based microwave radiometer (MWR) observations from the ARM Eastern North Atlantic (ENA) site at Graciosa Island to analyze correlations between CWP and WVP and test predictions of a humidity self-aggregation mechanism of MCC proposed by Zhou and Bretherton (2019a). The ground-based MWR measurements are useful for case studies, but they suffer from sensor wetting during and after precipitation (Turner et al., 2007) and are limited to only a few locations.

In this study, we take a wider satellite-based view of the mesoscale variability within closed and open MCC using two instruments on NASA's Aqua satellite. The backbone of our analysis is the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E, Wentz & Meissner, 2004), used to derive CWP, WVP, and rain rate (RR). This is coupled with cell type classifications from a machine learning algorithm using visible satellite imagery from the NASA Moderate Resolution Imaging Spectroradiometer (MODIS, King et al., 1992).

Here, we apply for the first time a two-dimensional wavelet analysis (e.g., Lau & Weng, 1995; Torrence & Campo, 1998) of AMSR-E derived CWP to identify a dominant local time and space dependent MCC scale where detectable. The analysis is also used to partition the co-variability between CWP and WVP between scales. The wavelet analysis is superior to the conventional windowed Fourier analysis in that the wavelet analysis can (a) exactly reconstruct the original data from the wavelet coefficients; (b) naturally remove constant, linear, and quadratic trends from the wavelet details with a suitable wavelet; (c) computationally much simpler; and (d) wavelet power spectra are less noisy than Fourier power spectra, unless the latter are octave-averaged.

Section 2 describes the data sets used. Section 3 details the wavelet analysis, including a demonstration of an application of two-dimensional discrete wavelet analysis to AMSR-E CWP. In Section 4, we show the dependence of MCC properties on cellular scale. Section 5 discusses the relationship between CWP and RR. Environmental controls on the MCC scale and type are analyzed in Section 6. Section 7 synthesizes results of Sections 4 and 5 into a mesoscale cell composite. Section 8 concludes with a discussion and summary.

## 2. Data

### 2.1. Regions

Three subtropical stratocumulus and downstream cloudiness transition regions selected by Eastman and Wood (2016) are used in this study, the Northeast (NE) Pacific (15°–30°N, 155°–115°W), the Southeast (SE) Pacific (30°–5°S, 105°–70°W), and the Southeast (SE) Atlantic (30°–5°S, 15°W–15°E). We leave out the eastern Indian Ocean included in Eastman and Wood (2016) because it has less stratocumulus coverage than the other three regions. All available data from 2008 are used.

### 2.2. AMSR-E

CWP and WVP in this study are sourced respectively from the columnar cloud liquid water and columnar water vapor in the AMSR-E (Kawanishi et al., 2003) Aqua L2B global swath ocean products derived from the Wentz Algorithm (version 2, Wentz & Meissner, 2004). The Aqua satellite is part of the A-Train satellite constellation that crosses the equator at 0130 and 1330 local times. The AMSR-E has a swath width of 1,445 km. The retrievals are provided on a non-uniform grid within the swath with pixel resolution of ~10 km at center-track, but the footprints of some of the sampled wavelengths are at a coarser resolution. The CWP retrieval relies heavily on the 37 GHz channel with a 14 × 8 km footprint, and the WVP retrieval

relies mostly on the 19 GHz channel with a  $27 \times 16$  km footprint. This means that WVP can only be resolved at scales exceeding 20 km. Estimated root mean square (RMS) error for the CWP and WVP retrievals are 0.017 and 0.57 mm respectively (Wentz & Meissner, 2004).

For calculating the MCC cell scale and other wavelet-based statistics, AMSR-E pixels are divided into “boxes” of  $16 \times 16$  grid points ( $\sim 160 \times 160$  km<sup>2</sup> near nadir). This size allows two-dimensional wavelet decompositions across four wavelength octaves (see Section 3) to capture the mesoscale variability within each box. To ensure a robust scale estimate from wavelet analysis, which requires nearly evenly distributed points in the sample boxes, the first 99 pixels from the right of each arcing swath are excluded from the analysis, as they are severely distorted.

### 2.3. Light RR Retrieval

The light RR in this study is retrieved from AMSR-E 89-GHz brightness temperature  $T_b$  following Eastman et al. (2019). This has advantages over the AMSR-E Aqua level-2B instantaneous surface rain product (AE\_Rain; Kummerow et al., 2015), which is based on the same two frequencies (37 and 19 GHz) used to derive CWP and WVP. The 89 GHz channel gives an independent retrieval that is more sensitive to drizzle, has a smaller footprint, and better matches high-resolution CloudSat retrievals (Eastman et al., 2019). The RR data are spatially interpolated and grouped into the same  $16 \times 16$  grid-box scenes used for CWP and WVP.

### 2.4. Correction of Wentz CWP

Rain can complicate microwave retrieval techniques by requiring a priori partitioning of total water into cloud and rain components. From the Wentz algorithm, rain is indicated when the total liquid water path exceeds a threshold of 0.18 mm, above which the algorithm assigns a certain fraction of liquid water content to rain assuming that cloud liquid water is a function of RR and depth of the rain column (Wentz & Spencer, 1998).

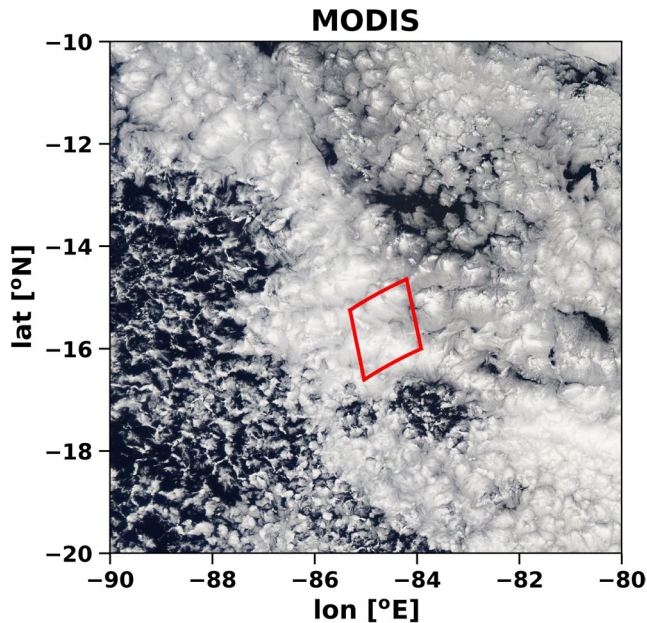
Previous studies (e.g., Seethala & Horváth, 2010; Lebsock & Su, 2014) have found that CWP retrievals from AMSR-E using the Wentz algorithm are fairly good where cloud cover is overcast and nonprecipitating, but subject to retrieval biases in areas with broken cloud cover and in the presence of precipitation, such as those found in the trade cumulus regions. The bias related to the partitioning between cloud water and rain is estimated to be the greatest among others (Lebsock & Su, 2014). Here we correct the AMSR-E  $CWP_{Wentz}$  following Eastman et al. (2021). They compared MODIS and AMSR-E CWP values within bins of constant RR. They fit the difference between the AMSR-E and MODIS CWP using a 7<sup>th</sup> polynomial fit in terms of the Eastman-derived RR in mm hr<sup>-1</sup>, and subtracted this fit from the  $CWP_{Wentz}$  to get a corrected CWP in mm

$$CWP = CWP_{Wentz} - \left( 0.5431 \cdot rr^7 - 2.6016 \cdot rr^6 + 5.3307 \cdot rr^5 - 6.1265 \cdot rr^4 + 4.3332 \cdot rr^3 - 1.9501 \cdot rr^2 + 0.5953 \cdot rr^{-0.0089} \right). \quad (1)$$

The correction reduces  $CWP_{Wentz}$  by less than 0.02 mm for  $rr < 0.1$  mm hr<sup>-1</sup>, 0.05 mm for  $rr = 0.2$  mm hr<sup>-1</sup>, and as much as 0.1 mm for  $rr = 0.7$  mm hr<sup>-1</sup> (see Figures A1 and A2 in Eastman et al., 2021). We will use this corrected CWP throughout this paper.

### 2.5. Mesoscale Cellular Convection Identification

The warm MCC cloud morphology identification data set used in this study is derived from the application of the neural net (NN) defined in Wood and Hartmann (2006) to a full year of satellite-derived CWP in 2008. CWP is estimated based on Collection 6 Level 2 daytime marine cloud retrievals (cloud optical depth and effective radius) at 1 km (nadir) resolution from MODIS on the Aqua satellite (Wood & Hartmann, 2006). MODIS CWP is divided into scenes of  $256 \times 256$  km<sup>2</sup> in size oversampled by 128 km in each direction. Each overlapping scene is classified into one of the three morphological types: closed MCC, open MCC, and cellular but disorganized clouds. The corresponding cloud types are placed at the center of each



**Figure 1.** MODIS  $10^\circ \times 10^\circ$  scene ( $\sim 1,110 \times 1,110$  km) at 1-km resolution at nadir over the SE Pacific on August 11, 2008, taken from NASA Worldview. An example closed-cell region of  $\sim 160 \times 160$  km (discussed later) is denoted by a red box.

overlapping scene with 128 km resolution. The morphology classification accuracy for the trained network is estimated to be 85%–90% (Wood & Hartmann, 2006). Sometimes, cells we might subjectively classify as closed are classified as open cells by the automated algorithm, especially in transition regions from closed to open cells. Details and examples of the cloud type classification are given by Wood and Hartmann (2006).

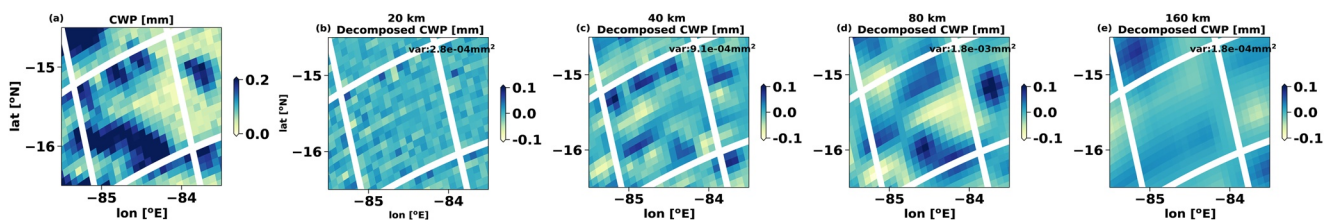
To adopt this cloud identification data set to our analysis, we re-grid the  $(128 \text{ km})^2$  MCC data onto AMSR-E CWP resolution ( $\sim 10$  km) using nearest-neighbor interpolation. The re-gridded MCC cloud types are then regrouped to the  $16 \times 16$  AMSR-E grid-box scenes. If over 50% of the pixels in an AMSR-E scene have been assigned the same MCC cloud type, the scene is considered to have that cloud type, otherwise it is classified as “Mixed pattern MCC.” Using this definition, of the 403,904 collected scenes, 7% are closed cell cases, 8% are open cell cases, 30% are disorganized cases, and the rest are mixed pattern MCC cases. A stricter 90% classification threshold of like-classified pixels in a scene reduces the number of samples of each MCC cloud type but has little effect upon the derived statistics.

## 2.6. Meteorological Control Metrics

Four environmental variables are examined in this study: sea surface temperature (SST), 10-m wind speed, estimated inversion strength (EIS), and planetary boundary layer (PBL) depth. SST and 10-m wind speed are taken from the ERA-Interim reanalysis on a  $1^\circ \times 1^\circ$  latitude-longitude grid (Dee et al., 2011). EIS is calculated following Wood and Bretherton (2006) using fields from ERA-Interim. Following Eastman et al. (2017), the PBL depth is estimated from the difference between the SST and cloud top temperature (CTT) along with a parametrized lapse rate (Wood & Bretherton, 2004). Here CTT is sourced from the MODIS Aqua joint CTT histograms on the  $1^\circ \times 1^\circ$  L3 grid (King et al., 2003). The CTT is corrected for partial cloudiness when cloud amount is below 90% (Eastman et al., 2017). PBL depths are only estimated when cloud amount is greater than 30%. All data are interpolated to match the AMSR-E data grid and regrouped into  $16 \times 16$  grid boxes.

## 3. Mesoscale Decomposition Using Wavelet Analysis

We apply a 2D discrete wavelet transform (DWT) as a band-pass filter bank in space to filter out synoptic-scale variability of CWP, WVP, and RR, and to segregate the gridded data into mesoscale wavelet octaves. To demonstrate the application of 2D DWT, we present a  $10^\circ \times 10^\circ$  Aqua MODIS scene over the SE Pacific on August 11, 2008 that includes various cloud types and scales (Figure 1). A  $16 \times 16$  AMSR-E grid box ( $\sim 160 \times 160$  km) in a region of closed cells is highlighted in red. We begin with applying 2D DWT to the CWP field in the highlighted region (Figure 2a).



**Figure 2.** (a) Advanced Microwave Scanning Radiometer for Earth Observing System measured cloud water path (CWP) in the example region shown in the red box of Figure 1. The white strips mark the border of the box and neighboring  $16 \times 16$  pixel regions. (b–e) CWP fluctuations decomposed into wavelength octaves of (b) 20, (c) 40, (d) 80, and (e) 160 km. CWP variances of the four wavelength octaves are indicated on the upper right of each panel.

### 3.1. Key Aspects of 2D DWT

The wavelet transform is a powerful mathematical tool for analysis of multi-scale features based on an octave band decomposition of the space-wavelength plane (Kumar & Foufoula-Georgiou, 1997). We use a 2D multilevel DWT, mathematically described in Mallat (1989a, 1989b), Daubechies (1988), Daubechies and Bates (1993), and Meyer (1992). This partitions a spatial field into “details” (expressing variability of the field) at multiple “levels” (scale octaves)  $p = 1, \dots, P$ , by successive application of wavelet transforms of levels 1, 2, ... $P$ . The maximum level  $P$  is chosen by the user based on the largest scales of interest for the analysis.

Here we apply a 2D DWT of the data with a nominal 10 km grid scale to segregate the gridded microwave-derived CWP into  $P = 4$  octaves characterizing spatial variability at mesoscale wavelengths. These octaves cover wavelengths in the broad ranges of  $20(2)^{-1/2} - 20 \times 2^{1/2}$  km,  $40(2)^{-1/2} - 40 \times 2^{1/2}$  km,  $80(2)^{-1/2} - 80 \times 2^{1/2}$  km,  $160(2)^{-1/2} - 160 \times 2^{1/2}$  km. We refer to these octaves by the weighted averages of their wavelength ranges, which are 20, 40, 80, and 160 km, respectively, as derived in Appendix A.

In each wavelength octave, the wavelet transform computes three “details” (representing variability in the  $x$ ,  $y$ , and diagonal direction respectively) and an “average” representing variability at coarser scales, which can iteratively be decomposed using wavelet transforms of higher wavelength octaves. In this study, we apply this process successively to levels 1 (20 km wavelength) to 4 (160 km wavelength). In the  $16 \times 16$  grid box of interest (Figure 2a), there are 256 wavelet coefficients:  $3(2^3)^2$  level 1 (20 km wavelength) details,  $3(2^2)^2$  level 2 (40 km wavelength) details,  $3 \cdot 2^2$  level 3 (80 km wavelength) details, 3 level 4 (160 km wavelength) details, and one level 4 average. The level 4 average and the details from levels 1 to 4 can be concatenated into a wavelet-transformed matrix with the same size as the original data.

It is straightforward to invert the 2D wavelet transform to recreate the original field from this matrix. The field can be filtered to a particular scale level  $p$  by zeroing out all of the coefficients of the wavelet-transformed matrix except for the details of level  $p$ ; the same principle can be applied to efficiently band-pass filter the field to a range of octave scales from  $p_1$  to  $p_2$ . In this study, we use a 2D inverse wavelet transform to obtain the local fluctuations of these fields at the corresponding four scales or combinations thereof, the wavelet-reconstructed CWP contributions in the highlighted  $16 \times 16$  box from level 1 (20 km wavelength) up to level 4 (160 km wavelength) are shown in Figures 2b–2e.

The variance of the field can also be wavelet-decomposed by scale level using the squares of the details. This decomposition is approximately spatially local, so that within an individual  $2^p \times 2^p$  block of grid points, the variance at scale  $p < P$  is the squared sum of the level- $p$  details in that block. In this study, we calculate the CWP variances of the four wavelength octaves in each  $16 \times 16$  box, and we identify the cell scale as the wavelength of the peak variance in this spectrum. In the highlighted box, CWP variance is the largest at level 3 (80 km wavelength), hence the cloud scale in this box is determined as 80 km, consistent with visual inspection of Figure 2a. Based on the RMS sampling errors cited in Section 2.2, only a CWP standard deviation exceeding 0.017 mm can be detected above the background noise of the microwave retrieval for MCC cells, so only  $16 \times 16$  blocks meeting this criterion are retained in the study. This leaves 29,042 closed cell cases and 31,083 open cell cases, which are the majority of the available closed and open scenes identified by the MCC classifier.

A similar analysis could be done using windowed Fourier analysis but wavelet analysis is simpler and much more numerically efficient. For a field such as we see in MCC, with variability across a range of wavelengths rather than a single wavelength that is constant across many cell widths, the natural octave partitioning of scales provided by wavelet analysis provides a spatially localized estimate of cell wavelength that is just as reliable as a Fourier-based method.

### 3.2. Choice of a Suitable Wavelet

We use Daubechies’ orthogonal wavelet (Daubechies, 1988; hereafter  $dbN$  where the index number  $N$  refers to the number of coefficients (filter length), or the order of wavelet) as our analyzing mother wavelet. The  $dbN$  wavelet is designed to zero-out all smooth polynomial functions of degree less than  $N$ , so they only contribute to average coefficients but not detail coefficients. The choice of wavelet order is a trade-off between two desirable characteristics: “locality” (keeping sharp discontinuities localized to details in just

a few spatial blocks, and minimizing artifacts near the edge of the spatial domain) and “spectral accuracy” (making the details into more accurate octave band-pass filters of the original spatial field with less spectral leakage). As the order  $N$  increases, better spectral accuracy is obtained at the expense of less locality. In this study, the  $db3$  wavelet is adopted, since it ensures details are not contaminated by large-scale linear and quadratic variations prominent in fields such as WVP, yet it is sharp enough to capture abrupt variations within individual mesoscale cells. Note that the wavelet-reconstructed field using  $db3$  uses data from neighboring boxes as well as the box of interest.

### 3.3. Application of 2D DWT to the Satellite Microwave Data Set

Figures 3a–3c shows the CWP, WVP, and RR over the same region as in the MODIS scene of Figure 1. Clouds with various scales as seen on the MODIS image are fairly well detected by AMSR-E CWP, although it is difficult to distinguish the boundary between closed and open cells that is obvious in the MODIS scene. The thicker clouds (high CWP) to the southwest and the center of the region are associated with more higher RR than their northeast counterpart (Figure 3c). Figure 3d indicates that the southwest region has patches of open cells with cell scales (wavelengths) of 20 or 40 km. The center region is filled with patches of closed cells with scales ranging from 20 to 80 km. The north and northeast regions contain thin clouds with predominantly closed-cell organization as identified by the MCC classification algorithm. The CWP variability in this region is only barely detectable by AMSR-E (Figure 3a), even though it is obvious in the MODIS visible image (Figure 1). Hereafter, we focus on closed and open cell cases only, for which the classifications are the most reliable.

Boxes labeled with 160 km scale should be interpreted with caution, since they can also include cells with wavelengths greater than 160 km and transitions between closed cells and open cells (or clear sky) that are falsely identified by the NN algorithm as large closed cells. For instance, in the  $10^\circ \times 10^\circ$  scene presented in Figures 1 and 3, the 160 km scale MCCs are detected in the northeast region (around  $84^\circ\text{W}$ ,  $14^\circ\text{S}$ ;  $81^\circ\text{W}$ ,  $10^\circ\text{S}$ ;  $80^\circ\text{W}$ ,  $16^\circ\text{S}$ ) where there are mesoscale patches of clear sky embedded in the closed cell region.

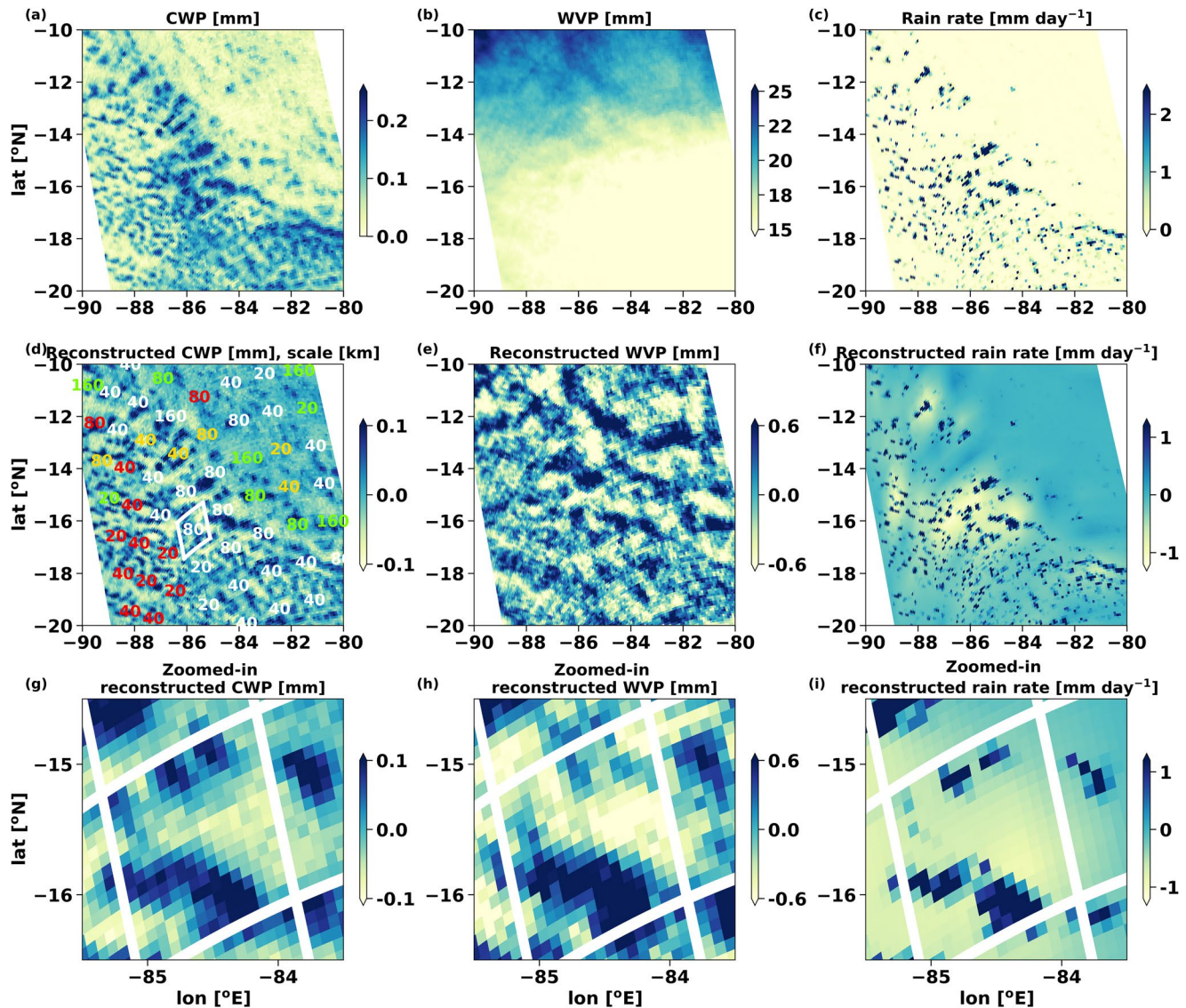
There is a latitudinal gradient of column moisture with the more tropical northern region  $\sim 8$  mm moister than the south (Figure 3b). This likely comes primarily from the free troposphere, with some contribution from gradients in boundary-layer depth and humidity. The mesoscale variability of WVP is more clearly seen in Figure 3e, which shows its reconstruction from an inverse wavelet transform of detail levels 1–4, which filters out synoptic scale variability. The reconstructed fluctuation of WVP correlates fairly well with that of CWP (Figures 3d and 3f), and with that of RR when it is present (Figure 3f), except that there is considerable WVP variability at 160 km and WVP noise at 20 km in the northeast region despite the weak cellularity in CWP.

The same example box shown in red in Figures 1 and 2 is indicated in white in Figure 3d. Figures 3g–3i show the reconstructed CWP, WVP, and RR fluctuations in and around this box of  $16 \times 16$  AMSR-E pixels. Combining the four wavelength octaves reliably reproduces the multi-scale cloud pattern as shown in Figure 2a. In this box, the reconstructed WVP pattern closely resembles that of reconstructed CWP but with a larger fluctuation amplitude. The scale of reconstructed RR fluctuation follows that of CWP but with much more concentrated peaks (Figure 3i), implying a non-linear relationship between CWP and RR as will be discussed in Section 5.

Figures 1–3 show that the  $db3$  wavelet transform is capable of decomposing and reconstructing the multi-scale variability of CWP, WVP, and RR. A lower order wavelet such as  $db1$  does not adequately remove large-scale WVP gradients, so it proves unsuitable for this analysis.

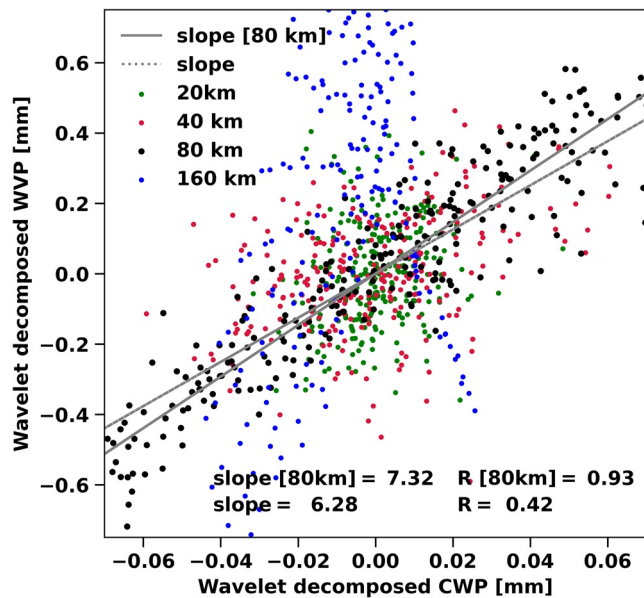
## 4. Dependence of MCC Properties on Cellular Scale

Wavelet decomposition enables the isolation of mesoscale CWP and WVP fluctuations at each wavelength octave. Zhou and Bretherton (2019a, 2019b) suggested that the growth and maintenance of MCC is tightly connected to mesoscale moisture anomalies, so one would expect a strong positive correlation between WVP and CWP when filtered to the dominant MCC scale (80 km in this box). Figure 4 shows a scatterplot of wavelet-decomposed WVP versus CWP fluctuations of the four wavelength octaves (20, 40, 80, and 160 km)



**Figure 3.** (a) The Advanced Microwave Scanning Radiometer for Earth Observing System retrieved (a) cloud water path (CWP), (b) water vapor path (WVP), and (c) rain rate over the same scene as in Figure 1. The white gaps at the bottom left corners fall outside of the swath; those at the upper right corners are the result of the data truncation. High-pass (d) CWP, (e) WVP, and (f) rain rate reconstructed from the detail coefficients of wavelength octaves 20–160 km. The computed cloud scales in each  $16 \times 16$  grid ( $\sim 160 \times 160$  km) box are indicated in (d). White, red, yellow, and green indicate closed, open, disorganized, and mixed pattern MCC respectively. In (d), the example closed-cell box from Figure 1 is indicated in white. (g–i) Same as (d–f) but zoomed into the example closed-cell box. The white strips mark the border of the box and its neighbors.

in the  $16 \times 16$  example box (256 grid-points) shown in Figure 1. Indeed, in this box, chosen as a particularly nice example, the decomposed WVP fluctuation shows a strong linear relationship with its CWP counterpart at the diagnosed MCC scale of 80 km (level 3) with a correlation coefficient of 0.92 and a slope near 6. This correlation slope is very close to the median slope for all closed cells of 80 km scale (6.7) as will be shown in Figure 5. The correlations at the other three wavelengths (20, 40, and 160 km) are more scattered. Although there are 256 grid-points, the octave-filtering reduces the effective degrees of freedom, which are controlled by the number of detail coefficients that contribute to each plot (192, 48, 12, and 3 for levels 1–4, respectively). Even so, the WVP-CWP relationship at 80 km is striking and drives most of the linear relation between WVP and CWP filtered to all four levels combined (compare the dashed line in Figure 4 with the solid line).



**Figure 4.** Scatterplots of wavelet-decomposed cloud water path and water vapor path fluctuations for octaves of 20 (green), 40 (red), 80 (black), and 160 km (blue). Least squares fits through the origin (0, 0) between fluctuations at wavelength 80 km and between fluctuations at all wavelengths are shown as gray solid and dashed lines, respectively. Their slopes and correlation coefficients are shown near the bottom of the plot.

lowest CWP standard deviations for closed and open cells occur at 20 km, where they are usually comparable to the noise RMS of 0.017 mm, making the MCC difficult to distinguish from a quasi-homogeneous cloudy layer with microwave data, even if it is evident on visible imagery. Overall, 20 km MCC scale usually occurs in cloud layers with low CWP and weak precipitation.

The CWP standard deviation is slightly higher in open than closed cells of all scales, implying that open cells are more inhomogeneous and are locally thicker. This generates patches of more intense precipitation which leads to the higher area-mean RR noted in Figure 5c.

The WVP standard deviation has a much redder distribution across scales than the CWP standard deviation. It is significantly less than the detection threshold of 0.57 mm for MCC scales of 20 and 40 km. Indeed, since the AMSR-E WVP has a 27 km footprint (Section 2.2), WVP perturbations at 20 km scale are meaningless, and at 40 km scale, they are still not well-resolved. WVP standard deviation increases at larger MCC scales, but this is not primarily a consequence of MCC for which the CWP standard deviation maintains similar (Figure 5d). In general, WVP variability is larger at longer wavelengths (e.g., 160 km), regardless of the MCC cell scale (not shown), suggesting that it is also responding to other large-scale dynamical processes in the free troposphere and boundary layer.

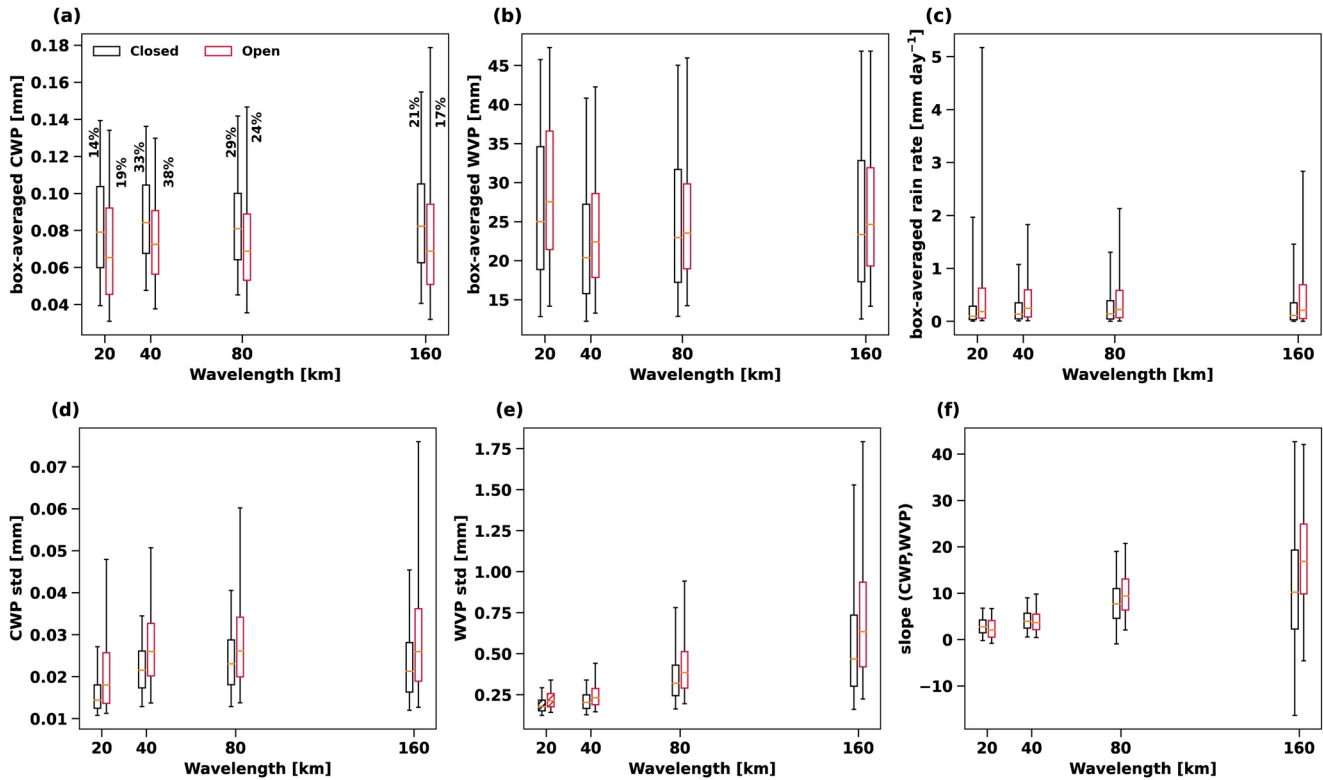
For each  $16 \times 16$  box, we perform a linear regression of the WVP perturbations on the CWP perturbations at the MCC scale, obtaining a regression slope for that box. We choose CWP as the predictor for this regression because its horizontal variability is more predominantly due to the MCC than for WVP. A box-whisker plot of the resulting slopes binned by the MCC scale is shown in Figure 5f. The regression slope  $dWVP'/dCWP'$  (WVP' and CWP' are the wavelet-filtered WVP and CWP fluctuations at a particular scale) is low ( $\sim 4$ ) at 20 and 40 km. As discussed above, the WVP perturbations are unreliable at those scales, but the slope is nevertheless usually positive. At an MCC scale of 80 km, the slope increases to  $\sim 8$ , which is very close to that found in numerical simulations by Zhou and Bretherton (2019a). The slope slightly increases to  $\sim 10$  at MCC scales of 160 km, but it is much more scattered, because the WVP perturbations are dominated by other processes and the CWP variability is not predominantly associated with individual mesoscale cells.

To get a more robust statistical perspective, we classify the closed and open cell properties by the dominant MCC scale (identified from CWP wavelet variance) within each  $16 \times 16$  box in our three regional data samples. Figure 5 presents this analysis using box-and-whisker plots separated by the MCC scale. The numbers along the whiskers in Figure 5a indicate the frequency of occurrence of closed and open cells of the different scales. Figure 5a shows that 40 km MCC scales occur most frequently. Since the CWP data has an  $14 \times 8$  km footprint, it is under-resolved at the 10 km pixel scale of our gridded AMSR-E data, so variability and frequency occurrence in the 20 km wavelength octave may be slightly underestimated. For WVP ( $27 \times 16$  km footprint), variability in the 20 km octave is primarily unphysical noise (hatched in Figure 5e), and variability in the 40 km octave may be underestimated. The RR retrieval has a footprint smaller than the pixel scale and does not suffer from these issues.

There is surprisingly little difference in the PDFs of box-mean CWP, WVP, and RR between boxes with MCC of scales 40, 80, and 160 km (Figures 5a–5c). For each MCC scale, the differences between the PDFs of box-mean CWP and box-mean WVP for closed versus open cells are modest but clear. Closed cells of each scale tend to have slightly higher mean CWP than open cells, but their area-mean RR tends to be lower than for open cells (as explained below). For 20 km MCC scale, the CWP and RR tend to be slightly lower compared to greater scales.

We filter the CWP and WVP data within each box to the MCC scale and perform statistics on those perturbations. The CWP standard deviation is similar for MCC scales of 40–160 km (Figure 5d), for which it corresponds to a typical CWP perturbation amplitude of 0.02–0.035 mm. The





**Figure 5.** The interquartile range (IQR; as indicated by the box plot) of the (a) cloud water path (CWP), (b) water vapor path (WVP), (c) rain rate, (d) CWP standard deviation, (e) WVP standard deviation, and (f) slope of the least squares linear fits of WVP fluctuations onto CWP fluctuations for closed and open MCC with different scales, across all 2008 data in the three study regions. The CWP standard deviation, WVP standard deviation, and the slope are computed from the wavelet-decomposed WVP and CWP fluctuations in the octave of the wavelength. The orange bar inside each box indicates a bin-median value and the whiskers indicate a range of 5th to 95th quartile. The frequency of occurrence of each IQR box is shown in (a). The hatching in (e) indicates that the WVP standard deviation in the 20 km octave is primarily unphysical (see text).

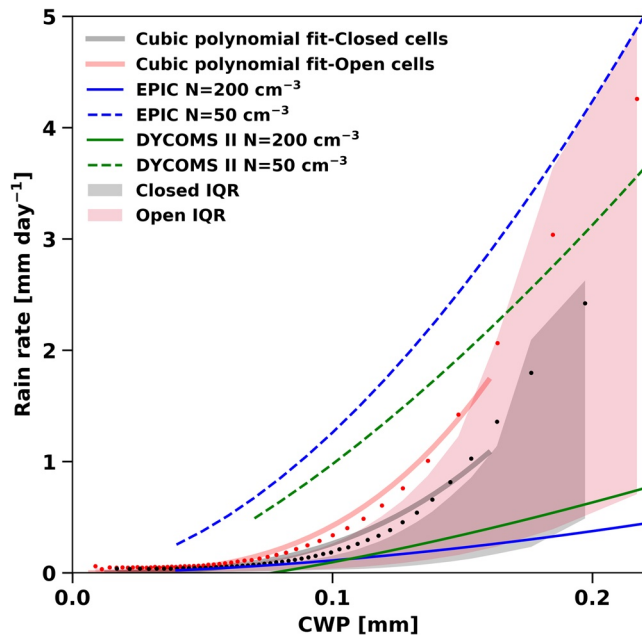
### 5. Nonlinear Relationship Between CWP and RR

Comparison of maps of CWP (e.g., Figure 3a) with RR (Figure 3c) suggest a strong but nonlinear relationship, with rain restricted to regions of high CWP as physically expected. Note that CWP and our RR estimate are derived from different microwave bands, so this is not an artifact of the sensors and retrieval. Figure 6 quantifies this relationship, showing the quartiles of CWP in bins of RR for AMSR-E  $\sim 10 \times 10$  km grid-points in closed and open MCC over the NE Pacific, SE Pacific, and SE Atlantic for 2008. The RR (in  $\text{mm day}^{-1}$ ) tends to increase cubically with CWP (in mm), with the composite fit shown in Figure 6:

$$\text{For closed cells: } RR_{\text{comp}} \sim 260 \cdot \text{CWP}^3 \quad (2.1)$$

$$\text{For open cells: } RR_{\text{comp}} \sim 420 \cdot \text{CWP}^3 \quad (2.2)$$

Figure 6 also presents the interquartile ranges of the RR-CWP relation. For both closed and open cells, the variability of RR increases with CWP. This could reflect the influence of cloud droplet number concentration ( $N_d$ ) on the RR-CWP relation, although there is no obvious way of testing this using the AMSR data alone. The variability of RR is broadly consistent with the nonlinear fitting envelopes derived from the ship-based remote sensing measurements during the Eastern Pacific Investigation of Climate Processes in the Coupled Ocean-Atmosphere System (EPIC; Comstock et al., 2004) between  $N_d = 50 \text{ cm}^{-3}$  and  $N_d = 200 \text{ cm}^{-3}$  (blue lines in Figure 6), and the envelopes from the flight-based remote sensing measurements during the second Dynamics and Chemistry of Marine Stratocumulus cloud experiment (DYCOMS II; Geoffroy et al., 2008; van Zanten et al., 2005) between  $N_d = 50 \text{ cm}^{-3}$  and  $N_d = 200 \text{ cm}^{-3}$  (green lines in Figure 6). The minimum and maximum  $N_d$  here are chosen based on the minimum and maxi-



**Figure 6.** Quartiles of cloud water path (CWP) versus rain rate for closed and open MCC clouds over NE Pacific, SE Pacific, and SE Atlantic for 2008. Dots indicate the mean rain rate within CWP bins (2%). The cubic polynomial fits through origin (0, 0) for the mean rain rate of closed and open cells are indicated by the solid red and gray lines, respectively. Shading indicates the interquartile range. The blue dashed and solid lines indicate the fits derived from the Eastern Pacific Investigation of Climate Processes in the Coupled Ocean-Atmosphere System (EPIC: Comstock et al., 2004) at cloud droplet number concentration of 50 and 200  $\text{cm}^{-3}$ , respectively. The green dashed and solid lines indicate the fits derived from the second Dynamics and Chemistry of Marine Stratocumulus cloud experiment (DYCOMS II; Geoffroy et al., 2008; van Zanten et al., 2005) at cloud droplet number concentration of 50 and 200  $\text{cm}^{-3}$ , respectively.

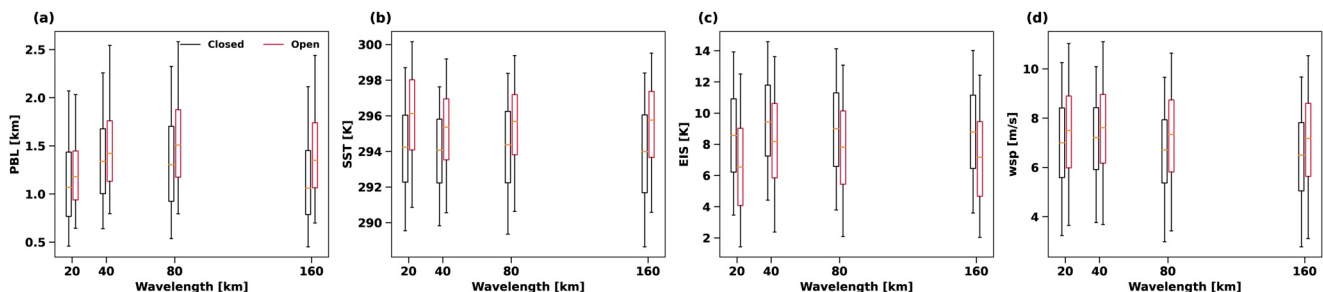
mum observed  $N_d$  during the field experiments (Geoffroy et al., 2008). Figure 6 shows that open cells create higher peak RRs compared to closed cells. This corroborates well-documented examples of pockets of open cells (e.g., Savic-Jovicic & Stevens, 2008; Wang & Feingold, 2009) in which a transition from closed to open cells is associated with dramatic precipitation-induced aerosol loss that causes much lower droplet concentrations in the pocket of open cells than in the surrounding overcast stratocumulus.

## 6. Environmental Control on MCC Scale

To investigate potential environmental controls on the MCC scale, we bin some candidate factors (PBL depth, SST, EIS, and 10 m wind speed) by the MCC scale for closed and open cells, using the same format as in Figure 5. Figure 7a shows that the bin-median PBL depth increases appreciably (from  $\sim 1$  to  $\sim 1.5$  km) from 20 to 40 km MCC, stays around 1.5 km for 80 km MCCs, and slightly lowers for 160 km scales. Our results corroborate Wood and Hartmann (2006)'s finding that small MCC tends to be associated with shallow PBLs, giving an approximate 40:1 cell aspect ratio, but this result breaks down for cells larger than 50 km wavelength. The slight decrease in PBL depth for 160 km scales may be an artifact; what is classified as 160 km scale MCC is often a boundary between patches of smaller cells in cloudier and clearer air masses rather than a single cell (see Section 3.2). The impacts of SST, EIS, and 10 m wind on the MCC scale are marginal (Figures 7b–7d). Furthermore, we found no significant geographical control on the cloud scale in the three regions (Figures S1 and S2).

Environmental controlling factors appear to be more important in distinguishing closed and open cells than in setting their scale. At all MCC scales, open cells are associated with deeper PBL, warmer SST, lower EIS and higher 10 m wind speed (Figures 7a–7d), suggesting that open cells tend to occur over warmer ocean farther away from the coast (as in Agee et al., 1973) where inversion strength is weaker (as in McCoy et al., 2017),

the boundary layer is deeper (as in Wood & Hartmann, 2006), and the trade winds are stronger. Open cells tend to also be associated with higher 10-m wind speed compared to closed cells at all MCC scales (Figure 7d), although the majority of the data between closed and open cells heavily overlap. This is consistent with Eastman et al. (2021), who found that open cells tend to occur in a higher wind-speed environment. Since warm SST and strong winds are primary drivers of surface fluxes, it is likely that closed to open cells transition is inclined to be facilitated in strong surface flux conditions. Unsurprisingly, open cells favor locations further downwind into the climatological transition from stratocumulus to cumulus (Figure S2), as also found by Wood and Hartmann (2006), Muhlbauer et al. (2014), and McCoy et al. (2017).



**Figure 7.** Same as Figure 5 but for the (a) planetary boundary layer height, (b) Sea surface temperature, (c) estimated inversion strength, and (d) 10-m wind speed.

## 7. Mesoscale Cell Composite

So far we have shown that within cells, CWP is correlated to RR and its background mesoscale-filtered WVP. Using the derived slope  $S$  of WVP fluctuations regressed onto CWP fluctuations in the  $16 \times 16$  pixel boxes, a WVP composite ( $WVP_{\text{comp}}$ ) is obtained as:

$$WVP_{\text{comp}} = S \cdot CWP' + \overline{WVP} \quad (3)$$

Here,  $CWP'$  is a representative CWP fluctuation across a cell, and  $\overline{WVP}$  is the domain-averaged WVP. We can also compute a RR composite using the cubic polynomial relation (Equation 2) between RR and CWP. This allows a “composite” WVP and RR for MCC of a particular scale to be estimated solely from typical CWP variations at that scale. We use a scale of 80 km as an example since it typically has large CWP variance, and it is well above the AMSR-E footprint resolution of CWP and WVP, and hence provides trustworthy CWP and WVP variability and co-variability.

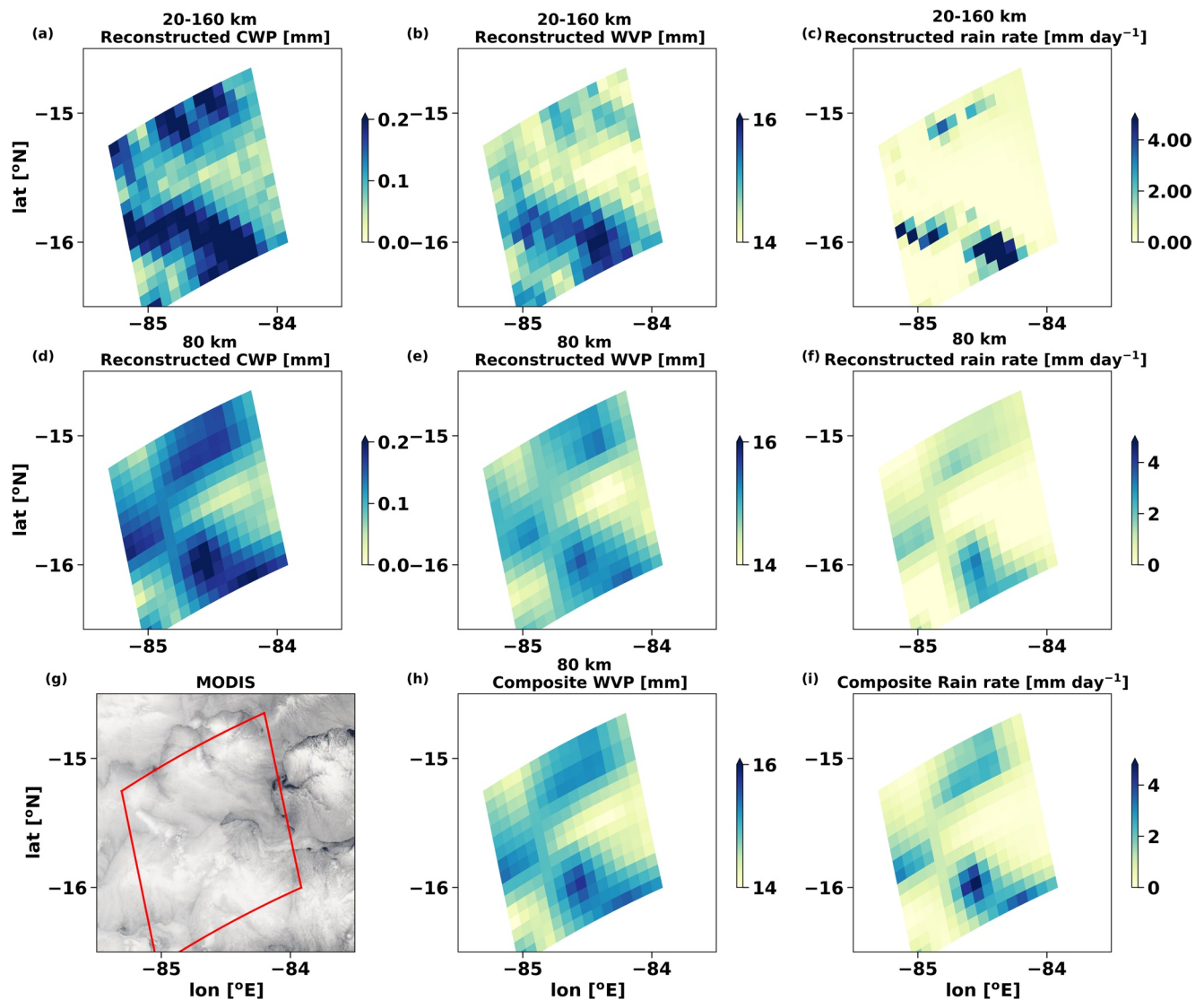
The WVP and RR composites are not just statistical summaries; they can also be good representations of WVP and RR for individual cases. Figure 8 shows a closed cell case (the same as the example box in Figure 1). The 20–160 km reconstructed CWP, WVP, and RR in Figures 8a–8c are computed as the sum of their box means and their wavelet-filtered fluctuations at the four finest octave scales. These realistically represent the cloud, humidity, and RR variability within the example box. The reconstructed CWP, WVP, and RR at 80 km in Figures 8d–8f are computed as the sum of the 80 km wavelet-filtered fluctuations and their box means. Comparing Figure 8a with Figure 8d shows that for this particularly clean closed-cell case, the 80 km reconstructed CWP, WVP, and RR are good representations of their overall mesoscale variations within the  $16 \times 16$  example box.

The WVP composite in Figure 8h is computed following Equation 3, using the median correlation slope derived for the 80-km closed cells (8, as shown in Figure 5f), and using the 80 km CWP perturbation in Figure 8d as a representative “composite” CWP for this typical closed cell. The composite WVP compares qualitatively well with the actual 80-km reconstructed WVP for the closed cell shown in Figure 8e. The RR composite in Figure 8i is computed from the 80 km reconstructed CWP using Equation 2; again it agrees well with the actual 80 km reconstructed RR for this case.

The correlations between reconstructed and composite WVPs and RRs are equally strong for open cells. An example of a box with 80-km scale open cell convection observed over the SE Pacific region on October 19, 2008 is shown in Figure 9, using the same format as in Figure 8. The polygonal patterns of cumulus characteristic open-cell organization are obvious on the MODIS image but not clear in the lower-resolution microwave data, even in the full 20–160 km reconstructed fields. The magnitudes of the CWP variations are similar to the closed-cell example. Both the reconstructed 80 km variability and the “composite” variability in WVP and RR also look qualitatively similar to the closed-cell example. The open-cell composite uses a slightly larger regression slope  $S = 9$  to relate 80 km WVP perturbations to CWP perturbations than for closed cells, based on Figure 5f.

Figures 8 and 9 demonstrate that the 80-km scale WVP and RR perturbations for a typical 80 km open or closed cell can be estimated from its 80 km CWP variations. Here we build a statistical composite based on all 80-km scale closed and open cells over three regions observed in 2008. We use 80 km cells because they are both common and well-resolved by all the AMSR-E retrievals, including WVP. We build idealized sinusoidal CWP and WVP sections across a typical 80-km scale closed and open cell. For each cell type, the wave mean is chosen as the median box mean and the wave amplitude (denoted  $\Delta CWP$  or  $\Delta WVP$ ) is  $2^{1/2}$  times the median standard deviation (since the standard deviation of a uniformly sampled sinusoid of amplitude  $a$  is  $2^{-1/2}a$ ). The maximum and minimum of the sinusoidal CWP wave represent the cloud centers and cloud edges respectively.

The bottom part of Figure 10 shows the resulting closed and open cell composites. The median composites, shown as the solid red lines, are computed from Figure 5 as follows. Figure 5a implies the bin-mean CWP is 0.08 mm for 80 km closed cells and 0.07 mm for 80 km open cells. Figure 5d implies that the median standard deviation of CWP is 0.023 and 0.026 mm for 80 km closed and open cells, so the CWP sinusoid amplitude is 1.4 times as large—0.032 mm for closed cells and 0.036 mm for open cells. Thus, the CWP



**Figure 8.** 20–160 km wavelet-reconstructed (a) cloud water path (CWP), (b) water vapor path (WVP), and (c) rain rate. 80 km wavelet-reconstructed (d) CWP, (e) WVP, and (f) rain rate for the closed MCC of scale 80 km in the example box in Figure 1. The corresponding MODIS scene is shown in (g). Composite (h) WVP and (i) rain rate computed from the 80 km CWP using Equations 2 and 3.

ranges from 0.048 to 0.112 mm across the composite closed cell and 0.034–0.106 mm across the composite open cell. Note that because we are only considering variability at a single scale (80 km), we do not expect a minimum composite CWP = 0 (no cloud) even in the driest part of open cells, which are usually nearly cloud-free.

The red shading in Figure 10 indicates the range of CWP composites obtained by using the envelope of interquartile ranges of CWP standard deviation from the boxes in Figure 5d (0.018–0.029 mm for closed cells and 0.02–0.034 mm for open cells) in place of its median, showing expected natural variability of MCC around this composite. Another source of variability is the box-mean CWP, which has an interquartile range of 0.06–0.1 mm for closed cells and 0.05–0.09 mm for open cells that affects both the composite mean and perturbation RR. We represent this source of variability in the CWP composites by adding the 25th to 50th quartile range of the box mean above the top of the shading, and the 50th to 75th quartile range of the box mean below the bottom of the shading (dashed lines in Figure 10, shown only where they are above zero, indicating some cloud is present). The dashed lines enclose an envelope including interquartile ranges of both amplitudes and box means.

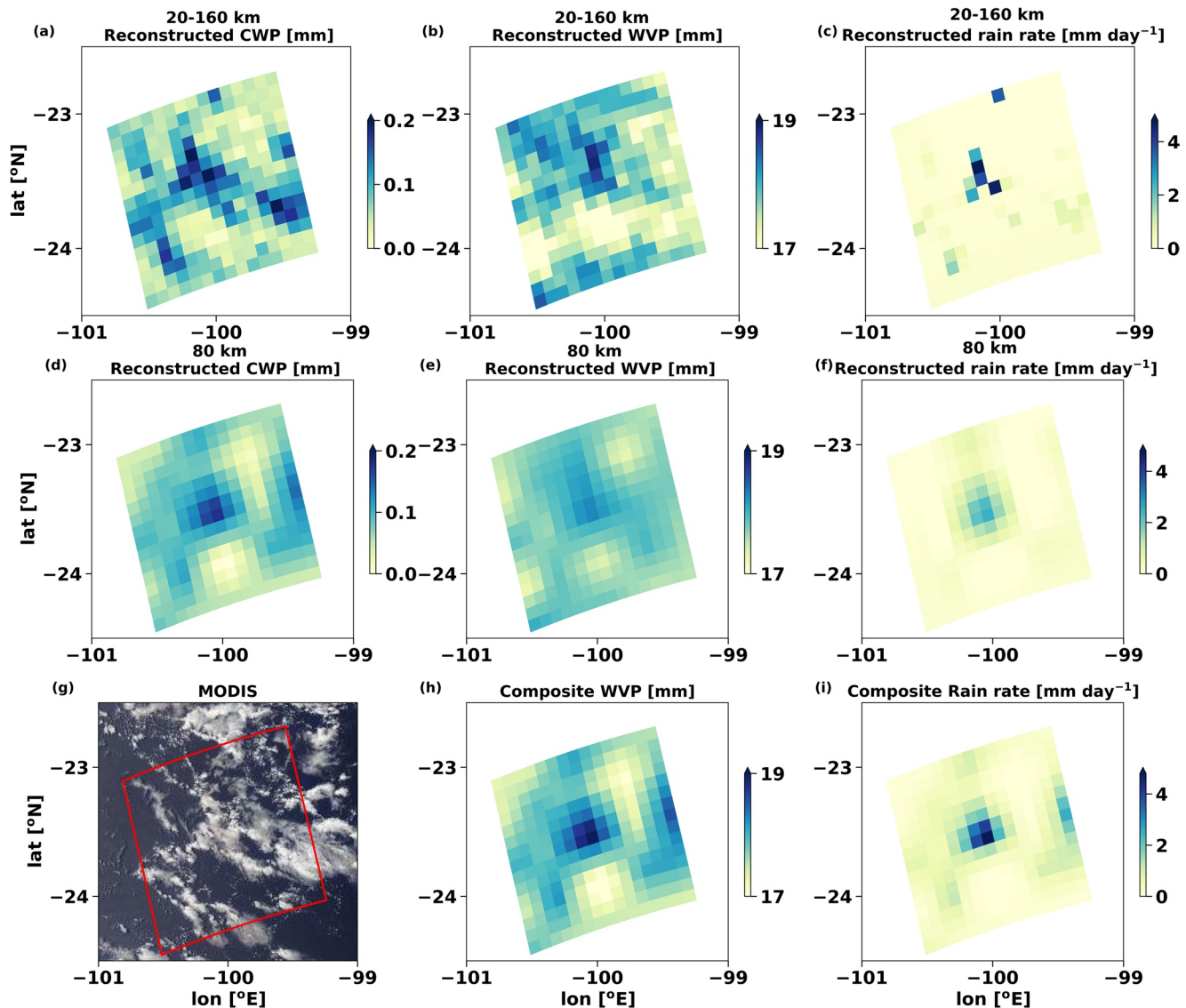
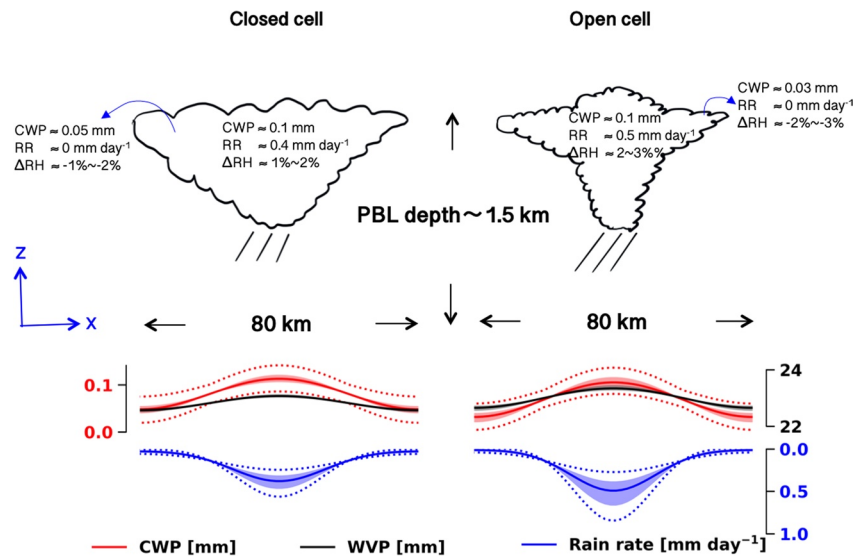


Figure 9. Same as Figure 8 but for an open cell case on October 19, 2008 over SE Pacific.

Multiplying the median CWP standard deviations by the median correlation slopes in Figure 5f (8 for 80 km closed cells and 9 for open cells), we get a median WVP standard deviation of 0.18 and 0.26 mm and hence sinusoid amplitudes of 0.25 and 0.4 mm for 80 km closed and open cells respectively. Adding the median of the box-mean WVP for each case gives the WVP composite (black curves). The natural variability of the WVP composite is marked by the gray shading, computed by multiplying the interquartile range of CWP standard deviations by the median correlation slopes in Figure 5f. The box-mean WVP varies so much across our sampling regions and seasons that we do not show this on Figure 10.

The RR composites are computed from the CWP composite following the cubic relation given in Equation 2. The median CWP sinusoid (red curve) generates the median rain-rate curve (blue curve), the red-shaded range of CWP sinusoids generates the blue-shaded range of RR, and the red dashed curve incorporating interquartile variability in both the CWP box-mean and cell-scale perturbation amplitude generates the blue dashed curve. The RR varies over a larger relative range than CWP due to its cubic dependence of CWP.

The CWP and RR are comparable in the thick part of the median open and closed cells, although there is a lot of variability around the median composites (Figure 10). The more intense open cells have walls



**Figure 10.** The median values (solid line) of Advanced Microwave Scanning Radiometer for Earth Observing System measured box-mean cloud water path (CWP), composite water vapor path (WVP) computed from Equation 3 at 80 km, and the composite rain rate computed from Equation 2 for closed and open cells. The shading indicates the envelope of interquartile ranges of amplitudes. The dashed lines mark a bigger envelope including interquartile ranges of both amplitudes and box means. The schematic of closed and open cells together with the median values of box-mean CWP and rain rate (RR) at the centers and edges of the cells are also shown. The relative humidity anomalies ( $\Delta RH$ ) are computed by assuming the saturation WVP in the PBL is 15 mm.

of thicker CWP with more intense precipitation than typically seen in closed cells (as seen from the largest shaded CWP and RRs in the bottom panel of Figure 10). The box-mean contribution to CWP is also an important source of variability, contributing especially to some cells raining a lot more than others (compare bottom and top blue dashed lines). In the thinnest part of the open cell CWP composites, the shaded region extends down nearly to CWP = 0 and when also considering the box mean variability, much of the drier portions of many open cells have CWP = 0 (cloud-free), consistent with visible satellite imagery.

The upper part of Figure 10 shows schematics for 80 km closed and open cell MCC. For each cell type, the CWP and RR at the cloud centers and edges in the schematics are computed from the maximum and minimum of the sinusoidal composites. The relative humidity perturbations ( $\Delta RH$ ) are computed as  $\Delta WVP / WVP_{sat}$ , where  $\Delta WVP$  is taken from the sinusoidal composite for that cell type and  $WVP_{sat}$  is the saturation WVP in the PBL (estimated as 15 mm or  $15 \text{ kg m}^{-2}$  of water, based on a representative PBL saturation mixing ratio of 10 g/kg and a PBL column dry air mass of  $1,500 \text{ kg m}^{-2}$ , corresponding to a 150 hPa pressure thickness). For closed and open cells,  $\Delta WVP$  is 0.2–0.3 mm and 0.25–0.43 mm respectively, so  $\Delta RH$  1%–2% for closed cells and 2%–3% for open cells. The sampled PBLs span a range of SSTs and depths, leading to a broad range of estimated saturation WVP of 10–20 mm. Across this range, we find MCC cell-scale perturbations of PBL  $\Delta RH \sim 1\%$ –4% for 80 km cells. This is why the composite is presented in terms of  $\Delta RH$  rather than  $\Delta WVP$ . On average, the median closed and open cell composites are nearly indistinguishable using only the AMSR-E data, even though open and closed cells are readily distinguishable in finer resolution MODIS visible imagery.

## 8. Discussion and Conclusion

We have applied a two-dimensional discrete wavelet transform to AMSR-E passive microwave data to study mesoscale cellular convection in subtropical marine low clouds and its relationship with environmental factors. The key AMSR-E fields used are CWP, WVP, and RR. The wavelet analysis partitions the variability and co-variability of these three fields within  $160 \times 160 \text{ km}$  subregions (“boxes”) of each data swath into

four octave scales, corresponding to representative wavelengths of 20, 40, 80, and 160 km. We use this partitioning to identify a dominant local scale of mesoscale CWP variability in each box, when it is present. The most common cell scales are 40–80 km. For cells of these scales, we find that cell-scale CWP perturbations are well correlated with WVP and RR perturbations, allowing us to create a composite cell structure. A machine learning scheme based on MODIS CWP is used to classify cells into open, closed or other, allowing us to compare the statistics and composite structure of open and closed cells.

The microwave composites are surprisingly similar for open and closed cells despite their clearly different appearance on the much finer resolution MODIS visible imagery. Closed cells have a slightly higher box-mean CWP and slightly less cell-scale CWP variability.

Our analysis suggests that MCC scales of 40 km and larger are uncorrelated with box-mean CWP, WVP, rain rate, or PBL depth, in contrast to a MODIS-based analysis by Wood and Hartmann (2006) showing a 40:1 ratio of cell scale to PBL depth. MCC of smaller 20 km scale usually occurs in a shallower boundary layer (~1 km) with clouds that are thinner, more homogeneous, and less drizzly. All cell scales have similar geographic distribution.

Our results indicate that SST, EIS, and 10-m wind speed do not play a major role in determining the scale of the MCC, which implies that the increase of the cell scale might stem more from its historical evolution than from the environmental conditions at the time of measurement. As in past studies (e.g., Agee et al., 1973; Eastman et al., 2021; McCoy et al., 2017; Muhlbauer et al., 2014; Wood & Hartmann, 2006), open cells favor somewhat deeper PBLs over warmer SST.

This study has several limitations. One of them is the broad  $27 \times 16$  km footprint of the AMSR-E WVP, which prevents us from resolving the moisture variability at small MCC wavelength (20 and 40 km). Future spaceborne differential absorption radar has the potential to provide column WVP at much higher spatial resolution than is currently possible (L. Millán et al., 2020). The WVP also cannot disentangle moisture variability in the boundary layer from that in the free troposphere, although multi-sensor studies combining MODIS and AMSR-E could be used in future to remove the free-tropospheric contribution to the column water vapor (L. F. Millán et al., 2019). The  $14 \times 8$  km CWP footprint is also much coarser than MODIS (~1 km) and cannot resolve narrow cloud bands typical of open cells (e.g., Figure 9). In addition, the WVP and (to a lesser extent) CWP retrievals are not sensitive enough to detect weak cellular variability and thin and/or broken cloud. Lastly, the current MCC classification is conservative in labeling some low-coverage closed cells as open cells. This might partially explain the resemblance of the statistics between closed and open cells in this study. It might be interesting to explore the use of this wavelet methodology on MODIS-based visible wavelength retrievals of CWP matched to the microwave data. It might also be interesting to use similar wavelet-based techniques on MCC associated with extratropical cyclones in the storm tracks to understand the scale-related problems.

## Appendix A: Computation of Octave Wavelengths

Consider a process sampled on a 2D grid with spacing  $(dx, dy = dx)$ . Let  $k_n = \pi / dx$  be the Nyquist wavenumber for this grid spacing. Then the range of 2D wavenumber space  $(k_x, k_y)$  is  $0 < |k_x|, |k_y| < k_n$ .

The level  $N$  wavelet details cover the wavenumber range of  $R_N: k_n / 2^N < |k_x|, |k_y| < k_n / 2^{N-1}$ .  $R_N$  consists of the square region  $S_1: |k_x|, |k_y| < k_n / 2^{N-1}$  with the square  $S_2: |k_x|, |k_y| > k_n / 2^N$  removed.

Let angle brackets  $\langle \cdot \rangle_{R_N}$  denote an area-integral over region  $R_N$ , and similarly for the regions  $S_1$  and  $S_2$ . Let  $A(R_N)$  denote the area of  $R_N$  and similarly for the other regions. Then the mean square wavenumber  $K_N^2 = k_x^2 + k_y^2$  sampled by the level  $N$  details is

$$K_N^2 = \frac{\langle k_x^2 + k_y^2 \rangle_{R_N}}{A(R_N)}$$

Here

$$A(R_N) = A(S_1) - A(S_2) = \left(\frac{2k_n}{2^{N-1}}\right)^2 - \left(\frac{2k_n}{2^N}\right)^2 = \frac{3k_n^2}{(2^{N-1})^2},$$

and

$$\langle k_x^2 + k_y^2 \rangle_{R_N} = \langle k_x^2 + k_y^2 \rangle_{S_1} - \langle k_x^2 + k_y^2 \rangle_{S_2} = \langle k_x^2 + k_y^2 \rangle_{S_1} (1 - 2^{-4})$$

Now

$$\langle k_x^2 + k_y^2 \rangle_{S_1} = \frac{\frac{k_n}{2^{N-1}}}{\frac{k_n}{2^{N-1}}} \iint (k_x^2 + k_y^2) dk_x dk_y$$

$$= 2 \int_{\frac{k_n}{2^{N-1}}}^{\frac{k_n}{2^{N-1}}} k_x^2 dk_x \int_{\frac{k_n}{2^{N-1}}}^{\frac{k_n}{2^{N-1}}} dk_y$$

$$= \frac{2}{3} k_x^3 \Big|_{\frac{k_n}{2^{N-1}}}^{\frac{k_n}{2^{N-1}}} \frac{2k_n}{2^{N-1}}$$

$$= \frac{8k_n^4}{3(2^{N-1})^4}$$

$$K_N^2 = \frac{8k_n^4}{3(2^{N-1})^4} \frac{15(2^{N-1})^2}{16 \cdot 3k_n^2} = \frac{5}{6} \frac{k_n^2}{(2^{N-1})^2}$$

or

$$K_N = \frac{k_n}{2^{N-1}} \left(\frac{5}{6}\right)^{1/2} \approx \frac{k_n}{2^{N-1}}$$

### Data Availability Statement

The mesoscale cellular convection identification data are processed using MODIS collection 6 L2 data ([dx.doi.org/10.5067/MODIS/MYD06\\_L2.006](https://dx.doi.org/10.5067/MODIS/MYD06_L2.006)). The light rain rate retrieval data are derived from AMSR-E 89 GHz brightness temperatures (<https://doi.org/10.5067/YL62FUZLAJUT>) and CloudSat Rain-Profile (<http://www.cloudsat.cira.colostate.edu/data-products/level-2c/2c-rain-profile>). The environmental variables in this study are derived from ERA-Interim reanalysis (<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim>) and MODIS L3 cloud top temperature (<https://atmosphere-imager.gsfc.nasa.gov/products/daily>).



**Acknowledgments**

We gratefully acknowledge funding from DOE Atmospheric System Research grant DE-SC0020134, NSF grant AGS-1660609, and NASA grant 80NS-SC19K1274. We thank Dr. Ping Yang and three anonymous reviewers for numerous suggestions and comments. We acknowledge the Aqua L2B global swath ocean products of AMSR-E from the National Snow and Ice Data Center (NSIDC) at [https://nsidc.org/data/AE\\_Ocean/versions/2](https://nsidc.org/data/AE_Ocean/versions/2).

**References**

Abel, S. J., Barrett, P. A., Zuidema, P., Zhang, J., Christensen, M., Peers, F., et al. (2020). Open cells exhibit weaker entrainment of free-tropospheric biomass burning aerosol into the south-east Atlantic boundary layer. *Atmospheric Chemistry and Physics*, 20.

Abel, S. J., Boutle, I. A., Waite, K., Fox, S., Brown, P. R., Cotton, R., et al. (2017). The role of precipitation in controlling the transition from stratocumulus to cumulus clouds in a Northern Hemisphere cold-air outbreak. *Journal of the Atmospheric Sciences*, 74, 2293–2314. <https://doi.org/10.1175/jas-d-16-0362.1>

Agee, E. M. (1984). Observations from space and thermal convection: A historical perspective. *Bulletin of the American Meteorological Society*, 65, 938–949. [https://doi.org/10.1175/1520-0477\(1984\)065<0938:ofsatc>2.0.co;2](https://doi.org/10.1175/1520-0477(1984)065<0938:ofsatc>2.0.co;2)

Agee, E. M., Chen, T. S., & Dowell, K. E. (1973). A review of mesoscale cellular convection. *Bulletin of the American Meteorological Society*, 54, 1004–1012. [https://doi.org/10.1175/1520-0477\(1973\)054<1004:aromcc>2.0.co;2](https://doi.org/10.1175/1520-0477(1973)054<1004:aromcc>2.0.co;2)

Atkinson, B. W., & Zhang, J. W. (1996). Mesoscale shallow convection in the atmosphere. *Reviews of Geophysics*, 34, 403–431. <https://doi.org/10.1029/96rg02623>

Berner, A. H., Bretherton, C. S., Wood, R., & Muhlbauer, A. (2013). Marine boundary layer cloud regimes and POC formation in a CRM coupled to a bulk aerosol scheme. *Atmospheric Chemistry and Physics*, 13. <https://doi.org/10.5194/acp-13-12549-2013>

Boutle, I. A., & Abel, S. J. (2012). Microphysical controls on the stratocumulus topped boundary-layer structure during VOCALS-REX. *Atmospheric Chemistry and Physics*, 12, 2849–2863. <https://doi.org/10.5194/acp-12-2849-2012>

Bretherton, C. S., & Wyant, M. C. (1997). Moisture transport, lower-tropospheric stability, and decoupling of cloud-topped boundary layers. *Journal of the Atmospheric Sciences*, 54, 148–167. [https://doi.org/10.1175/1520-0469\(1997\)054<0148:mtltsa>2.0.co;2](https://doi.org/10.1175/1520-0469(1997)054<0148:mtltsa>2.0.co;2)

Comstock, K. K., Wood, R., Yuter, S. E., & Bretherton, C. S. (2004). Reflectivity and rain rate in and below drizzling stratocumulus. *Quarterly Journal of the Royal Meteorological Society*, 130, 2891–2918. <https://doi.org/10.1256/qj.03.187>

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41, 909–996. <https://doi.org/10.1002/cpa.3160410705>

Daubechies, I., & Bates, B. J. (1993). *Ten lectures on wavelets*.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597. <https://doi.org/10.1002/qj.828>

Eastman, R., Lebsock, M., & Wood, R. (2019). Warm rain rates from AMSR-E 89-GHz brightness temperatures trained using CloudSat Rain-Rate observations. *Journal of Atmospheric and Oceanic Technology*, 36, 1033–1051. <https://doi.org/10.1175/jtech-d-18-0185.1>

Eastman, R., McCoy, I. L., & Wood, R. (2021). Environmental and internal controls on Lagrangian transitions from closed cell mesoscale cellular convection over subtropical oceans. *Journal of the Atmospheric Sciences*. <https://doi.org/10.1175/JAS-D-20-0277.1>

Eastman, R., & Wood, R. (2016). Factors controlling low-cloud evolution over the eastern subtropical oceans: A Lagrangian perspective using the A-Train satellites. *Journal of the Atmospheric Sciences*, 73, 331–351. <https://doi.org/10.1175/jas-d-15-0193.1>

Eastman, R., Wood, R., & Ting O, K. (2017). The subtropical stratocumulus-topped planetary boundary layer: A climatology and the Lagrangian evolution. *Journal of the Atmospheric Sciences*, 74, 2633–2656. <https://doi.org/10.1175/jas-d-16-0336.1>

Field, P. R., Cotton, R. J., McBeath, K., Lock, A. P., Webster, S., & Allan, R. P. (2014). Improving a convection-permitting model simulation of a cold air outbreak. *Quarterly Journal of the Royal Meteorological Society*, 140, 124–138. <https://doi.org/10.1002/qj.2116>

Fletcher, J., Mason, S., & Jakob, C. (2016). The climatology, meteorology, and boundary layer structure of marine cold air outbreaks in both hemispheres. *Journal of Climate*, 29, 1999–2014. <https://doi.org/10.1175/jcli-d-15-0268.1>

Geoffroy, O., Brenguier, J. L., & Sandu, I. (2008). Relationship between drizzle rate, liquid water path and droplet concentration at the scale of a stratocumulus cloud system. *Atmospheric Chemistry and Physics*, 8, 4641–4654. <https://doi.org/10.5194/acp-8-4641-2008>

Goren, T., & Rosenfeld, D. (2012). Satellite observations of ship emission induced transitions from broken to closed cell marine stratocumulus over large areas. *Journal of Geophysical Research - D: Atmospheres*, 117. <https://doi.org/10.1029/2012jd018878>

Kawanishi, T., Sezai, T., Ito, Y., Imaoka, K., Takeshima, T., Ishido, Y., et al. (2003). The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), NASA's contribution to the EOS for global energy and water cycle studies. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 184–194. <https://doi.org/10.1109/tgrs.2002.808331>

King, M. D., Kaufman, Y. J., Menzel, W. P., & Tanre, D. (1992). Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS). *IEEE Transactions on Geoscience and Remote Sensing*, 30, 2–27. <https://doi.org/10.1109/36.124212>

King, M. D., Menzel, W. P., Kaufman, Y. J., Tanré, D., Gao, B. C., Platnick, S., et al. (2003). Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 442–458. <https://doi.org/10.1109/tgrs.2002.808226>

Kumar, P., & Fofoula-Georgiou, E. (1997). Wavelet analysis for geophysical applications. *Reviews of Geophysics*, 35, 385–412. <https://doi.org/10.1029/97rg00427>

Kummerow, C., Ferraro, R., & Randel, D. (2015). *AMSR-E/Aqua L2B global swath surface precipitation GSFC profiling algorithm*.

Lau, K. M., & Weng, H. (1995). Climate signal detection using wavelet transform: How to make a time series sing. *Bulletin of the American Meteorological Society*, 76(12), 2391–2402. [https://doi.org/10.1175/1520-0477\(1995\)076<2391:csduwt>2.0.co;2](https://doi.org/10.1175/1520-0477(1995)076<2391:csduwt>2.0.co;2)

Lebsock, M., & Su, H. (2014). Application of active spaceborne remote sensing for understanding biases between passive cloud water path retrievals. *Journal of Geophysical Research - D: Atmospheres*, 119, 8962–8979. <https://doi.org/10.1002/2014jd021568>

Mallat, S. G. (1989a). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 674–693. <https://doi.org/10.1109/34.192463>

Mallat, S. G. (1989b). Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on Acoustics*, 37, 2091–2110. <https://doi.org/10.1109/29.45554>

McCoy, I. L., Wood, R., & Fletcher, J. K. (2017). Identifying meteorological controls on open and closed mesoscale cellular convection associated with marine cold air outbreaks. *Journal of Geophysical Research - D: Atmospheres*, 122, 11–678. <https://doi.org/10.1002/2017jd027031>

Meyer, Y. (1992). *Wavelets and Operators: Volume 1 (Vol. 37)*. Cambridge University Press.

Millán, L., Roy, R., & Lebsock, M. (2020). Assessment of global total column water vapor sounding using a spaceborne differential absorption radar. *Atmospheric Measurement Techniques*, 13, 5193–5205. <https://doi.org/10.5194/amt-13-5193-2020>

Millán, L. F., Lebsock, M. D., & Teixeira, J. (2019). Variability of bulk water vapor content in the marine cloudy boundary layers from microwave and near-infrared imagery. *Atmospheric Chemistry and Physics*, 19, 8491–8502. <https://doi.org/10.5194/acp-19-8491-2019>

Muhlbauer, A., McCoy, I. L., & Wood, R. (2014). Climatology of stratocumulus cloud morphologies: Microphysical properties and radiative effects. *Atmospheric Chemistry and Physics*, 14. <https://doi.org/10.5194/acp-14-6695-2014>

- Rothermel, J., & Agee, E. M. (1980). Aircraft investigation of mesoscale cellular convection during AMTEX 75. *Journal of the Atmospheric Sciences*, 37(5), 1027–1040. [https://doi.org/10.1175/1520-0469\(1980\)037<1027:aiomcc>2.0.co;2](https://doi.org/10.1175/1520-0469(1980)037<1027:aiomcc>2.0.co;2)
- Sandu, I., & Stevens, B. (2011). On the factors modulating the stratocumulus to cumulus transitions. *Journal of the Atmospheric Sciences*, 68, 1865–1881. <https://doi.org/10.1175/2011jas3614.1>
- Savic-Jovicic, V., & Stevens, B. (2008). The structure and mesoscale organization of precipitating stratocumulus. *Journal of the Atmospheric Sciences*, 65, 1587–1605. <https://doi.org/10.1175/2007jas2456.1>
- Seethala, C., & Horváth, Á. (2010). Global assessment of AMSR-E and MODIS cloud liquid water path retrievals in warm oceanic clouds. *Journal of Geophysical Research - D: Atmospheres*, 115. <https://doi.org/10.1029/2009jd012662>
- Sharon, T. M., Albrecht, B. A., Jonsson, H. H., Minnis, P., Khaiyer, M. M., van Reken, T. M., et al. (2006). Aerosol and cloud microphysical characteristics of rifts and gradients in maritime stratocumulus clouds. *Journal of the Atmospheric Sciences*, 63, 983–997. <https://doi.org/10.1175/jas3667.1>
- Stevens, B., Vali, G., Comstock, K., Wood, R., Van Zanten, M. C., Austin, P. H., et al. (2005). Pockets of open cells and drizzle in marine stratocumulus. *Bulletin of the American Meteorological Society*, 86, 51–58. <https://doi.org/10.1175/bams-86-1-51>
- Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79(1), 61–78. [https://doi.org/10.1175/1520-0477\(1998\)079<0061:apgtwa>2.0.co;2](https://doi.org/10.1175/1520-0477(1998)079<0061:apgtwa>2.0.co;2)
- Turner, D. D., Clough, S. A., Liljegren, J. C., Clothiaux, E. E., Cady-Pereira, K. E., & Gaustad, K. L. (2007). Retrieving liquid water path and precipitable water vapor from the atmospheric radiation measurement (ARM) microwave radiometers. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 3680–3690. <https://doi.org/10.1109/tgrs.2007.903703>
- van Zanten, M. C., Stevens, B., Vali, G., & Lenschow, D. H. (2005). Observations of drizzle in nocturnal marine stratocumulus. *Journal of the Atmospheric Sciences*, 62, 88–106. <https://doi.org/10.1175/jas-3355.1>
- Wang, H., & Feingold, G. (2009). Modeling mesoscale cellular structures and drizzle in marine stratocumulus. Part I: Impact of drizzle on the formation and evolution of open cells. *Journal of the Atmospheric Sciences*, 66, 3237–3256. <https://doi.org/10.1175/2009jas3022.1>
- Wentz, F. J., & Meissner, T. (2004). *AMSR-E/Aqua L2B global swath ocean products derived from Wentz algorithm, version 2. [Columnar cloud liquid water; columnar water vapor]*. NASA National Snow and Ice Data Center Distributed Active Archive Center. [https://doi.org/10.5067/AMSR-E/AE\\_OCEAN.002](https://doi.org/10.5067/AMSR-E/AE_OCEAN.002)
- Wentz, F. J., & Spencer, R. W. (1998). SSM/I rain retrievals within a unified all-weather ocean algorithm. *Journal of the Atmospheric Sciences*, 55, 1613–1627. [https://doi.org/10.1175/1520-0469\(1998\)055<1613:sirrrwa>2.0.co;2](https://doi.org/10.1175/1520-0469(1998)055<1613:sirrrwa>2.0.co;2)
- Wood, R. (2012). Stratocumulus clouds. *Monthly Weather Review*, 140(8), 2373–2423. <https://doi.org/10.1175/mwr-d-11-00121.1>
- Wood, R., & Bretherton, C. S. (2004). Boundary layer depth, entrainment, and decoupling in the cloud-capped subtropical and tropical marine boundary layer. *Journal of Climate*, 17, 3576–3588. [https://doi.org/10.1175/1520-0442\(2004\)017<3576:bldead>2.0.co;2](https://doi.org/10.1175/1520-0442(2004)017<3576:bldead>2.0.co;2)
- Wood, R., & Bretherton, C. S. (2006). On the relationship between stratiform low cloud cover and lower-tropospheric stability. *Journal of Climate*, 19, 6425–6432. <https://doi.org/10.1175/jcli3988.1>
- Wood, R., & Hartmann, D. L. (2006). Spatial variability of liquid water path in marine low cloud: The importance of mesoscale cellular convection. *Journal of Climate*, 19, 1748–1764. <https://doi.org/10.1175/jcli3702.1>
- Xue, H., Feingold, G., & Stevens, B. (2008). Aerosol effects on clouds, precipitation, and the organization of shallow cumulus convection. *Journal of the Atmospheric Sciences*, 65(2), 392–406. <https://doi.org/10.1175/2007jas2428.1>
- Yamaguchi, T., & Feingold, G. (2015). On the relationship between open cellular convective cloud patterns and the spatial distribution of precipitation. *Atmospheric Chemistry and Physics*, 15, 1237–1251. <https://doi.org/10.5194/acp-15-1237-2015>
- Zhou, X., & Bretherton, C. S. (2019a). Simulation of mesoscale cellular convection in marine stratocumulus: 2. Nondrizzling conditions. *Journal of Advances in Modeling Earth Systems*, 11, 3–18. <https://doi.org/10.1029/2018ms001448>
- Zhou, X., & Bretherton, C. S. (2019b). The correlation of mesoscale humidity anomalies with mesoscale organization of marine stratocumulus from observations over the ARM Eastern North Atlantic Site. *Journal of Geophysical Research - D: Atmospheres*, 124, 14059–14071. <https://doi.org/10.1029/2019jd031056>