ABSTRACT

RHODES, SPENCER. Large-scale Environments Associated with Southeast Atlantic Marine Stratocumulus Cloud-eroding Boundaries. (Under the direction of Sandra Yuter).

Expansive areas of marine low cloud over the subtropical eastern oceans contribute a strong cooling effect within the Earth's radiation balance. In the subtropical southeast Atlantic, rapidly moving cloud-eroding boundaries occur frequently and can reduce and remove cloud cover over areas the size of California in a matter of hours. Previous work suggests that the rapid cloud erosion along sharp transitions hundreds of km long is associated with atmospheric gravity waves. This study uses 10+ years of satellite data and global forecast model reanalysis to examine which aspects of the large-scale environment are more and less favorable for cloud-eroding boundaries on seasonal and day-to-day time scales in the southeast Atlantic and why similar cloud-eroding boundaries do not occur in the southeast Pacific.

Previous work using satellite data has shown that marine low clouds in the southeast Atlantic have lower liquid water paths, lower cloud tops, and are thinner as compared to the southeast Pacific. We build on these findings and demonstrate that within the range of conditions for the southeast Atlantic that shallower boundary layers with thinner clouds coincide with time periods and geographic locations of more frequent cloud-eroding boundaries.

We use unsupervised machine learning, self-organizing maps, to objectively determine the distribution of spatial patterns in global forecast model reanalysis fields and to relate the occurrence of cloud-eroding boundaries to different nodes within those distributions. Only the spatial pattern of low boundary layer heights near the convex west African coast (from 10° to 25°S) were consistently associated with cloud-eroding boundaries using different input data (normalized anomalies versus full fields), seasons, and times of day. The objectively determined distributions of spatial patterns of estimated inversion strength, SST, mean sea level pressure,

geopotential heights, relative humidity, and u and v winds did not reveal a robust association with the occurrence of cloud-eroding boundaries. A set of criteria was developed for determining whether large-scale differences identified by self-organizing maps are relevant to a given feature (in our case, cloud-eroding boundaries) which entails several adjacent nodes with incidence of the feature of interest at greater than 95% probability based on Monte Carlo random sampling.

Large-scale Environments Associated with Southeast Atlantic Marine Stratocumulus Cloud-eroding Boundaries

by
Spencer Rhodes

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Marine, Earth, and Atmospheric Sciences

Raleigh, North Carolina

2019

APPROVED BY:

Dr. Sandra Yuter Chair of Advisory Committee	Dr. David Mechem
Dr. Matthew Parker	

ACKNOWLEDGEMENTS

The author wishes to express many thanks to Dr. Sandra Yuter for all her time, energy, and support in the production of this thesis, and attributes much of his career success to the opportunities and mentorship she has provided over the years. The author also wishes to express gratitude to the committee members: Dr. David Mechem and Dr. Matthew Parker for their valuable feedback and helpful discussions regarding the thesis. Dr. Mechem is additionally thanked for providing skeleton Matlab code to run and analyze self-organizing maps. Special thanks go to Dr. Matthew Miller for all his insight, expertise, and code that was used throughout the thesis research. Laura Tomkins is also thanked for her collaboration and feedback as she and the author studied cloud-eroding boundaries in parallel at different institutions from modeling and observational perspectives, respectively. Members of the EnvironmentAnalytics research group that contributed to expansion of the manual cloud-eroding boundary dataset and helped with some last-minute requests are greatly thanked (Ronak Patel, Toby Peele, Daniel Hueholt, Lindsay Hochstatter, Dr. Matthew Miller). The author thanks John Hader for his helpful discussions of cloud-eroding boundaries and for access to code used during his thesis studies. The author also thanks Dr. Dave Rahn from Kansas University for lending his expertise in coastal meteorology, and Kevin Nelson at Texas A&M for his help dissecting the drawbacks of GPS radio occultations. This work was funded by NSF AGS grant 1656314.

TABLE OF CONTENTS

LIST OF TABLES viii

LIST OF FIGURES ix

1. INTRODUCTION 1

2. DATA AND METHODS 6

2.1. Cloud-eroding boundary dataset 6

2.2. Categorization and comparison of large-scale environments 8

2.3. Selecting the number of SOM nodes 12

2.4. ERA5 data in SOM analyses 14

2.5. Determining whether SOM results are relevant 17

2.6. MODIS liquid water path and derived fields 19

3. RESULTS 24

3.1. Cloud susceptibility to erosion 24

3.1.1. Time-series of MODIS liquid water path 24

3.1.2. Time-series of MODIS-derived cloud thickness and peak liquid water content 27

3.1.3. MODIS liquid water path maps 28

3.1.4. Relationships between low cloud liquid water path, coastline structure, and the Benguela coastal low-level jet 30

3.1.5. MODIS cloud top temperature and MODIS-derived cloud thickness and peak liquid water content on days with and without cloud-eroding boundaries 32

3.1.6. Role of lower boundary layer height in cloud susceptibility to erosion 34

3.1.7. Dry air required to evaporate liquid water at cloud top on days with and without cloud-eroding boundaries 37

3.2. Large-scale factors relevant for cloud erosion 38

3.2.1. Low boundary layer height 39

3.2.2. Boundary layer height in the southeast Pacific 41

3.2.3. Projections of MODIS data on boundary layer height SOMs 42

3.2.4. Possible factor: estimated inversion strength 43

4. CONCLUSIONS 44

REFERENCES 49

TABLES 57

FIGURES 61

APPENDICES 105

APPENDIX A Flight data from ORACLES 106

APPENDIX B GPS Radio Occultations 110

APPENDIX C Coastal highlands temperature analysis 116

APPENDIX D Deriving cloud thickness and peak liquid water content from MODIS 120

APPENDIX E SOM Node Configuration Analyses 122

APPENDIX F Supplemental MODIS Terra Figures 130

APPENDIX G Supplemental SOM Figures 143

LIST OF TABLES

Table 1: SOMs that were trained in the southeast Atlantic..... 58

Table 2: SOMs that were trained in the southeast Pacific 59

Table 3: Calculations of peak LWC required to maintain saturation after dry air entrainment..... 60

LIST OF FIGURES

Figure 1.1: Example of a large westward-moving cloud-eroding boundary 62

Figure 2.1: Classification example of days with and without a cloud-eroding boundary 63

Figure 2.2: Cloud-eroding boundary daily probability 66

Figure 2.3: Seasonal mean sea level pressure patterns in the southeast Atlantic 67

Figure 2.4: Seasonal pattern of Benguela coastal low-level jet 68

Figure 2.5: Elbow plots of SOM node testing 69

Figure 2.6: ORACLES lidar comparison with ERA5 boundary layer height and vertical profiles of relative humidity, specific humidity and potential temperature 70

Figure 2.7: Example of SOM statistical relevance test 74

Figure 2.8: Liquid water path dependence on cloud thickness and LWC and the adiabatic LWC profile of marine low clouds 75

Figure 2.9: CALIOP comparison to MODIS Collection 5 and 6 cloud top height 76

Figure 2.10: Idealized variability of cloud thickness and peak LWC for varying cloud top height and liquid water path 77

Figure 2.11: MODIS-derived cloud thickness using constant and varying adiabatic LWC change in cloud 78

Figure 3.1: MODIS conditional and non-conditional liquid water path compared to AMSR-E measurements 79

Figure 3.2: MODIS monthly time series 80

Figure 3.3: Spatial patterns of MODIS cloud-conditioned liquid water path 82

Figure 3.4: MODIS conditional liquid water path for different CF thresholds 83

Figure 3.5: Low-level coastal jet presence in the southeast Pacific and southeast Atlantic 84

Figure 3.6: Hydraulic expansion fan schematic 85

Figure 3.7: Effect of convex coastline on MSLP, boundary layer height, and low-level wind 86

Figure 3.8: ERA5 975-hPa divergence separated by cloud-eroding boundary days 87

Figure 3.9: ERA5 975-hPa divergence by season in the southeast Atlantic 88

Figure 3.10: MODIS spatial analysis in Apr-May-Jun separated by cloud-eroding boundary days 89

Figure 3.11: Spatial difference between MODIS-derived cloud thickness using constant and varying adiabatic LWC change in cloud 92

Figure 3.12: Dry air required to maintain saturation for given peak LWC and cloud top temperature 93

Figure 3.13: MODIS-derived dry air required to evaporate liquid water at cloud top in Apr-May-Jun 94

Figure 3.14: MODIS-derived dry air to evaporate peak LWC in Jan-Feb on days without a cloud-eroding boundary 96

Figure 3.15: SOM of ERA5 boundary layer height anomalies 97

Figure 3.16: SOM of ERA5 full boundary layer height field 98

Figure 3.17: ERA5 sea surface temperature separated by cloud-eroding boundary days in Jan-Feb and Apr-May-Jun 99

Figure 3.18: Southeast Pacific SOM of ERA5 boundary layer height anomalies 100

Figure 3.19: MODIS liquid water path projected on full boundary layer height SOM 101

Figure 3.20: Number of samples included in averages for Figure 3.19 102

Figure 3.21: SOM of ERA5 estimated inversion strength anomalies 103

Figure 3.22: SOM of ERA5 full estimated inversion strength field 104

Figure A1: ORACLES-ERA5 boundary layer height comparison (12 Sep 2016) 107

Figure A2: ORACLES-ERA5 boundary layer height comparison (16 Sep 2016) 107

Figure A3: ORACLES-ERA5 boundary layer height comparison (18 Sep 2016) 108

Figure A4: ORACLES-ERA5 boundary layer height comparison (20 Sep 2016) 108

Figure A5: ORACLES-ERA5 boundary layer height comparison (22 Sep 2016) 109

Figure A6: ORACLES-ERA5 boundary layer height comparison (24 Sep 2016)..... 109

Figure B1: Diagram of GPS radio occultation technique 110

Figure B2: Signal drift in GPS radio occultations 111

Figure B3: High-res sounding from St. Helena colocated with GPS radio occultation..... 113

Figure B4: Same as Figure B3 but for a different day 114

Figure B5: GPS radio occultation compared to ERA5 temperature and mixing ratio..... 114

Figure B6: Same as Figure B5 but for a different day 115

Figure C1: ERA5 minimum temperatures over the African highlands on days with and without a cloud-eroding boundary in January and May..... 117

Figure C2: Same as Figure C1 but for MERRA2 data..... 117

Figure C3: ERA5 temperature change from local sunset to local midnight over the African highlands on days with and without a cloud-eroding boundary in January and May..... 118

Figure C4: Same as Figure C3 but for MERRA2 data..... 118

Figure C5: ERA5 temperature change from local sunset to minimum temperature over the African highlands on days with and without a cloud-eroding boundary in January and May..... 119

Figure C6: Same as Figure C5 but for MERRA2 data..... 119

Figure E1: SOM errors for 22 UTC Apr-May-Jun boundary layer height..... 122

Figure E2: SOM errors for 22 UTC Aug-Sep boundary layer height..... 123

Figure E3: SOM errors for 22 UTC May boundary layer height 123

Figure E4: SOM errors for 22 UTC Oct-Nov-Dec boundary layer height..... 124

Figure E5: SOM errors for 22 UTC Jan-Feb mean sea level pressure 124

Figure E6: SOM errors for 22 UTC Oct-Nov-Dec mean sea level pressure 125

Figure E7: SOM errors for 22 UTC Aug-Sep sea surface temperature..... 125

Figure E8: SOM errors for 22 UTC May sea surface temperature..... 126

Figure E9: SOM errors for 22 UTC Aug-Sep 1000-hPa v-winds 126

Figure E10: SOM errors for 22 UTC May 1000-hPa v-winds 127

Figure E11: SOM errors for 22 UTC Oct-Nov-Dec 1000-hPa v-winds 127

Figure E12: SOM errors for 22 UTC Aug-Sep 850-hPa geopotential height 128

Figure E13: SOM errors for 22 UTC May 850-hPa geopotential height 128

Figure E14: SOM errors for 22 UTC Oct-Nov-Dec 850-hPa geopotential height 129

Figure F1: Monthly mean LWP using cloud fraction threshold of 0.8 131

Figure F2: Monthly mean LWP using cloud fraction threshold of 1.0 132

Figure F3: Monthly mean 3.7 μm LWP in the southeast Atlantic 133

Figure F4: Number of days satisfying averaging conditions by month for cloud fraction thresholds of 0.8, 0.9, and 1.0 in the southeast Atlantic 134

Figure F5: Number of days satisfying averaging conditions out of days with and without a cloud-eroding boundary in Jan-Feb and Apr-May-Jun 137

Figure F6: MODIS cloud top temperature and MODIS-derived fields on days without a cloud-eroding boundary in Jan-Feb 138

Figure F7: Number of days satisfying averaging conditions by month in the southeast Pacific 139

Figure F8: Monthly mean fractional entrainment to evaporate peak LWC in the southeast Pacific 140

Figure F9: Inner-quartile ranges of MODIS-derived cloud depth 141

Figure F10: Inner-quartile ranges of MODIS-derived peak LWC 141

Figure F11: Inner-quartile ranges of MODIS cloud top temperature 142

Figure G1: Apr-May-Jun SOM of boundary layer height normalized anomalies 144

Figure G2: Apr-May-Jun SOM of mean sea level pressure normalized anomalies 145

Figure G3: Boundary layer height projected on Jan-Feb SOM of boundary layer height normalized anomalies 146

Figure G4: Boundary layer height projected on May SOM of boundary layer height normalized anomalies 147

Figure G5: Boundary layer height projected on Oct-Nov-Dec SOM of boundary layer height normalized anomalies 148

Figure G6: Mean sea level pressure projected on Apr-May-Jun SOM of boundary layer height normalized anomalies 149

Figure G7: Mean sea level pressure projected on Apr-May-Jun SOM of full boundary layer height 150

Figure G8: Sea surface temperature projected on Apr-May-Jun SOM of boundary layer height normalized anomalies 151

Figure G9: Sea surface temperature projected on Apr-May-Jun SOM of full boundary layer height 152

Figure G10: Estimated inversion strength projected on Apr-May-Jun SOM of boundary layer height normalized anomalies 153

Figure G11: Estimated inversion strength projected on Apr-May-Jun SOM of full boundary layer height 154

Figure G12: Meridional wind at 975 hPa projected on Apr-May-Jun SOM of boundary layer height normalized anomalies 155

Figure G13: Meridional wind at 975 hPa projected on Apr-May-Jun SOM of full boundary layer height 156

Figure G14: Mean sea level pressure projected on Apr-May-Jun SOM of mean sea level pressure normalized anomalies 157

Figure G15: Sea surface temperature (SST) projected on Apr-May-Jun SOM of SST normalized anomalies 158

Figure G16: Boundary layer height projected on Apr-May-Jun SOM of SST normalized anomalies 159

Figure G17: Lower-tropospheric stability (LTS) projected on Apr-May-Jun SOM of LTS normalized anomalies 160

Figure G18: Meridional wind at 975 hPa projected on Apr-May-Jun SOM of 975-hPa meridional wind normalized anomalies 161

Figure G19: Boundary layer height projected on Apr-May-Jun SOM of 975-hPa meridional wind normalized anomalies..... 162

Figure G20: Zonal wind at 950 hPa projected on Apr-May-Jun SOM of 950-hPa zonal wind normalized anomalies 163

Figure G21: 950-hPa geopotential height (Z) projected on Apr-May-Jun SOM of 950-hPa Z normalized anomalies..... 164

Figure G22: 850-hPa geopotential height (Z) projected on Apr-May-Jun SOM of 850-hPa Z normalized anomalies..... 165

Figure G23: 500-hPa geopotential height (Z) projected on Apr-May-Jun SOM of 500-hPa Z normalized anomalies..... 166

Figure G24: 800-hPa relative humidity (RH) projected on Apr-May-Jun SOM of 800-hPa RH normalized anomalies 167

Figure G25: 750-hPa relative humidity (RH) projected on Apr-May-Jun SOM of 750-hPa RH normalized anomalies 168

1. Introduction

Subtropical marine stratocumulus clouds are a vital component of the Earth's climate system. These low clouds are great reflectors of solar radiation while also emitting thermal radiation at a temperature close to the underlying sea-surface temperature (SST), resulting in net cooling (Manabe and Strickler, 1964; Hartmann et al. 1992; Klein and Hartmann, 1993). Stratocumulus clouds blanket approximately one fifth of the Earth annually and almost a quarter of the ocean surface (Hahn and Warren, 2007). Slingo (1990) found that in some climate models the predicted temperature increase from a doubling of carbon dioxide (CO_2) concentration would be offset by a 20% increase in the area of the globe covered by low-level stratus clouds. Three areas of the most persistent marine stratocumulus throughout the year, the northeast Pacific, southeast Pacific, and southeast Atlantic subtropical ocean basins, are responsible for a quarter of the net cooling effect of all clouds (Hahn and Warren, 2007).

Rapid and abrupt cloud-clearing events have been recently documented in the southeast Atlantic marine stratocumulus cloud deck (Hader, 2016; Yuter et al. 2018; example shown in Figure 1.1). The cloudiness transitions first become apparent off the west African coast (10° to 25°S) around local midnight, propagating predominately westward at ~8–12 m/s well into the next day. The transition can yield partial or complete clearing over a span of a few to tens of kilometers and extend hundreds to thousands of kilometers in length, making them capable of clearing out large portions of the marine stratocumulus deck in just a few hours. An abrupt change in regional albedo will considerably alter the solar radiation absorbed by the ocean, which has major implications for climate sensitivity as well as marine ecology and biogeochemistry (Brunet et al. 2008). Herein, we refer to these abrupt cloudiness transitions as cloud-eroding boundaries following the terminology used by Yuter et al. (2018).

Sharp transitions in space between regions of overcast and clear skies are not new in the literature on marine low clouds. Previous studies have documented marine low cloudiness transitions associated with equatorward advection from extratropical cyclones (Klein, 1997; Rozendaal and Rossow, 2003; George and Wood, 2010). Synoptic scale baroclinic weather systems can modulate cloud coverage, type, and height in the midlatitudes (Lau and Crane, 1995; Klein, 1999; Stephens, 2005). The influence of higher-latitude synoptic storms can also lead to inversion strength anomalies which influence the marine low cloud deck (de Szoeke et al. 2016). Off the coast of California, large, diurnally-varying areas of clear skies can develop where marine low clouds are usually present (Kloesel, 1992). These multi-day clearing episodes, which happen most often in the summer, are likely due to synoptic-scale perturbations in the alignment and strength of the northeast Pacific ridge and the subsequent interactions with coastal topography (Crosbie et al. 2016).

The underlying mechanism of these cloudiness transitions in the northeast Pacific is advection, which is inconsistent with the characteristics of cloud-eroding boundaries in the southeast Atlantic. Low-level winds in the southeast Atlantic are prevailing south-southeasterly, whereas the boundaries propagate predominately westward. Hader (2016) provides a clear example when the boundary propagation is perpendicular to the mean low-level flow (see his Animation 3.9). As well, advection does not account for the frequently observed gravity wave trains at the leading edge of cloud-eroding boundaries, shown close up by Yuter et al. (2018, their Figure 1). Propagation of cloud-eroding boundaries against the mean cloud movement, wave-like features along and near the cloud-eroding boundaries, and the presence of a persistent capping inversion atop the marine low clouds which can serve as a waveguide (Klein and

Hartmann, 1993) strongly suggests cloud erosion is caused primarily by atmospheric gravity waves (Yuter et al. 2018).

Propagating atmospheric gravity waves can go largely undetected (that is, without a high-resolution pressure sensor) if not for their effects on cloud fields. In the southeast Pacific, diurnally generated gravity waves from heating of the elevated Andes terrain propagate westward and modulate liquid water path in the marine low cloud deck (Garreaud and Muñoz, 2004; O'Dell et al. 2008; Wood et al. 2009). Additionally, intermittent atmospheric gravity wave trains emanating from a disturbed subtropical jet can interact with marine low clouds in the southeast Pacific, modulate liquid water path, and reduce cloud fraction during the day (Allen et al. 2013; Connolly et al. 2013).

Numerous studies have documented cloud generation by atmospheric gravity waves (Clarke et al. 1981; Christie et al. 1981; Désalmand et al. 2003; da Silva and Magalhães, 2009; Birch and Reeder, 2013; Lutzak, 2013) with agreement on the physical mechanisms responsible. Generally atmospheric gravity waves are thought to have a reversible effect on clouds. Assuming the moisture content in the parcel is unchanged and air temperature decreases with increasing height, cloud cover and liquid water path increases with upward motion, decreases with downward motion, and returns to the pre-wave cloud state after wave passage. Therefore, another localized mechanism to irreversibly reduce cloudiness on short (tens of minutes) time scales must be associated with the cloud-eroding waves in the subtropical southeast Atlantic.

Yuter et al. (2018) hypothesize that atmospheric gravity waves cause the abrupt cloudiness transitions in the subtropical southeast Atlantic, triggered by interaction between offshore flow and the stable marine boundary layer in a fashion akin to processes that generate cloud-forming atmospheric gravity waves (Clarke et al. 1981; Désalmand et al. 2003; da Silva

and Magalhães, 2009; Birch and Reeder, 2013). The offshore flow is likely a combination of land breeze circulation and downslope flow from the African highlands. If advection of offshore flow were solely responsible for the cloud-eroding boundaries, we would expect the cloud-eroding boundaries to slow their westward movement with distance from the coast, which is not the case (Yuter et al. 2018, their Figure 3). To account for the swift removal of cloud observed along cloud-eroding boundaries, Yuter et al. (2018) further hypothesize that the atmospheric gravity waves in cloud-eroding boundaries causes rapid entrainment of warm, dry air from the free troposphere into the cloud layer by enhanced turbulence. Locally enhanced turbulent kinetic energy and mixing across the inversion has been documented within solitary wave packets in the central United States (Koch et al. 2008).

For cloud-eroding boundaries to prevail and reduce cloud, the clouds must be susceptible to this kind of erosion. The southeast Pacific has a generally similar environmental setup to the southeast Atlantic, featuring persistent marine stratocumulus with high terrain to the east, and yet cloud-eroding boundaries were rarely found there by Hader (2016) and Yuter et al. (2018) when examining the recent satellite record. Based on the work of Hader (2016), cloud-eroding boundaries in the southeast Atlantic occur most often in May (as much as 2 out of 3 days) with virtually no occurrence in January and February (1 in 10 days on average).

The primary goal of this study is to determine what environmental conditions are more favorable for cloud-eroding boundaries on seasonal and day-to-day time scales in the southeast Atlantic.

We investigate the large-scale factors that are and are not relevant for cloud-eroding boundaries. In the southeast Atlantic, limited in-situ observations are available, so the bulk of our observational results come from analysis of satellite data. We also seek out important large-

scale differences in model reanalysis data with the use of machine learning. Additionally, our findings help explain why similar cloud-eroding boundaries are not also present in the subtropical southeast Pacific.

2. Data and Methods

2.1. Cloud-eroding boundary dataset

A key component of this analysis is a data set classifying days on which cloud-eroding boundaries do and do not occur in the subtropical southeast Atlantic. Hader (2016) conducted manual, visual analysis of MODIS Terra and Aqua corrected reflectance images from 8 May 2012 to 1 November 2015 to record whether a westward-moving cloud-eroding boundary was present in the subtropical southeast Atlantic. Yuter et al. (2018) extended this dataset to 31 July 2017 and found a nearly identical monthly distribution. In this study, we further extend the dataset back to 1 April 2007 to obtain 10+ years of recorded cloud-eroding boundaries.

Following Hader (2016), we examine the same region over the subtropical southeast Atlantic, from 10° to 20°S and 0° to 12°E. Days are marked as having a cloud-eroding boundary if the following criteria are met:

- A cloud boundary is present, with a transition on the order of a few to a few tens of kilometers
- Complete or partial clearing behind the boundary is visually obvious in both satellite images
- The boundary is at least 200 km (2° latitude/longitude) in length
- The boundary has a noticeable, predominately westward component of motion

Boundary movement was detected by flipping between Terra and Aqua images using the NASA Earth Observing System Data and Information System (EOSDIS) worldview tool (https://worldview.earthdata.nasa.gov). Most days can be quickly classified as having or not having a cloud-eroding boundary, which we refer to herein as 'Yes' and 'No' days. On some days, apparent cloud boundaries in the region only satisfy a subset of the criteria above. In such

cases, we log the day as *possibly* having a cloud-eroding boundary, referred to as 'Maybe' days. In cases when the swath of the MODIS instrument misses a portion of the cloud deck or boundaries, a determination of Yes, No, or Maybe is still made with the available information.

To suppress potential classification biases, Spencer Rhodes trained three additional analysts to conduct classification. Each person was given the criteria for classification and trained with real examples from the classifications by Hader (2016). Once all analysts had gone through 1 April 2007 to 8 May 2012, decisions by the analysts for each day were compared. A day was marked Yes if at least three analysts reported a Yes day or if two reported Yes and two reported Maybe. A day was marked No if all analysts reported No or if three reported No and one reported Maybe. The remaining days were sorted into probable and unlikely to have a cloud-eroding boundary. The unlikely days were ones that reported, of the four analysts, 2 No, 1 Yes and 1 Maybe; 3 No and 1 Yes; 3 Maybe and 1 No; or 2 Maybe and 2 No. The unlikely days were then reexamined by Spencer Rhodes to give a final classification. The probable days were ones that reported, of the four analysts, 4 Maybe; 1 Yes and 3 Maybe; 2 Yes, 1 Maybe and 1 No; 1 Yes, 2 Maybe and 1 No; or 2 Yes and 2 No. The probable days were then evaluated by Spencer Rhodes and Dr. Matthew Miller (a satellite data expert) to make a final classification of Yes, No or Maybe (examples shown in Figure 2.1).

With the resulting 10 year and 4 month dataset (1 April 2007 to 31 July 2017), we plot the daily probability of Yes, No and Maybe days for each calendar month (Figure 2.2). As noted by Hader (2016), the average Yes days per month is a robust underestimate of the actual number of cloud-eroding boundaries given that our classification of Yes means *at least one* clearly delineated cloud-eroding boundary is present. The basis of our subjective analysis of satellite images is similar to what other studies have done to estimate the occurrence frequency of

gravity-wave-related phenomena (e.g. Désalmand et al., 2003; da Silva and Magalhães, 2009; Magalhães et al. 2011; Birch and Reeder, 2013).

2.2. Categorization and comparison of large-scale environments

We examine the distribution of several types of synoptic scale spatial patterns and how these align with the observed occurrence of cloud-eroding boundaries. The subtropical southeast Atlantic region is relatively quiescent compared to the ITCZ to the north and the mid-latitude storm track to the south. The mean synoptic pattern in the southeast Atlantic is dominated by a semi-permanent high pressure center located near 30°S and 5°W (Figure 2.3) and the Benguela coastal low level jet (Patricola and Chang, 2017). The subtropical high pressure is strongest in austral winter (Jun-Jul-Aug) and weakest towards the end of austral summer. The seasonal variability of the Benguela jet is described by Patricola and Chang (2017, see Figure 2.4). The distribution of synoptic spatial patterns around the seasonal means and how these coincide with cloud-eroding boundaries may provide information on environments which are more and less conducive to cloud-eroding boundaries.

Several objective techniques exist to determine the distribution of 2D spatial patterns. Linear decomposition methods include empirical orthogonal functions (EOF), principal component analysis (PCA), *k*-means clustering, and self-organizing maps (SOMs). SOMs are an automated neural-network technique that produces a user-defined number of data states (synoptic patterns, in our case), continuously distributed and spanning the parameter space of the data (Hewitson and Crane 1992; Kohonen 2001). SOMs were first introduced by Kohonen (1981) and have experienced an uptick in usage with the recent availability of open source SOM toolkits. In atmospheric science, SOMs have been applied to synoptic and climate classification

(Cavazos, 2000; Reusch et al. 2007; Bailey et al. 2011; Kennedy et al. 2016; Mechem et al. 2018), cloud classification (Ambroise et al. 2000), and extreme weather (Cassano et al. 2006; Andersen-Frey et al. 2017). Liu and Weisberg (2005, 2007) compared the ability of EOF and SOMs to extract ocean current patterns given the same input data and found the SOM patterns to be more accurate and intuitive than leading-mode EOF patterns (Liu and Weisberg, 2011). SOMs also better capture asymmetric features (Liu and Weisberg, 2005), provide more robust feature extraction in cases of complex datasets with multiple overlapping patterns (Reusch et al. 2005; Annas et al. 2007; Astel et al. 2007), and are less sensitive to noise (Liu et al. 2006a) compared to EOF and PCA. Lin and Chen (2006) found that SOMs better determined cluster membership than *k*-means. Furthermore, a traditional cluster analysis is insufficient in our context as it does not represent the full multidimensional distribution of the entire dataset as a continuum. Analyzing the continuum of 2D spatial patterns is especially relevant as we are also interested in the transitions between nodes with high data density.

Following the approach of Hewitson and Crane (2002), we employ SOMs to categorize the large-scale environments in the southeast Atlantic and Pacific for different times of day and periods during the year. With the use of SOMs, we seek to identify *whether* there are any large-scale synoptic factors favorable for cloud-eroding boundaries. The SOM algorithm yields a distribution of the input large-scale synoptic spatial patterns. For this study, we used monthly or seasonal sets of synoptic variables for a given time of day as inputs to the SOMs. For example, 1,001 maps were input to the SOM of mean sea level pressure at 22 UTC for April-May-June 2007-2017. We then check the set of days within each SOM output node against our cloud-eroding boundary data set (Section 2.1, Figure 2.2) and count the number of yes, no and maybe days that occurred in each node. This method aids in determining which distinct patterns

(nodes), if any, more often coincide with cloud-eroding boundaries in the full context of the distribution of patterns. In comparison, a composite map of days with cloud-eroding boundaries could average together potentially disparate distinct environments which would smooth out the key features of interest.

When the SOM is initialized, the nodes begin in a default configuration irrespective of the parameter space of the input data. A competitive learning process begins, iterating through randomly selected records of the input dataset. On each iteration, the node that best matches the data (or winning node) is nudged closer to the data. The nodes adjacent to the winning node (or neighborhood nodes) also adjust towards the winning node. The neighborhood function of a SOM is what distinguishes it from a *k*-means clustering approach (Kohonen, 2001; Kennedy et al. 2016; Mechem et al. 2018). The adjustment of nodes and competitive learning is solely based on the data with no prior knowledge or preconceptions of how the large-scale environments should be distributed. SOMs represents a form of unsupervised machine learning, yielding an objective categorization of the large-scale environments (Mechem et al. 2018).

We utilized the open source SOM Toolbox for MATLAB (www.cis.hut.fi/somtoolbox/) to create SOMs trained on a variety of atmospheric fields that we deemed potentially relevant for cloud-eroding boundaries (see Tables 1 and 2). Although SOMs are technically unsupervised, knowing the nature of the answer expected from the SOM helps in choosing appropriate training variables, data filters (e.g. seasonal averages or normalized anomalies), and domain size and dimension.

Mechem et al. (2018) trained their SOMs on normalized anomalies to suppress influence from increased synoptic variability over poleward latitudes. In our study, we trained SOMs separately on normalized anomalies and on the full atmospheric fields. Training on the

normalized anomalies ensures consistency for training on different times of year. We train SOMs on the full fields to test the robustness of any potentially relevant results deduced from SOMs of normalized anomalies. But whether the SOM was trained on the anomaly or field full, we present the findings in terms of full synoptic patterns which best match each node. The reason for this choice of presentation is that the full synoptic pattern shows the actual structure of the atmosphere. For example, winds are a function of the actual pressure field, not the anomaly pressure field. Additionally, depiction of the full field yields a better comparison of seasonal and regional differences (e.g. southeast Pacific vs. southeast Atlantic, EIS in May vs. EIS in December). Examples showing the patterns of normalized anomalies in each node are provided in Appendix G.

The equation for computing the normalized anomalies of a given field (Z) is shown in Equation 2.1, where Z is the full field at a given time, $\bar{Z}$ is the 10-day centered mean, and σ_Z is the standard deviation of the 10-day centered field.

$$\langle Z \rangle = \frac{(Z-\bar{Z})}{\sigma_Z} \tag{2.1}$$

The main difference between training on normalized anomalies versus full fields lies in the different context afforded. Normalized anomalies encapsulate information about that day and the surrounding 10 days but leave out the details of the full and mean fields. As such, two days with similar patterns of normalized anomalies may in fact have quite different full fields, but they are similar in terms of their deviation from their 10-day context. On the other hand, full fields are exactly as they appear, so similar-looking patterns are indeed close to one another. However, the 10-day context is left out, which could mean that the similar patterns are alike for different reasons. In short, both provide a unique way to interrogate synoptic differences with SOMs.

To avoid capturing diurnal processes, data was taken from the same hour each day. Since the boundaries begin to propagate away from the African coast at midnight, we analyzed data from 11pm local time, which is roughly 22 UTC in the southeast Atlantic and 04 UTC on the following day in the southeast Pacific. We also ran SOMs on data from 11am local time to see if there was agreement of cloud-eroding boundary association (or lack thereof) between the two times for each atmospheric field. A matrix showing the complete set of SOMs that were run in each domain is shown in Tables 1 and 2.

To assess the large-scale environments in which the boundaries occur, we ran the SOMs over the region from 5° to 35°S and 10°W to 20°E. This region encompasses the footprint of cloud-eroding boundaries in the southeast Atlantic described by Yuter et al. (2018) which stretches from the African coast (~11°E) to as far as 4°W between roughly 10° and 20°S. The SOM analysis region additionally includes a portion of the African landmass to the east and goes well past the furthest westward extent of the cloud-eroding boundaries. The north side is equatorward of 10°S but remains largely outside the ITCZ, while the south side extends poleward of where the subtropical high-pressure center tends to be situated around 30°S in order to fully capture the high-pressure center regardless of season. For the southeast Pacific domain, we used the same latitude range for consistency and selected 95° to 65°W for longitude bounds to get roughly the same proportion of land to ocean in the domain as in the southeast Atlantic.

2.3. Selecting the number of SOM nodes

The number of nodes is an important user-defined quantity for SOM analysis. Using too few nodes will overgeneralize the dataset continuum, potentially combining unique synoptic states to yield physically implausible states. Conversely, using too many nodes will make the

SOM analysis difficult to decipher, with similar synoptic states being represented by multiple nodes. For this form of unsupervised machine learning, the number of nodes chosen for the SOM is equally as important as the data being used for training.

To determine the optimal number of nodes we follow the method of Mechem et al. (2018). We ran the SOM algorithm on several large-scale atmospheric fields for a range of SOM node array configurations (2x2 to 10x10) and evaluate error metrics for each node configuration. Two error metrics are examined. The quantization error, a measure of how close each synoptic setting is to its best-matching node, decreases with increasing number of nodes. The elbow criterion (Tibshirani 2001) describes the point at which adding more nodes "fail[s] to add a significant amount of information" (Schuenemann et al. 2009) to the SOM. If a SOM has more nodes but roughly the same quantization error, then the SOM with less nodes represents the different synoptic settings of the underlying dataset equally as well. The elbow criterion would then be the node configuration at which additional nodes results in a minimal reduction of quantization error, justifying a lower bound for the number of nodes. The topographic error, a ratio of the number of synoptic settings whose second best-matching node is nonadjacent to its best-matching node, will increase with additional nodes as similar synoptic settings become more likely to be mapped to nonadjacent nodes. The node configuration at which additional nodes dramatically increases topographic error marks the upper bound for the number of nodes (Mechem et al., 2018).

Figure 2.5 shows the quantization and topographic error for varying SOM node configurations trained on mean sea level pressure and estimated inversion strength. The same analyses were performed for other time periods and fields in the southeast Atlantic (see Appendix E). Although the analyses do not consistently agree on a single node configuration, a

5 x 5 node configuration (25 total nodes) minimizes quantization and topographic error, provides a center node about which symmetry in the synoptic states can be seen, and keeps the cognitive load required to digest the SOM information to a relative minimum (as opposed to using a higher number of nodes).

2.4. ERA5 data in SOM analyses

The model reanalysis we elect to use for training SOMs is the fifth major global Re-Analysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), known as ERA5. The ERA-Interim reanalysis dataset was discontinued as of August 2019, leaving ERA5 as the de facto reanalysis product from ECMWF until the release of ERA6 around 2020. ERA5 offers several advantages over its predecessor as well as other model reanalysis datasets. ERA5 relies on the ECMWF Integrated Forecast System (cycle 41R2) and uses 4D-Var data assimilation of a vast array of observations (24 million per day as of 2018, for the full list see https://confluence.ecmwf.int/display/CKB/ERA5+data+documentation#ERA5datadocumentation-Observations). Compared to ERA-Interim, ERA5 additionally incorporates aircraft and surface pressure data, fixes an issue with inconsistent SSTs over climatic time periods, runs at a higher resolution in both the horizontal (31-km) and vertical (137 vertical levels, up to 0.01 hPa), and produces hourly reanalysis fields. Hourly reanalysis provides a considerable advantage for the purposes of our study, allowing us to train SOMs on data just before local midnight when cloud-eroding boundaries first become apparent off the coast of W. Africa, and to train on the same local time in the southeast Pacific. Few, if any, other model reanalysis datasets currently offer hourly output. As well, the 31-km horizontal resolution yields a well-resolved 0.25x0.25 grid, which is better suited to capturing complex

variability along the coast in the southeast Atlantic (Patricola and Chang, 2017) and southeast Pacific.

ERA5 data are obtained from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/) from 1 January 2007 to 31 December 2017 for the southeast Atlantic and southeast Pacific domains described in Section 2.2. We use this 11-year range as it encompasses our entire cloud-eroding boundary dataset. We also derive some fields for training with SOMs which are not included in ERA5, namely lower tropospheric stability (LTS) and estimated inversion strength (EIS). LTS is defined as the difference between the potential temperature (θ) of the free troposphere (700 hPa) and the surface, $\mathrm{LTS} = \theta_{700} - \theta_0$ (Slingo, 1987; Klein and Hartmann, 1993; Klein, 1997). While LTS is highly correlated with cloud fraction in the tropics and subtropics, the correlations fall off for midlatitude stratus. EIS was devised by Wood and Bretherton (2006) to define a more general, regime independent, parameter of inversion strength that correlates well with stratus cloud fraction in the tropics, subtropics and midlatitudes. Furthermore, LTS does not specifically target the strength of the capping inversion over marine stratocumulus, as it includes potential temperature lapse rates in the free troposphere and potentially in the decoupled layer. The surface mixed layer (below the decoupled layer) is presumed to have no vertical gradient of potential temperature. We calculate the estimated inversion strength (EIS) described by Wood and Bretherton (2006) from ERA5 data using:

$$\mathrm{EIS} = \mathrm{LTS} - \Gamma_{\mathrm{m}}^{850}(z_{700} - \mathrm{LCL}) \tag{2.2}$$

where $\Gamma_{\mathrm{m}}^{850}$ is the moist adiabatic potential temperature gradient at 850 hPa in Kelvin per meter, z_{700} is the geopotential height at 700 hPa in meters, and LCL is the lifted condensation level in meters.

The height of the marine boundary layer is represented by the top of the capping inversion that overlies marine stratocumulus, and hence is a proxy for the top of stratocumulus clouds. For this reason, we make extensive use of the boundary layer height diagnostic available in ERA5. Boundary layer height in ERA5 is determined as the level where the bulk Richardson number exceeds 0.25 (ECMWF, 2016), interpolated between model levels as necessary. Seidel et al. (2012) tested several methods for estimating boundary layer height, including ones based on parcel methods, height of a surface-based inversion, first temperature or humidity discontinuity, and large change in potential temperature. Seidel et al. (2012) found the bulk Richardson number method, originally formulated by Vogelezang and Holstag (1996), to be most suitable for a global dataset consisting of model reanalysis and radiosonde data, given that it performs well in both stable and convective boundary layers and is less sensitive to sounding vertical resolution. The Integrated Forecast System used in ERA5 computes the bulk Richardson number using Equation 3.90 of ECMWF (2016).

According to Seidel et al. (2012), relative uncertainty of boundary layer height using the bulk Richardson number method can be larger than 50% for estimated values less than 1 km. However, they did not examine regional differences in the uncertainty, specifically over the ocean near coastlines. Near-coast boundary layer height in ERA5 may be subject to this high uncertainty, but we are most interested in relative seasonal and day-to-day differences of ERA5 boundary layer height which would likely overshadow the uncertainty associated with using the bulk Richardson number method. More robust methods of boundary layer height detection based on ERA5 profiles at each grid point and time are left to future work. A comparison of ERA5 data to airborne high spectral resolution lidar backscatter from the ORACLES field campaign (Dzambo et al. 2019) reveals that a relative humidity threshold (e.g. 70% as suggested by

Shinozuka et al. 2019) provides and improved estimation of the boundary layer depth in these cases (Figure 2.6, panels a and d) compared to the ERA5 boundary layer height product. We also show the comparison of ERA5 potential temperature and specific humidity profiles to ORACLES lidar data (Figure 2.6, panels b, c, e, f), both of which may aid estimation of the boundary layer in future work.

In this study, we also investigate the slope of boundary layer depth away from the coast as potentially relevant for cloud-eroding boundaries. We compute the zonal gradient of boundary layer height (Δz) between adjacent pixels in the 0.25° × 0.25° grid in meters per km and account for the change in distance per degree longitude with latitude, as shown by the equation below.

$$\Delta \mathrm{z} = \frac{\left(\mathrm{z_{i+1,j}} - z_{i,j}\right)}{d * \left(\mathrm{lon}_{i,j} - \mathrm{lon}_{i,j}\right)}$$
$$d = \frac{2\pi r_e \cos\left(\mathrm{lat}_j\right)}{360} \tag{2.3}$$

In the equation above (2.3), $z_{i,j}$ is the boundary layer height in meters on a longitude-latitude (i,j) grid and d is the distance in km per degree longitude at a given latitude (lat_j), where r_e is the radius of earth in km. The zonal gradient is computed such that positive values represent boundary layer height sloping up towards the east. Therefore, the difference in longitude between adjacent grid points $\left(\mathrm{lon}_{i,j} - \mathrm{lon}_{i,j}\right)$ is always positive. The results of this calculation are presented in Section 3.2.4.

2.5. Determining whether SOM results are relevant

Usage of SOMs to elucidate any conditions in the large-scale environment which are conducive to cloud-eroding boundaries is incomplete without some significance test. Previous

studies have utilized the number of records which map into each node (also known as the percentage of hits) to identify the spatial patterns which describe most of the multidimensional distribution (e.g. Mechem et al. 2018). Our study focuses more on nodes with a common spatial pattern and higher counts of Yes or No days that map into those nodes. We propose a method for testing the relevance of spatial patterns identified by SOM nodes inspired by Monte-Carlo significance testing.

For a given SOM, we randomly assign the Yes and No days in the time period to the SOM nodes for 10,000 iterations, preserving the chance that a day will fall into a given node using the percentage of days in the time period that matched the node. Figure 2.7 shows the distributions of randomly assigned Yes and No days (panels b and c) across the 10,000 iterations for a node of the SOM trained on normalized anomalies of ERA5 boundary layer height (see Figure 3.15). We use the distributions to decide whether the actual number of matching Yes/No days was unlikely to be random using a 95% confidence interval. For example, in Figure 2.7 (panel a) the SOM node had 30 matching Yes days. The 95th percentile in the distribution of randomly assigned Yes days (panel b) is 29 Yes days, so the 30 matching Yes days in this node is unlikely to be random. In contrast, the 5th and 95th percentiles in the distribution of randomly assigned No days (panel c) are 31 and 52, respectively, so the occurrence of 35 matching No days in this node could just be by chance.

We then define a spatial pattern as relevant if at least three adjacent nodes in a 5 x 5 SOM has unlikely random counts of Yes or No days of the same type, since adjacent nodes in a SOM are most similar. By same type we are referring to, for example, the nodes all reporting unlikely random high counts of Yes days or, equivalently, unlikely random low counts of No days. The reason for this 2-step criterion is due to the nature of a SOM. A single node with unlikely

random counts of Yes or No days is not by itself indicative of a meaningful spatial pattern. Three or more nodes in a 5 x 5 SOM with a similar spatial pattern and unlikely random counts of Yes or No days is very unlikely to happen by chance, implying the spatial pattern is relevant for cloud-eroding boundaries. Further relevance is demonstrated when adjacent nodes in a corner of the SOM report unlikely random high counts of Yes days (or low counts of No days) while the nodes in the opposite corner report unlikely random low counts of Yes days (or high counts of No days). To our knowledge, this is the first evaluation of the statistical meaningfulness of different spatial patterns in a SOM with respect to some feature of interest.

2.6. MODIS liquid water path and derived fields

We compute monthly maps and time series of *conditionally averaged* daytime MODIS Terra Collection 6.1 Level-3 1° × 1° satellite data (MOD08_D3) in the southeast Atlantic and southeast Pacific over the 18-year period from 2000 to 2018. Specifically, we make use of the daytime liquid water path (LWP), cloud fraction and cloud top temperature, pressure and height.

Liquid water path retrievals from MODIS are unreliable at cloud peripheries and/or where a pixel is only partially filled with cloud (Coakley et al. 2005). Additionally, MODIS 5-km cloud top properties, which are fed into Level-3 products, suffer from beam-filling effects in regions of high spatial heterogeneity (e.g. broken low cloud fields), documented in Quality Statement #3 of 06_L2 MODIS cloud issues. For our conditional averages of liquid water path, cloud top temperature, and other MODIS-derived quantities, we only used values with corresponding cloud top $\leq$ 2.5 km and cloud fraction thresholds of 80, 90 and 100% at 1° × 1° pixel scale to test the sensitivity of our results. The cloud top threshold was used to minimize the inclusion of clouds other than marine stratocumulus.

The full sequence of images for each cloud fraction threshold can be found in Appendix F. Our findings vary minimally using the different cloud fraction thresholds. In our results, we highlight the 90% cloud fraction threshold as it is a sound compromise between filtering out pixels with potentially high sub-pixel heterogeneity and gathering a robust sample size.

MODIS liquid water path retrievals can be obtained from three different channels: 1.6, 2.1 and 3.7 μm channels. The default channel for MODIS liquid water path retrievals is 2.1 μm, but studies have shown that retrievals using the 3.7 μm channel are less sensitive to pixel heterogeneity (Zhang and Platnick 2011; Painemal et al. 2013). We examined the 3.7 μm liquid water path retrievals (see Appendix F) and found minimal differences between the 2.1 and 3.7 μm retrievals, with liquid water path tending to be only slightly higher in the 2.1 μm retrievals, consistent with previous findings (Painemal et al. 2016).

We converted MODIS liquid water path into cloud thickness and a linear liquid water content (LWC) profile assuming an adiabatic water-mixing ratio (Cotton and Yuter, 2008; Miller, 2010). Wood (2012) also indirectly calculated cloud thickness from climatological mean liquid water path using this assumption and found qualitative consistency with field studies in terms of the range of values in the tropics, subtropics, and mid-latitudes. The dependence of liquid water path on cloud layer average LWC and cloud thickness is shown in Figure 2.8 (panel a). These conditions encompass clouds with vertically uniform LWC profiles to linear, adiabatically stratified LWC. The latter profile (Figure 2.8 panel b) of increasing LWC with height is the more common case in marine stratiform clouds as verified by aircraft observations (Nicholls and Leighton, 1986; Gerber, 1996; Miles et al. 2000; Wood, 2005a) and surface-based remote sensing (Albrecht et al. 1990; Zuidema et al. 2005). Despite this, MODIS liquid water path retrievals are based on a vertically uniform LWC profile assumption (King et al. 1998,

Hubanks et al. 2019). The equation for liquid water path (LWP) used by MODIS is shown below:

$$\mathrm{LWP} = \gamma r_e \tau_c \tag{2.4}$$

where r_e is the droplet effective radius near cloud top, derived from the 2.1μm channel, τ_c is cloud optical thickness at 0.86μm wavelength, and γ is a constant value of either $\frac{2}{3}$ for vertically uniform (Stephens, 1978) or $\frac{5}{9}$ for adiabatically stratified (Nicholls and Leighton, 1986). Based on this relationship, adiabatically stratified LWP would be 17% less than LWP from MODIS using the vertically uniform model, creating a small positive bias in regions dominated by marine stratocumulus (Zhou et al. 2016). We tested the effect of modifying MODIS liquid water path to be based on the adiabatically stratified assumption and found only minute changes in spatial patterns which do not affect our results.

Our methodology for deriving the profile of cloud liquid water content and cloud thickness from MODIS requires liquid water path as well as cloud top temperature and pressure. Previous MODIS collections (i.e. data versions) have struggled to adequately estimate cloud top properties (Menzel et al. 2008; Baum et al. 2012), with cloud top heights in marine stratocumulus biased high by more than 2 km with respect to CALIPSO cloud products (Holz et al. 2008) as recently as Collection 5. With the advent of Collection 6, a revised method is used that performs much better in situations where a temperature inversion is present which plagued Collection 5 (Baum et al. 2012), as shown by Figure 2.9.

The algorithm we employed to derive the LWC profile with liquid water path equal to the MODIS liquid water path retrieval is documented extensively in Appendix D. The resulting LWC profile is used to estimate cloud thickness, peak LWC at cloud top, and cloud layer average LWC. The exact relation between cloud top temperature and cloud top height above sea

level is tricky in regions with strong temperature inversions such as marine stratocumulus regions. Hence, we focus on the cloud thickness, the relative difference between the cloud top and cloud base height, rather than cloud top height itself.

Wood and Taylor (2001) estimated stratocumulus cloud thickness directly from liquid water path using:

$$\mathrm{LWP}_{adiab} = \Gamma_{adiab} h^2/2 \tag{2.5}$$

where LWP_{adiab} is the liquid water path, h is cloud depth, and Γ_{adiab} is a constant adiabatic rate of change of LWC with height in the cloud. To use Equation 2.5, an estimate must be made regarding what value to use for Γ_{adiab} which in turn is based upon an assumption of a cloud top temperature and pressure (i.e. altitude).

We consider the implications of varying cloud top temperature on both derived thickness and peak LWC using values common in the southeast Atlantic and southeast Pacific (275 to 285 K). Figure 2.10 compares the LWC profiles derived using our varying Γ_{adiab} and two constant Γ_{adiab} treatments. At 293 K and 900 m altitude, adiabatic change of LWC in cloud is 2 g/kg per km of ascent (Cotton and Yuter, 2008), or 2.4 g/m^3 per km assuming a density of 1.2 kg/m^3. If the cloud falls mostly below this in altitude (higher temperature), it will be biased *thick* compared to our more accurate treatment, whereas if it falls mostly above this altitude (lower temperature) it will be biased thin. The biases are on the order of tens of meters. To provide reasonable bounds in our comparison, we used Γ_{adiab} = 2$\pm$0.2 g/kg per km. Figure 2.11 compares the monthly time series of MODIS-derived cloud thickness over the southeast Atlantic using constant Γ_{adiab} = 2.4 g/m^3 per km and varying Γ_{adiab}. Using Γ_{adiab}= 2.16 g/m3 per km does give a thickness line with no more than 5 meters of bias (not shown).

Since cloud top heights vary among the ocean basins, with distance from the coastline within marine low cloud decks, and diurnally, use of a constant Γ_{adiab} is akin to assuming those variations in cloud top height do not matter to the topic of interest. As will be shown in Section 3, the boundary layer height which is the cloud top height appears to be an important factor for the occurrence of cloud-eroding boundaries. Additionally, it is infeasible to accurately estimate both thickness and peak LWC using a constant Γ_{adiab}. Although the magnitudes of the differences in cloud thickness and peak LWC between use of a constant Γ_{adiab} versus a varying Γ_{adiab} are small, for our purposes, the more precise calculation we outline in Appendix D with varying Γ_{adiab} is warranted.

3. Results

3.1. Cloud susceptibility to erosion

In this section, we document the spatial distribution and seasonality of liquid water path in the southeast Atlantic and southeast Pacific. We also convert liquid water path to an estimated cloud thickness and a profile of liquid water content to constrain the degree to which a lower liquid water path may make clouds more susceptible to erosion. The liquid water path is the total amount of liquid in a column, measured in g/m^2. Previous work by Zuidema et al. (2016) showed that the daily mean liquid water path in the subtropical southeast Atlantic is lowest in May (see Figure 3.1), coinciding with the peak month of cloud-eroding boundaries. Assuming a single cloud layer of marine stratocumulus, liquid water path would be lower either because the clouds are thinner or higher in altitude. Since liquid water path is more dependent on cloud thickness than height (Albrecht et al. 1990) it is likely that the clouds are thinner. A thinner cloud is easier to evaporate through entrainment of unsaturated air than a thicker cloud, as is a cloud with less liquid water content available to counteract entrainment. In our domain, the entraining unsaturated air is the subsiding warm, dry air above the marine stratocumulus cloud layer. Compared to Zuidema et al. (2016), our analysis does not include cloud-free pixels (i.e. LWP = 0).

3.1.1. Time-series of MODIS liquid water path

The seasonal variability of conditionally area-averaged MODIS liquid water path is shown in Figure 3.2 (panels a and b). The monthly data are spatially averaged within the pink boxes in Figure 3.2 panels e and f utilizing only pixels over the ocean. The box in the southeast Pacific is slightly poleward (2.5 degrees) of the one in the southeast Atlantic in order to better capture the concave coastline of western South America. The percent of days satisfying the

conditions (cloud fraction ≤ 90% and cloud top ≥ 2.5 km) in the month of May is shown in panels e and f (Figure 3.2). Most of the pixels in the pink regions have 117 or more days satisfying the conditions out of 588 total days in May for the 19-year dataset. A small pocket of pixels in the southeast Atlantic polygon southwest of the coastal bulge of Angola and Namibia (12° to 22°S) only has 58 to 117 days (10-20%) satisfying the conditions, but we still deem this a reasonable sample size. May is the month when cloud-eroding boundaries are most frequent, and as a result has the fewest areal average samples with cloud fraction ≥ 90%. See Appendix F for the full set of monthly maps showing the number of days satisfying the averaging conditions in each pixel.

In the southeast Pacific over our area of interest, conditional liquid water path transitions from lowest in austral summer (Dec-Feb) to highest in austral winter (Jun-Aug). Many studies have identified Sep-Oct-Nov, the period of maximum low cloud *coverage* in the southeast Pacific and southeast Atlantic (Klein and Hartmann, 1993; Wood and Hartmann, 2006; Toniazzo et al. 2011; Muhlbauer et al. 2014), as primarily associated with the maximum in lower-tropospheric stability (LTS). LTS is the difference between potential temperature at 700 mb and at the surface. The high liquid water path in July in the southeast Pacific coincides with increased frequency of closed cellular convection (Muhlbauer et al. 2014, see their Figure 5). Wood and Hartmann (2006) studied the relationship between cloud fraction and liquid water path. In their Figure 9, closed cellular convection coincides with significantly higher liquid water path at 1 km^2 pixel scale.

Areal average liquid water path of low clouds in the southeast Pacific is higher throughout the year than in the southeast Atlantic (Figure 3.2 panels a and b), with the only time when the monthly average liquid water paths are similar during austral summer (Dec-Feb) when

virtually no cloud-eroding boundaries occur. Whereas the southeast Pacific has a single peak of liquid water path in July, the southeast Atlantic has two local maxima. Liquid water path is highest in August, decreasing through to November, then briefly increases in December through February, before heading to the yearly minimum in May, the peak month of cloud-eroding boundaries.

Our seasonal trends in liquid water path over the southeast Pacific and southeast Atlantic generally agree with the results from Zuidema et al. (2016), with discrepancies mainly due to the analysis domain, time of satellite retrievals, and averaging conditions. The unconditional averages used in Zuidema et al. (2016) are based on Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) retrievals aboard the Aqua satellite with a 1:30 pm local time overpass. The diurnal cycle of marine stratocumulus cloudiness in the southeast Atlantic yields the lowest liquid water paths in the afternoon (Burleyson and Yuter, 2015). Compared to the analysis domains we used in Figure 3.2 (panels e and f), Zuidema's et al. (2016, their Figure 4a) were further offshore (90° to 80°W in the southeast Pacific and 0° to 10°E in the southeast Atlantic), further north (both at 10° to 20°S), and smaller (10° × 10° compared to our 15° × 15°).

To remove the impact of the different domains and to separate out the differences accounted for by conditional averaging, a different sensor, and time of day, two versions of the monthly time series of liquid water path were computed using the same domains as Zuidema et al. (2016) in the southeast Pacific and southeast Atlantic (Figure 3.1). MODIS conditional (cloud fraction ≥ 0.9) liquid water path means (2000 to 2018) from the Terra (10:30 am LT) satellite are about 3 to 22 g/m^2 higher (on average 11.3 g/m^2 higher) but generally parallel the non-conditional means (cloud fraction > 0). This is consistent with previous work by Greenwald et al. (1995) which found higher cloud-conditional mean liquid water path as compared to

unconditional averages. The larger time of day differences combined with the difference between the sensors (MODIS versus AMSR-E) are illustrated by comparing the unconditional MODIS values from 10:30 am local time to the unconditional AMSR-E values from 1:30 pm local time. These differences account for 13 to 36 g/m^2 (on average 22.6 and 24.1 g/m^2 in the southeast Pacific and southeast Atlantic, respectively) lower liquid water path in unconditional AMSR-E. Future work will also compare MODIS liquid water path from Aqua (1:30 pm LT). The results from Seethala and Horváth (2010) help to explain the lower AMSR-E liquid water path relative to MODIS. The key feature of lowest liquid water path in May in the southeast Atlantic is consistent across the monthly time series based on different sensors, time of day, domains, and whether a conditional average is employed.

3.1.2. Time-series of MODIS-derived cloud thickness and peak liquid water content

To diagnose the cloud characteristics associated with the liquid water path values and spatial patterns, we estimate cloud thickness and liquid water content profiles using the MODIS retrieved liquid water path and cloud top properties (temperature, pressure and height). The details of this calculation are fully laid out in Section 2.6. Area-averaged monthly time series of estimated cloud thickness (gray) and mean cloud liquid water content (blue) are shown in Figure 3.2, panels c and d, with the area used for averaging indicated by the pink polygons in panels e and f.

The time series of cloud thickness mirrors that of liquid water path, consistent with liquid water path being proportional to the square of cloud thickness (e.g. Figure 2.10). Clouds are thinner in the southeast Atlantic (216 m to 269 m) throughout the year than in the southeast Pacific (272 m to 317 m), consistent with adiabatic thickness derived from MODIS Terra by Wood (2012, see his Figure 5a). Clouds are also thinnest in May in the southeast Atlantic (216

m compared to the annual mean of 249 m), coinciding with the peak month of cloud-eroding boundaries. A secondary local minimum in cloud thickness occurs in November, driven by a corresponding local minimum in liquid water path (Figure 3.2 panels a and c) and associated with the small local maximum in occurrence of cloud-eroding boundaries according to our 10-year dataset (Figure 2.2).

The seasonal variability of *mean* cloud liquid water content (not shown) has a relatively small range in the southeast Atlantic (~0.24 g/kg in May compared to ~0.28 g/kg in Jan-Feb and Aug-Oct). Given the small range, and the fact that liquid water content near cloud top would be first to interact with entraining dry air from above the cloud, we instead look at the *peak* liquid water content (Figure 3.2, panels c and d). Peak liquid water content is highest in Jan-Feb (~0.54 g/m^2), the months when cloud-eroding boundaries are least frequent. As expected from the low liquid water path values, the month of May has the lowest peak liquid water content (~0.47 g/kg). The seasonal peak liquid water content over the southeast Atlantic qualitatively mirrors our analysis of cloud-eroding boundary occurrence.

3.1.3. MODIS liquid water path maps

Figure 3.3 shows spatial patterns of conditional time-averaged liquid water path for a selection of months, including the month of May (panels c and i), using the cloud fraction ≥ 0.9 and cloud top height ≤ 2.5 km conditions. In the southeast Atlantic, liquid water path in clouds is lowest in May especially around the convex coastline of Angola and Namibia (10° to 22.5°S), the same near-shore area where cloud-eroding boundaries begin their westward propagation. In the same month over the southeast Pacific, liquid water path values are considerably higher, even near the western coastline of South America. During months of low cloud-eroding boundary occurrence in the southeast Atlantic (February and September), liquid water path is ~40 g/m^2

higher in the area from 5° to 10°E and 10° to 20°S as compared to May. The westward extent of low liquid water path values (< 60 g/m^2) is also further west from the coast in May compared to all other months. The typical footprint of large cloud-eroding boundaries (e.g. Figure 1.1) covers the spatial pattern of low liquid water path extending west from the coast of Angola and Namibia from 10° to 22.5°S.

It is possible that some 1° × 1° pixels are on cloud-eroding boundaries, such that 10% of the pixels are completely eroded, resulting in lower liquid water path due to the cloud-eroding boundary rather than preconditioning of the cloud field by the background environment. To examine the sensitivity to our cloud fraction threshold of 0.9, Figure 3.4 compares the spatial patterns of conditional time-averaged liquid water path using cloud fraction thresholds of 0.9 and 1.0 (while still applying the cloud top height ≤ 2.5 km condition). With complete cloudiness (cloud fraction = 100%), the southeast Atlantic in May still has lower liquid water paths around the convex coastline where cloud-eroding boundaries occur.

Also evident from the spatial maps of conditional liquid water path in Figure 3.3 are the higher values in the southeast Pacific, consistent with the time series examined in the previous section (3.1.1). While there is some decline in liquid water path near the coast in both regions, consistent with previous studies (e.g. Wood and Hartmann, 2006), the reduction is much less in the southeast Pacific than in the southeast Atlantic. November is the only month in the southeast Pacific when liquid water path is notably lower (< 60 g/m^2 from 14° to 25°S) near the coast as compared to the same latitudes at 80°W, and in that month the low values only extend 3 to 4 degrees offshore (i.e. around 350 km). In contrast, liquid water path < 60 g/m^2 extends 10 degrees offshore (> 1000 km) at its widest point (15°S) in May over the southeast Atlantic.

3.1.4. Relationships between low cloud liquid water path, coastline structure, and the Benguela coastal low-level jet

The two annual local minima in liquid water path over the southeast Atlantic in May and November (see Figure 3.2 panel a) may be related to the interrelations between low marine clouds, the convex coastline of West Africa, and the Benguela low-level coastal jet. The southerly Benguela jet exhibits two maxima relative to the geography (Figure 3.5), one positioned around 25–27.5°S and a northward maximum at 17.5°S, near the Angola-Benguela SST front and the crest of the convex coastline. Based on Patricola and Chang's (2017) analysis of the fine-resolution (0.25° × 0.25°) satellite-based Scatterometer Climatology of Ocean Winds (SCOW) (Risien and Chelton, 2008) 10-meter monthly meridional (v) wind from September 1999 to October 2009, the northward maximum is the sole maximum during March-April-May (Figure 2.4). The interaction of the south-southeasterly jet with the convex coastline gives rise to a hydraulic expansion fan downwind of the coastal bulge (Patricola and Chang 2017).

Off the west African coast in the subtropical southeast Atlantic, the Froude number (the ratio of flow velocity to gravity wave speed) is likely greater than one in the shallow marine boundary layer (Winant et al. 1988). A hydraulic expansion takes place when the vertical boundaries (in our case, the African highlands) of the channel flow expand, allowing the flow to accelerate downstream. The hydraulic expansion fan implies lateral spreading of the flow and thus thinning of the marine boundary layer (Winant et al. 1988; Samelson, 1992). A sketch adapted from Winant et al. (1988, their Figure 8) is presented in Figure 3.6 for visual aid which includes the effects of friction. Patricola and Chang (2017) conducted simulations using the Weather Research and Forecasting (WRF) model version 3.5.1 with the original, convex coastline and a modified, straight north-south coastline to test the sensitivity of effects

characteristic of hydraulic expansion fans to the convex coastline. The simulation with the original coastline exhibited a stronger northward maximum of the Benguela jet in concert with a stronger zonal pressure gradient. The marine boundary layer was also lower in the downwind vicinity of the northward Benguela jet maximum (Figure 3.7).

The existence of the thinnest clouds in May (and the surrounding months, April and June) aligns with the more prevalent hydraulic expansion fan effect on the marine boundary layer during this season. Beginning in March and continuing through May, area-averaged cloud top temperature rises from ~276 K to 283 K, signaling the shallower marine boundary layer when the northern Benguela jet maximum is most active, according to Patricola and Chang (2017, see Figure 2.4). However, the spatial extent of thinner clouds in Apr-May-Jun is much further offshore than the expected area of influence by the hydraulic expansion fan.

Furthermore, the impact of the hydraulic expansion on the marine boundary layer is unclear when looking at 975-hPa horizontal divergence, the approximate level of the Benguela jet. Figure 3.8 shows much stronger convergence along the convex coastline from 10° to 23°S on the days with a cloud-eroding boundary (Yes days) in Apr-May-Jun compared to No days in Jan-Feb. There is a small pocket of divergence right at the convex bend near 17.5°S on the Yes days, but the same signature is also seen for the No days. A seasonal analysis of the low-level divergence does not indicate any important differences in the vicinity of the convex coastline nor where the cloud-eroding boundaries tend to occur (Figure 3.9). Given these findings, we conclude that the convex coastline structure of west Africa in the southeast Atlantic does not appear to have a major role in creating favorable environments for cloud-eroding boundaries.

3.1.5. MODIS cloud top temperature and MODIS-derived cloud thickness and peak liquid water content on days with and without cloud-eroding boundaries

We further examine the spatial patterns of low clouds in the southeast Atlantic in Apr-May-Jun, specifically looking at their peak liquid water content, thickness and cloud top. As in Section 3.1.4, peak liquid water content and cloud thickness are derived from MODIS liquid water path and cloud top property retrievals (see Section 2.6 for details). Plots of conditional (cloud fraction ≥ 0.9 and cloud top height ≤ 2.5 km) time-averaged liquid water path, cloud top temperature, peak liquid water content and cloud depth in Apr-May-Jun over the southeast Atlantic are presented in Figure 3.10. To assess differences of the day-to-day environment which may be a factor for cloud-eroding boundaries, we separately average the MODIS-derived fields by cloud-eroding boundary day. As in Section 2.1, we define "Yes days" as days when a cloud-eroding boundary was observed (based on our 11-year dataset, Figure 2.2) and "No days" as days absent of cloud-eroding boundaries. "Maybe" days (Section 2.1) are not included in this portion of the analysis.

On days when a cloud-eroding boundary is present, the clouds are thinner with lower peak liquid water content, particularly around the convex coastline of Angola and Namibia (10° to 22.5°S) where cloud-eroding boundaries occur (Figure 3.10 panels e–h). As expected, this result mimics what is seen for liquid water path in Section 3.1.3 and again in panels a and b of Figure 3.10. On Yes days, 200-m thin and less than 0.45 g/kg peak liquid water content clouds extend out to ~2°E from 10° to 15°S and out to 5°E at 20°S, with clouds thinner than 160 meters and with peak liquid water content around 0.3 g/kg in a 2–3° corridor around the convex coastline. On No days, clouds in that 2–3° corridor tend to be closer to 200 meters thick and 0.4 g/kg peak liquid water content, with clouds thicker than 200 meters and higher than 0.45 g/kg

peak liquid water content extending westward from ~7.5°E. The differences between peak liquid water content and cloud depth on Yes and No days in the area around the convex coastline (eastward of 5°E from 10° to 20°S) is around 0.05–0.1 g/kg and 40 meters, respectively. In the same area, cloud tops on Yes days are 2–3 K warmer (Figure 3.10, panels c and d). To summarize, clouds on Yes days are in a shallower marine boundary layer (evidenced by warmer cloud tops) and as a result are thinner with lower peak LWC due to lower liquid water path.

For comparison to previous work, we show the difference between cloud thickness derived using our true adiabatic treatment and using Wood and Taylor's (2001) linear adiabatic treatment (see Section 2.6 for details). Figure 3.11 shows the difference between our true adiabatic thickness (Z_{ad}) and thickness derived assuming a constant change of LWC with height in cloud (Z_{lin}). The Z_{ad} cloud depths are higher than those derived assuming a constant adiabatic water content change in cloud, but only by tens of meters. As expected, the discrepancy is smallest around the convex coastline where liquid water path is low. These clouds with lower liquid water path are thinner so using the actual adiabatic water-mixing ratio does not accumulate as much of a bias.

Inner quartile ranges for the fields plotted in Figure 3.10 (liquid water path, cloud top temperature, peak liquid water content, and cloud thickness) are included in Appendix F. The inner quartile ranges are smaller on Yes days in the area around the convex coastline where the cloud-eroding boundaries occur. The same tendency in the inner quartile ranges is not seen on No days. Hader (2016) reasoned that peak occurrence of cloud-eroding boundaries in May could be the result of "Goldilocks" conditions, or the overlap of environmental conditions which are favorable for initiation and maintenance of abrupt cloud erosion in long, continuous lines. The

results shown agree with this premise, leading us to infer that cloud susceptibility is a necessary but insufficient condition for cloud-eroding boundaries.

3.1.6. Role of lower boundary layer height in cloud susceptibility to erosion

Hader (2016) and Yuter et al. (2018, their Figure 4) documented the at times pervasive nature of gravity wave trains in the low cloud field over the southeast Atlantic. It is possible that the mechanism(s) which perturb cloud, such as gravity waves, may be present quite frequently but lack the ability to remove cloud in the southeast Pacific and during times of year in the southeast Atlantic when clouds are more resilient.

For cloud to be removed, air in the cloud must either remain sub-saturated after rapid entrainment or remain displaced downward. The hypothesized mechanism for abrupt, rapid cloud erosion in cloud-eroding boundaries is an enhanced entrainment event of the warm, dry air above cloud associated with gravity wave passage (Yuter et al. 2018; Hader, 2016). In marine low clouds, peak liquid water content is at cloud top owing to a near adiabatic water-mixing ratio (Cotton and Yuter, 2008; Miller, 2010).

We consider what happens when dry air is mixed into the two reservoirs of moisture within cloud; water vapor represented by the saturation mixing ratio and cloud liquid water content represented by estimated peak liquid water content (Section 2.6). When dry air is mixed into cloud, relative humidity immediately drops below 100% and cloud droplets evaporate. If the amount of dry air is small enough and the mass of cloud droplets is large enough, recovery to 100% relative humidity can be accomplished by evaporating only a portion of the initial cloud's liquid water content. In this case, the cloud remains after the entrainment, albeit with an initially lower liquid water content of cloud droplets which can recover as the cloud-boundary layer system reestablishes equilibrium with time. Retaining the cloud preserves the amplifying

feedbacks such as cloud top radiative cooling and boundary layer mixing that help move water vapor up from the ocean surface to cloud level. In contrast, total removal of cloud is a threshold condition that favors decoupling of the boundary layer and restricting moisture flux upwards to cloud level from the surface. If the evaporated mass of cloud drops needed to maintain saturation (100% relative humidity) is greater than or equal to the initial liquid water content, then the cloud will dissipate. Given the sensitive nature of the cloud system to dry air entrainment, any enhanced entrainment mechanism will make it easier to evaporate the cloud, and even more so for a thinner cloud with less liquid water. Furthermore, cloud top entrainment is strongly aided by negative buoyancy associated with evaporational cooling of condensate at cloud top (e.g. Lilly 1968; Deardorff 1980a; Randall 1980; Nicholls and Turton 1986). Since lower altitude clouds have lower peak LWC, all other things being equal, they would have less entrainment at cloud top.

We quantify roughly how much dry air would be needed to evaporate all liquid water at cloud top (i.e. the peak LWC). Given the disparity of cloud top temperatures between Yes and No days, we also want to test whether more or less dry air would be required to evaporate cloud top for a warmer cloud top. We use a range of cloud top temperatures from 270 to 290 K and a cloud top pressure of 925 hPa to explore the sensitivity of cloud susceptibility to boundary layer height. In these idealized calculations, we do not account for factors which would dictate the entrainment rate as we are solely focused on how the vertical placement of clouds may make them more or less susceptible to removal. Figure 3.12 shows the minimum peak liquid water content that would be required to maintain saturation after some fractional entrainment of dry air relative to the moist air at cloud top. For simplicity, the dry air is assumed to have a mixing ratio of 0 g/kg. The equations used to compute the values in Figure 3.12 are shown below:

$$w_{mixed} = \frac{(1\ m^3)w_{top}\ +\ V_{dry}\left(0\frac{g}{kg}\right)}{1 + V_{dry}} = \frac{w_{top}}{1 + V_{dry}} \tag{3.1}$$

$$\mathrm{LWC}_{mixed} = \frac{(1\ m^3)\mathrm{LWC}\ +\ V_{dry}\left(0\frac{g}{kg}\right)}{1 + V_{dry}} = \frac{\mathrm{LWC}}{1 + V_{dry}} \tag{3.2}$$

$$\text{Minimum peak LWC} = \left(1 + V_{dry}\right)\left(w_{s,top} - w_{mixed}\right) = V_{dry}w_{s,top} \tag{3.3}$$

where $w_{s,top}$ is the saturation mixing ratio at cloud top (a function of the temperature and pressure), w_{top} is the water vapor mixing ratio at cloud top, assumed to be equal to $w_{s,top}$ (i.e. RH = 100%), LWC is taken to be the peak liquid water content, V_{dry} is the volume of dry air being entrained, and w_{mixed} and LWC_{mixed} are the water vapor and liquid water mixing ratios after entrainment, respectively. Note that this formulation neglects the expected warming from mixing in the free tropospheric air above cloud, which would act to further dry out the cloud top since the resulting $w_{s,top}$ after mixing would be larger.

As expected, mixing in a greater portion of dry air requires more mass of cloud droplets to be evaporated to compensate. Furthermore, warmer cloud tops require more liquid water content to maintain saturation and are more sensitive to the amount of dry air entrained. Opposite to this effect of warmer cloud tops is higher cloud top pressure, though the sensitivity is much less. Increasing cloud top temperature by 2 K increases the liquid water content required to maintain saturation approximately five times more than decreasing cloud top pressure by 25 hPa. The impact of cloud top pressure becomes slightly more important with increasing cloud top temperature, but even at 288 K the impact of $\pm$25 hPa is still more than 4.5 times less than the impact of $\pm$2 K cloud top temperature.

3.1.7. Dry air required to evaporate liquid water at cloud top on days with and without cloud-eroding boundaries

We further characterize the differences between the environments on days when cloud-eroding boundaries do and do not occur by plotting the fractional entrainment of dry air (RH = 0%) that would be required to evaporate all liquid water at cloud top (the peak liquid water content). For this calculation, we use Equation 3.3 but solve for V_{dry} using peak LWC derived from MODIS and the MODIS cloud top temperature and pressure to compute the saturation mixing ratio at cloud top. We again represent V_{dry} as fractional entrainment, that is, the amount of dry air to entrain as a fraction of a unit volume of moist air at cloud top. Figure 3.13 shows the results of this calculation, conditioned on cloud fraction thresholds of 0.9 and 1.0 (as well as cloud top height $\leq$ 2.5 km). Exact values from the calculation are presented in Table 3. Where and when cloud-eroding boundaries occur, only 3.5% fractional entrainment is needed to evaporate the water at cloud top. The fractional entrainment is about 1% higher on days when cloud-eroding boundaries do not occur in the same area. Both are relatively small amounts of dry air, so a 1% difference between the Yes and No days means that 23% less dry air is needed to evaporate the peak liquid water content on Yes days relative to No days.

We also show the peak liquid water content and fractional entrainment required to evaporate it for No days in January and February for comparison (Figure 3.14). In accordance with the spatial pattern of liquid water path in February (Figure 3.3 panel a), peak liquid water content is a bit lower (< 0.5 g/kg) close to the coast (east of 10°E from 10° to 20°S) with less fractional entrainment (< 4.5%) required to evaporate the liquid water at cloud top there~. However, the values are not as low as for Apr-May-Jun and westward of 9°E the peak liquid

water content is greater than 0.5 g/kg and more than 5% fractional entrainment is required to evaporate the liquid water at cloud top.

Furthermore, while liquid water path is equally high in Aug-Sep-Oct as in Jan-Feb, cloud top temperature is roughly 5 K warmer in Aug-Sep-Oct, meaning those clouds require less dry air entrainment to be removed. Unsurprisingly, cloud-eroding are extremely rare in Jan-Feb (~1 in 50 days) whereas they are only uncommon in Aug-Sep-Oct, occurring roughly every 1 in 10 days. There may yet be other dynamical reasons, possibly related to the mechanism(s) of cloud erosion, why cloud-eroding boundaries are uncommon in Aug-Sep-Oct. Nonetheless, the evidence shown suggests that clouds which succumb to the boundaries tend to be more susceptible to cloud-eroding boundaries.

3.2. Large-scale factors relevant for cloud erosion

Previous work by Hader (2016) examined differences in some large-scale fields between times of year when cloud-eroding boundaries are more and less frequent (namely May and January, respectively). Hader (2016) was unable to identify a clear large-scale signal which coincided with the peak occurrence of cloud-eroding boundaries in May but did make the following observations. Stability is stronger in January, which favors wave propagation while inhibiting dry air entrainment. The estimated inversion strength (EIS) and subsidence suggest a less robust cloud deck in May, consistent with our findings of cloud susceptibility in Section 3.1. Finally, the semi-permanent high pressure is roughly 5 longitude closer to the coast in May than in January. In this study, we analyze many more large-scale fields through the lens of self-organizing maps (SOMs) and apply our test for relevance (see Section 2.5) to identify any large-scale patterns that may be more or less favorable for cloud-eroding boundaries.

The full set of fields for various months and times of day we used to train each SOM is shown in Tables 1 and 2. In total, over 200 SOMs were trained and analyzed. Most of the fields that were analyzed did not exhibit any relevance for cloud-eroding boundaries. As described in Section 2.5, we deem a spatial pattern, as analyzed by the SOM, relevant if at least 3 adjacent nodes all have unlikely to be random high/low counts of Yes or No boundary days (i.e. days with and without a cloud-eroding boundary, respectively). Some SOMs we ran had very few (if any) nodes with unlikely random counts of Yes or No days, while others would show one to two nodes, sometimes next to one another, as having unlikely random counts of Yes or No days. We leave perusal of these SOMs to Appendix G which includes an archive of selected SOMs from the 200+ that were trained. In the remainder of this section, we will focus on fields which do appear relevant for cloud-eroding boundaries, along with some that were not clearly relevant contrary to our expectations.

3.2.1. Low boundary layer height

We begin by presenting SOMs of ERA5 boundary layer height, the only field that consistently passed our statistical relevance test for SOMs of different time periods *and* different training data (normalized anomalies vs. the full field). We first look at the results from SOMs trained on Apr-May-Jun, since this time period gives us a robust sample size of both Yes and No days. Figure 3.15 depicts the averaged best-matching boundary layer height fields projected onto the SOM nodes trained on boundary layer height anomalies. We also show the SOM trained on the full field of boundary layer height in Figure 3.16. Depicting a SOM trained on the full field does not require projecting averages of the best-matching records in the full field into each node, since the weights in the SOM are trained to the full values (rather than anomalies). As a result, the spatial patterns in Figure 3.16 are smoother than in Figure 3.15.

Both SOMs reveal statistical relevance for the spatial pattern of low ERA5 boundary layer height (< 500 m) around the convex coastline between 10° and 22.5°S and out to 5°E over the southeast Atlantic. The pattern has unlikely random high counts of Yes days in nodes 6, 11, 16, 17, and 21 of the SOM trained on normalized anomalies (Figure 3.15) and in nodes 1, 2, 16, 21, 22 and 24 of the SOM trained on the full fields (Figure 3.16).

In both SOMs, nodes on the opposite side (10 and 20 in Figure 3.15 and nodes 3, 4, 5, 10, 15, and 20 in Figure 3.16) have unlikely random low counts of Yes days and/or high counts of No days and exhibit higher boundary layer height around the convex coastline (10° to 20°S) and a sharper increase of boundary layer height away from the coast. Note that models, including ERA5 reanalysis, often underestimate boundary layer height near the coast (e.g. Rahn and Garreaud, 2010). Errors in the absolute values of near-coast boundary layer height should not detract from the comparative results herein for Yes days with relatively lower boundary layer height around the convex coastline and a more gradual rise in boundary layer height moving away from the coast as compared to No days.

While the lower boundary layer height is amplified downwind of the convex coastline bend at 17.5°S, the overall signature of low boundary layer extending out from the coastal bulge (approx. 12° to 22°S) has a much larger footprint than could be explained by characteristics of a hydraulic expansion fan (Section 3.1.4). Furthermore, in the nodes with unlikely random high Yes days, the low boundary layer height extends southward to 30°S. The height of the marine boundary layer depends on several factors, including the large-scale subsidence and sea surface temperatures. Apr-May-Jun marks the time of year when the subtropical southeast Atlantic is shifting to cooler SSTs. In ERA5, the SST along the coast from 17.5° to 30°S is about 2 K cooler on the Yes days in Apr-May-Jun compared to No days in Jan-Feb (Figure 3.17). The

dominating factors driving low boundary layer height out over the ocean, especially from 10° to 20°S, remains an active area of investigation.

An additional SOM trained on normalized anomalies of boundary layer height only in May (see Appendix G) is noisier than the SOM for Apr-May-Jun owing to the smaller sample size, but shows a similar spatial pattern of lower boundary layer heights around the convex coastline for Yes Days.

3.2.2. Boundary layer height in the southeast Pacific

We evaluate differences in ERA5 boundary layer height between the southeast Atlantic and southeast Pacific using SOMs. Figure 3.18 shows averaged best-matching boundary layer height patterns for the SOM trained on normalized anomalies of boundary layer height in the southeast Pacific at the same local time as Figure 3.15 (~11 pm).

As expected, boundary layer height increases westward away from the coasts of Chile and Peru (Mechoso et al. 2014). A key difference in the spatial patterns of boundary layer height changes between the southeast Atlantic and southeast Pacific is that the pattern in the southeast Pacific is relatively consistent north to south as a function of distance offshore from the coast, whereas the southeast Atlantic features north to south asymmetry in the zonal boundary layer height variations with lower boundary heights around the convex coastline where the cloud-eroding boundaries occur (Section 3.2.1). The close packing of contours illustrates the relatively sharp gradient of boundary layer height perpendicular to the coastline along the entire coast of Chile and Peru in all nodes of the SOM shown in Figure 3.18.

It appears that the southeast Pacific favors deeper boundary layers than the southeast Atlantic, in agreement with lower cloud top temperatures, higher liquid water path, and thicker clouds (Figure 3.2). Furthermore, the sharp increase of boundary layer height away from the

coast is present in all southeast Pacific boundary layer height SOM nodes throughout the year (see Appendix G), in stark contrast with the pattern of lower boundary layer height around the convex coastline in the southeast Atlantic when cloud-eroding boundaries are more frequent.

3.2.3. Projections of MODIS data on boundary layer height SOMs

We can also project observed variables from MODIS onto the SOM nodes by averaging the spatial patterns from the days which best match each node. Since we only trained SOMs on ERA5 data from 2007 to 2017 (the current extent of the cloud-eroding boundary dataset), we are only able to include MODIS data from this time range. The key point of this sub-section is to provide another assessment of cloud susceptibility between Yes and No days through the lens of the boundary layer height SOM, since liquid water path and other cloud top properties retrieved by MODIS are a function of cloud top height (i.e. boundary layer height).

Figure 3.19 shows the conditional (cloud fraction ≥ 0.9 and cloud top ≤ 2.5 km) MODIS liquid water path projected onto the SOM trained on normalized anomalies of boundary layer height. The spatial patterns of liquid water path are noisier than the boundary layer height patterns reviewed in the previous sub-sections (Secs. 3.2.1 and 3.2.2). This is mostly due to small sample sizes (Figure 3.20). Of importance are the pattern similarities in liquid water path between nodes which all have unlikely random high counts of Yes days. In most of these nodes (1, 2, 6, 16, 21, 22), liquid water path is less than 50 g/m^2 around the convex coastline of Angola and Namibia (10° to 22.5°S, Figure 3.19), in the same place where low boundary layer heights are found in the trained SOM. The association of lower liquid water path with lower boundary layer height is expected and gives us more confidence that ERA5 boundary layer height despite its known errors in absolute magnitude (Section 2.4) is at least qualitatively capturing the spatial patterns in the southeast Atlantic.

3.2.4. Possible factor: estimated inversion strength

Given the findings of Hader (2016) regarding weaker stability in May compared to January, as determined by estimated inversion strength (EIS), we expected to find relevance of EIS for cloud-eroding boundaries. While a stronger inversion is more favorable for gravity wave propagation, a weaker inversion makes it easier to entrain subsiding dry air from the above the inversion into the cloud. Using EIS derived from ERA5 data (see Section 2.4), Figure 3.21 shows average EIS projected onto the SOM trained on normalized anomalies of EIS, while Figure 3.22 shows the SOM trained on the full field of EIS, both for Apr-May-Jun in the southeast Atlantic at 22 UTC.

In both SOMs of EIS, node 1 exhibits the weakest stability and, in the opposite corner, node 25 has the strongest stability. While there are associations between a few nodes with unlikely random counts of Yes days and weaker stability (nodes 6, 11, and 21 in Figure 3.21 and nodes 1 and 23 in Figure 3.22) and the converse of unlikely random counts of No days and stronger stability (nodes 5, 20, and 25 in Figure 3.21 and nodes 5, 10 in Figure 3.22), neither the anomaly or full field EIS SOMs pass our tests for statistical relevance. Myers and Norris (2013) found that EIS is the dominant factor contributing to climatological variability of cloud fraction and EIS and subsidence contribute about equally to climatological variability in liquid water path. They also found that for conditions of uniform subsidence, enhanced EIS reduces cloud top height but increases cloud fraction.

4. Conclusions

Analysis of the large-scale environments in which cloud-eroding boundaries do and do not occur has been conducted. Observations suggest that the rapid cloud erosion along sharp transitions 100s of km long is associated with atmospheric gravity waves (Hader, 2016; Yuter et al, 2018). Building on the work of Hader (2016) and Yuter et al. (2018), satellite data were used to expand the climatology of cloud-eroding boundaries from 5 to 11 years (now 2007–2017). Satellite data were also interrogated to address cloud susceptibility to erosion in the southeast Atlantic and southeast Pacific, as well as on days when cloud-eroding boundaries do and do not occur in the southeast Atlantic. In addition to satellite data, model reanalysis was explored with self-organizing maps (SOMs) to tease out any large-scale factors which may be relevant for cloud-eroding boundaries. SOMs were trained on various ERA5 reanalysis fields (e.g. mean sea level pressure, geopotential heights, SST, estimated inversion strength, lower tropospheric stability, boundary layer height, relative humidity, and low-level u and v winds) and visualized to look for consistent spatial patterns among neighboring nodes with high or low numbers of matching days with or without cloud-eroding boundaries. This study also developed objective criteria for determining relevance from SOM results that involves three or more adjacent nodes with incidence of the feature of interest at greater than 95% probability based on Monte Carlo random sampling.

The key findings from the study are:

- Marine low clouds in the southeast Atlantic during April-May-June often have markedly lower liquid water paths and are thinner than clouds in the southeast Atlantic during other seasons and throughout the year in the southeast Pacific.

- During May in the southeast Atlantic, the peak month of cloud-eroding boundaries, monthly average liquid water path conditioned on cloud being present is at the annual minimum and cloud top temperature is at its annual maximum (lowest altitude). Consequently, conditional peak liquid water content is lowest (~0.47 g/kg) and clouds are thinnest (216 m).
- Within the month of May, cloud liquid water paths and cloud top heights conditioned on the presence of cloud fractions ≥ 0.9 (e.g. not including clear skies) are lower on days with cloud-eroding boundaries than those without.
- Only small amounts of entrained dry air (on the order of 10^{-2} m^3 per 1 m^3 of cloudy air) are needed to remove marine low cloud when and where cloud-eroding boundaries typically occur in the southeast Atlantic.
- Self-organizing maps of ERA5 reanalysis show that lower boundary layer heights near the convex west African coast (from 10° to 25°S) were consistently relevant for cloud-eroding boundaries using different input data (normalized anomalies vs. full data), seasons and times of day.
- An important negative result from the SOM analysis is a lack of robust associations between the occurrence of cloud eroding boundaries and the environmental factors of estimated inversion strength, mean sea level pressure, sea surface temperature, lower tropospheric stability, geopotential height fields at 1000, 950, 900, 850 and 500 hPa, relative humidity at 800, 750 and 700 hPa, and low-level *u* and *v* winds at 1000, 950 and 900 hPa.

The reason why cloud-eroding boundaries happen in the southeast Atlantic, and more often in April, May and June, appears two-fold. Both the southeast Atlantic and southeast

Pacific have strong land-breezes and downslope flow associated with coastal topography, which is hypothesized to be related to triggering the cloud-eroding waves. The key differences we have identified between the two regions is the lower liquid water path, thinner clouds, and lower boundary layer height in the southeast Atlantic. The combination of lower liquid water path and a shallower boundary layer to the north of the convex coastline (17.5°S) yield clouds more susceptible to removal by cloud-eroding boundaries. In addition to cloud susceptibility to entrainment of dry air, low boundary layer height may be important for some other dynamical reason related to the atmospheric gravity waves, or other factors may be playing a larger role in the cloud erosion. Although sustainment of cloud-eroding boundaries is not specifically addressed in this study, the westward extent (out to near 0°E) of thin and low clouds is likely related to how far west rapid cloud erosion associated with atmospheric gravity waves can occur.

Differences in the nighttime temperature depression over the African coastal highlands, used as a proxy for downslope flow strength, do distinguish between days with and without cloud-eroding boundary occurrence (Appendix C). Further investigation of the dynamical aspects of large-scale environments that favor generation of cloud-eroding waves is merited and will likely require high resolution in-situ measurements. The SOMs trained on mean sea level pressure do not indicate a preference for the location of the semi-permanent subtropical high pressure when cloud-eroding boundaries occur. Previous studies have found low-cloud cover variations associated with changes in the position of the subtropical high (Klein et al. 1995; Garreaud et al. 2001). A seasonal shift in the central location of the high pressure may also be important for modulating the low cloud deck, but we leave this investigation to future work.

When the vertical resolution of model output is too coarse to resolve the shallow inversion at the top of the subtropical marine stratocumulus boundary layer, accurate estimation

of the height of the boundary layer from a model can be difficult (Section 2.4). The ERA5 boundary layer height field utilized in this study is taken as the vertical level where the Richardson number exceeds 0.25 (ECMWF, 2016). Our plots of ORACLES Lidar data overlaid with ERA5 boundary layer height and potential temperature contours in Appendix A confirm a low bias by ERA5 boundary layer height compared to the cloud top sensed by airborne lidar. Despite this low bias, the relative patterns in spatial variations in boundary layer heights from ERA5 appear to be physically reasonable and the sanity check of comparing the independent MODIS liquid water paths with ERA5 boundary layer heights in Figure 3.19 shows relative consistency. Instead of the ERA5 boundary layer height field, a better measure to use for training SOMs may be based on the vertical gradient of potential temperature or the level where relative humidity drops below 70% (Shinozuka et al. 2019). Near future work will compare the ERA5 boundary layer height with the relative humidity criteria for the ORACLES sample to determine if redoing the SOM boundary layer height analysis is warranted.

Questions remain regarding the exact mechanisms which govern the rapid cloud erosion along cloud-eroding boundaries. What is the strength of the gravity wave perturbations which lead to erosion? To what degree is entrainment of dry air from above the boundary layer enhanced by gravity waves? What is the amount of dry air mixed into cloud? What amplifying feedbacks are involved? Why do some cloud-eroding boundaries have wave trains (Yuter et al. 2018) while others do not? A field campaign aims to address these questions using airborne cloud radar, lidar, and numerous dropsondes. A feature that field project scientists can check for in 3-day model forecasts to predict days in May that are more likely to have cloud-eroding boundaries are an especially shallow marine boundary layer around the convex coastline of west Africa.

In addition, there are still questions about the radiative impact of cloud-eroding boundaries, their sensitivity to climate change and interdependence with large-scale circulations (e.g. ENSO or the Indian monsoon). An active area of research is automated image processing of visible and infrared satellite imagery to identify cloud-eroding boundaries and quantify the amount of cloud removed over a 30-year span. At this point, like Hader (2016), we are still unable to comment on whether rapid clearing of marine stratocumulus decks in the southeast Atlantic would become more frequent in a warming climate.

REFERENCES

Albrecht, B. A., C. W. Fairall, D. W. Thomson, A. B. White, J. B. Snider, and W. H. Schubert, 1990: Surface-based remote-sensing of the observed and the adiabatic liquid water content of stratocumulus clouds. *Geophys. Res. Lett.*, **17**, 89–92.

Allen, G., G. Vaughan, T. Toniazzo, H. Coe, P. Connolly, S. E. Yuter, C. D. Burleyson, P. Minnis, and J. K. Ayers, 2013: Gravity-wave-induced perturbations in marine stratocumulus. *Q.J.R. Meteorol. Soc.*, **139**, 32-45.

Anderson-Frey, A. K., Y. P. Richardson, A. R. Dean, R. L. Thompson, and B. T. Smith, 2017: Self-Organizing Maps for the Investigation of Tornadic Near-Storm Environments. *Wea. Forecasting,* **32**, 1467–1475.

Annas, S., T. Kanai, and S. Koyama, 2007: Principal component analysis and self-organizing map for visualizing and classifying fire risks in forest regions, *Agricultural Information Research*, **16**, No. 2, 44–51, ISSN 1881-5219

Astel, A., S. Tsakouski, P. Barbieri, and V. Simeonov, 2007: Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets, *Water Research,* **41**, 4566–4578, ISSN 0043-1354

Anderson-Frey, A. K., Y. P. Richardson, A. R. Dean, R. L. Thompson, and B. T. Smith, 2017: Self-Organizing Maps for the Investigation of Tornadic Near-Storm Environments. *Wea. Forecasting,* **32**, 1467–1475, https://doi.org/10.1175/WAF-D-17-0034.1

Ao, C. O., D. E. Waliser, S. K. Chan, J.-L. Li, B. Tian, F. Xie, and A. J. Mannucci, 2012: Planetary boundary layer heights from GPS radio occultation refractivity and humidity profiles, *J. Geophys. Res.*, **117**, D16117.

Bailey, A., T. N. Chase, J. J. Cassano, and D. Noone, 2011: Changing Temperature Inversion Characteristics in the U.S. Southwest and Relationships to Large-Scale Atmospheric Circulation. *J. Appl. Meteor. Climatol.,* **50**, 1307–1323.

Baum, B. A., W. P. Menzel, R. A. Frey, D. C. Tobin, R. E. Holz, S. A. Ackerman, A. K. Heidinger, and P. Yang, 2012: MODIS Cloud-Top Property Refinements for Collection 6. *J. Appl. Meteor. Climatol.,* **51**, 1145–1163.

Birch, C. E. and M. J. Reeder, 2013: Wave-cloud lines over northwest Australia. *Q.J.R. Meteorol. Soc.*, **139**, 1311-1326.

Burleyson, C. D. and S. E. Yuter, 2015: Patterns of Diurnal Marine Stratocumulus Cloud Fraction Variability. *J. Appl. Meteor. Climatol.,* **54**, 847–866.

Brunet C., R. Casotti, V. Vantrepotte, 2008: Phytoplankton diel and vertical variability in photobiological responses at a coastal station in the Mediterranean Sea, *Journal of Plankton Research*, **30**, No. 6, 645–654, https://doi.org/10.1093/plankt/fbn028.

Cavazos, T., 2000: Using Self-Organizing Maps to Investigate Extreme Climate Events: An Application to Wintertime Precipitation in the Balkans. *J. Climate,* **13**, 1718–1732.

Chao, S., 1985: Coastal Jets in the Lower Atmosphere. *J. Phys. Oceanogr.,* **15**, 361–371.

Chen, T., W. B. Rossow, and Y. C. Zhang, 2000: Radiative effects of cloud-type variations. *J. Climate*, **13**, 264–286.

Christie, D. R., K. J. Muirhead, and R. H. Clarke, 1981: Solitary waves in the lower atmosphere, *Nature*, **293**(5827), 46-49.

Clarke, R. H., R. K. Smith, and D. G. Reid, 1981: The Morning Glory of the Gulf of Carpentaria: An Atmospheric Undular Bore. *Mon. Wea. Rev.,* **109**, 1726–1750.

Connolly, P. J., G. Vaughan, P. Cook, G. Allen, H. Coe, T. W. Choularton, C. Dearden, and A. Hill, 2013: Modelling the effects of gravity waves on stratocumulus clouds observed during VOCALS-UK. *Atmos. Chem. Phys.*, **13**, 7133-7152.

Cotton, W. and S. E. Yuter, 2008: Ch 2—Principles of Cloud and Precipitation Formation in Aerosol pollution impacts on precipitation: A scientific review. WMO/IUGG International Aerosol Precipitation Science Assessment Group, *Springer*, 13-43.

Crosbie, E., Z. Wang, A. Sorooshian, P. Y. Chuang, J. S. Craven, M. M. Coggon, M. Brunke, X. Zeng, H. Jonsson, R. K. Woods, R. C. Flagan, and J. H. Seinfeld, 2016: Stratocumulus Cloud Clearings and Notable Thermodynamic and Aerosol Contrasts across the Clear–Cloudy Interface. *J. Atmos. Sci.,* **73**, 1083–1099.

da Silva, J. C. B. and J. M. Magalhães, 2009: Satellite observations of large atmospheric gravity waves in the Mozambique Channel, *Int. J. Remote Sens.*, **30**(5), 1161-1182.

Deardorff, J. W., 1980: Cloud top entrainment instability. *J. Atmos. Sci.*, **37**, 561–563.

Désalmand, F., A. Szantai, L. Picon, and M. Desbois, 2003: Systematic observation of westward propagating cloud bands over the Arabian Sea during Indian Ocean Experiment (INDOEX) from Meteosat-5 data, *J. Geophys. Res.*, **108**, 8004.

de Szoeke, S. P., K. L. Verlinden, S. E. Yuter, and D. B. Mechem, 2016: The Time Scales of Variability of Marine Low Clouds. *J. Climate,* **29**, 6463–6481.

Dzambo, A. M., T. L'Ecuyer, O. O. Sy, and S. Tanelli, 2019: The Observed Structure and Precipitation Characteristics of Southeast Atlantic Stratocumulus from Airborne Radar during ORACLES 2016–17. *J. Appl. Meteor. Climatol.,* **58**, 2197–2215.

ECMWF, 2016: Turbulent transport and interactions with the surface, In *IFS documentation CY41r2 Part IV: Physical Processes*, ECMWF: Reading, UK, pp. 52, https://www.ecmwf.int/node/16648.

Garreaud, R. and R. Muñoz, 2004: The Diurnal Cycle in Circulation and Cloudiness over the Subtropical Southeast Pacific: A Modeling Study. *J. Climate,* **17**, 1699–1710.

Garreaud, R., J. Rutllant, J. Quintana, J. Carrasco, and P. Minnis, 2001: CIMAR-5: A snapshot of the lower troposphere over the subtropical southeast Pacific. *Bull. Amer. Meteor. Soc.*, **82**, 2193–2207.

George, R. C. and R. Wood, 2010: Subseasonal variability of low cloud radiative properties over the southeast Pacific Ocean, *Atmos. Chem. Phys.*, **10**, 4047-4063, https://doi.org/10.5194/acp-10-4047-2010.

Gerber, H., 1996: Microphysics of marine stratocumulus with two drizzle modes. *J. Atmos. Sci.*, **53**, 1649–1662.

Greenwald, T. J., G. L. Stephens, S. A. Christopher, and T. H. VonderHaar, 1995: Observations of the global characteristics and regional radiative effects of marine cloud liquid water. *J. Climate*, **8**, 2928–2946.

Hader, J. D., 2016: Propagating, Cloud-eroding Boundaries in Southeast Atlantic Marine Stratocumulus, Master's thesis, *N.C. State Univ.*, Raleigh, NC.

Hahn, C. J., and S. G. Warren, 2007: A gridded climatology of clouds over land (1971–96) and ocean (1954–97) from surface observations worldwide. Numeric Data Package NDP-026EORNL/CDIAC-153, CDIAC, Department of Energy, Oak Ridge, TN.

Hartmann, D. L., M. E. Ockert-Bell, and M. L. Michelson, 1992: The effect of cloud type on Earth's energy balance: Global analysis, *J. Climate*, **5**, 1281–1304.

Holz, R. E., S. A. Ackerman, F. W. Nagle, R. Frey, S. Dutcher, R. E. Kuehn, M. A. Vaughan, and B. Baum, 2008: Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP, *J. Geophys. Res.*, **113**, D00A19.

Hubanks, P., S. Platnick, M. King, and B. Ridgway, 2019: MODIS Algorithm Theoretical Basis Document No. ATBD-MOD-30 for Level-3 Global Gridded Atmosphere Products (08_D3, 08_E3, 08_M3) and Users Guide. *NASA Goddard Space Flight Center*, Sep 2019, https://atmosphere-imager.gsfc.nasa.gov/sites/default/files/ModAtmo/L3_ATBD_C6_C61_2019_02_20.pdf.

Kennedy, A. D., X. Dong, and B. Xi, 2016: Cloud fraction at the ARM SGP site: reducing uncertainty with self-organizing maps, *Theor. Appl. Climatol.*, **124**, 43-54, https://doi.org/10.1007/s00704-015-1384-3.

King, M. D., S.-C. Tsay, S. E. Platnick, M. Wang, and K.-N. Liou, 1998: MODIS Algorithm Theoretical Basis Document No. ATBD-MOD-06, MOD06 – Cloud product. *NASA Goddard Space Flight Center*, Sep 2019, https://atmosphere-imager.gsfc.nasa.gov/sites/default/files/ModAtmo/atbd_COP.pdf.

Klein, S. A. and C. Jakob, 1999: Validation and Sensitivities of Frontal Clouds Simulated by the ECMWF Model. *Mon. Wea. Rev.,* **127**, 2514–2531.

Klein, S. A., 1997: Synoptic Variability of Low-Cloud Properties and Meteorological Parameters in the Subtropical Trade Wind Boundary Layer. *J. Climate,* **10**, 2018–2039.

Klein, S. A., D. L. Hartmann, and J. R. Norris, 1995: On the relationships among low-cloud structure, sea surface temperature, and atmospheric circulation in the summertime northeast Pacific. *J. Climate*, **8**, 1140–1155.

Klein, S. A. and D. L. Hartmann, 1993: The Seasonal Cycle of Low Stratiform Clouds. *J. Climate,* **6**, 1587–1606.

Kloesel, K.A., 1992: Marine Stratocumulus Cloud Clearing Episodes Observed during FIRE. *Mon. Wea. Rev.,* **120**, 565–578.

Koch, S. E., W. Feltz, F. Fabry, M. Pagowski, B. Geerts, K. M. Bedka, D. O. Miller, and J. W. Wilson, 2008: Turbulent Mixing Processes in Atmospheric Bores and Solitary Waves Deduced from Profiling Systems and Numerical Simulation. *Mon. Wea. Rev.,* **136**, 1373–1400.

Kohonen, T., 2001: Self-organizing maps. 3rd edition Berlin, *Springer*.

Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. R. Hardy, 1997: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System, *J. Geophys. Res.*, **102**(D19), 23429–23465.

Kursinski, E. R., 1997: The GPS radio occultation concept: theoretical performance and initial results, Ph.D. thesis, *Calif. Inst. Of Technol.*, Pasadena, CA.

Kursinski, E. R., G. A. Hajj, K. R. Hardy, L. J. Romans, and J. T. Schofield, 1995: Observing tropospheric water vapor by radio occultation using the Global Positioning System, *Geophys. Res. Lett.*, **22**(17), 2365-2368.

Lau, N. and M. W. Crane, 1995: A Satellite View of the Synoptic-Scale Organization of Cloud Properties in Midlatitude and Tropical Circulation Systems. *Mon. Wea. Rev.,* **123**, 1984–2006.

Lilly, D. K., 1968: Models of cloud-topped mixed layers under a strong inversion. *Quart. J. Roy. Meteor. Soc.*, **94**, 292–309.

Lin, G.-F. and L.-H. Chen, 2006: Identification of homogeneous regions for regional frequency analysis using the self-organizing map, *Journal of Hydrology,* **324**, 1-9, ISSN 0022-1694

Liu, Y. and R. H. Weisberg, 2011: A Review of Self-Organizing Map Applications in Meteorology and Oceanography, *Self Organizing Maps – Applications and Novel Algorithm Design*, Dr Josphat Igadwa Mwasiagi (Ed.), ISBN: 978-953-307-546-4, InTech, Available from: http://www.intechopen.com/books/selforganizing-maps-

applications-and-novel-algorithm-design/a-review-of-self-organizing-map-applications-inmeteorology-and-oceanography

Liu, Y. and R. H. Weisberg, 2007: Ocean currents and sea surface heights estimated across the West Florida Shelf. *Journal of Physical Oceanography,* **37**, 1697-1713, ISSN 0022- 3670

Liu, Y., R. H. Weisberg, and C. N. K. Mooers, 2006a: Performance evaluation of the self-organizing map for feature extraction, *Journal of Geophysical Research,* **111**, C05018, doi:10.1029/2005JC003117, ISSN 0148-0227

Liu, Y., R. H. Weisberg, and R. He, 2006b: Sea Surface Temperature Patterns on the West Florida Shelf Using Growing Hierarchical Self-Organizing Maps. *J. Atmos. Oceanic Technol.,* **23**, 325–338.

Liu, Y. and R. H. Weisberg, 2005: Patterns of ocean current variability on the West Florida Shelf using the self-organizing map*, Journal of Geophysical Research,* **110**, C06003, doi:10.1029/2004JC002786, ISSN 0148-0227

Lutzak, P. A., 2013: A Proposal for Analyzing and Forecasting Lower-Atmospheric Undular Bores in the Western Gulf of Mexico Region. *Wea. Forecasting,* **28**, 55–76.

Manabe, S., and R. F. Strickler, 1964: Thermal equilibrium of the atmosphere with a convective adjustment, *J. Atmos. Sci.*, **21**, 361–385.

Mapes, B. E., 1993: Gregarious Tropical Convection. *J. Atmos. Sci.,* **50**, 2026–2037.

Mechem, D. B., C. S. Wittman, M. A. Miller, S. E. Yuter, and S. P. de Szoeke, 2018: Joint Synoptic and Cloud Variability over the Northeast Atlantic near the Azores. *J. Appl. Meteor. Climatol.,* **57**, 1273–1290, https://doi.org/10.1175/JAMC-D-17-0211.1.

Mechoso, C. R., R. Wood, R. Weller, C. S. Bretherton, A. D. Clarke, H. Coe, C. Fairall, J. T. Farrar, G. Feingold, R. Garreaud, C. Grados, J. McWilliams, S. P. de Szoeke, S. E. Yuter, and P. Zuidema, 2014: Ocean–Cloud–Atmosphere–Land Interactions in the Southeastern Pacific: The VOCALS Program. *Bull. Amer. Meteor. Soc.,* **95**, 357–375.

Menzel, W. P., R. A. Frey, H. Zhang, D. P. Wylie, C. C. Moeller, R. E. Holz, B. Maddux, B. A. Baum, K. I. Strabala, and L. E. Gumley, 2008: MODIS Global Cloud-Top Pressure and Amount Estimation: Algorithm Description and Results. *J. Appl. Meteor. Climatol.,* **47**, 1175–1198.

Miles, N. L., J. Verlinde, and E. E. Clothiaux, 2000: Cloud droplet size distributions in low-level stratiform clouds. *J. Atmos. Sci.*, **57**, 295–311.

Miller, M. A., 2010: Satellite Observations of Low Marine Clouds, Ph.D. thesis, *N.C. State Univ.*, Raleigh, NC.

Muhlbauer, A., I. L. McCoy, and R. Wood: Climatology of stratocumulus cloud morphologies: microphysical properties and radiative effects, *Atmos. Chem. Phys.*, **14**, 6695-6716, https://doi.org/10.5194/acp-14-6695-2014.

Muñoz, R. C. and R. Garreaud, 2005: Dynamics of the Low-Level Jet off the West Coast of Subtropical South America. *Mon. Wea. Rev.,* **133**, 3661–3677.

Myers, T.A. and J.R. Norris, 2013: Observational Evidence That Enhanced Subsidence Reduces Subtropical Marine Boundary Layer Cloudiness. *J. Climate,* **26**, 7507–7524.

Nicholls, S. and J. Leighton, 1986: An observational study of the structure of stratiform cloud sheets: Part I. Structure. *Q.J.R. Meteorol. Soc.*, **112**, 431-460.

Nicholls, S. and J. D. Turton, 1986: An observational study of the structure of stratiform cloud sheets: Part II. Entrainment. *Quart. J. Roy. Meteor. Soc.*, **112**, 461–480.

O'Dell, C. W., F. J. Wentz, and R. Bennartz, 2008: Cloud Liquid Water Path from Satellite-Based Passive Microwave Observations: A New Climatology over the Global Oceans. *J. Climate,* **21**, 1721–1739.

Patricola, C. M. and P. Chang, 2017: Structure and dynamics of the Benguela low-level coastal jet. *Clim Dyn*, **49**, 2765, https://doi.org/10.1007/s00382-016-3479-7.

Platnick, S., M. King, P. Hubanks, 2017: MODIS Atmosphere L3 Daily Product. NASA MODIS Adaptive Processing System, *Goddard Space Flight Center*, doi:10.5067/MODIS/MOD08_D3.061.

Rahn, D. A. and R. Garreaud, 2010: Marine boundary layer over the subtropical southeast Pacific during VOCALS-Rex – Part 1: Mean structure and diurnal cycle, *Atmos. Chem. Phys.*, **10**, 4491-4506, https://doi.org/10.5194/acp-10-4491-2010.

Randall, D. A., J. A. Coakley, C. W. Fairall, R. A. Kropfli, and D. H. Lenschow, 1984: Outlook for Research on Subtropical Marine Stratiform Clouds, *Bulletin of the American Meteorological Society*, **65**(12), 1290-301.

Randall, D. A., 1980: Conditional instability of the first kind upside down. *J. Atmos. Sci.*, **37**, 125–130.

Reusch, D. B., R. B. Alley, and B. C. Hewitson, 2007: North Atlantic climate variability from a self-organizing map perspective, *J. Geophys. Res.*, **112**, D02104.

Reusch, D. B., R. B. Alley, and B. C. Hewitson, 2005: Towards ice-core-based synoptic reconstructions of west 54ntarctic climate with artificial neural networks. *International Journal of Climatology,* **25**, 581-610, ISSN 0899-8418.

Rozendaal, M.A. and W.B. Rossow, 2003: Characterizing Some of the Influences of the General Circulation on Subtropical Marine Boundary Layer Clouds. *J. Atmos. Sci.,* **60**, 711–728.

Samelson, R. M., 1992: Supercritical marine-layer flow along a smoothly Varying Coastline. *J Atmos Sci.*, **49**, 1571–1584.

Seethala, C., and Á. Horváth, 2010: Global assessment of AMSR-E and MODIS cloud liquid water path retrievals in warm oceanic clouds, *J. Geophys. Res.*, **115**, D13202.

Seidel, D., Y. Zhang, A. Beljaars, J.-C. Golaz, and B. Medeiros, 2012: Climatology of the planetary boundary layer over continental United States and Europe. *J. Geophys. Res.*, **117**, D17106.

Schuenemann, K. C., J. J. Cassano, and J. Finnis, 2009: Synoptic forcing of precipitation over Greenland: Climatology for 1961–99. *J. Hydrometeor.*, **10**, 60–78, https://doi.org/10.1175/2008JHM1014.1.

Shinozuka, Y., P. E. Saide, G. A. Ferrada, S. P. Burton, R. Ferrare, S. J. Doherty, H. Gordon, K. Longo, M. Mallet, Y. Feng, Q. Wang, Y. Cheng, A. Dobracki, S. Freitag, S. G. Howell, S. LeBlanc, C. Flynn, M. Segal-Rosenhaimer, K. Pistone, J. R. Podolske, E. J. Stith, J. R. Bennett, G. R. Carmichael, A. da Silva, R. Govindaraju, R. Leung, Y. Zhang, L. Pfister, J.-M. Ryoo, J. Redemann, R. Wood and P. Zuidema, 2019: Modeling the smoky troposphere of the southeast Atlantic: a comparison to ORACLES airborne observations from September of 2016, Submitted to *Atmos. Chem. Phys.*

Slingo, J. M., 1987: The Development and Verification of A Cloud Prediction Scheme For the Ecmwf Model. *Q.J.R. Meteorol. Soc.*, **113**, 899-927.

Stephens, G. L., 2005: Cloud Feedbacks in the Climate System: A Critical Review. *J. Climate,* **18**, 237–273.

Stephens, G. L., 1978: Radiation Profiles in Extended Water Clouds. I: Theory. *J. Atmos. Sci.,* **35**, 2111–2122.

Tibshirani, G., G. Walther, and T. Hastie, 2001: Estimating the number of clusters in a data set via the gap statistic. *J. Roy. Stat. Soc.*, **63**, 411–423, https://doi.org/10.1111/14679868.00293.

Tomkins, L. M., 2019: Regional flow conditions associated with stratocumulus cloud-clearing events over the southeast Atlantic, Master's thesis, *Kansas Univ.*, Lawrence, KS.

Toniazzo, T., S. J. Abel, R. Wood, C. R. Mechoso, G. Allen, and L. C. Shaffrey, 2011: Large-scale and synoptic meteorology in the south-east Pacific during the observations campaign VOCALS-Rex in austral Spring 2008. *Atmos. Chem. Phys.*, **11**, 4977–5009, doi:10.5194/acp-11-4977-2011.

Vogelezang, D. and A. Holtslag, 1996: Evaluation and model impacts of alternative boundary-layer height formulation. *Boundary-Layer Meteorol.*, **81**, 245–269.

Winant, C. D., C. E. Dorman, C. A. Friehe, and R. C. Beardsley, 1988: The marine layer off Northern California: an example of supercritical channel flow. *J Atmos Sci.* **45**, 3588–3605.

Wood, R., 2012: Stratocumulus Clouds. *Mon. Wea. Rev.,* **140**, 2373–2423.

Wood, R., M. Köhler, R. Bennartz, and C. O'Dell, 2009: The diurnal cycle of surface divergence over the global oceans. *Q.J.R. Meteorol. Soc.*, **135**, 1484-1493.

Wood, R. and C. S. Bretherton, 2006: On the Relationship between Stratiform Low Cloud Cover and Lower-Tropospheric Stability. *J. Climate,* **19**, 6425–6432.

Wood, R. and D. L. Hartmann, 2006: Spatial Variability of Liquid Water Path in Marine Low Cloud: The Importance of Mesoscale Cellular Convection. *J. Climate,* **19**, 1748–1764.

Wood, R., 2005a: Drizzle in stratiform boundary layer clouds. Part I: Vertical and horizontal structure. *J. Atmos. Sci.*, **62**, 3011–3033.

Wood, R. and C. S. Bretherton, 2004: Boundary layer depth, entrainment, and decoupling in the cloud-capped subtropical and tropical marine boundary layer. *J. Climate*, **17**, 3576–3588.

Xie, F., D. L. Wu, C. O. Ao, A. J. Mannucci, and E. R. Kursinski, 2012: Advances and limitations of atmospheric boundary layer observations with GPS occultation over southeast Pacific Ocean, *Atmos. Chem. Phys.*, **12**, 903-918.

Yuter, S. E., J. D. Hader, M. A. Miller, and D. B. Mechem, 2018: Abrupt cloud clearing of marine stratocumulus in the subtropical southeast Atlantic, *Science*, **361**(6403), 697-701.

Zhou, L., Q. Liu, D. Liu, L. Xie, L. Qi, and X. Liu, 2016: Validation of MODIS liquid water path for oceanic nonraining warm clouds: Implications on the vertical profile of cloud water content, *J. Geophys. Res. Atmos.*, **121**, 4855–4876.

Zuidema, P., P. Chang, B. Medeiros, B.P. Kirtman, R. Mechoso, E.K. Schneider, T. Toniazzo, I. Richter, R.J. Small, K. Bellomo, P. Brandt, S. de Szoeke, J.T. Farrar, E. Jung, S. Kato, M. Li, C. Patricola, Z. Wang, R. Wood, and Z. Xu, 2016: Challenges and Prospects for Reducing Coupled Climate Model SST Biases in the Eastern Tropical Atlantic and Pacific Oceans: The U.S. CLIVAR Eastern Tropical Oceans Synthesis Working Group. *Bull. Amer. Meteor. Soc.,* **97**, 2305–2328, https://doi.org/10.1175/BAMS-D-15-00274.1

Zuidema, P., E. R. Westwater, C. Fairall, and D. Hazen, 2005: Ship-based liquid water path estimates in marine stratocumulus. *J. Geophys. Res.*, **110**, D20206.

TABLES

Table 1: SOMs that were run for the southeast Atlantic. Each column is a period the SOM was trained on, with cells indicating what model time(s) of day were used in Greenwich mean time (Z). A SOM is run for each individual time of day. Asterisks (*) indicate that SOMs were also run for just the days on which cloud-eroding boundaries did or did not occur (Yes and No days, see Section 2.1).

	Southeast Atlantic				
Training Var	Jan-Feb	May	Apr-May-Jun	Aug-Sep	Oct-Nov-Dec
500mb Z	10Z, 22Z	10Z, 22Z	10Z, 22Z*	10Z, 22Z	10Z, 22Z
850mb Z	10Z, 22Z	10Z, 22Z	10Z, 22Z*	10Z, 22Z	10Z, 22Z
900mb Z	10Z, 22Z	10Z, 22Z	10Z, 22Z		
950mb Z	10Z, 22Z	10Z, 22Z	10Z, 22Z		
1000mb Z	10Z, 22Z	10Z, 22Z	10Z, 22Z		
MSLP	10Z, 22Z	10Z, 22Z*	10Z, 22Z*	10Z, 22Z	10Z, 22Z
BLH	10Z, 22Z	10Z, 22Z*	10Z, 22Z*	10Z, 22Z	10Z, 22Z
SST	22Z	22Z	22Z	22Z	22Z
LTS	22Z	22Z	22Z	22Z	22Z
EIS	10Z, 22Z	10Z, 22Z*	10Z, 22Z*	10Z, 22Z	10Z, 22Z
700mb RH	22Z	22Z	22Z	22Z	22Z
750mb RH	22Z	22Z	22Z	22Z	22Z
800mb RH	22Z	22Z	22Z	22Z	22Z
900mb U	10Z, 22Z	10Z, 22Z	10Z, 22Z		
900mb V	10Z, 22Z	10Z, 22Z	10Z, 22Z		
925mb V	22Z	22Z	22Z		
950mb U	10Z, 22Z	10Z, 22Z	10Z, 22Z		
950mb V	10Z, 22Z	10Z, 22Z	10Z, 22Z		
975mb V	22Z	22Z	22Z		
1000mb U	10Z, 22Z	10Z, 22Z	10Z, 22Z		
1000mb V	10Z, 22Z	10Z, 22Z	10Z, 22Z		

Table 2: SOMs that were run for the southeast Pacific. Each column is a period the SOM was trained on, with cells indicating what model time(s) of day were used in Greenwich mean time (Z). A SOM is run for each individual time of day. Asterisks (*) indicate that SOMs were also run for just the days on which cloud-eroding boundaries did or did not occur (Yes and No days, see Section 2.1).

	Southeast Pacific				
Training Var	Jan-Feb	May	Apr-May-Jun	Aug-Sep	Oct-Nov-Dec
500mb Z	16Z, 04Z	16Z, 04Z	16Z, 04Z	16Z, 04Z	16Z, 04Z
850mb Z	16Z, 04Z	16Z, 04Z	16Z, 04Z	16Z, 04Z	16Z, 04Z
MSLP	16Z, 04Z	16Z, 04Z	16Z, 04Z	16Z, 04Z	16Z, 04Z
BLH	16Z, 04Z	16Z, 04Z	16Z, 04Z	16Z, 04Z	16Z, 04Z
SST	04Z	04Z	04Z	04Z	04Z
LTS	04Z	04Z	04Z	04Z	04Z
EIS	16Z, 04Z	16Z, 04Z	16Z, 04Z	16Z, 04Z	16Z, 04Z
700mb RH	04Z	04Z	04Z	04Z	04Z
750mb RH	04Z	04Z	04Z	04Z	04Z
800mb RH	04Z	04Z	04Z	04Z	04Z

Table 3: Peak liquid water content (LWC) required at cloud top to maintain saturation after mixing 0.05 m^3 of dry air (0 g/kg) with 1 m^3 of moist air at cloud top using Equation 3.3. Moist air at cloud top has mixing ratio equal to the saturation mixing ratio. The amount of liquid water at cloud top is the peak LWC, since the liquid water present at any level in the cloud (assuming adiabatic water-mixing ratio) is the saturation mixing ratio at that level minus the saturation mixing ratio at cloud base. See Section 3.1.6 for more details.

Cloud top temp. (K)	**Cloud top pressure (hPa)**	**Cloud top mixing ratio (g/kg)**	**Mixing ratio after entrainment (g/kg)**	**Peak LWC required to maintain saturation**
280	950	6.5578	6.2455	0.3279
280	925	6.7369	6.4161	0.3368
281.5	925	7.4718	7.1160	0.3736
283	925	8.2777	7.8835	0.4139
284.5	925	9.1607	8.7245	0.4580
286	925	10.1274	9.6451	0.5064
286	900	10.4134	9.9175	0.5207

FIGURES

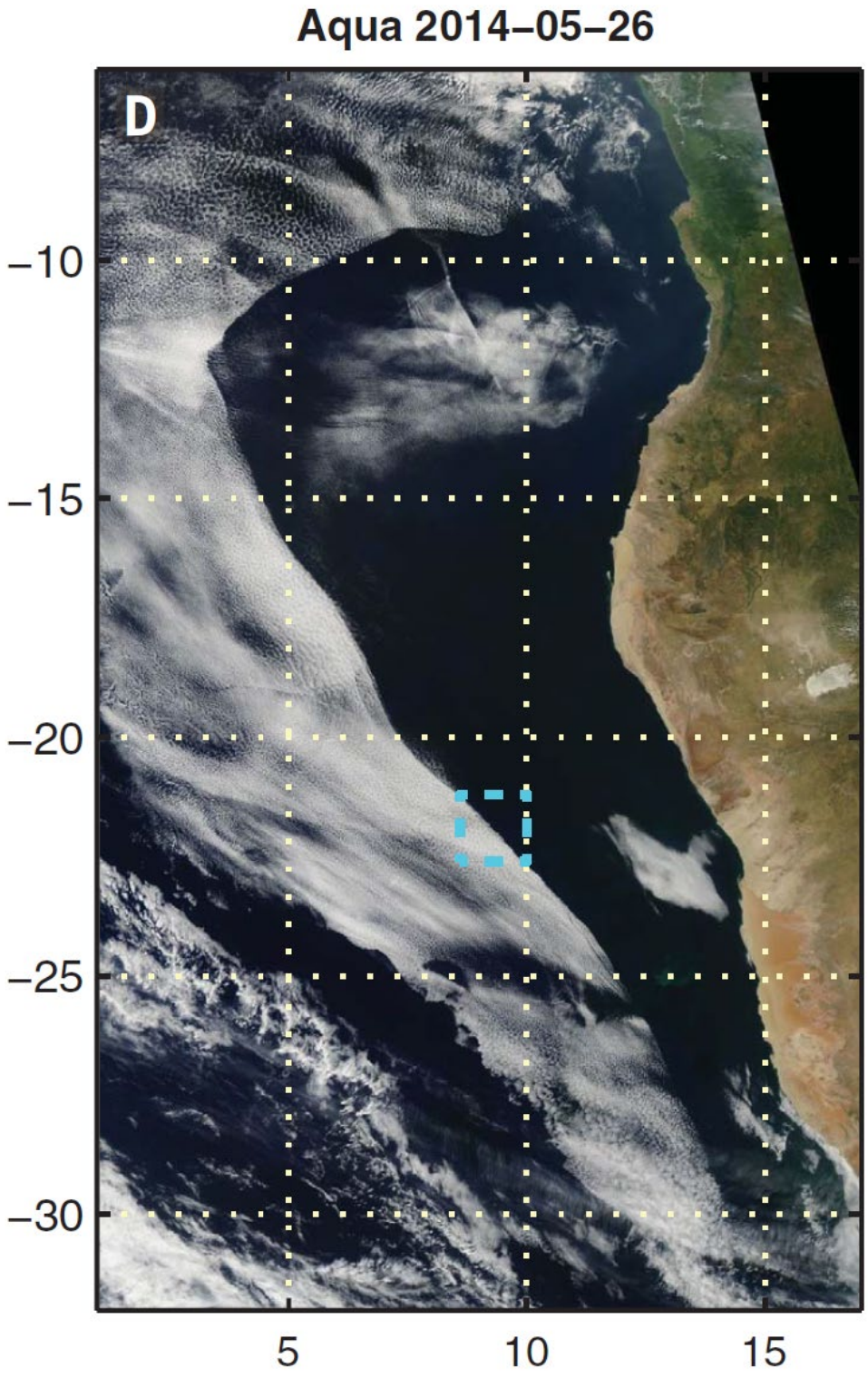


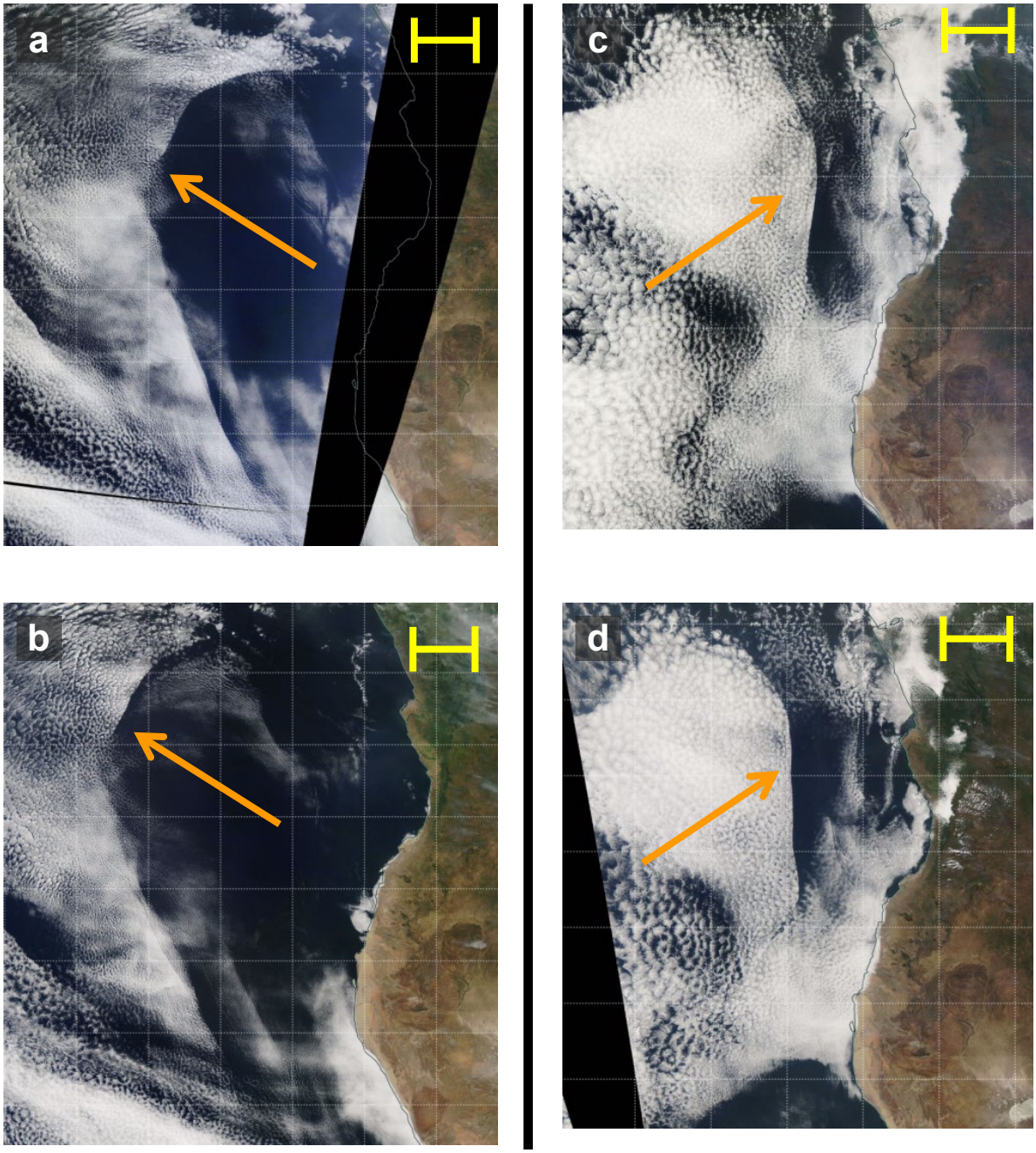
Figure 1.1: Taken from Figure 1 of Yuter et al. (2018). Example of a large westward-moving cloud-eroding boundary off the coast of Africa in MODIS corrected reflectance imagery from Aqua (1:30 pm LT). Latitude and longitudeFigure 0.1 lines shown by dotted lines.

Figure 2.1: Examples of cloud-scene pair classifications in the cloud-eroding boundary dataset (Section 2.1). Panels a, c, e, and g are corrected reflectance images from MODIS Terra (10:30 am LT) and panels b, d, f, and h are from MODIS Aqua (1:30 pm LT). Panels a–d show examples of cloud scenes categorized as having a sharp cloud boundary (Yes day) with orange arrows pointing to the boundaries and their westward propagation. Panels e–h show examples of boundaries (indicated by orange circles) which do not completely meet the criteria for a cloud-eroding boundary (see Section 2.1). The yellow (or black in e, f) scale bars in each panel represent 200 km. Note this is only an illustration, and that the manual nature of this categorization does employ some level of subjectivity.

Examples of Yes days

May 25, 2013

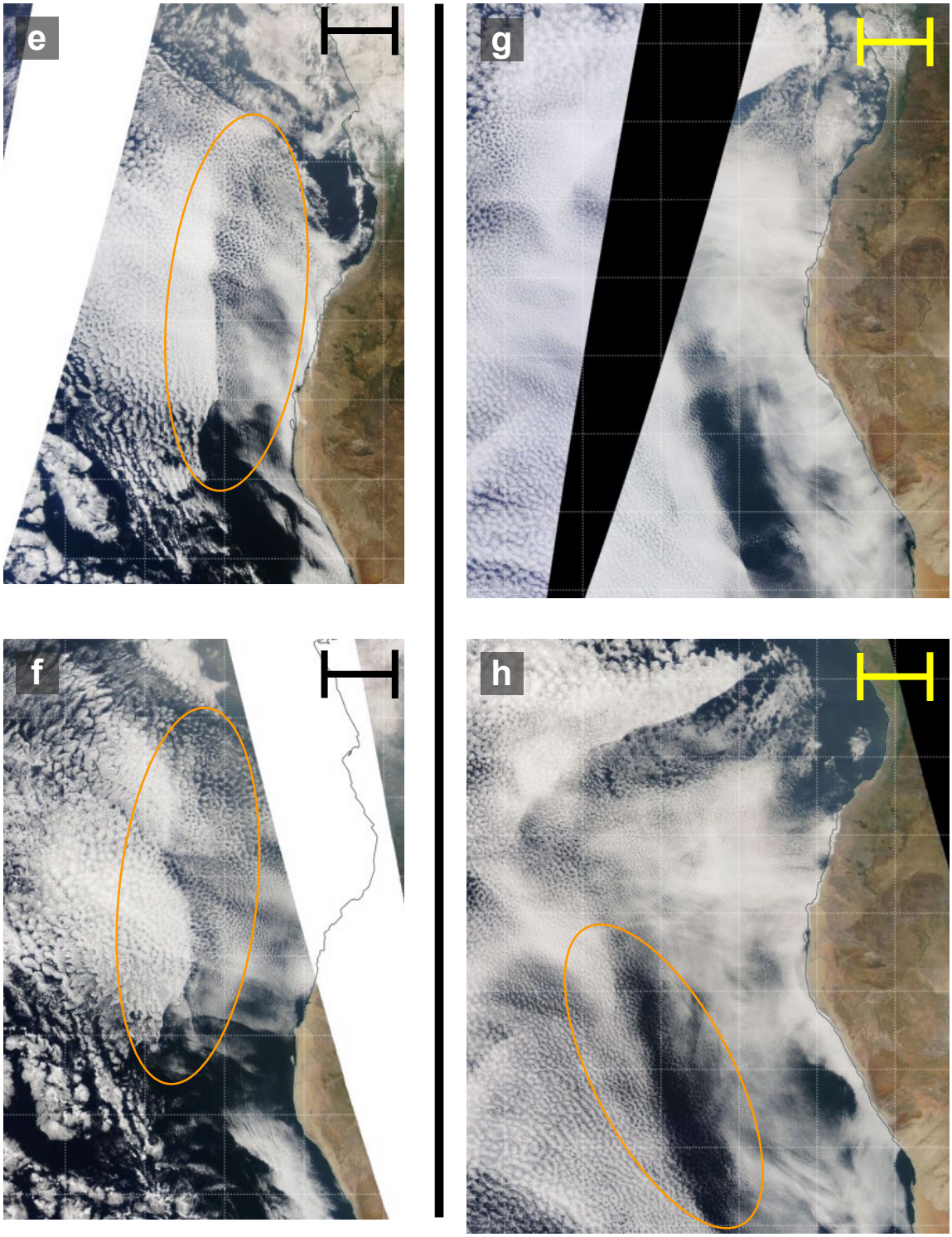
June 13, 2015



Examples of Maybe days

May 25, 2012

June 21, 2015



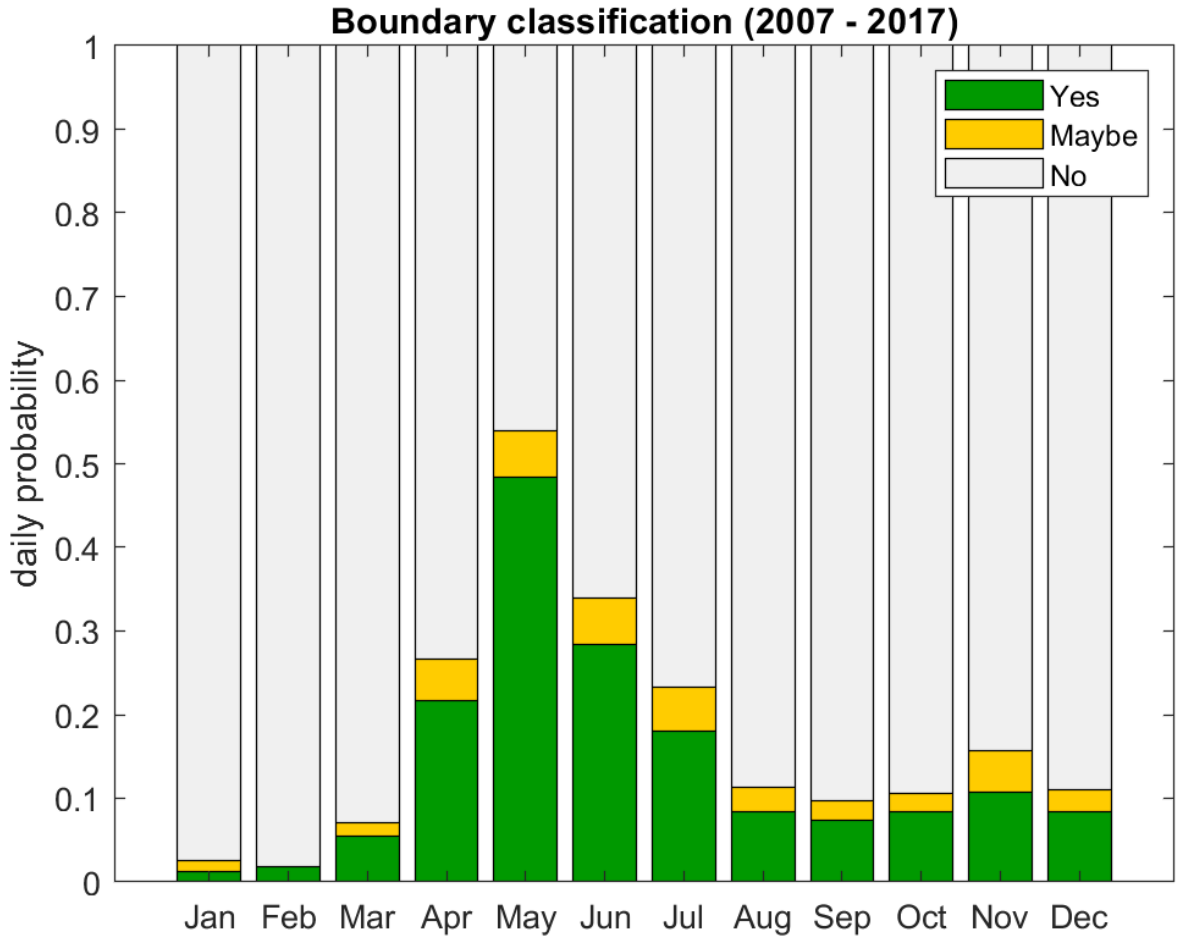


Figure 2.2: Probability of a cloud-eroding boundary occurring on a given day in the southeast Atlantic (10° to 20°S, 0° to 12°E) by month. Probabilities based on manual analysis of MODIS Terra and Aqua satellite image pairs from 2007 to 2017. *Maybe* days are ones when there may appear to be a cloud-eroding boundary but it does not meet all criteria outlined in Section 2.1.

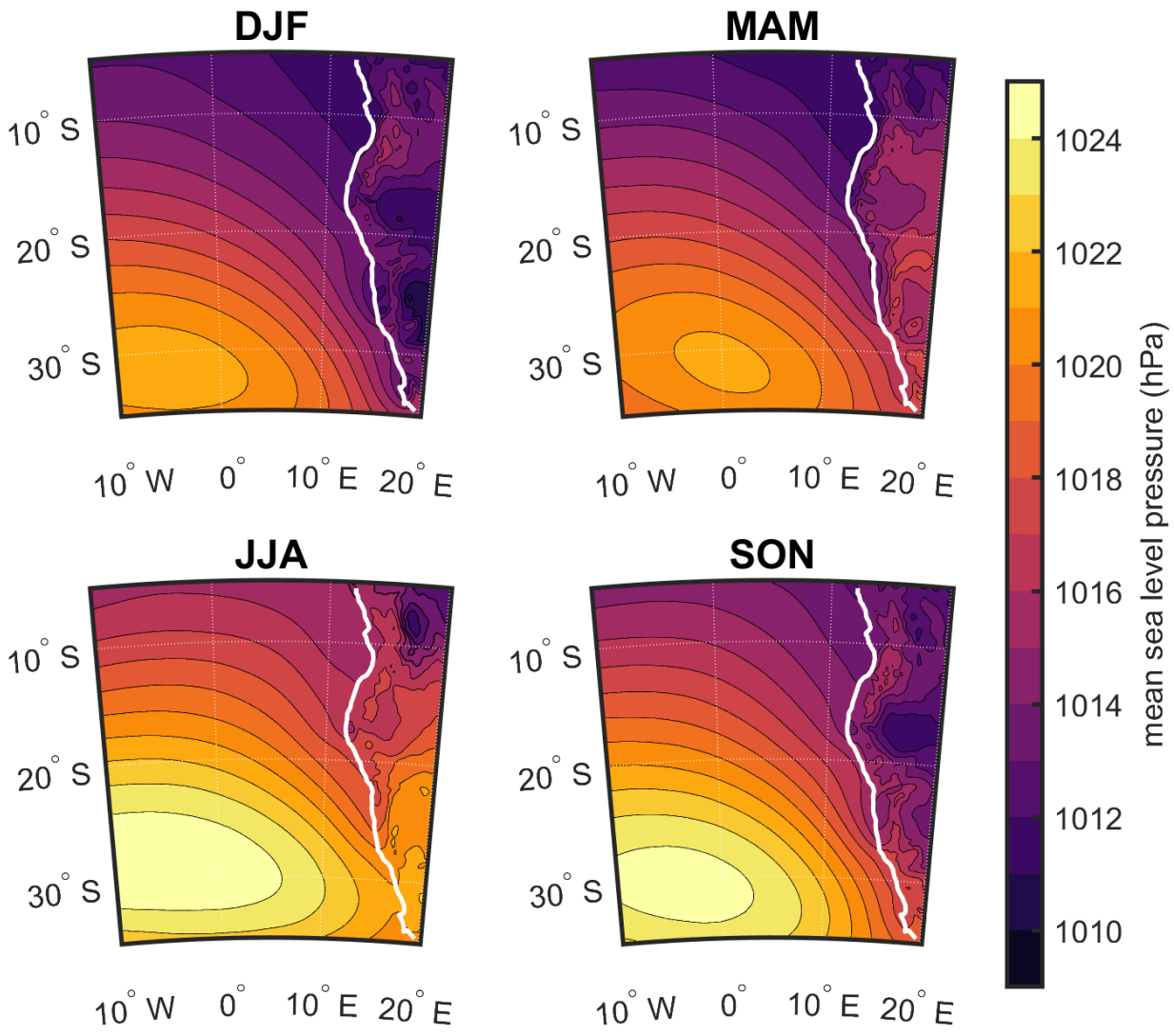


Figure 2.3: Mean sea level pressure at 22 UTC over the southeast Atlantic by season, contoured every 1 hPa. Title over each panel indicates months included in average: **DJF** = Dec-Jan-Feb, **MAM** = Mar-Apr-May, **JJA** = Jun-Jul-Aug, **SON** = Sep-Oct-Nov.

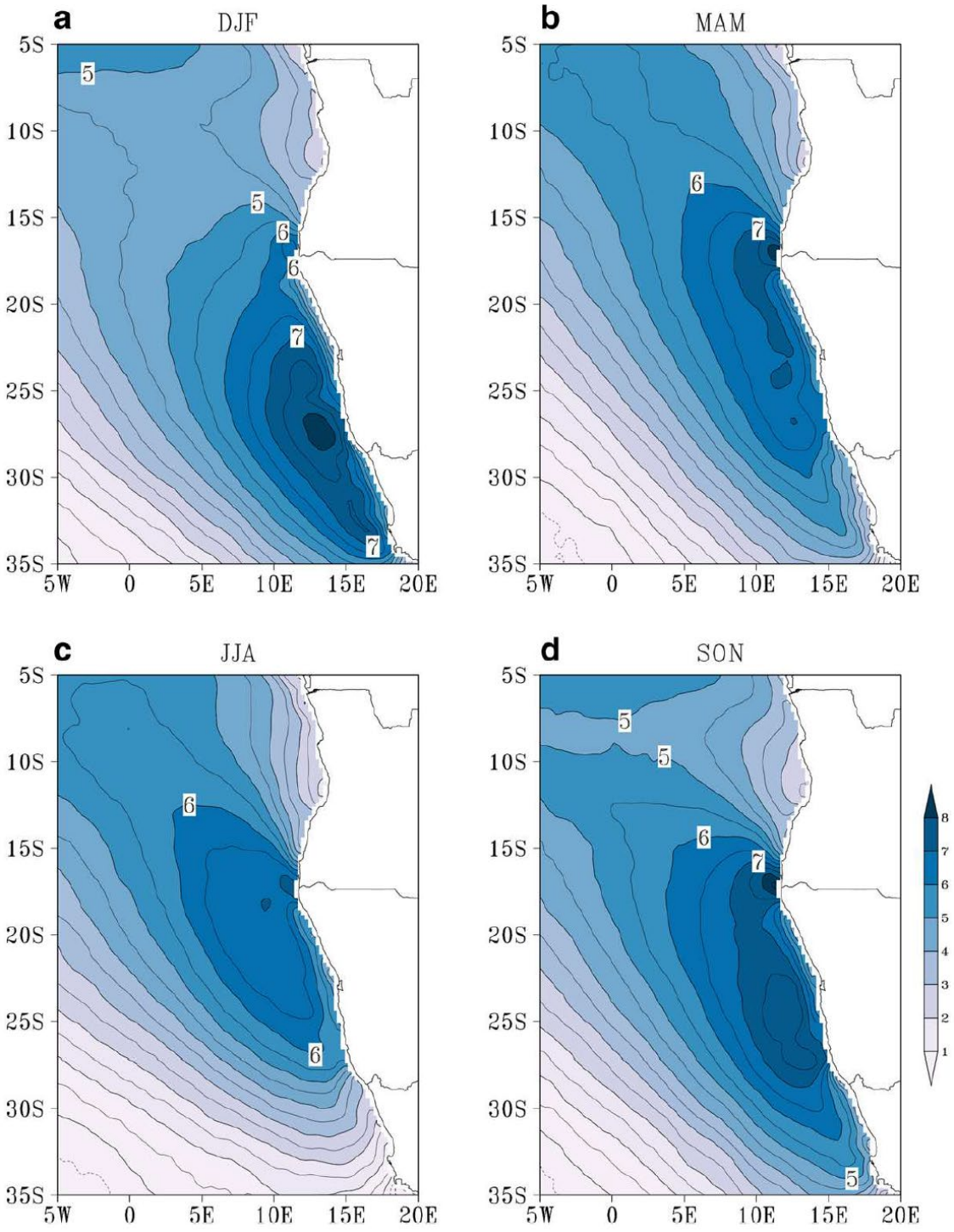


Figure 2.4: Taken from Patricola and Chang (2017, their Figure 2). 10-meter meridional wind (m/s) from the SCOW climatology averaged for **a** Dec-Jan-Feb, **b** Mar-Apr-May, **c** Jun-Jul-Aug, and **d** Sep-Oct-Nov. Areas with no data are indicated in white. Contours of 5+ m/s are labeled.

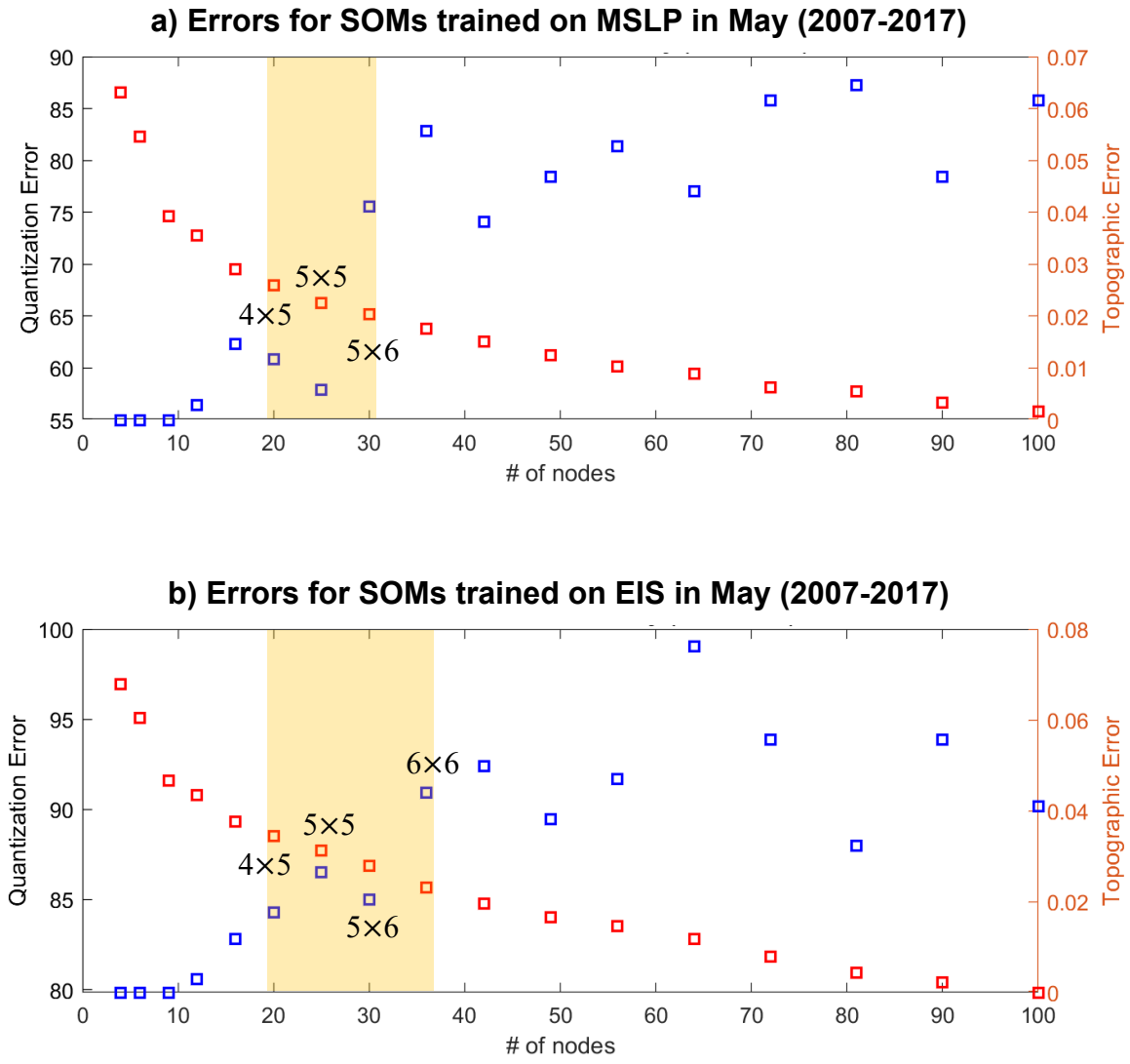
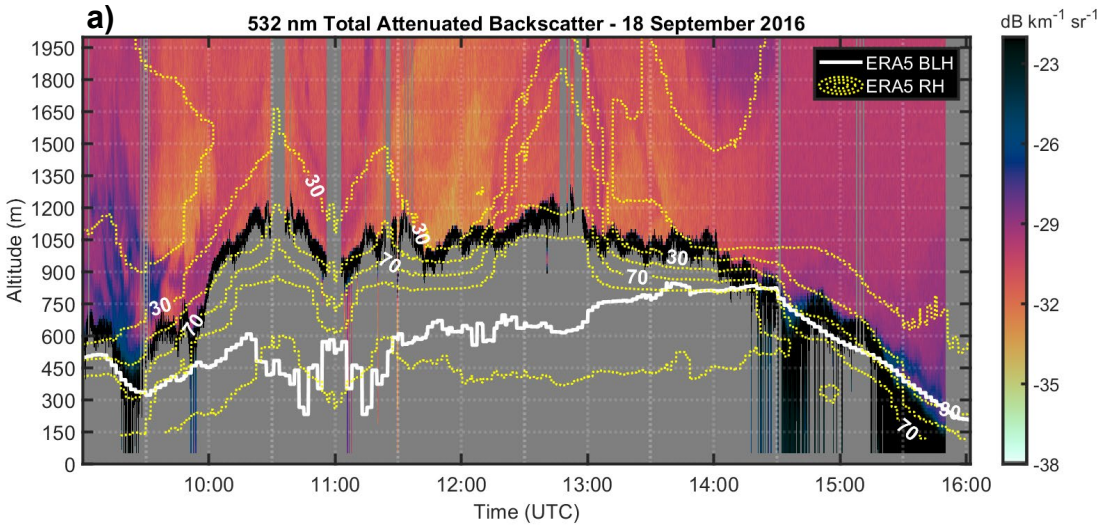
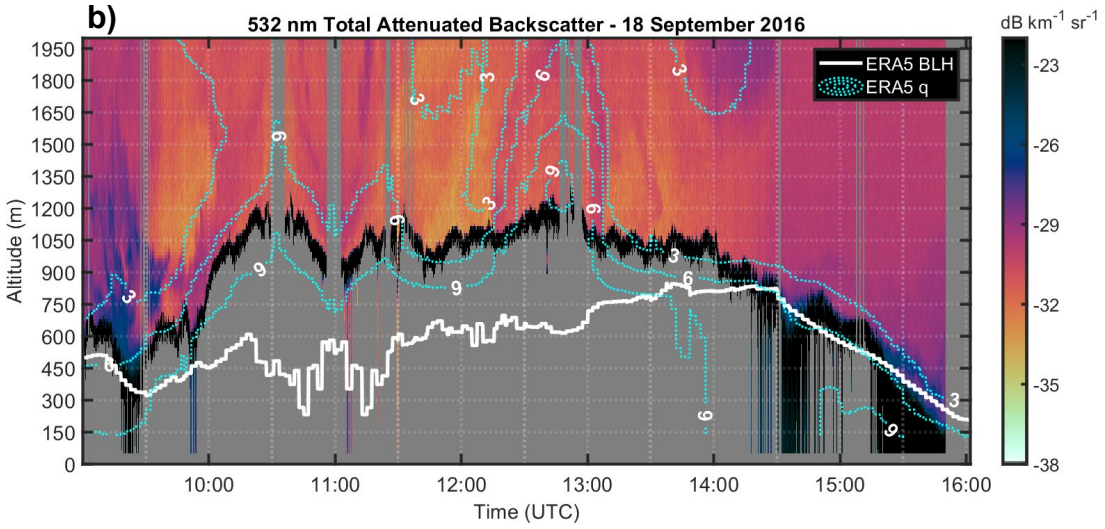
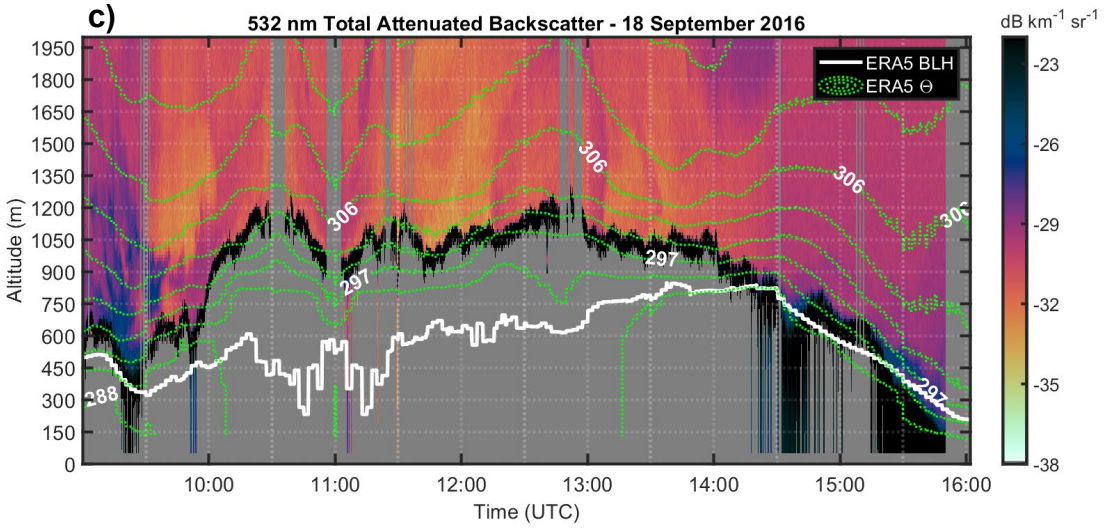


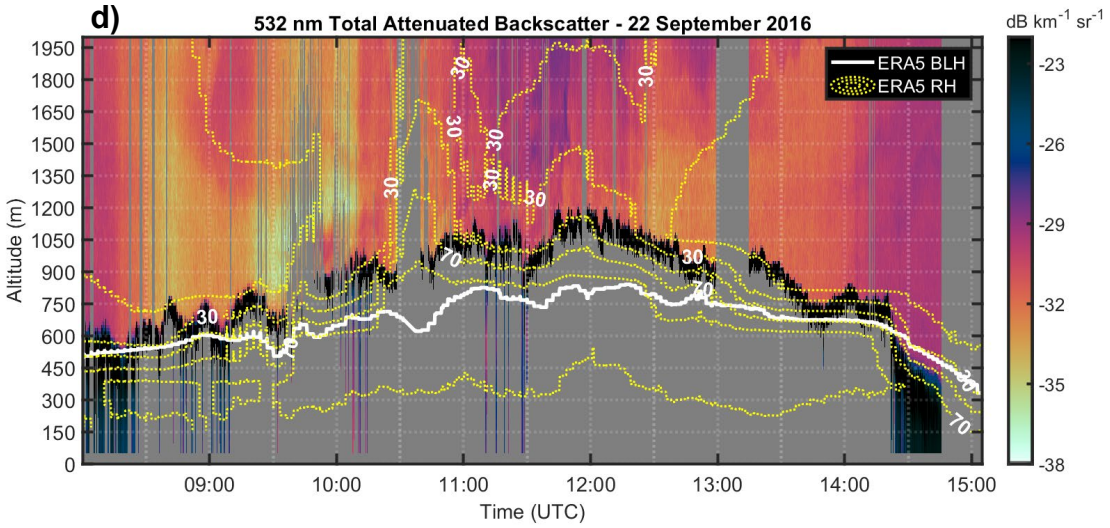
Figure 2.5: Elbow plots used to determine the optimal number of SOM nodes. All SOMs were trained on data over the southeast Atlantic (5° to 30°S, -10° to 20°E) in May at 22Z from 2007 to 2017, where (**a**) shows the error for SOMs trained on mean sea level pressure and (**b**) on estimated inversion strength. Map-sizes used were 22, 23, 33, …, up to 1010. The optimal number of nodes chosen is 25, highlighted by the region shaded in yellow in both plots.

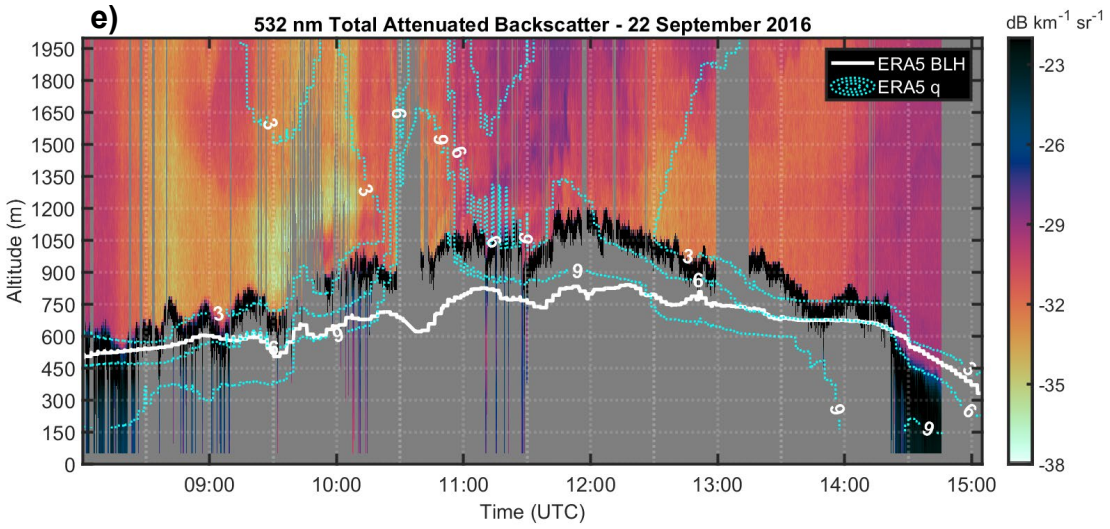
Figure 2.6: Total 532 nm attenuated backscatter from lidar aboard NASA ER-2 aircraft during the ORACLES field campaign. ER-2 flight tracks are shown in panels g and h (green) with gray dots every 5 minutes, blue circles around each hour, and X's indicating the (blue) start and (red) end of the flight. Coincident hourly ERA5 data from the closest time and grid box are overlaid with contours of (a,d) relative humidity every 20%, (b,e) specific humidity every 3 g/kg, and (c,f) potential temperature every 3 K. ERA5 data come from 25-hPa pressure levels in the lowest 300 hPa of the troposphere.

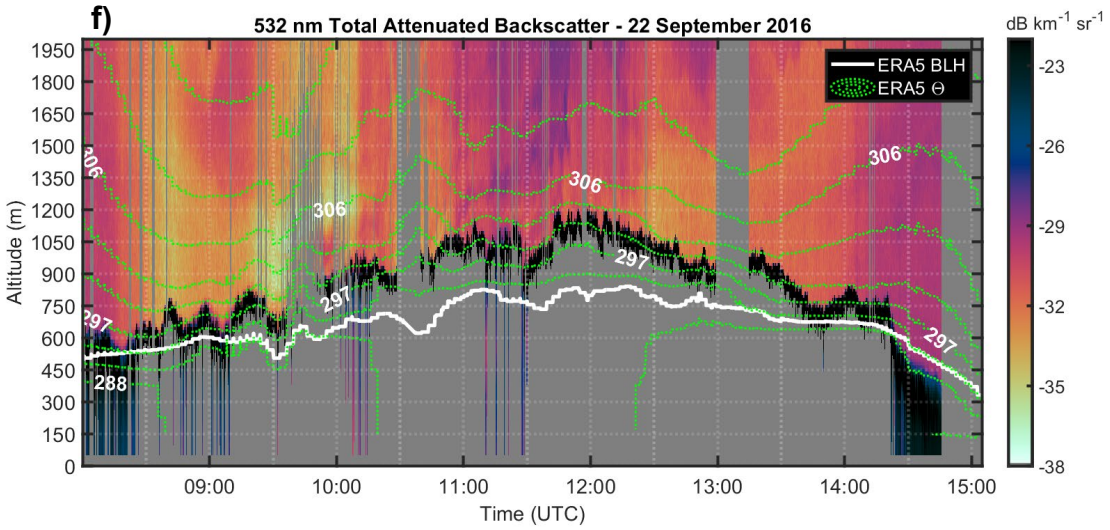

a)
532 nm Total Attenuated Backscatter - 18 September 2016
dB km-1 sr-1
ERA5 BLH
ERA5 RH
Altitude (m)
Time (UTC)


b)
532 nm Total Attenuated Backscatter - 18 September 2016
dB km-1 sr-1
ERA5 BLH
ERA5 q
Altitude (m)
Time (UTC)

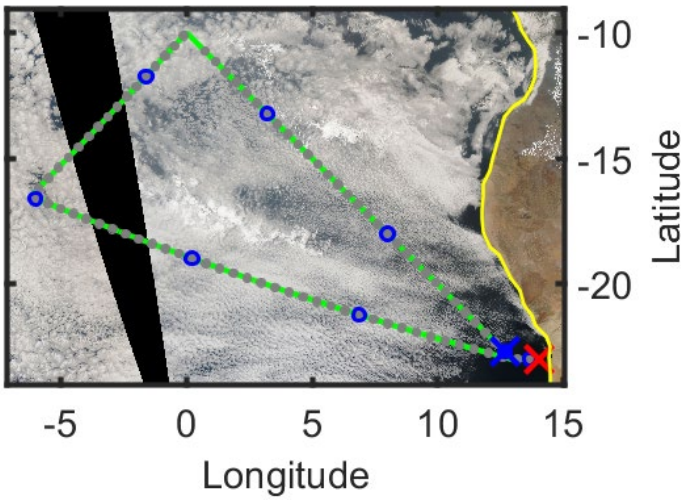

c)
532 nm Total Attenuated Backscatter - 18 September 2016
dB km-1 sr-1
ERA5 BLH
ERA5 Θ
Altitude (m)
Time (UTC)


d)
532 nm Total Attenuated Backscatter - 22 September 2016
ERA5 BLH
ERA5 RH
Altitude (m)
Time (UTC)
dB km-1 sr-1

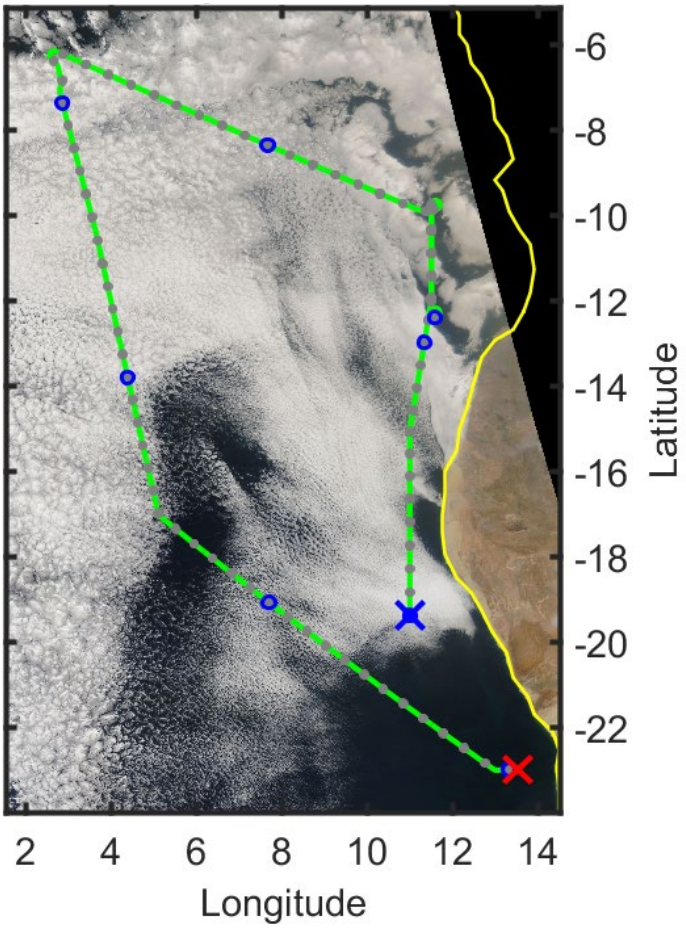

e)
532 nm Total Attenuated Backscatter - 22 September 2016
ERA5 BLH
ERA5 q
Altitude (m)
Time (UTC)
dB km-1 sr-1


f)
532 nm Total Attenuated Backscatter - 22 September 2016
ERA5 BLH
ERA5 Θ
Altitude (m)
Time (UTC)
dB km-1 sr-1

g) MODIS Aqua (18 Sep 2016)



h) MODIS Aqua (22 Sep 2016)



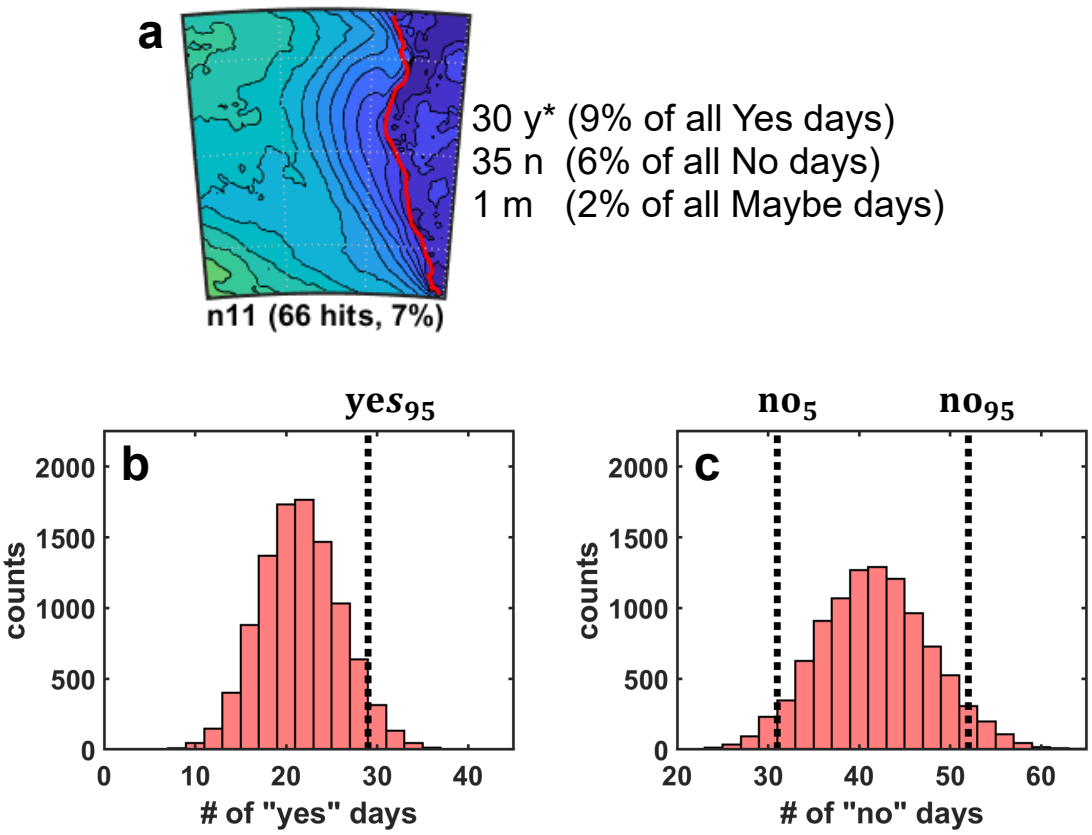


Figure 2.7: Example of how matching Yes, No and Maybe days are evaluated for randomness. In panel **a**, node 11 from the SOM trained on normalized anomalies of boundary layer height (Figure 3.15) is shown with number of matching Yes, No and Maybe days annotated to the right (along with the percentage of matching days for comparison to Figure 3.15). Histograms depict the distribution of how many (**b**) Yes and (**c**) No days were randomly assigned to the SOM node, with the bold dotted lines indicating 5th and 95th percentiles (5th percentile not shown in **b**). The matching Yes days (30) is unlikely to be random because it exceeds the 95th percentile in **b**. The matching No days (35) may just be by chance since it falls between the 5th and 95th percentiles in **c**.

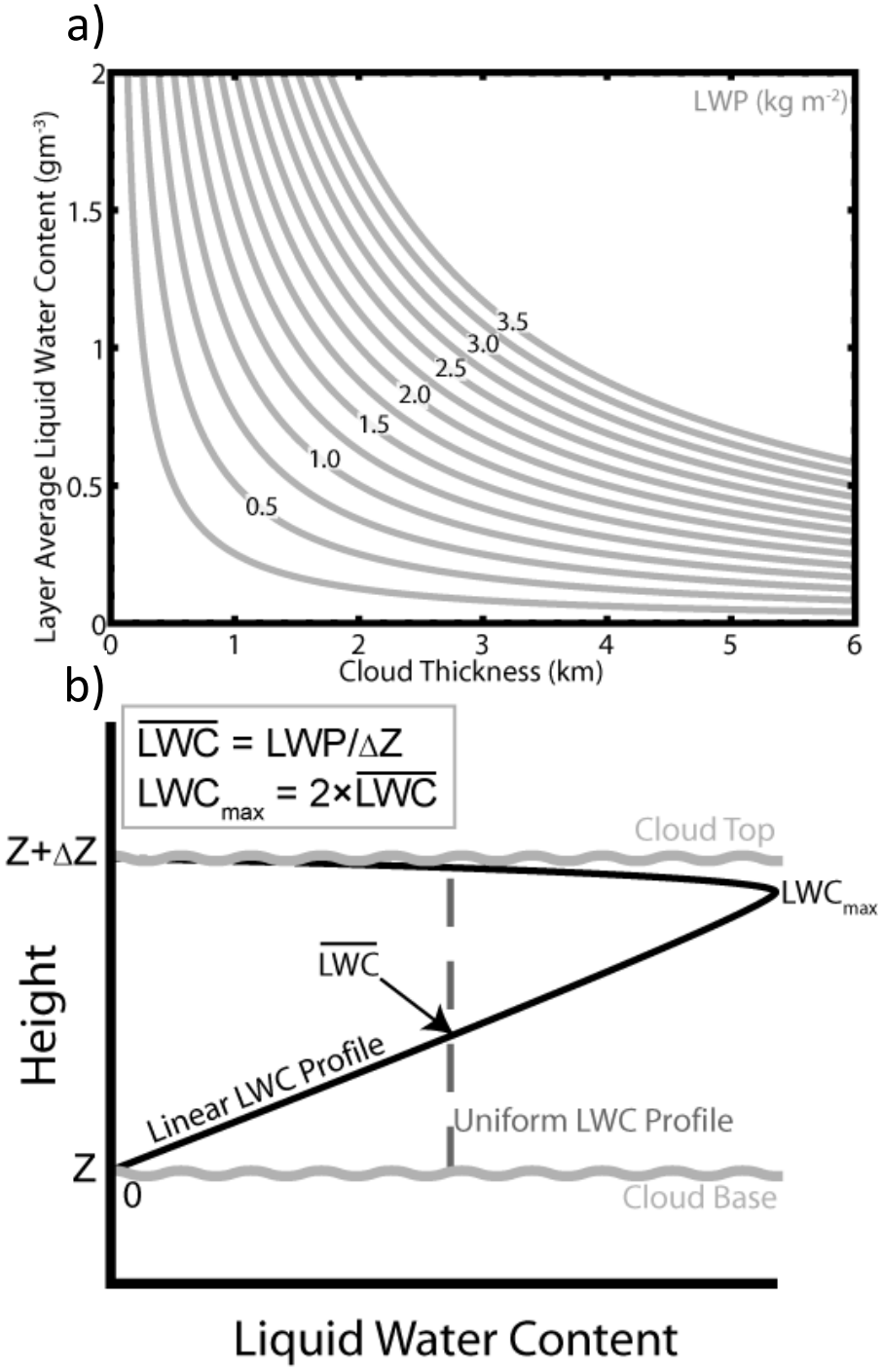


Figure 2.8: Taken from Figure 2.12 of Miller (2011). Panel **a** shows the dependence of liquid water path (LWP) in gray contours on the mean cloud liquid water content (LWC) and cloud thickness. Panel **b** shows the idealized LWC profile in marine stratocumulus.

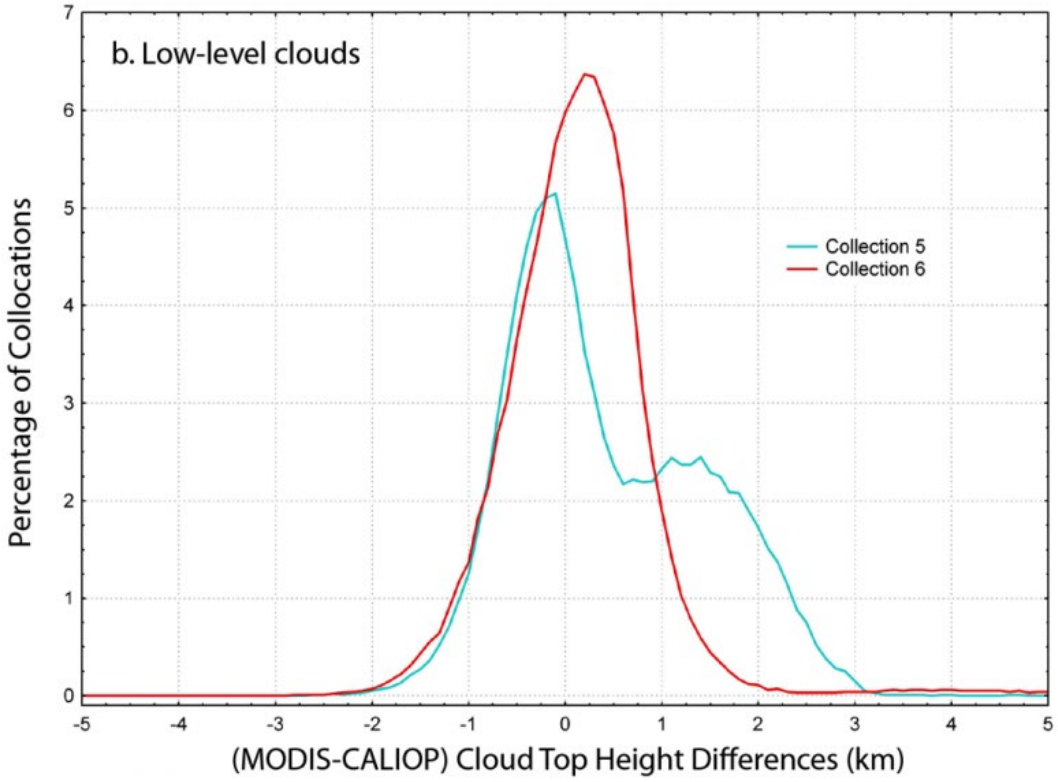


Figure 2.9: Taken from Figure 12b of Baum et al. (2012). Low-level clouds were determined with the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) satellite platform. 259,209 total collocations between CALIOP and MODIS were found for August, 2006. The distributions of cloud top height difference between CALIOP and MODIS Collection 5 (red) and Collection 6 (blue) data are shown. Percentages are calculated for 0.1-km cloud top height difference bins.

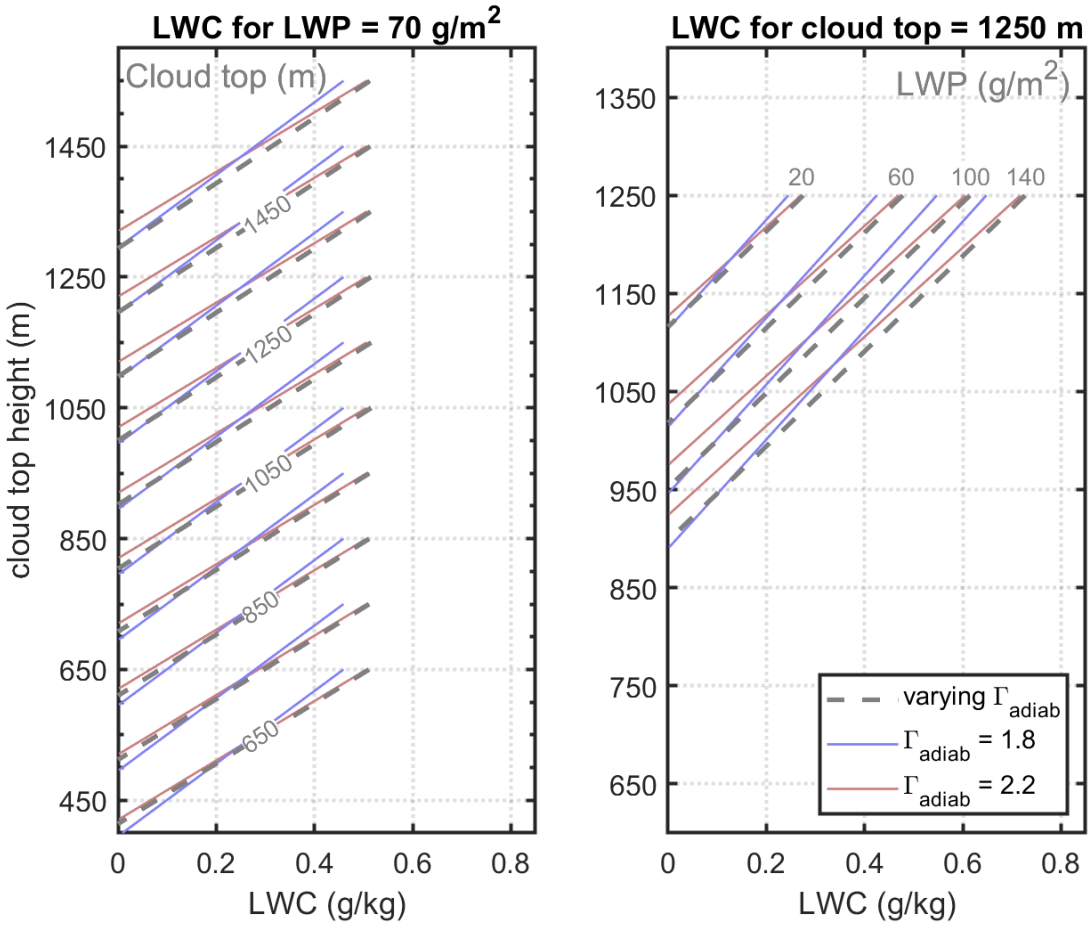


Figure 2.10: Idealized profiles of liquid water content in marine low clouds with varying adiabatic LWC change with height (dashed gray) and constant 1.8 g/kg per km (blue) and 2.2 g/kg per km (red). Panel **a** varies cloud top height while holding liquid water path (LWP) constant at 70 g/m^2, and panel **b** varies LWP while holding cloud top height constant at 1.25 km. Temperature at cloud top is derived using a lapse rate of 7 K/km and 295 K for the surface temperature. Pressure at cloud top is derived according to hydrostatic balance. See Section 2.6 for details.

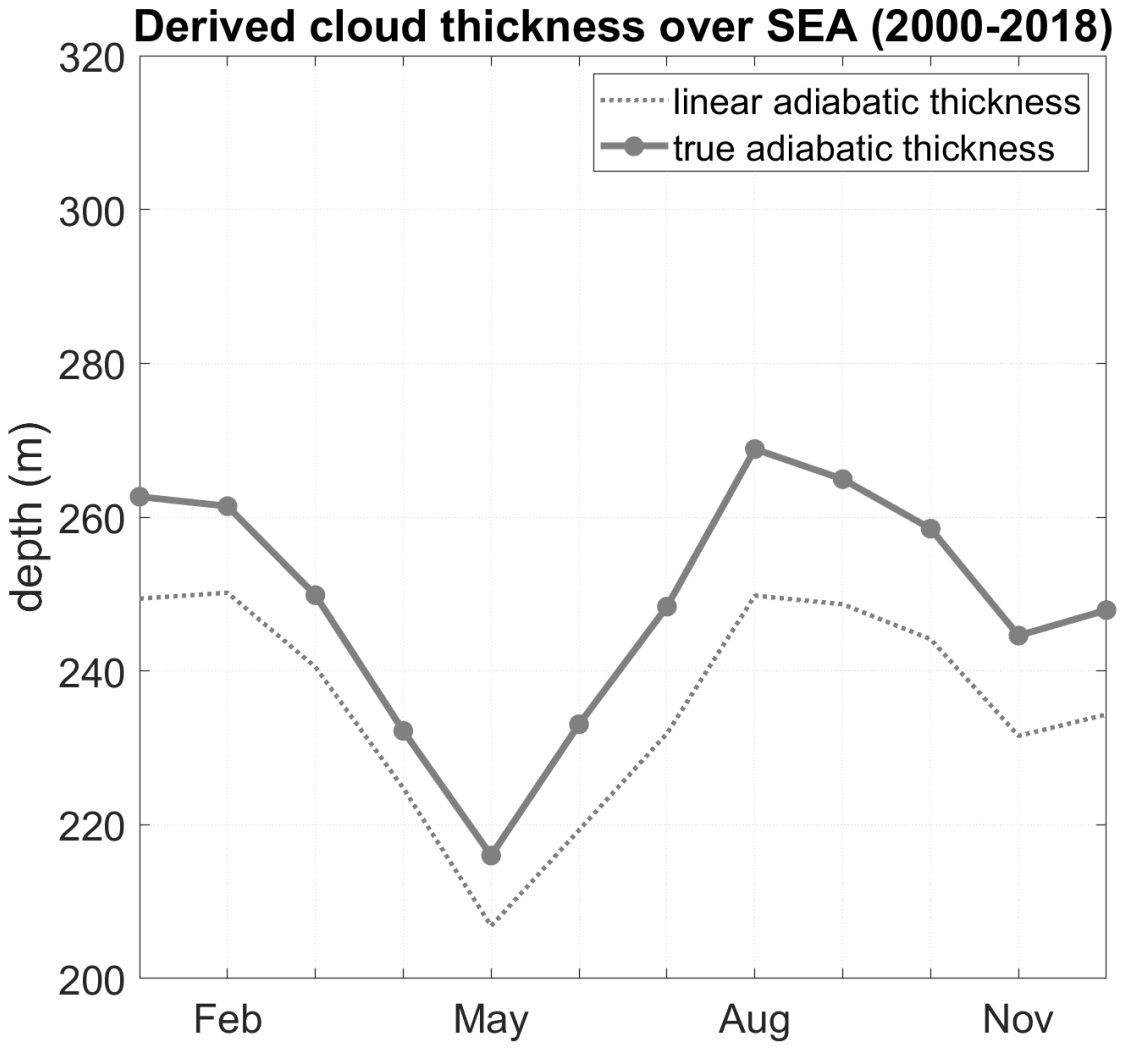


Figure 2.11: MODIS-derived cloud thickness from Terra (10:30 am LT) using constant adiabatic change of liquid water content with height in cloud (dashed) and varying adiabatic change of liquid water content based on the derived temperature profile in cloud (bold line). The constant used for the linear adiabatic treatment was 2.4 g/m^3 per km.

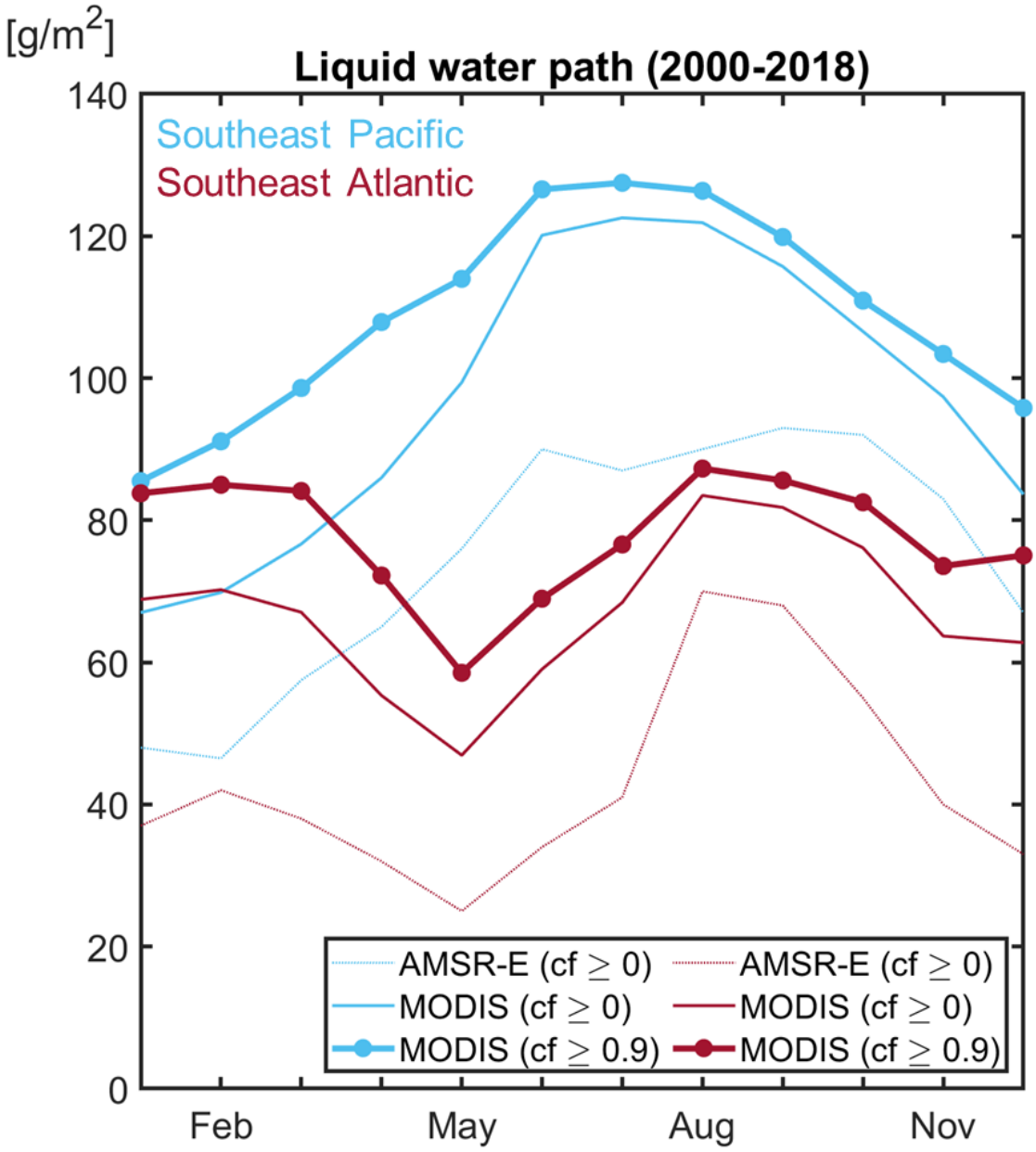
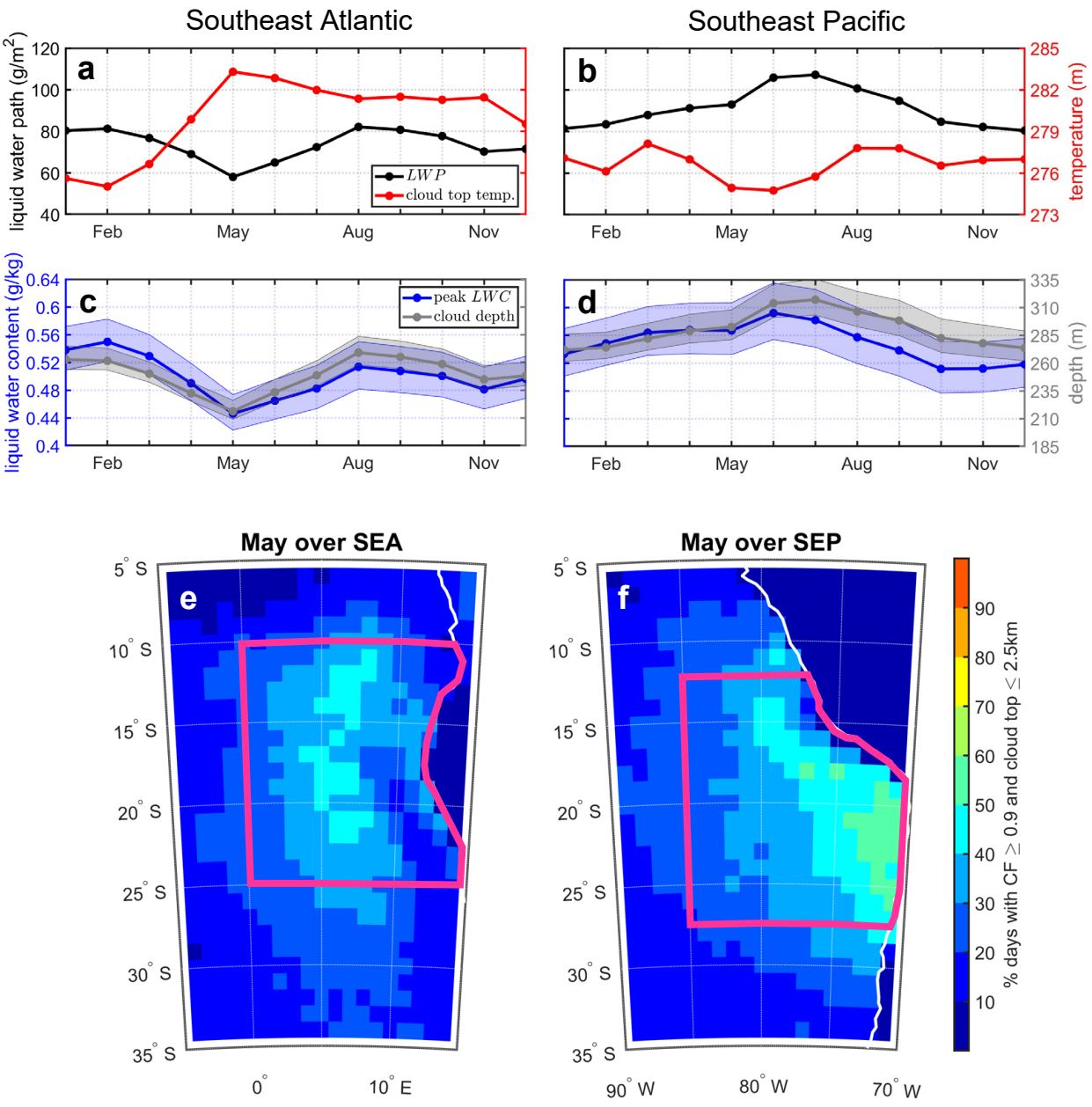


Figure 3.1: Seasonal daily mean liquid water path (g/m^2) for the southeast Atlantic (maroon) and southeast Pacific (light blue) using the same domains as Zuidema et al. (2016). Both domains span 10° to 20°S, with the southeast Pacific going from 90° to 80°W and the southeast Atlantic encompassing 0° to 10°E. Dotted lines depict non-conditional mean liquid water path from Zuidema et al. (2016, their Figure 4d), based on 2002–2011 Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) retrievals from the Aqua satellite 1:30 pm overpass. Solid lines are conditional (bold, cloud fraction ≥ 0.9) and non-conditional (thin) mean liquid water path based on 2000–2018 MODIS retrievals from the Terra (10:30 am LT) satellite.

Figure 3.2: Comparison of spatially averaged monthly time series based on 1° × 1° MODIS Terra (10:30 am LT) data with cloud fraction at least 90% and cloud top at or below 2.5 km for the southeast Atlantic and southeast Pacific for (**a**, **b**) MODIS liquid water path (black) and cloud top temperature (red), and (**c**, **d**) estimated cloud thickness (gray) and peak liquid water content (blue). The sensitivity of the cloud depth and and peak LWC calculations to changes in cloud top temperature +/- 5K is indicated by the upper and lower bounds of the shaded regions in **c** and **d**. The geographic regions for **e**) the southeast Atlantic and **f**) the southeast Pacific used for averaging are outlined with a thick pink line and overlaid on top of the percent of 588 days in May from 2000 to 2018 meeting the criteria of cloud fraction ≥ 90% and cloud top at or below 2.5 km. Coastlines are shown by a thin white line.


Southeast Atlantic
Southeast Pacific
a
b
c
d
e
f
liquid water path (g/m^2)
temperature (m)
liquid water content (g/kg)
depth (m)
LWP
cloud top temp.
peak LWC
cloud depth
Feb
May
Aug
Nov
May over SEA
May over SEP
% days with CF ≥ 0.9 and cloud top ≤ 2.5km

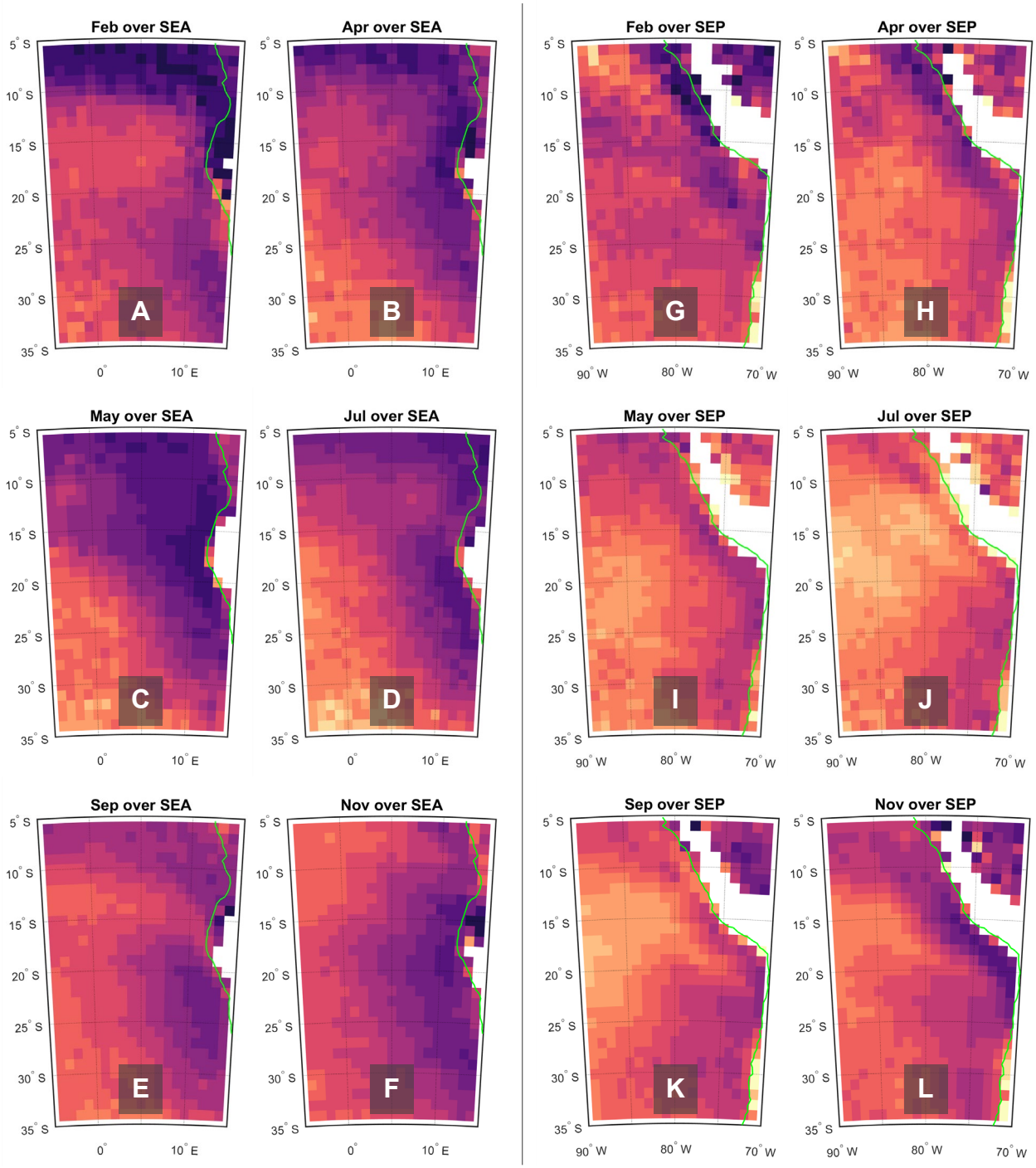


Figure 3.3: Conditionally averaged MODIS Terra (10:30 am LT) daily liquid water path (shaded) for selected months (Feb, Apr, May, Jul, Sep, Nov), where cloud fraction is at least 0.9 at 1 1 pixel scale and cloud top height is at or below 2.5 km, from 2000 to 2018. Plots **A**–**F** are over the southeast Atlantic and **G**–**L** are over the southeast Atlantic. Coastlines are shown in each plot with a thin line.

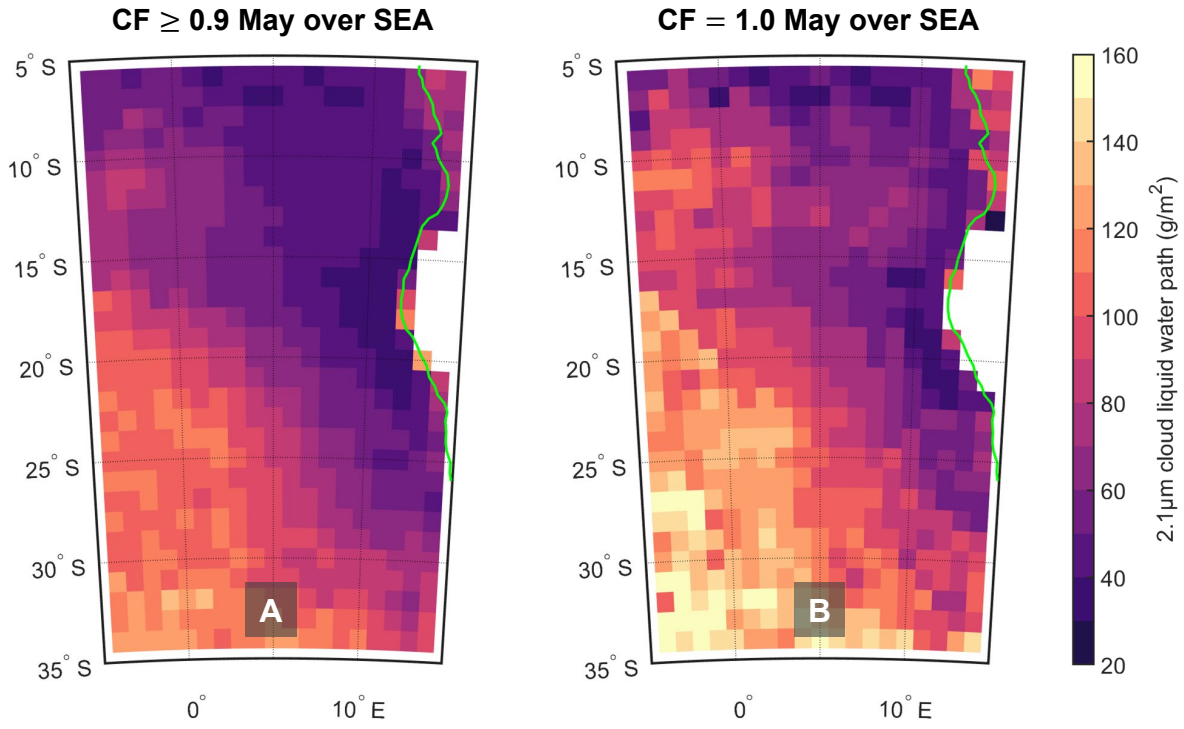


Figure 3.4: Conditionally averaged MODIS Terra (10:30 am LT) daily liquid water path for cloud top ≤ 2.5 km altitude with cloud fraction of 1.0 (**A**) and cloud fraction of at least 0.9 (**B**). Coastline shown by thin red line. See Appendix F for plots of other months of liquid water path conditional upon 0.8 and 1.0 cloud fraction thresholds.

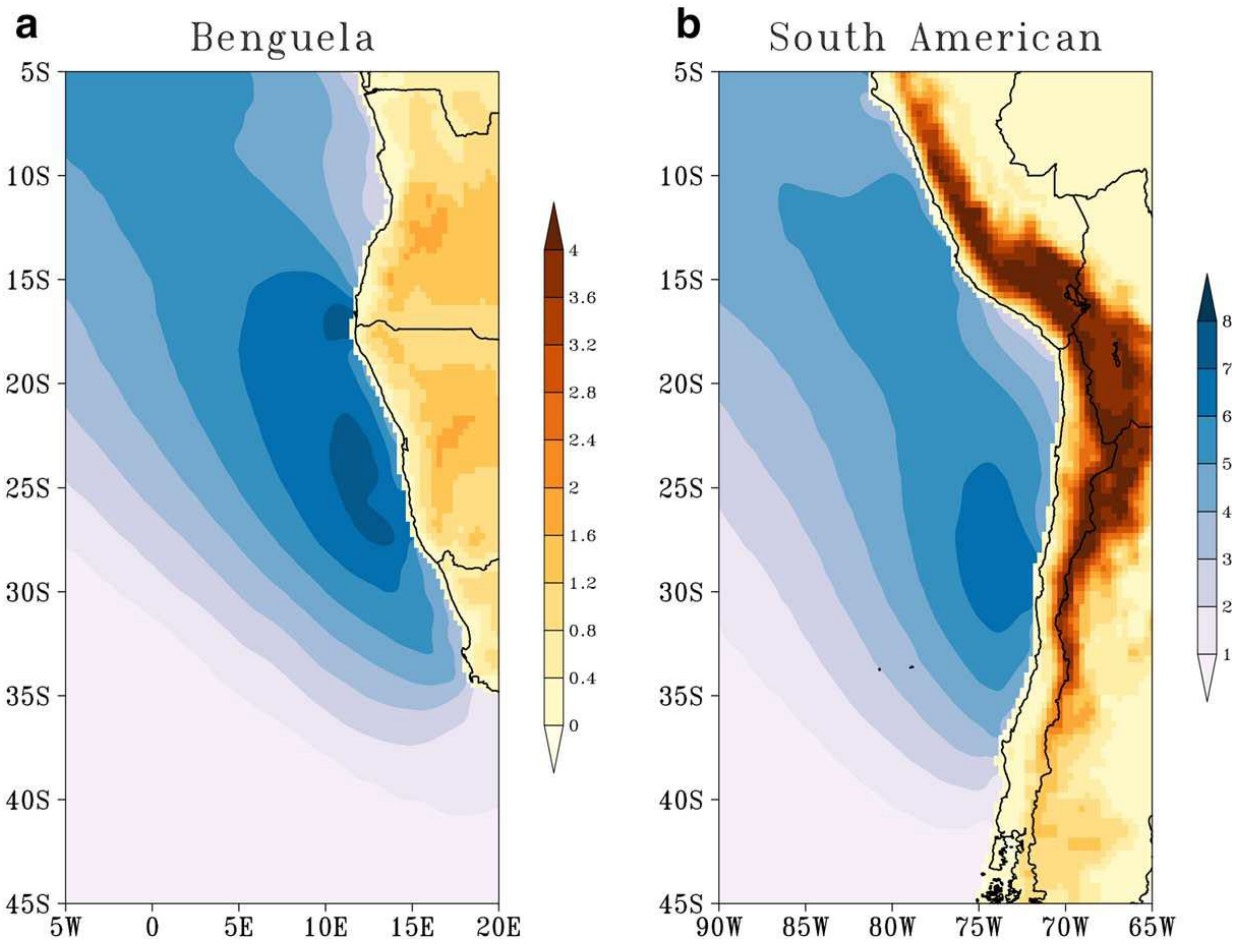


Figure 3.5: Taken from Patricola and Chang (2017, their Figure 1). 10-meter meridional wind (m/s) in blue and terrain height (km) in orange over the **a** Benguela and **b** South American coastal upwelling regions. Winds obtained from Scatterometer Climatology of Ocean Winds (SCOW) and elevation data from NCEP Climate Forecast System Reanalysis.

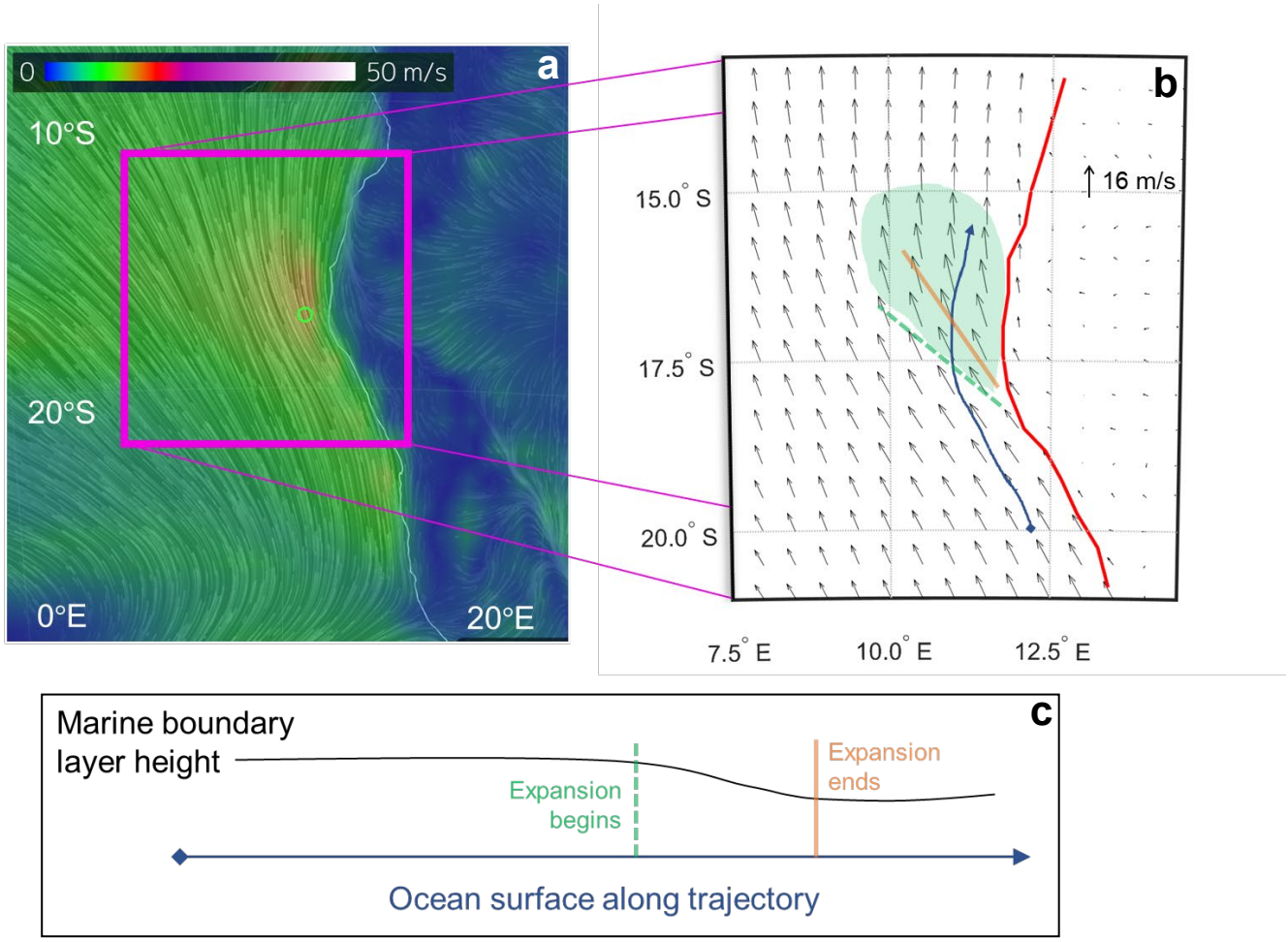


Figure 3.6: Detailed schematic of hydraulic expansion fan effect on offshore winds at 975 hPa (approximate level of the Benguela jet) adapted from Winant et al. (1988). Panel **a** shows GFS reanalysis of 1000-hPa wind (shading) with streamlines overlaid from 00 UTC on 8 May 2018. Orange shading around convex coastline is ~14 m/s. Panel **b** is a close-up view of **a** showing where the expansion begins (dashed green line), where it ends (solid orange) and the region that experiences stronger winds as a result (green shading). Wind vectors are taken from ERA5 at 975-hPa from the same time. Shown at the bottom is **c**) relative marine boundary layer height along the blue trajectory in **b**, with the beginning and end of expansion annotated.

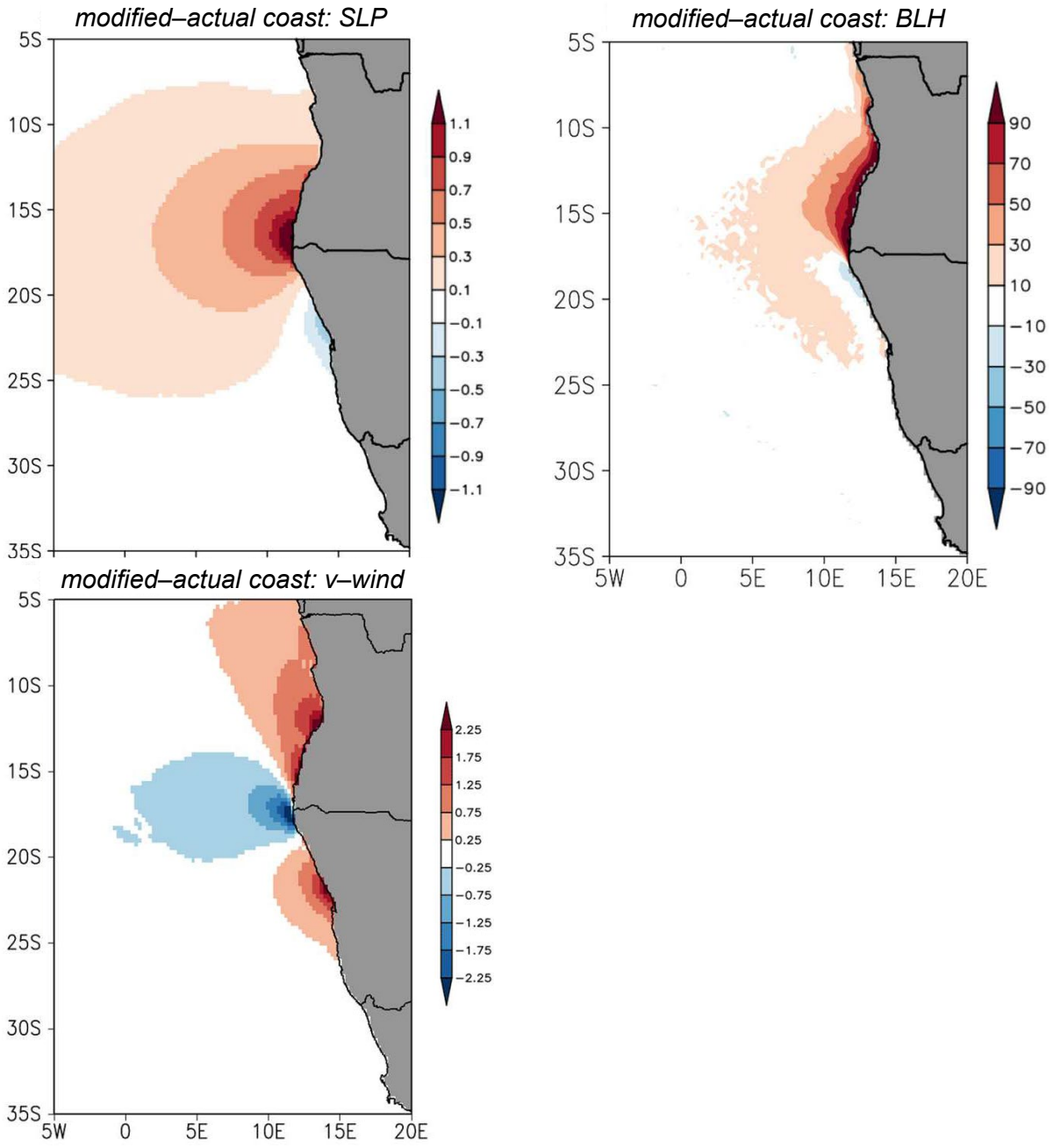


Figure 3.7: Taken from Patricola and Chang (2017, their Figures 13f, 14c, 14i). Difference of average sea level pressure (top left, in hPa), boundary layer height (right, in meters), and *v*-wind (bottom left, in m/s) between modified and actual coastline scenarios in 27-km WRF simulations for January 2005 to December 2009.

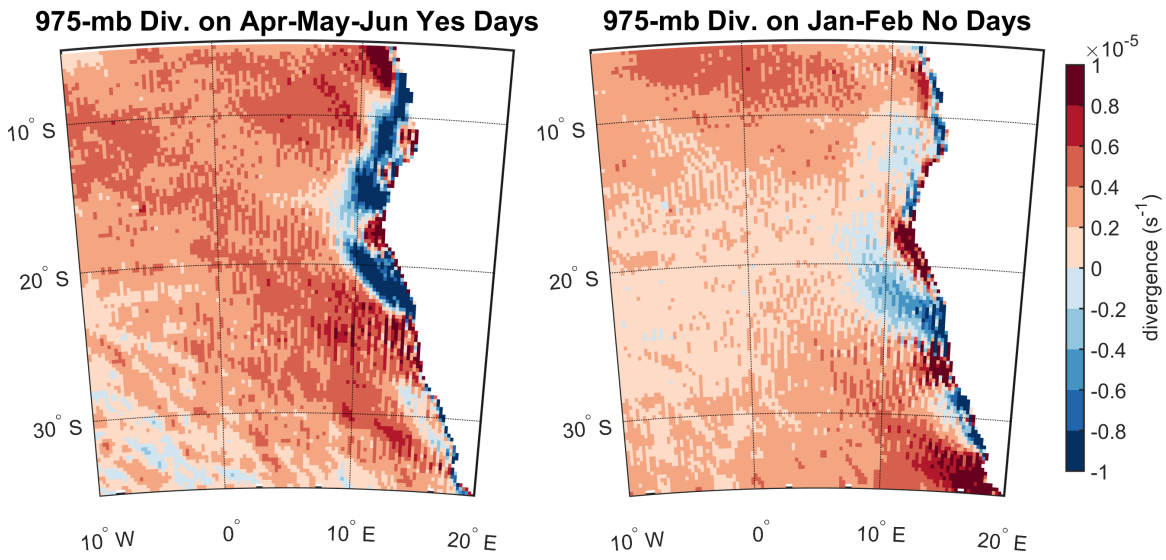


Figure 3.8: ERA5 975-hPa divergence (s^{-1}) at 22 UTC over the southeast Atlantic averaged by (**a**) days with a cloud-eroding boundary (Yes days) in Apr-May-Jun and (**b**) No days in Jan-Feb from 2007 to 2017. The positive values (red shading) indicate horizontal divergence, with blue shading for horizontal convergence.

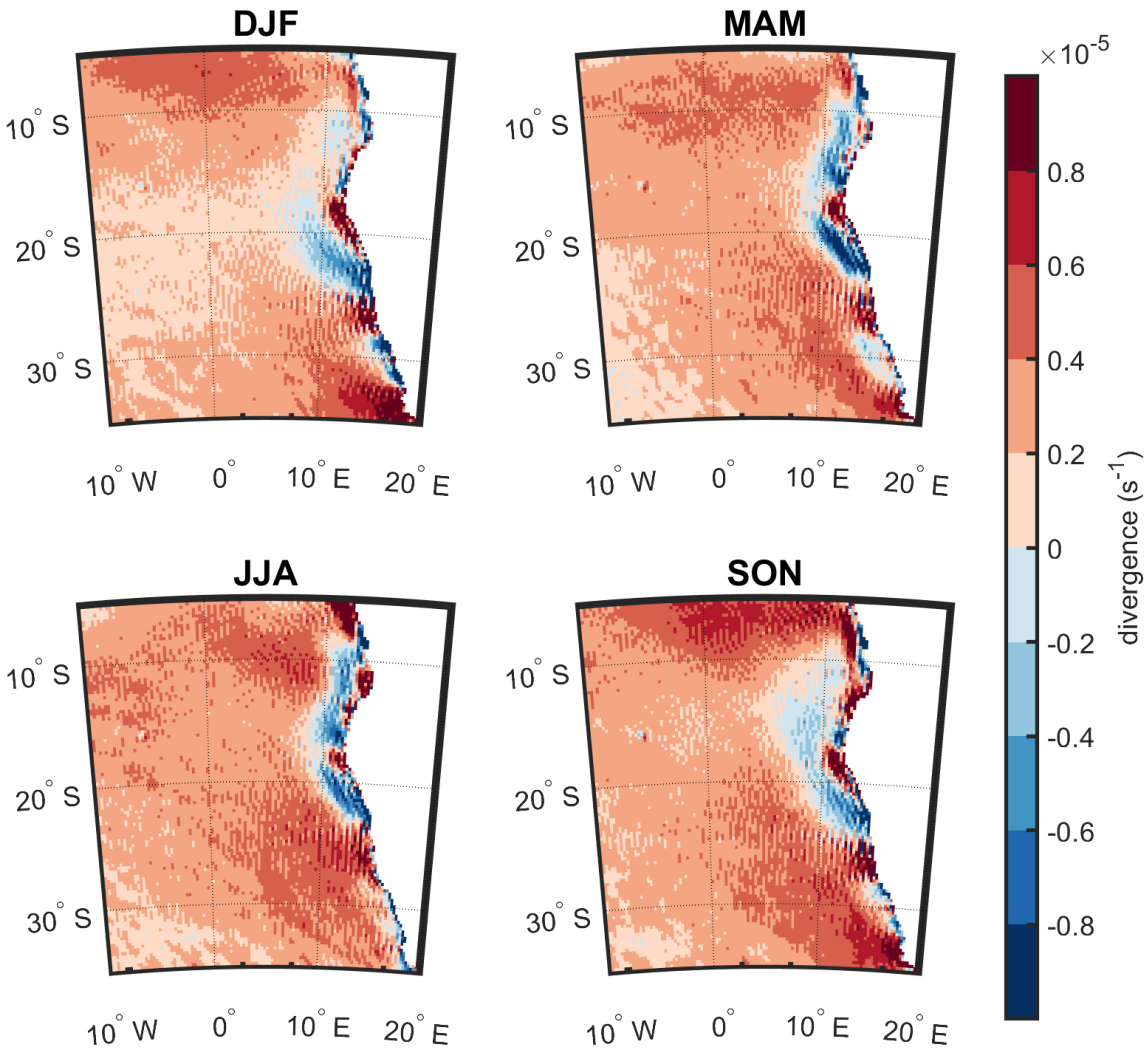
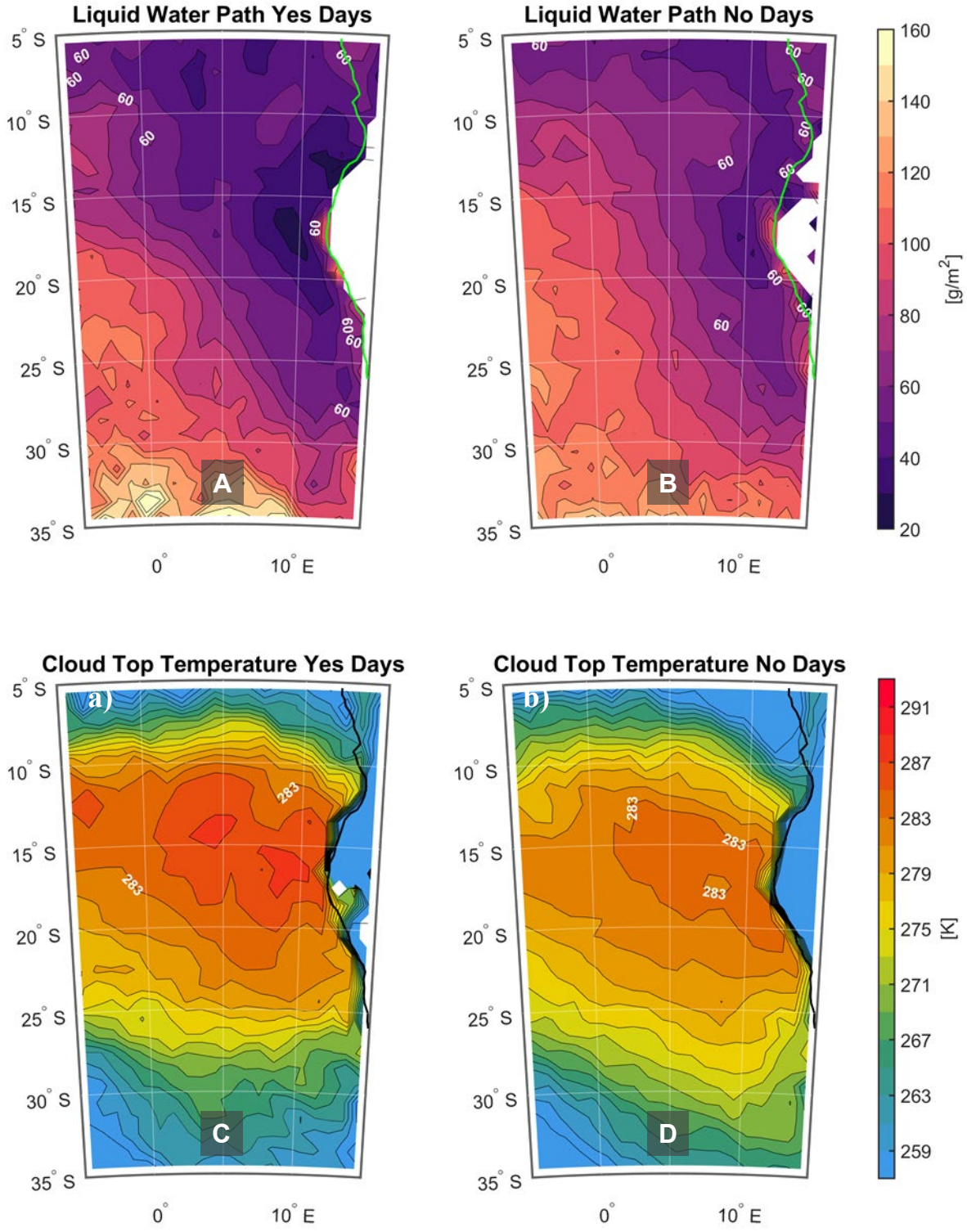
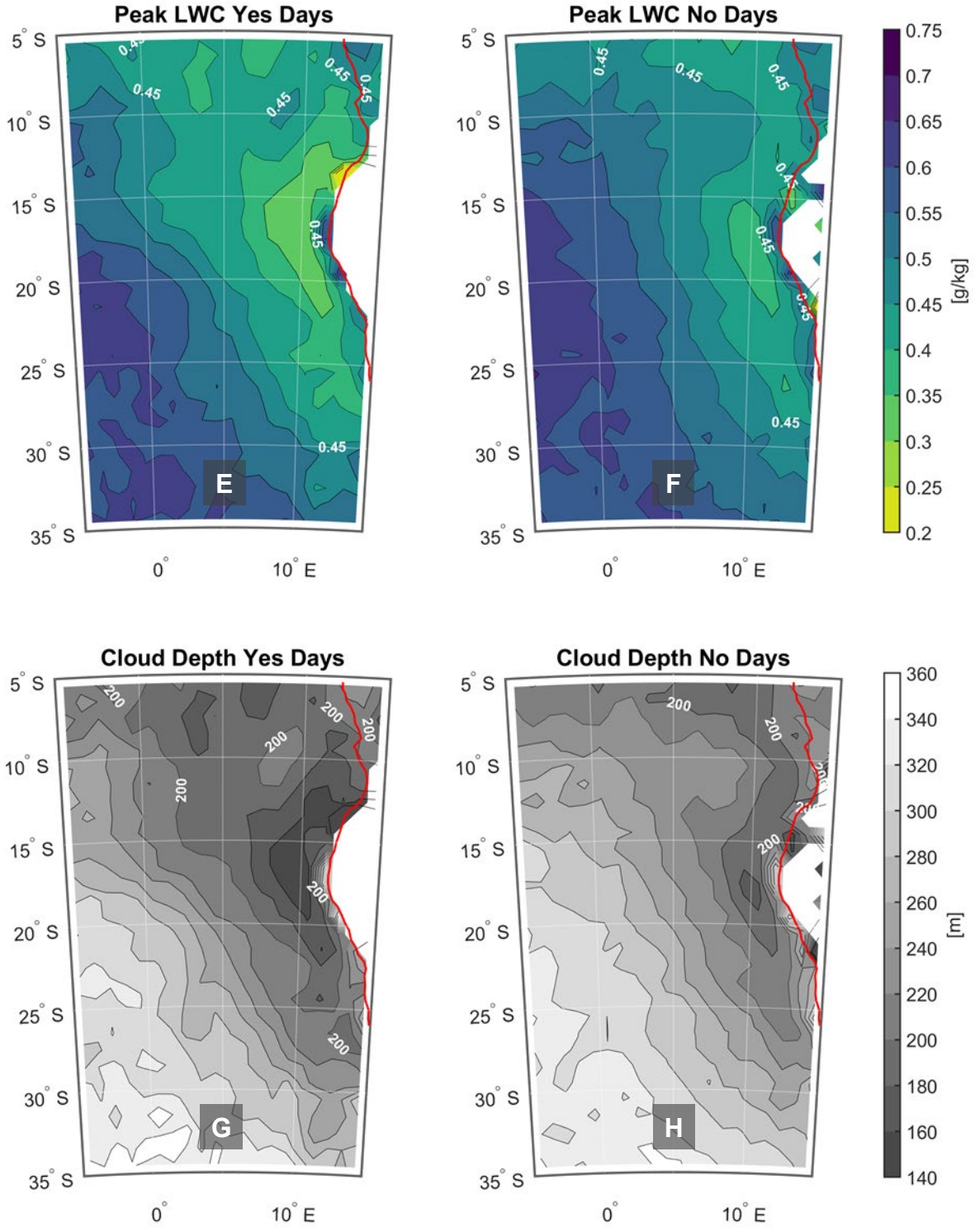


Figure 3.9: ERA5 975-hPa divergence (s^{-1}) at 22 UTC over the southeast Atlantic by season from 2007 to 2017. Red shading indicates horizontal divergence, with blue shading for horizontal convergence. Title over each panel indicates months included in average: **DJF** = Dec-Jan-Feb, **MAM** = Mar-Apr-May, **JJA** = Jun-Jul-Aug, **SON** = Sep-Oct-Nov.

Figure 3.10: Conditionally (cloud fraction ≥ 90%, cloud top ≤ 2.5 km) averaged MODIS (**A**/**B**) liquid water path and (**C**/**D**) cloud top temperature and MODIS-derived (**E**/**F**) peak liquid water content and (**G**/**H**) cloud thickness over the southeast Atlantic for days with a cloud-eroding boundary ("Yes Days", **A**/**C**/**E**/**G**) and days without a cloud-eroding boundary ("No Days", **B**/**D**/**F**/**H**) in April-May-June from 2007 to 2017. Liquid water path contoured every 10 g/m^2 with 60 g/m^2 labeled, cloud top temperature contoured every 2 K with 283 K labeled, thickness contoured every 20 meters with 200 m labeled, and peak LWC contoured every 0.05 g/kg with 0.45 g/kg labeled. Coastline is shown by the thin green (**A**/**B**), black (**C**/**D**), or red line (**E**–**H**). MODIS retrievals are from the Terra (10:30 am LT) satellite.


Liquid Water Path Yes Days
Liquid Water Path No Days
Cloud Top Temperature Yes Days
Cloud Top Temperature No Days
5° S
10° S
15° S
20° S
25° S
30° S
35° S
0°
10° E
A
B
C
D
a)
b)
60
283
160
140
120
100
80
60
40
20
[g/m2]
291
287
283
279
275
271
267
263
259
[K]


Peak LWC Yes Days
Peak LWC No Days
Cloud Depth Yes Days
Cloud Depth No Days
5° S
10° S
15° S
20° S
25° S
30° S
35° S
0°
10° E
E
F
G
H
0.45
200
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
[g/kg]
360
340
320
300
280
260
240
220
200
180
160
140
[m]

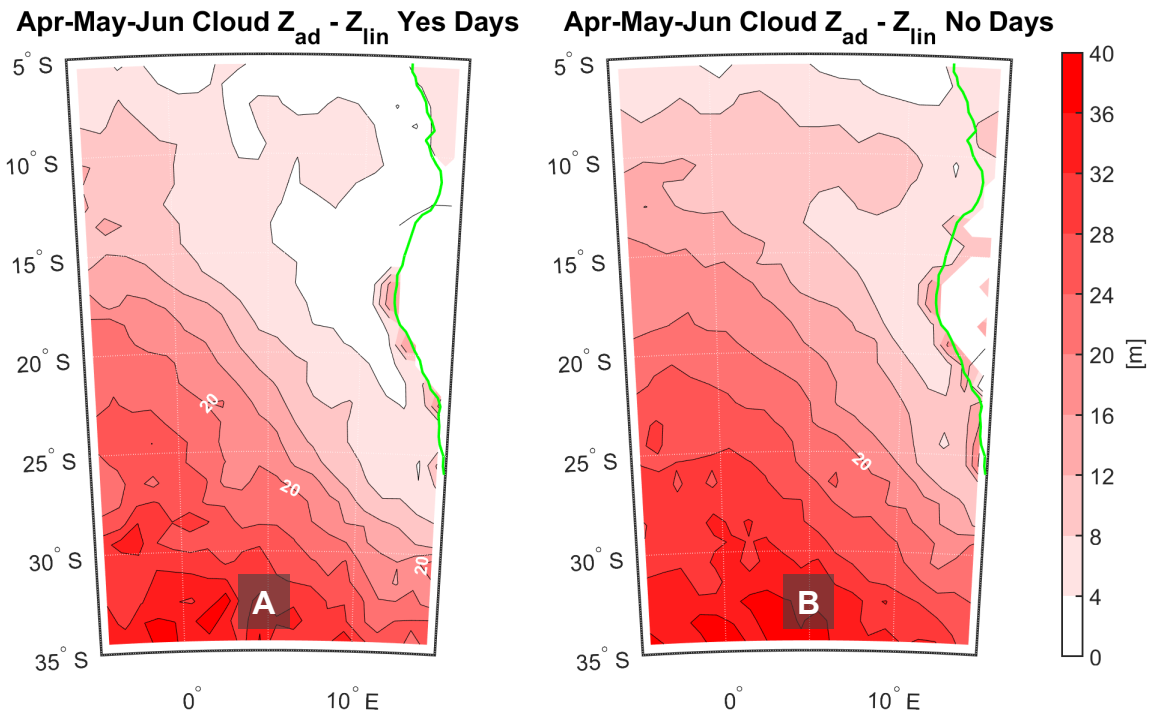


Figure 3.11: Mean difference between adiabatic cloud thickness derived from MODIS Terra (10:30 am LT) using a varying adiabatic change of liquid water content in cloud (Z_{ad}) and a constant adiabatic change of liquid water content (Z_{lin}) of 2.4 g/m^3 per km. Differences are contoured every 4 meters, with $Z_{ad} \geq Z_{lin}$ everywhere. Data are averaged by (**A**) days with a cloud-eroding boundary ("Yes" days) and (**B**) No days in Apr-May-Jun from 2000 to 2018.

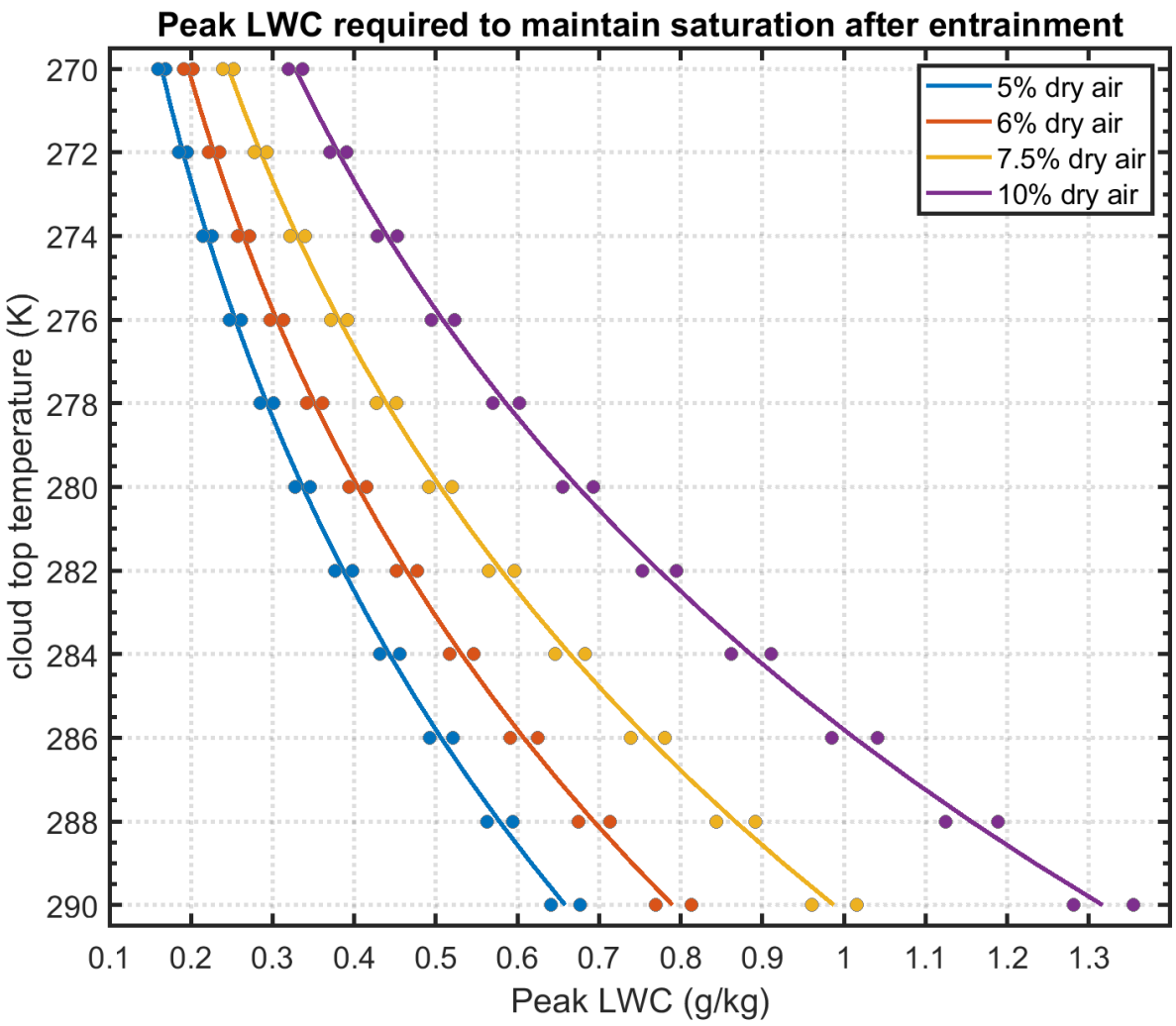
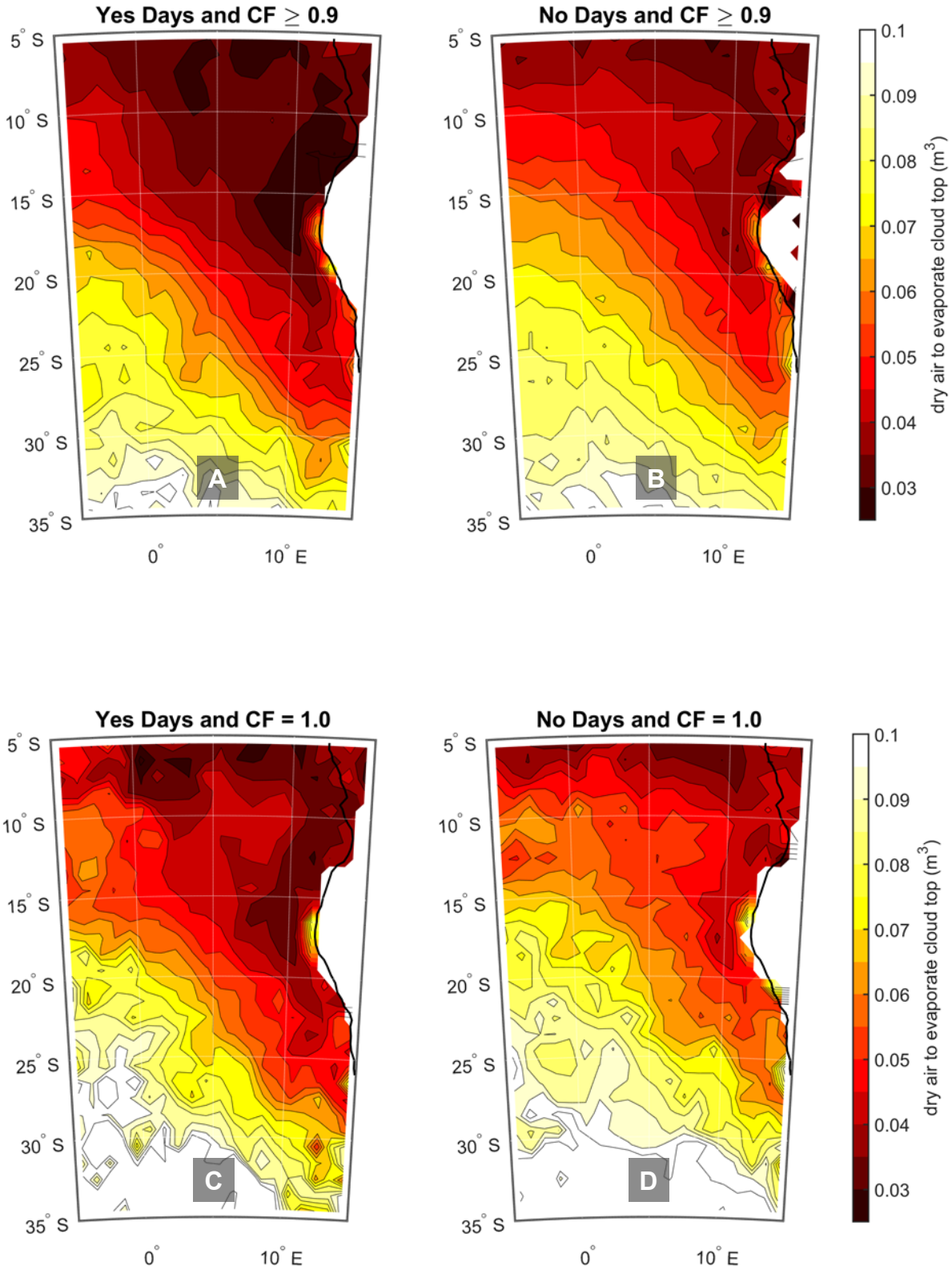


Figure 3.12: The amount of liquid water (in g/kg) that would be required to maintain saturation after entraining completely dry air (0 g/kg) with 1 m^3 of air at cloud top. Mixing ratio at cloud top is the saturation mixing ratio, calculated from cloud top temperature (varied from 270 to 290 K) and pressure (925 hPa). The sensitivity of the calculation to changing cloud top pressure by +/-25 hPa is shown by dots along each curve. Calculations for 950 hPa are to the left of each curve and for 900 hPa to the right. With higher pressure, less peak LWC is needed to maintain saturation.

Figure 3.13: The volume (in m^3) of dry air (0 g/kg) required to completely evaporate liquid water in 1 m^3 of air at cloud top, based on MODIS cloud top properties (temperature and pressure) and MODIS-derived peak liquid water content separated by Yes (**A**/**C**) and No (**B**/**D**) days. All panels are conditioned on cloud top $\leq$ 2.5 km altitude, with panels **A** and **B** also conditioned on cloud fraction $\geq$ 0.9 while panels **C** and **D** are also conditioned on cloud fraction = 1.0. MODIS retrievals are from the Terra (10:30 am LT) satellite.


Yes Days and CF ≥ 0.9
No Days and CF ≥ 0.9
Yes Days and CF = 1.0
No Days and CF = 1.0
5° S
10° S
15° S
20° S
25° S
30° S
35° S
0°
10° E
A
B
C
D
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
dry air to evaporate cloud top (m3)

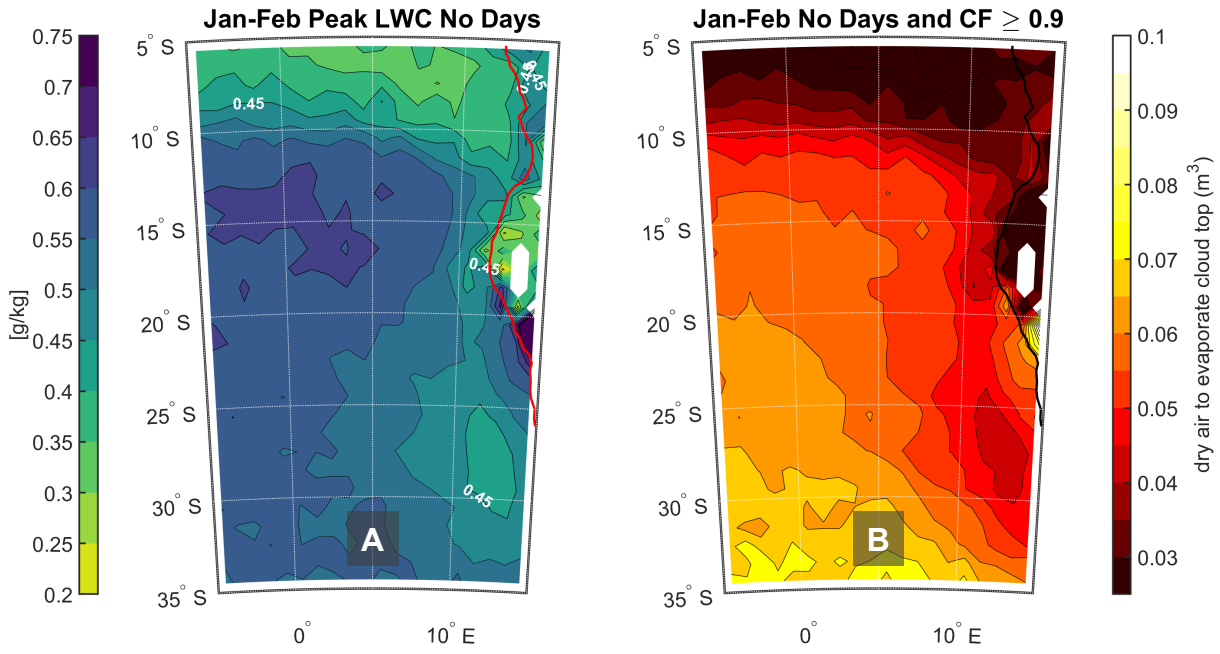


Figure 3.14: Conditionally (cloud fraction ≥ 90%, cloud top ≤ 2.5 km) averaged MODIS-derived (**A**) Peak liquid water content and (**B**) volume (in m^3) of dry air (0 g/kg) required to completely evaporate liquid water in 1 m^3 of air at cloud top for No days in Jan-Feb (2000-2018). MODIS retrievals are from the Terra (10:30 am LT) satellite.

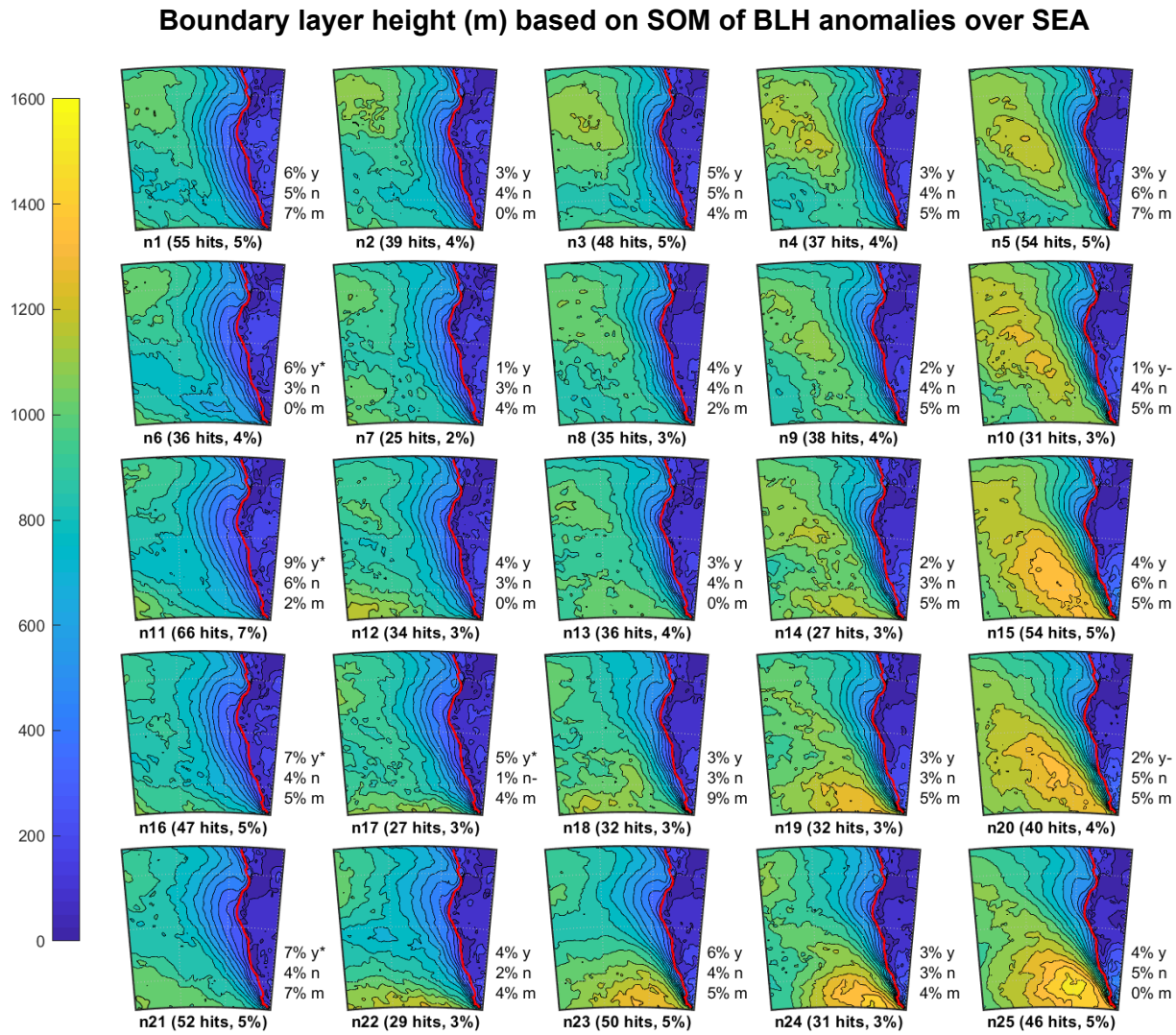


Figure 3.15: SOM trained on normalized anomalies of ERA5 boundary layer height in Apr-May-Jun at 22 UTC over the southeast Atlantic. Nodes depict average of boundary layer height fields that best match that node, contoured every 80 meters. Label below each node indicates number of days that best match that node. Percent of best matching yes (y), no (n), and maybe (m) boundary days out of all days in Apr-May-Jun from 2007-2017 are annotated to the right of each node. Asterisks (*) indicate unlikely to be random high counts, and dashes (-) indicate unlikely to be random low counts.

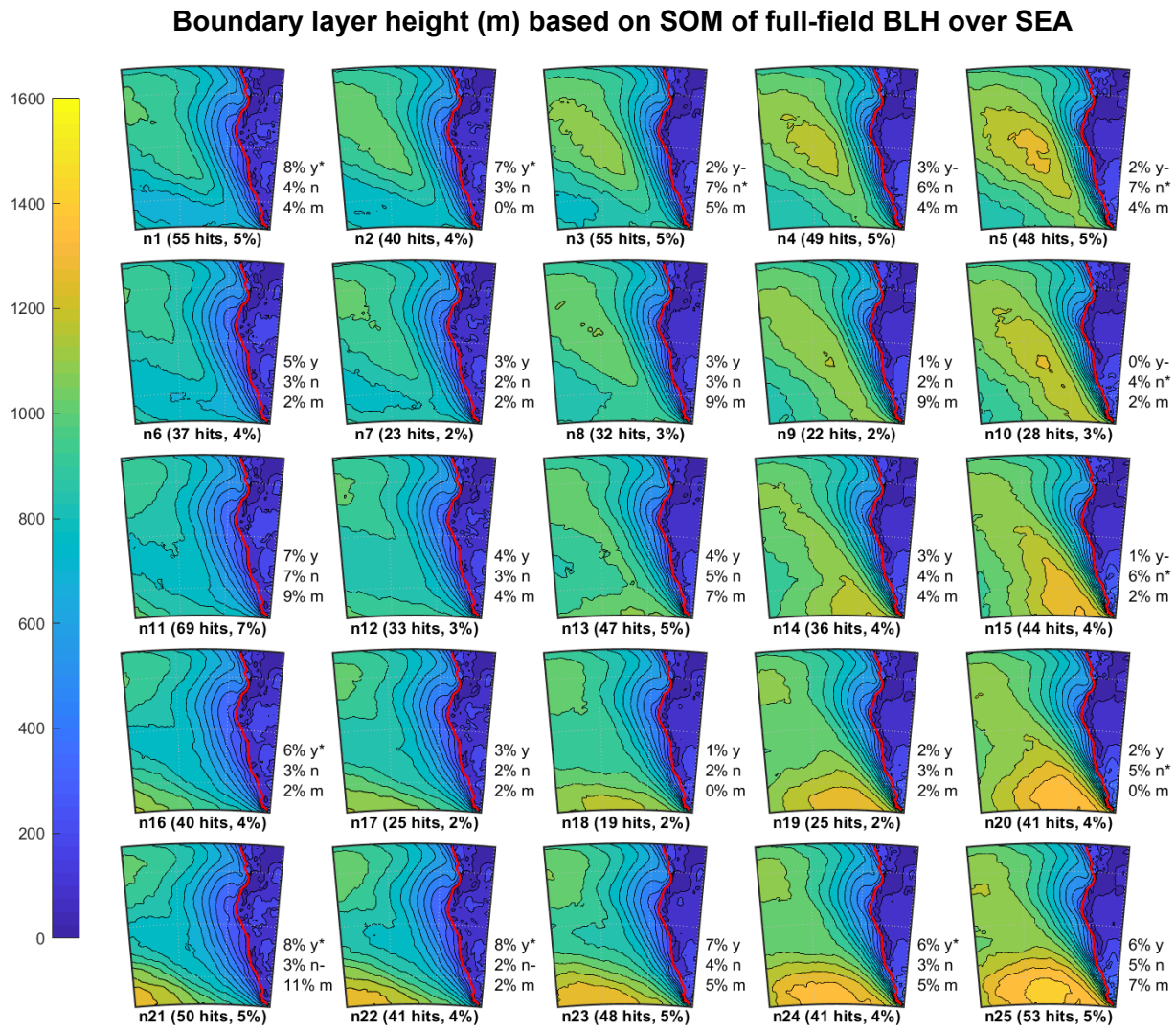


Figure 3.16: SOM trained on full ERA5 boundary layer height in Apr-May-Jun at 22 UTC over the southeast Atlantic. Nodes depict result of SOM trained on full field of boundary layer height, contoured every 80 meters. Node labels and annotations are the same format as Figure 3.15.

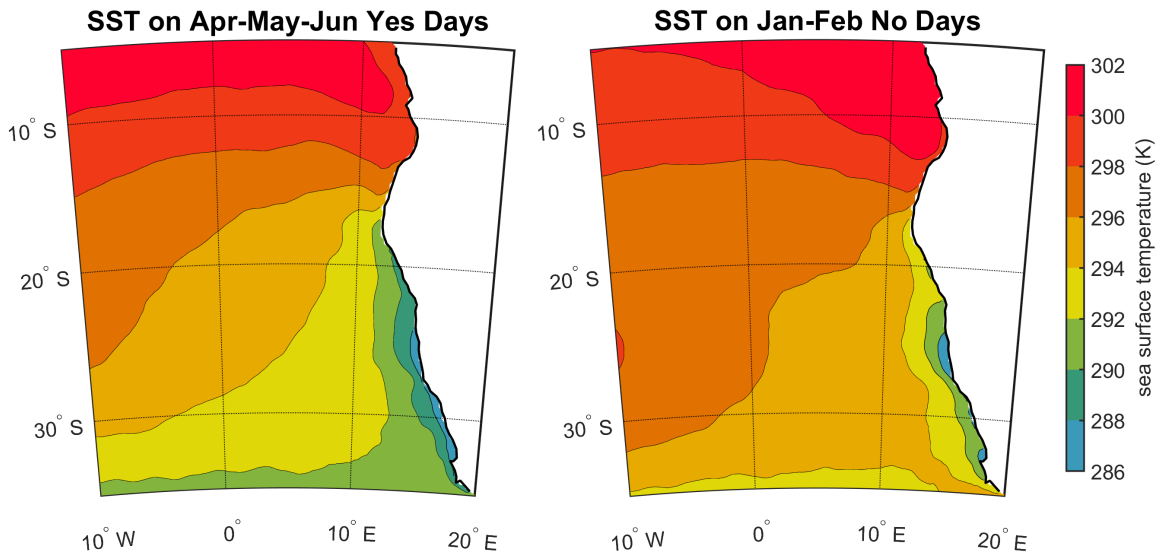


Figure 3.17: ERA5 sea surface temperature in K from 2007 to 2017 averaged by (**a**) days with a cloud-eroding boundary ("Yes" days) in Apr-May-Jun and (**b**) No days in Jan-Feb. Contours are every 2 K, with the green line indicating the west African coast.

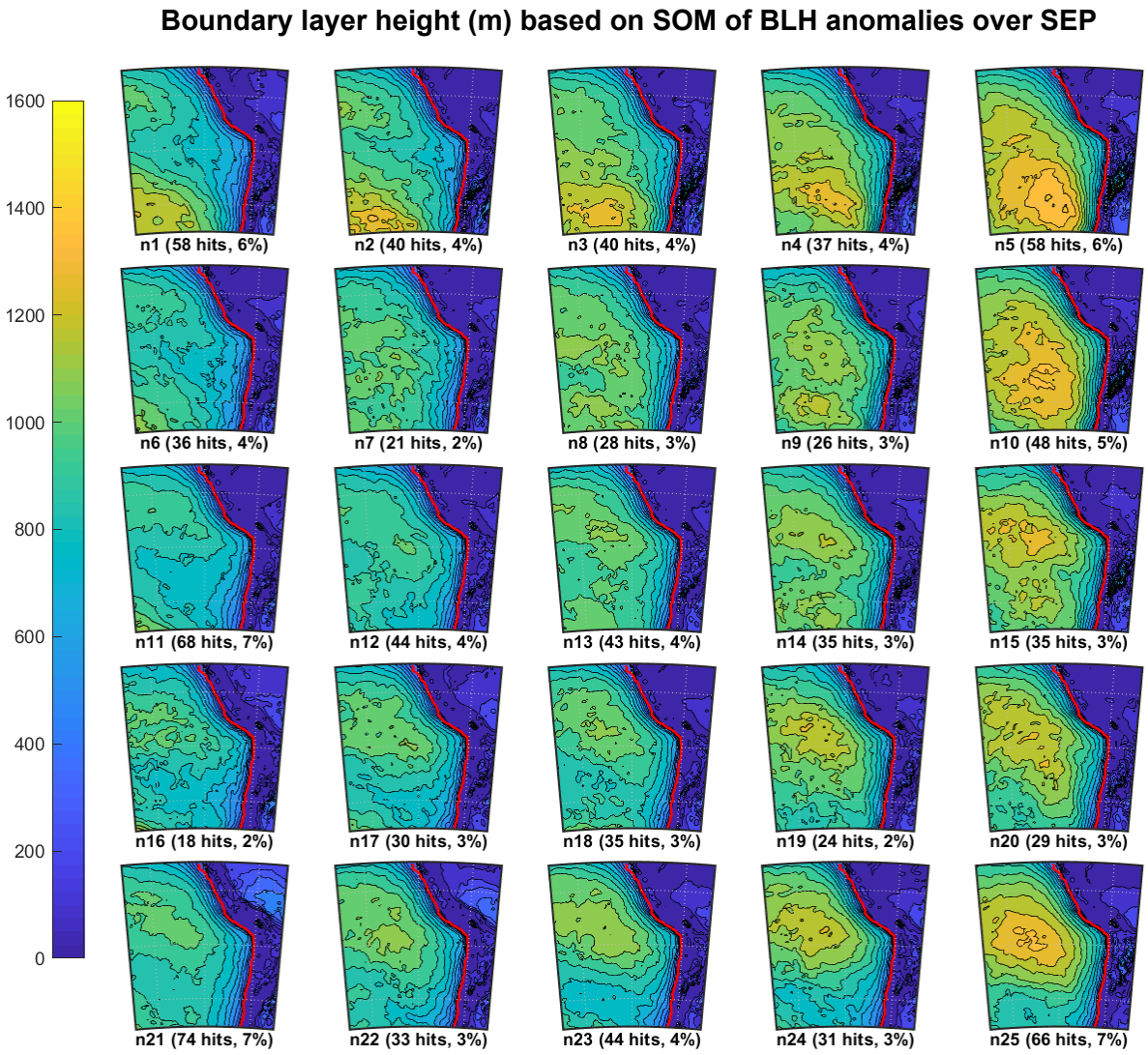


Figure 3.18: SOM trained on normalized anomalies of ERA5 boundary layer height in Apr-May-Jun at 22 UTC over the southeast Pacific. Nodes depict average of boundary layer height fields that best match that node, contoured every 80 meters. Node labels are the same format as Figure 3.15.

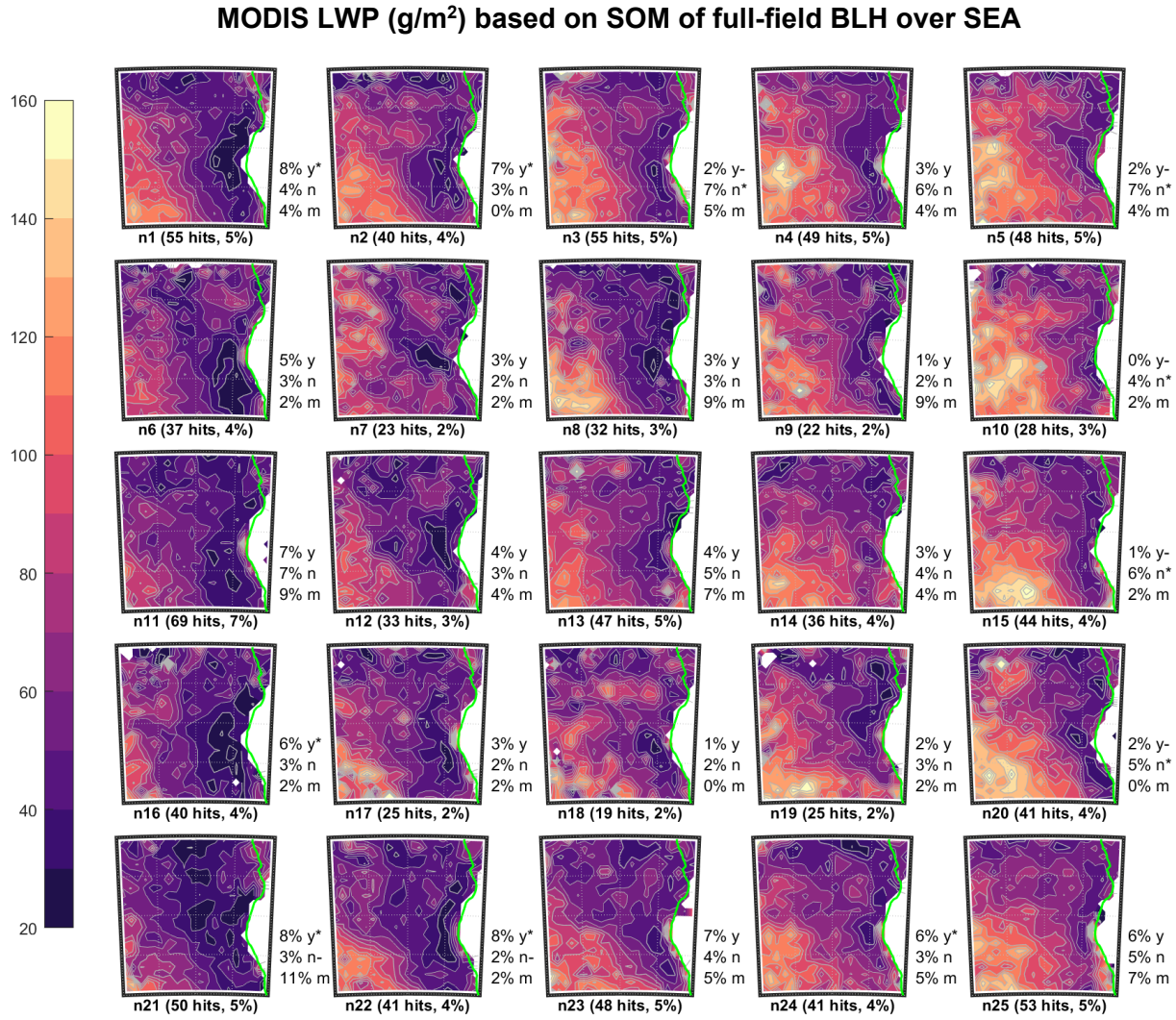


Figure 3.19: Conditionally averaged (cloud fraction ≥ 0.9 and cloud top height ≤ 2.5 km) MODIS Terra (10:30 am LT) liquid water path by best-matching node in the SOM trained on the full field of ERA5 boundary layer height. Liquid water path contoured every 10 g/m^2, and node labels are the same format as Figure 3.15.

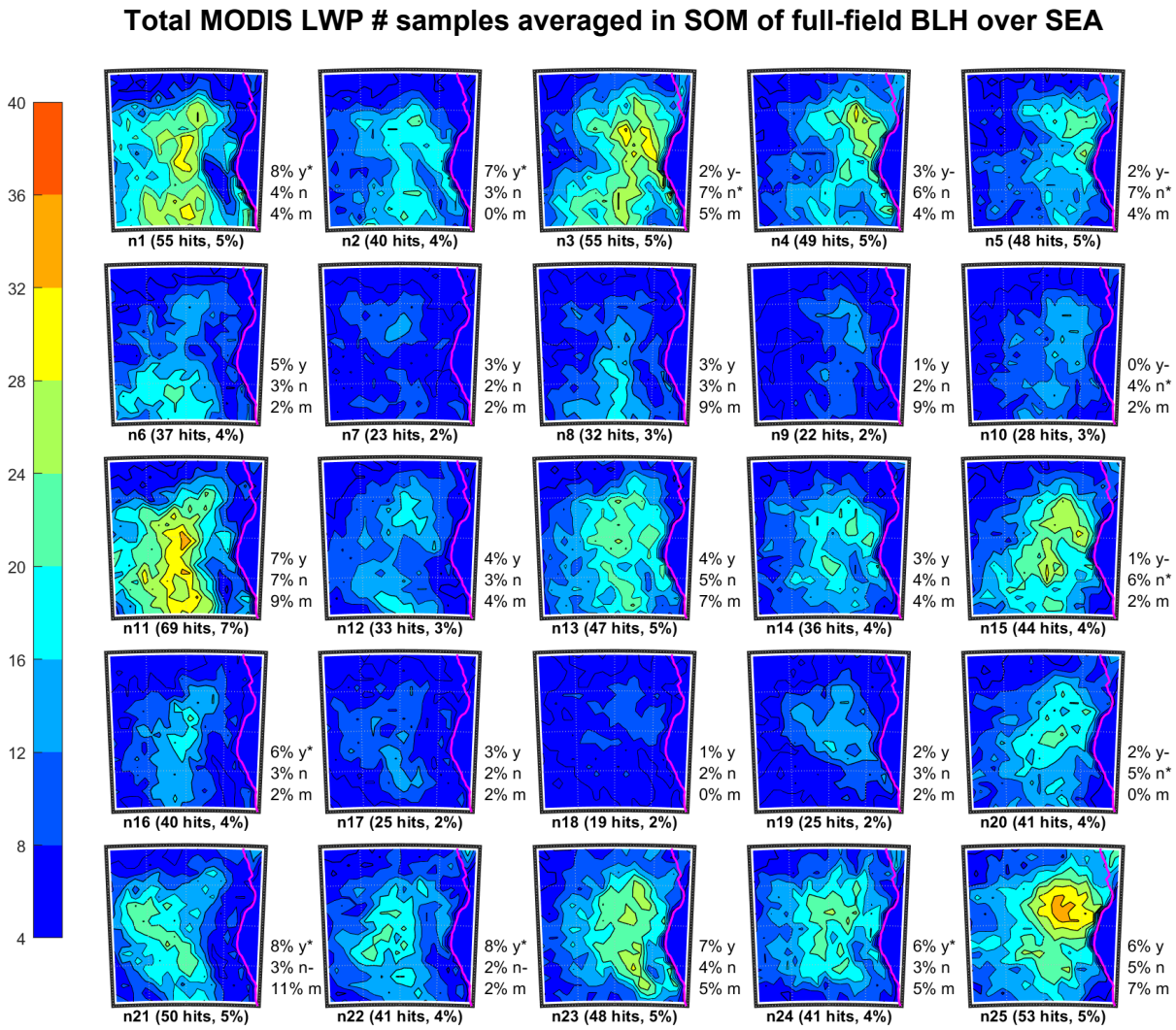


Figure 3.20: In the SOM trained on the full field of ERA5 boundary layer height, the number of samples per pixel included in the averaging for each node, conditional on cloud fraction ≥ 0.9 and cloud top ≤ 2.5 km altitude.

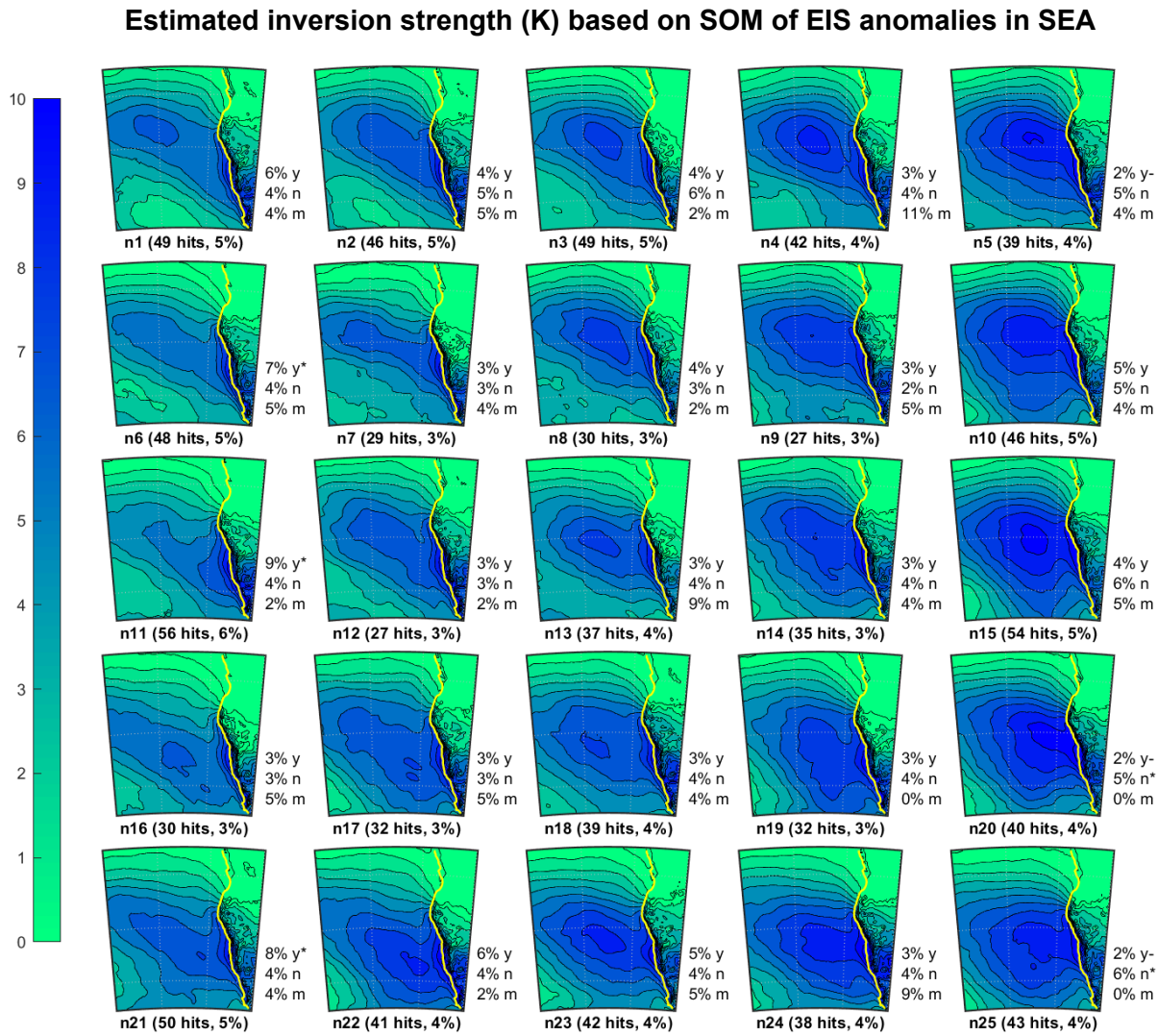


Figure 3.21: SOM trained on normalized anomalies of EIS derived from ERA5 in Apr-May-Jun at 22 UTC. Nodes depict average of EIS fields that best match that node, contoured every 0.5 K. Node labels and annotations are the same format as Figure 3.15.

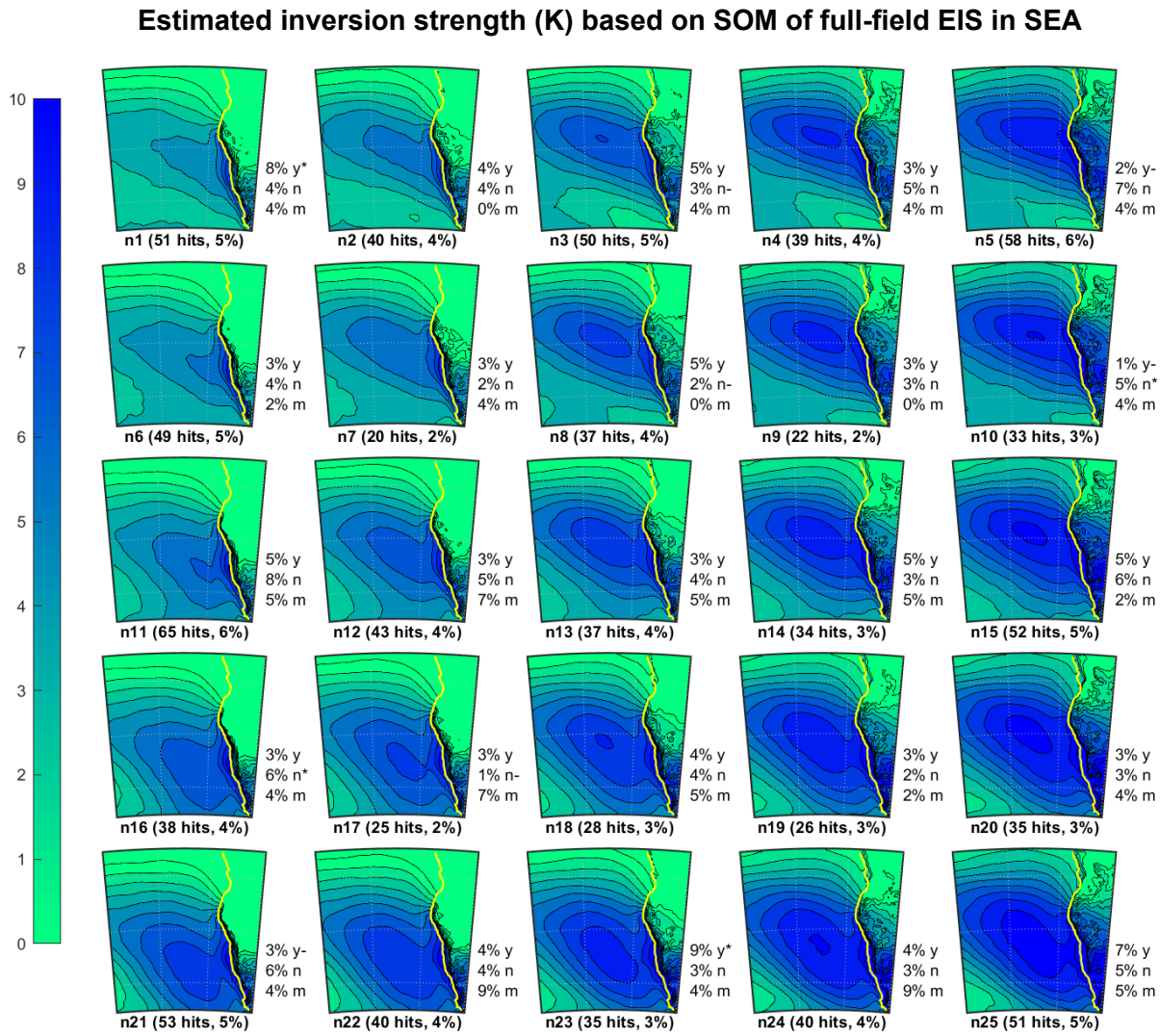


Figure 3.22: SOM trained on full EIS in Apr-May-Jun at 22 UTC. Nodes depict result of SOM trained on full field of EIS, contoured every 0.5 K. Node labels and annotations are the same format as Figure 3.15.

APPENDICES

APPENDIX A – Flight data from ORACLES

A recent field campaign, ObseRvations of Aerosols above Clouds and their intEractionS (ORACLES, Dzambo et al. 2019), provides some data in a region deprived of quality in-situ measurements of key atmospheric properties. The field observations were obtained during September and October, a time when cloud-eroding boundaries are uncommon (Section 2.1). However, of interest in the context of our study is the accuracy of ERA5 boundary layer height. Airborne high spectral resolution lidar data captured from ER-2 aircraft during September 2016 allows for evaluation of ERA5 boundary layer height. Figures A1–6 compare the attenuated backscatter at 532 nm wavelength from the lidar to ERA5 boundary layer height and vertical profile of potential temperature. ERA5 data are taken from the pixel that is collocated with the aircraft's location at the closest valid time. Lidar does not penetrate cloud, so where there is no lidar echo below a certain height there is cloud, effectively marking the observed boundary layer depth. An example of broken cloud is seen in Figure A6 by radar echo reaching the surface between 1000 and 1030 UTC.

Evident from the figures is a low (up to 600 m) bias of ERA5 boundary layer height relative to airborne lidar-sensed boundary layer height, consistent with known issues of errors ± 50% for boundary layers at or below 1 km altitude (Seidel et al. 2012, Section 2.4). The vertical gradient of potential temperature, which increases drastically at the inversion which traps marine stratocumulus, better estimates the lidar-sensed boundary layer height but is also biased low. Shinozuka et al. (2019) suggest that a relative humidity threshold (70%) better approximates the height of the inversion and thus the boundary layer height, consistent with the results shown in Figure 2.6 of this thesis. Near future work will compare ERA5 relative humidity to the

ORACLES lidar data to verify the utility of this approach before potentially computing relative-humidity-based boundary layer height for analysis using SOMs.

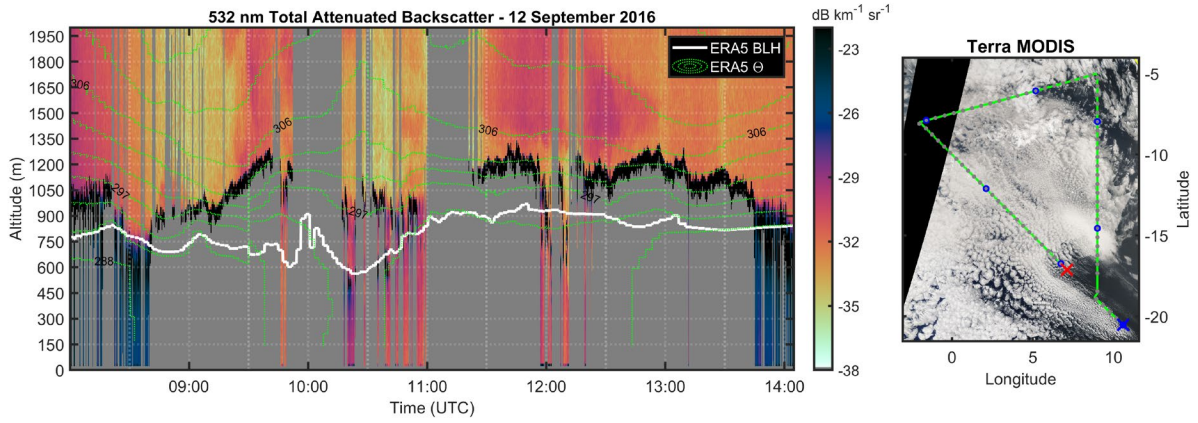


Figure A1: September 12, 2016: Attenuated backscatter (dB km^{-1} sr^{-1}, shading) from ORACLES ER-2 lidar compared to ERA5 boundary layer height (white) and potential temperature (green, contoured every 3K). Corrected reflectance image from closest MODIS Terra/Aqua overpass also shown with flight path overlaid, with the (blue) starting and (red) ending locations marked by X's and blue dots at each hour.

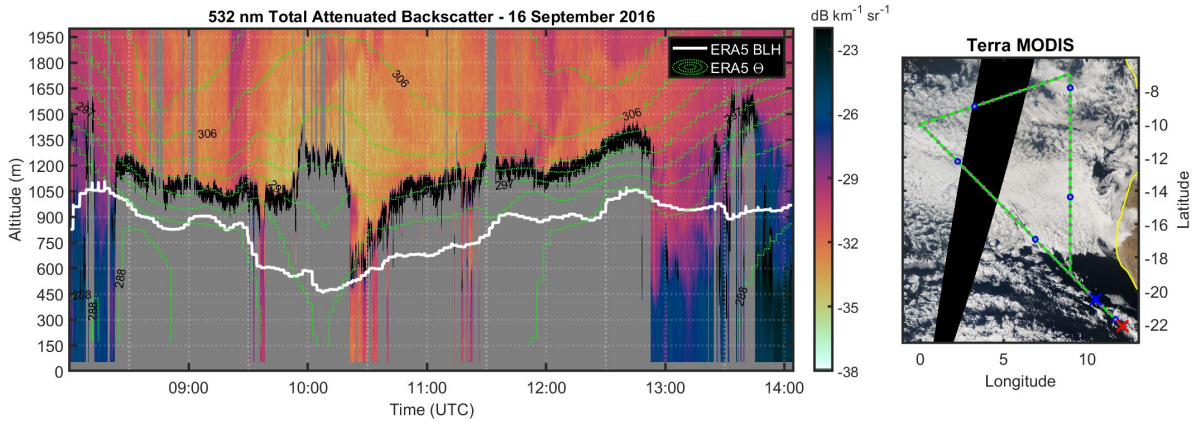


Figure A2: September 16, 2016: Attenuated backscatter (dB km^{-1} sr^{-1}, shading) from ORACLES ER-2 lidar compared to ERA5 boundary layer height (white) and potential temperature (green, contoured every 3K). Corrected reflectance image from closest MODIS Terra/Aqua overpass also shown with flight path overlaid, with the (blue) starting and (red) ending locations marked by X's and blue dots at each hour.

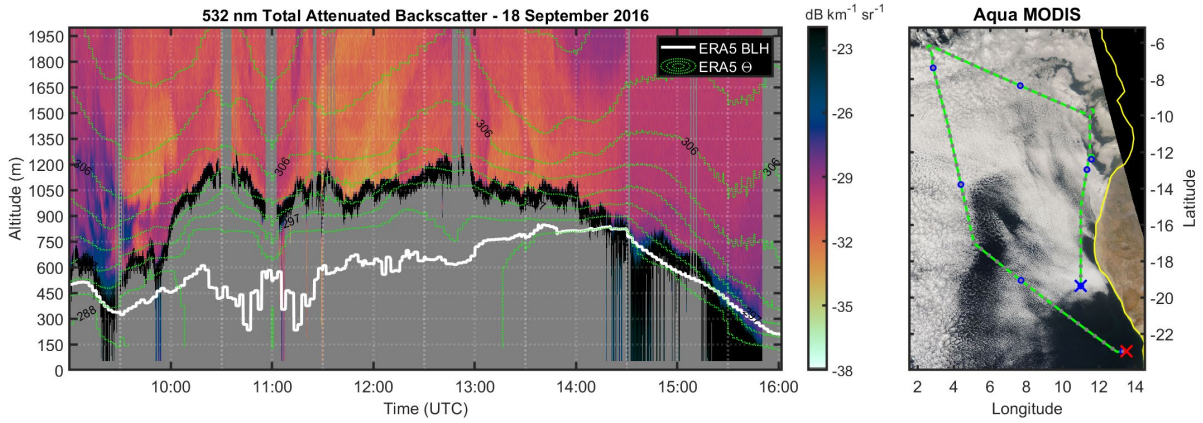


Figure A3: September 18, 2016: Attenuated backscatter (dB km^{-1} sr^{-1}, shading) from ORACLES ER-2 lidar compared to ERA5 boundary layer height (white) and potential temperature (green, contoured every 3K). Corrected reflectance image from closest MODIS Terra/Aqua overpass also shown with flight path overlaid, with the (blue) starting and (red) ending locations marked by X's and blue dots at each hour.

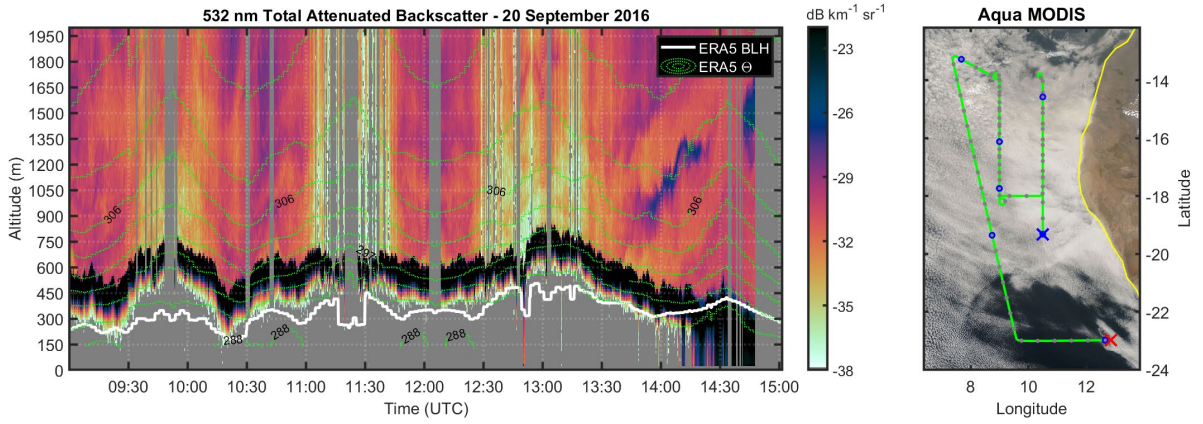


Figure A4: September 20, 2016: Attenuated backscatter (dB km^{-1} sr^{-1}, shading) from ORACLES ER-2 lidar compared to ERA5 boundary layer height (white) and potential temperature (green, contoured every 3K). Corrected reflectance image from closest MODIS Terra/Aqua overpass also shown with flight path overlaid, with the (blue) starting and (red) ending locations marked by X's and blue dots at each hour.

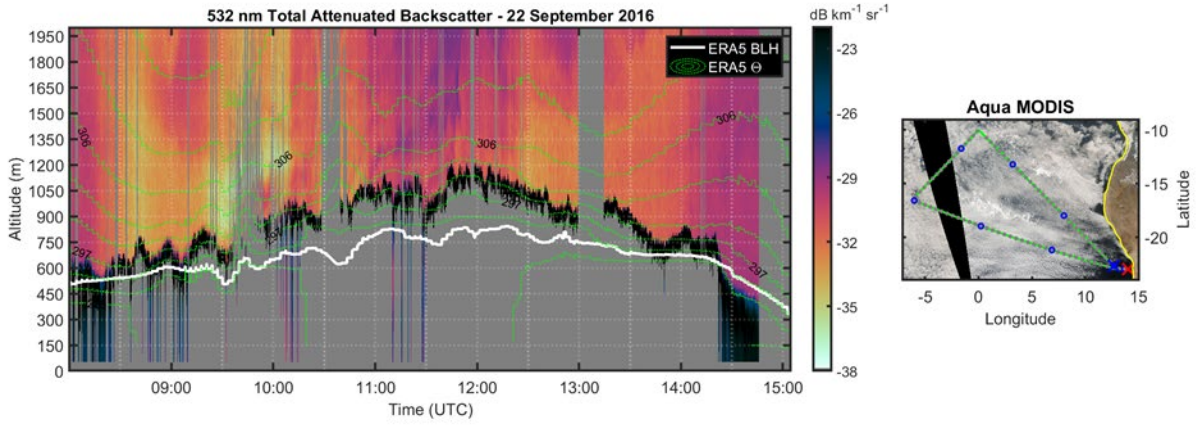


Figure A5: September 22, 2016: Attenuated backscatter (dB km^{-1} sr^{-1}, shading) from ORACLES ER-2 lidar compared to ERA5 boundary layer height (white) and potential temperature (green, contoured every 3K). Corrected reflectance image from closest MODIS Terra/Aqua overpass also shown with flight path overlaid, with the (blue) starting and (red) ending locations marked by X's and blue dots at each hour.

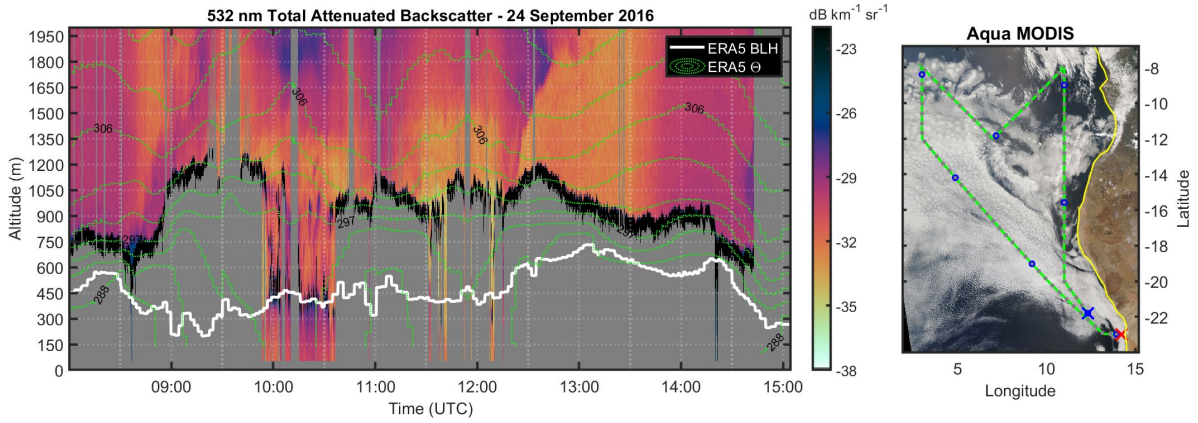


Figure A6: September 22, 2016: Attenuated backscatter (dB km^{-1} sr^{-1}, shading) from ORACLES ER-2 lidar compared to ERA5 boundary layer height (white) and potential temperature (green, contoured every 3K). Corrected reflectance image from closest MODIS Terra/Aqua overpass also shown with flight path overlaid, with the (blue) starting and (red) ending locations marked by X's and blue dots at each hour.

APPENDIX B – GPS Radio Occultations

In this appendix, we elaborate on how we chose to use GPS radio occultation data to aid in our analysis of the large-scale factors relevant to cloud-eroding boundaries in the southeast Atlantic. GPS radio occultations obtain estimates of dry temperature (where moisture is small or not present) and pressure by assessing the angle at which a beam from a GPS satellite bends as it passes through different vertical levels of the atmosphere. GPS satellites orbit at about 22,000 km above the Earth's surface, sending continuous signals out which can be received and interpreted by low-Earth orbiting (LEO) satellites (orbiting ~800 km above ground) to obtain a profile through the atmosphere as the satellites move (see Figure B1 below). In this study, we examine GPS radio occultations obtained by the COSMIC mission (consisting of the FORMOSAT constellation of LEO satellites) and by the EUMETSAT Polar System (consisting

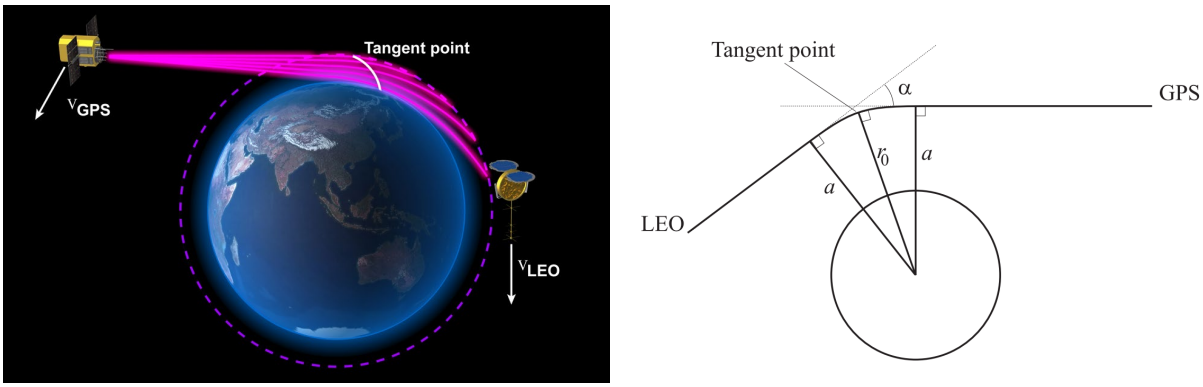


Figure B1: Diagrams illustrating the GPS radio occultation technique. An occultation profile is obtained by combining the tangent point at multiple levels in conjuncture with the LEO and GPS satellites' movement. The tangent point is demonstrated more concisely in the right panel, while the left panel shows the succession of occultations which produce a vertical profile with varying location with height. Note in the left panel that the horizontal distance covered by the vertical profile is not to scale. The horizontal position of the profile typically varies from 50-100 km through the depth of the troposphere (see Figure B2). (left panel from UCAR; right panel is Figure 1.4 from Syndergaard, 1999)

of the MetOp constellation of LEO satellites). According to Snell's law, the amount of ray bending is proportional to vertical gradients in density and hence refractive index. The details of GPS signal processing to obtain a vertical profile of refractivity are outlined by Kursinski et al. (1997). Refractivity (N) is then related to temperature (T; K), pressure (p; hPa), water vapor pressure (e; hPa), and electron density (, electrons per cubic meter) by:

where scattering from suspended liquid or ice particles is neglected (Kursinski, 1997) and f is the frequency of the GPS carrier signal in Hz. The 3rd term in this equation is the ionospheric term and is only pertinent above 60-90 km altitude. This leaves the first two terms, which are denoted the dry and moist terms, respectively.

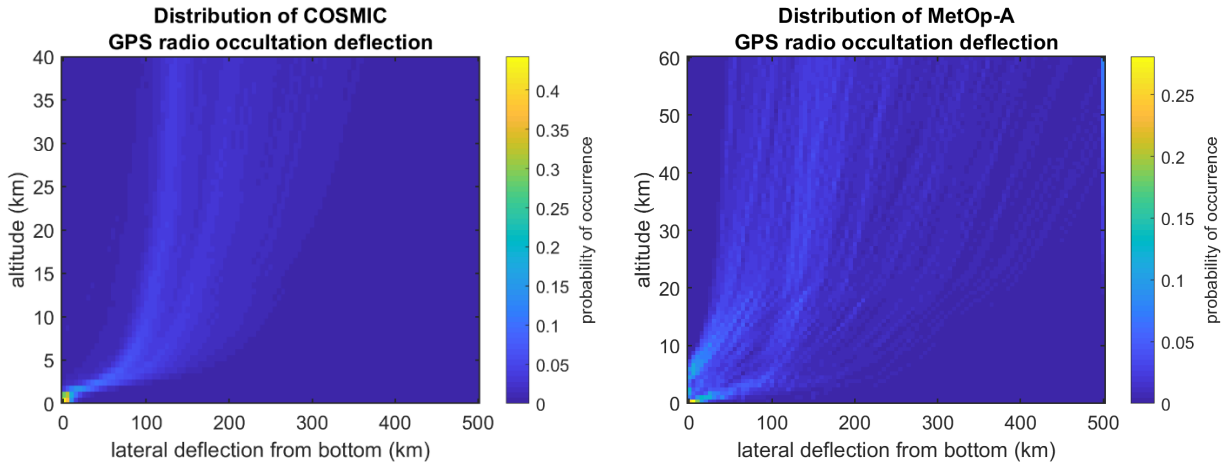


Figure B2: Plots showing how the location of GPS radio occultations vary with height. The left panel is a distribution of GPS radio occultations from COSMIC during the period 2007-2017 (n = 8976 occultations). The right panel is the same as the left but with GPS radio occultations taken from MetOp-A during the period 2007-2012 (n = 769 occultations).

Pressure is derived from density according to the ideal gas law and hydrostatic equilibrium, starting high in the atmosphere using a temperature guess. Pressure errors from the temperature guess diminish as the integration continues downward through the atmosphere

(Kursinski et al. 1997). Thus, with pressure derived on each level and refractivity obtained from the GPS radio occultation, the vertical profile of *either* temperature or water vapor pressure can be derived. Where water vapor pressure is relatively small (< 0.01 Pa), temperature can be derived with high accuracy (root-mean-square temperature errors of 0.25 K or less above 7 km altitude; Kursinski et al. 1997). However, where water vapor pressure is non-negligible (i.e. in the mid- to lower-troposphere, especially at lower latitudes), either temperature *or* water vapor pressure can be estimated, since deriving one requires independent observations of the other. Using independent observations of water vapor pressure prove ineffective for accurate estimation of temperature below 7 km altitude, especially in lower latitudes and where observational data is sparse (e.g. over the southeast Atlantic), as shown in our Figures B3 and B4 and in Figure 23 from Kursinski et al. (1997). Kursinski et al. (1995) estimated the accuracy of water vapor estimates and found that humidity profiles can be accurate to 5% in the convective boundary layer if independent temperature is accurate to ±1.5 K. However, given what we found when comparing the radio occultation-derived profiles of mixing ratio to interpolated ECMWF reanalysis (see Figures B5 and B6), we cannot justify the usage of this data in our study.

Given these drawbacks of radio occultations, their primary use in our study is to provide an estimate for planetary boundary layer height. One of the most skillful algorithms for estimating boundary layer height from radio occultation data is by finding the level of the minimum (most negative) vertical gradient of refractivity (Ao et al. 2012; Xie et al. 2012). Since refractivity is strongly affected by vertical gradients in moisture and temperature, and since refractivity increases towards the surface, the level where refractivity decreases the most with height provides a good estimate of boundary layer height. Moreover, vertical gradients in temperature and moisture atop marine stratocumulus are especially large. The resulting sharp

refractivity gradient thus yields the best possible boundary layer height estimates, conveniently in our domain of interest. Figures B3 and B4 illustrate how the minimum refractivity gradient of the GPS radio occultation, or even a threshold on the refractivity gradient, yields a reasonable estimate of the marine boundary layer height.

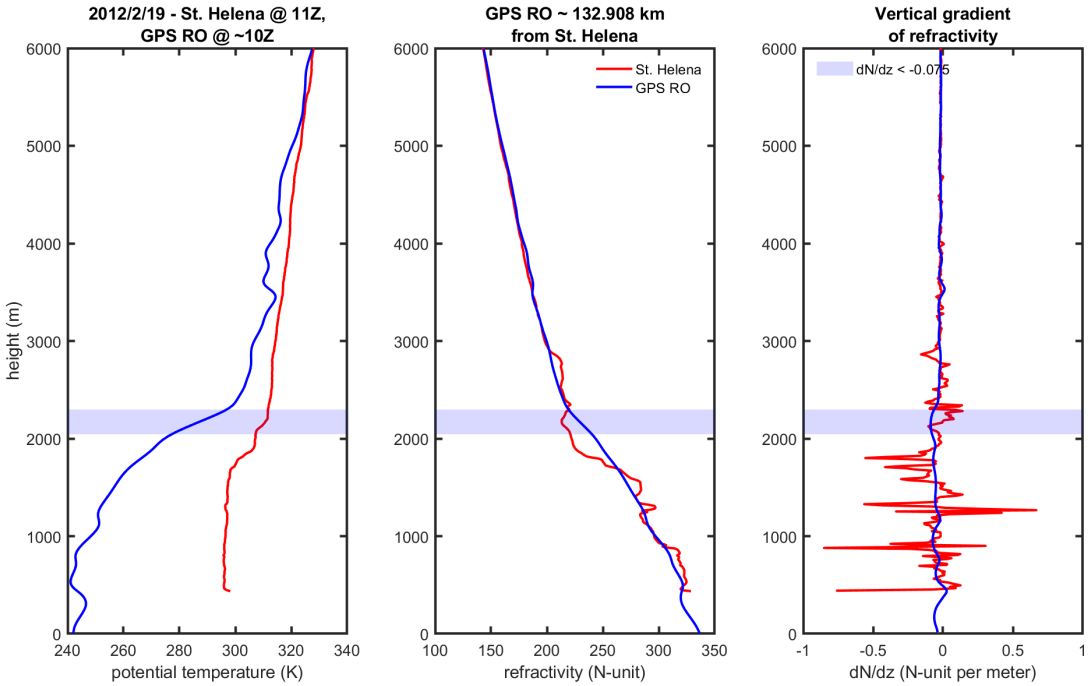


Figure B3: Comparison of a GPS radio occultation (RO) from the MetOp-A LEO satellite and a high-resolution sounding from St. Helena Island. Blue shaded area represents where refractivity from GPS RO is less than -0.075 N-unit per meter. Difference between observed potential temperature and dry potential temperature from the GPS radio occultation increases drastically at the top of the marine boundary layer.

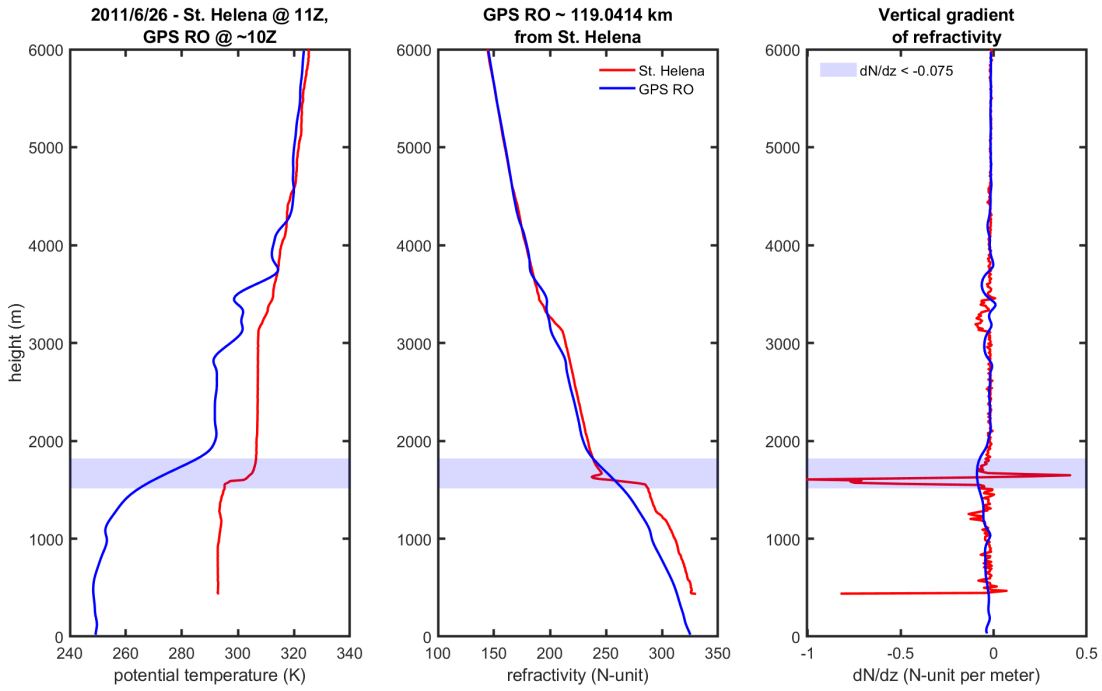


Figure B4: As in Figure B3, but for a different day (26 June 2011).

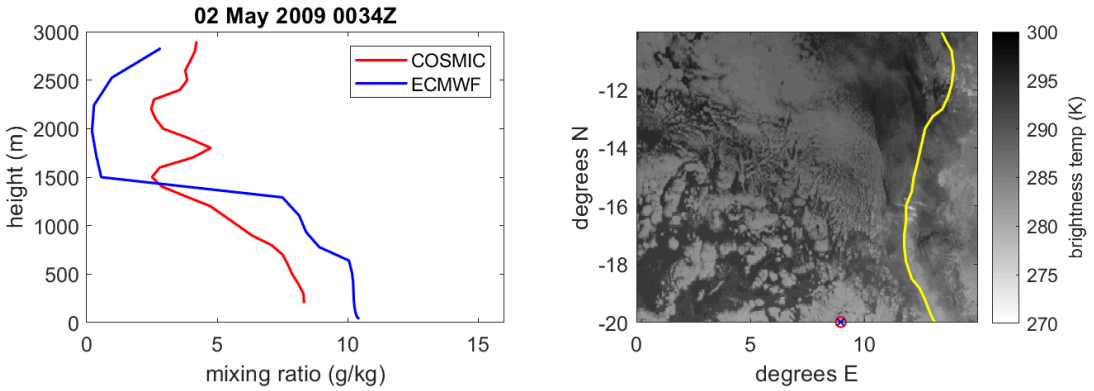


Figure B5: Mixing ratio from a GPS radio occultation from a COSMIC LEO satellite collocated with ECMWF reanalysis. The right panel plots IR brightness temperature, with the yellow line indicating the African coast. The red circle and blue x denote the location of the COSMIC radio occultation. Gridded ECMWF reanalysis is interpolated to this point.

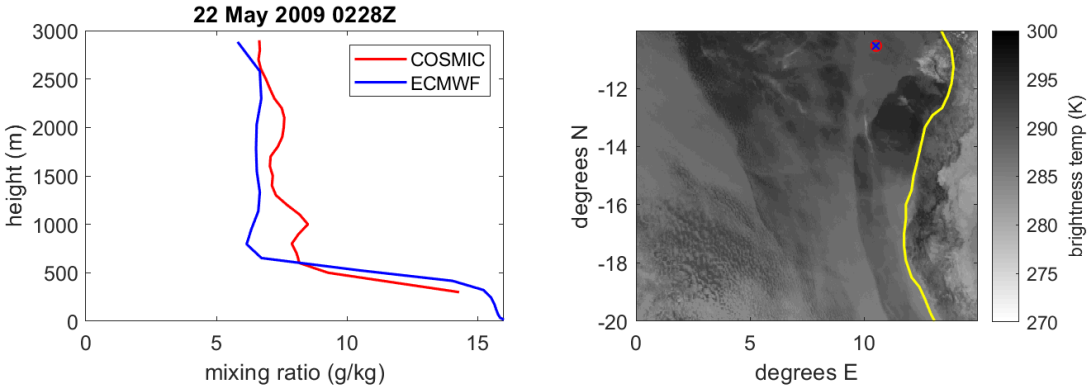


Figure B6: As in Figure B5, but for a different time and location in the southeast Atlantic.

APPENDIX C – Coastal highlands temperature analysis

We looked at minimum temperatures and temperature change from sunset over the coastal highlands in model reanalysis as a proxy to strength of the downslope flow. A stronger downslope flow may be more effective at generating atmospheric gravity waves or may excite gravity waves with a larger amplitude that could erode cloud more efficiently.

We examined temperatures from ERA5 reanalysis as well as the Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA2). MERRA2 mean temperatures are colder in January than ERA5 (Figure C1, C2). There is separation between January and May minimum temperatures in ERA5 and later in the month of May in MERRA2, but there is no clear separation for days with and without a cloud-eroding boundary (i.e. Yes and No days) in May. This is consistent with analyses of temperature change from sunset to local midnight and the minimum temperature (Figures C3 to C6). ERA5 shows substantial difference in the temperature change from sunset between January and May, but there are many instances of Yes days with relatively small temperature depression and No days with relatively large temperature depression. The results from MERRA2 further suggest a lack of dependence of cloud-eroding boundaries on the temperature depression from sunset. If the strength of the downslope flow is important for determining the existence of cloud-eroding boundaries, it is not indicated by this proxy.

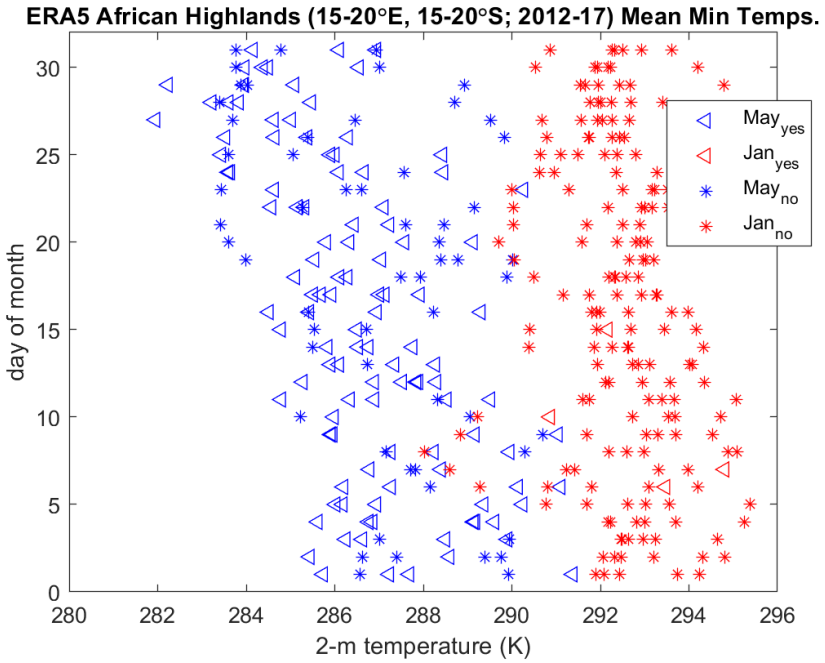


Figure C1: Average ERA5 minimum temperature over the coastal highlands of Africa (15° to 20°E, 15° to 20°S) in January (red) and May (blue). Days with a cloud-eroding boundary are marked with triangles, with asterisks marking days without a cloud-eroding boundary.

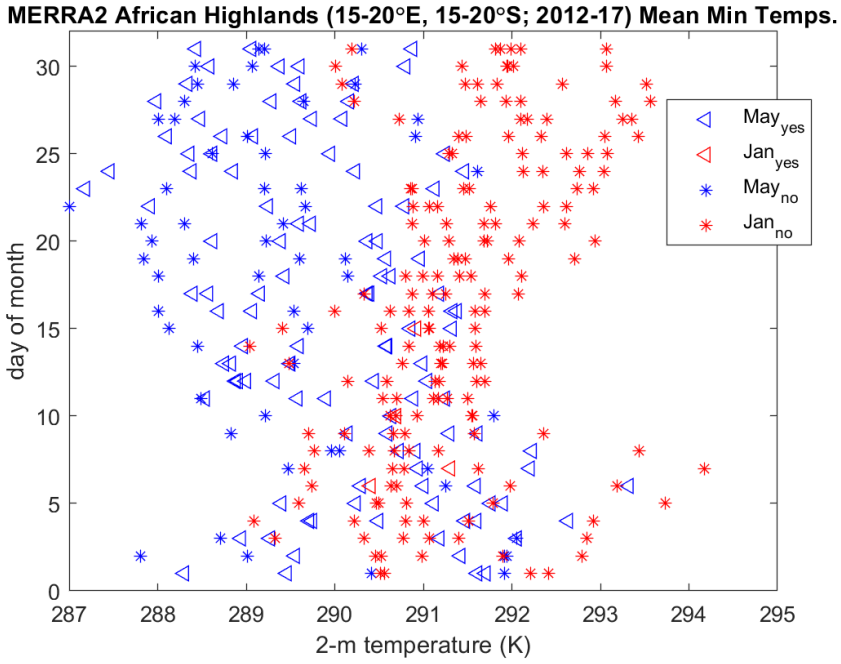


Figure C2: As in Figure C1 but for MERRA2 data.

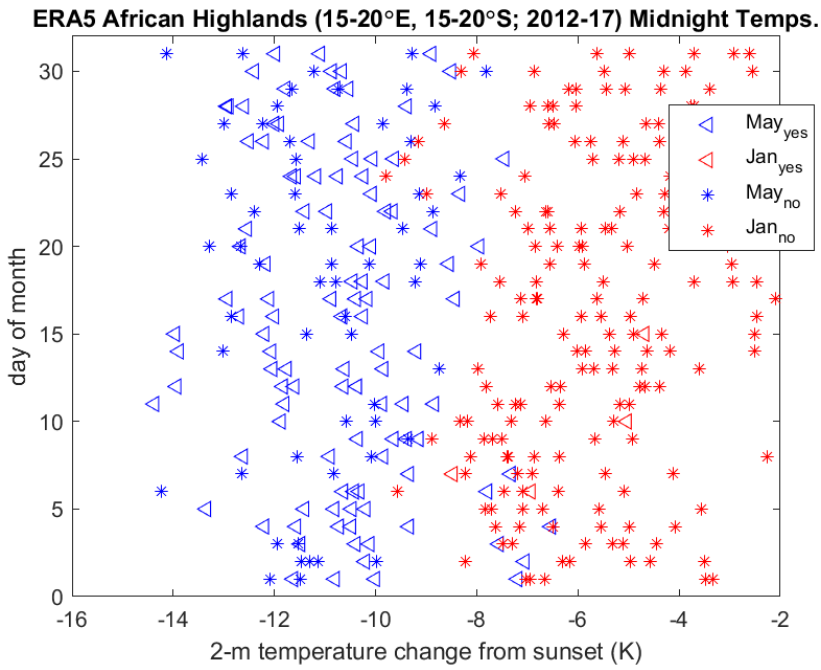


Figure C3: Average ERA5 temperature depression from sunset to local midnight over the coastal highlands of Africa (15° to 20°E, 15° to 20°S) in January (red) and May (blue). Yes days are marked with triangles and No days are marked with asterisks.

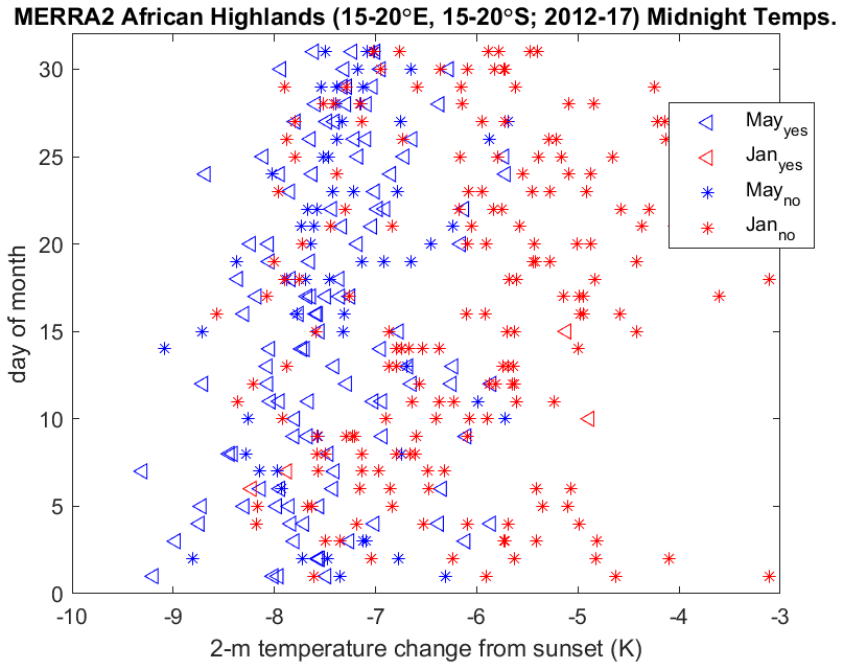


Figure C4: As in Figure C3 but for MERRA2 data.

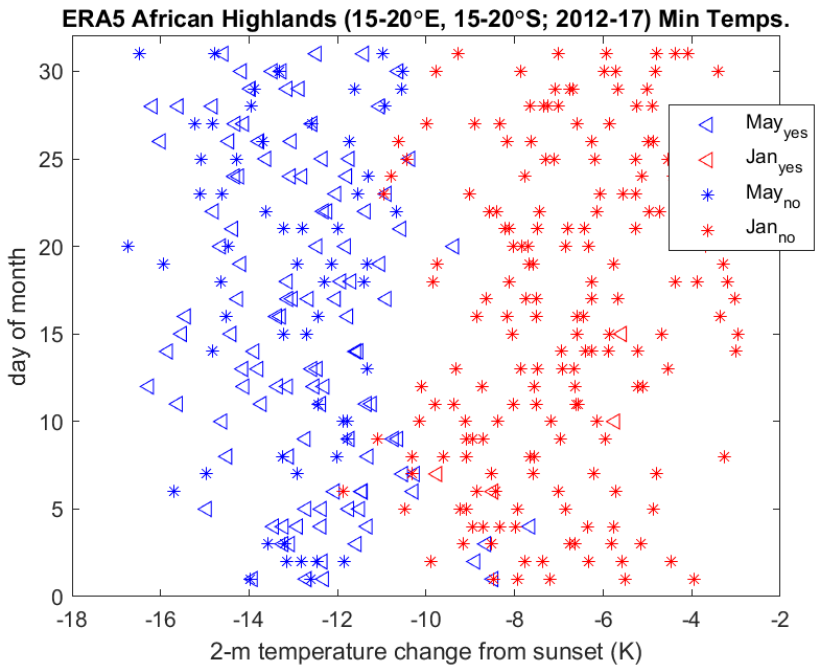


Figure C5: Average ERA5 temperature depression from sunset to the minimum temperature over the coastal highlands of Africa (15° to 20°E, 15° to 20°S) in January (red) and May (blue). Yes days are marked with triangles and No days are marked with asterisks.

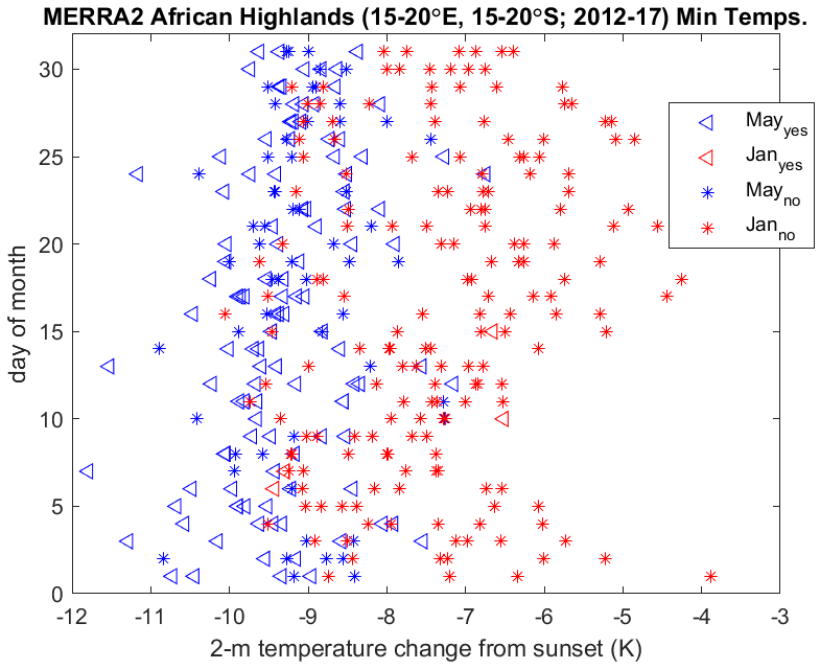


Figure C6: As in Figure C5 but for MERRA2 data.

APPENDIX D – Deriving cloud thickness and peak liquid water content from MODIS

In this appendix, we document the methodology (including equations) used to derive the liquid water content (LWC) profile from MODIS liquid water path and cloud top temperature and pressure retrievals using a varying adiabatic change of LWC with height in cloud. We use an iterative method to find the LWC profile which integrates to the MODIS liquid water path retrieval. The vertical extent of the derived LWC profile is the cloud thickness.

Starting at cloud top, we calculate the saturation mixing ratio using MODIS cloud top temperature and pressure according to the simplified Clausius-Clapeyron relationship:

$$e_s = 6.112 * \exp\left(17.67 * \frac{T}{T + 243.5}\right) \tag{D1}$$

$$w_s = 621.97 \frac{e_s}{p - e_s} \tag{D2}$$

where e_s is the saturation vapor pressure (in hPa), T is temperature (in °C), p is pressure (in hPa), and w_s is the saturation mixing ratio in g/kg. We also calculate the moist adiabatic lapse rate Γ_m at cloud top:

$$\Gamma_m = \frac{g}{c_p}\left(\frac{1 + \frac{L_v w_s}{R_d T}}{1 + \frac{L_v^2 w_s}{R_v c_p T^2}}\right) \tag{D3}$$

where g is the gravitational acceleration (~9.81 m/s), c_p is specific heat at constant pressure (in J kg^{-1} K^{-1}), L_v is the latent heat of vaporization (in J kg^{-1}), R_d is the dry air gas constant (in J kg^{-1} K^{-1}), R_v is the water vapor gas constant (in J kg^{-1} K^{-1}), and T is temperature (in K). For this equation (D3), w_s is converted to kg/kg.

We now begin to iterate down into the cloud to find the cloud base which yields the LWC profile with liquid water path equal to that retrieved from MODIS. In our algorithm, we move down through the cloud 2 meters at a time. The temperature at the new level (2 meters below the

previous level) is computed using Γ_m from the previous level. Pressure at the new level (p_2) is then calculated with the hypsometric equation:

$$p_2 = p_1 * \exp\left(\frac{g\Delta z}{R_d T}\right) \quad \text{(D4)}$$

where p_1 is pressure at the previous level (in hPa), Δz is the layer thickness (in our algorithm, 2 meters), and T is temperature (in K). We now have temperature and pressure at the new level, so we can calculate the saturation mixing ratio using Equation D1 and D2, along with the density at the new level using the ideal gas law. Density is required in order to compute liquid water path (g/m^2) from a LWC profile in g/kg.

The new level is the presumed cloud base. As such, the LWC profile can be computed by subtracting the saturation mixing ratio at each level from the saturation mixing ratio at cloud base (Cotton and Yuter, 2008), yielding a profile of increasing LWC with height in the cloud. We compute the liquid water path of this estimated LWC profile of the cloud to compare with the MODIS liquid water path (LWP) retrieval:

$$\text{LWP}_{est} = \sum_i \left(w_{s,base} - w_{s,i}\right) \rho_i \Delta z \quad \text{(D5)}$$

where i is each 2-meter level in the estimated profile, $w_{s,base}$ is saturation mixing ratio at the base of the estimated profile (in g/kg), $w_{s,i}$ is the saturation mixing ratio at each level i (in g/kg), ρ_i is the density at each level i (in kg m^{-3}), and Δz = 2 meters. If LWP_{est} is less than the MODIS-retrieved LWP, we continue the iteration. If LWP_{est} exceeds the MODIS LWP, we interpolate between that cloud base and the previous estimated cloud base (2 meters above) to get a more exact cloud thickness and peak LWC. Note that if the estimated cloud base goes below ground according to the MODIS cloud top height, we mark that cloud thickness and peak LWC as not-a-number (i.e. missing).

APPENDIX E – SOM Node Configuration Analyses

This appendix documents the SOM error analyses for various ERA5 fields (e.g. boundary layer height, mean sea level pressure, SST, *v*-wind, 850-hPa geopotential height) and times of year in the southeast Atlantic. The range of node configurations from 4x5 to 5x6 consistently stands out as the ideal number of nodes for this region (5° to 35°S, 10°W to 20°E). Details about the choice of domain size are provided in Section 2.2.

In the 14 figures below, the quantization error (blue) and topographic error (red) are plotted. The optimal node location is determined according to the elbow criteria and where topographic error begins to sizably increase (Section 2.3). Variable abbreviations in the figure titles are as follows: boundary layer height (BLH), mean sea level pressure (MSLP), sea surface temperature (SST), meridional wind (V), and geopotential height (Z).

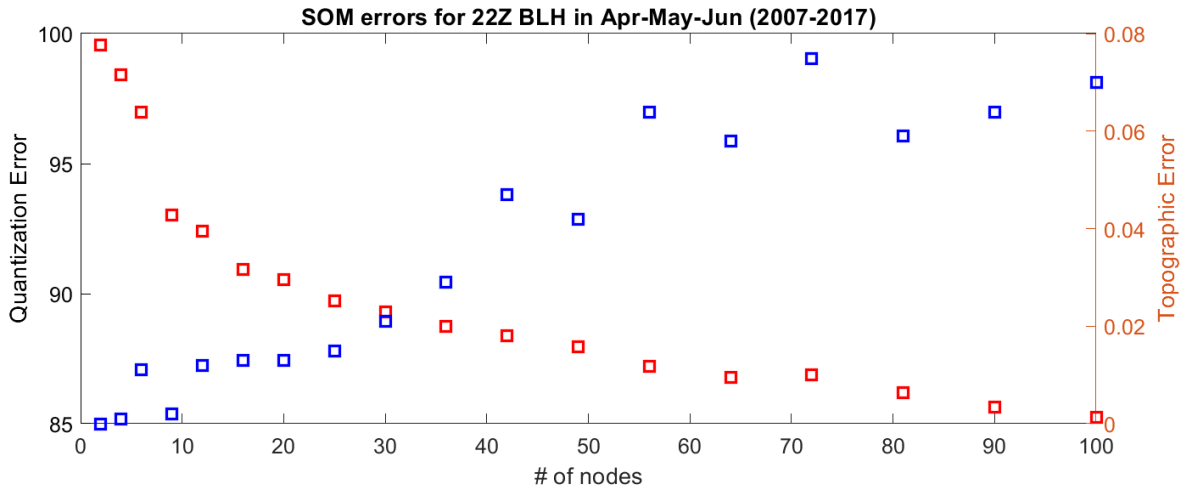


Figure E1: SOM errors for 22 UTC Apr-May-Jun boundary layer height.

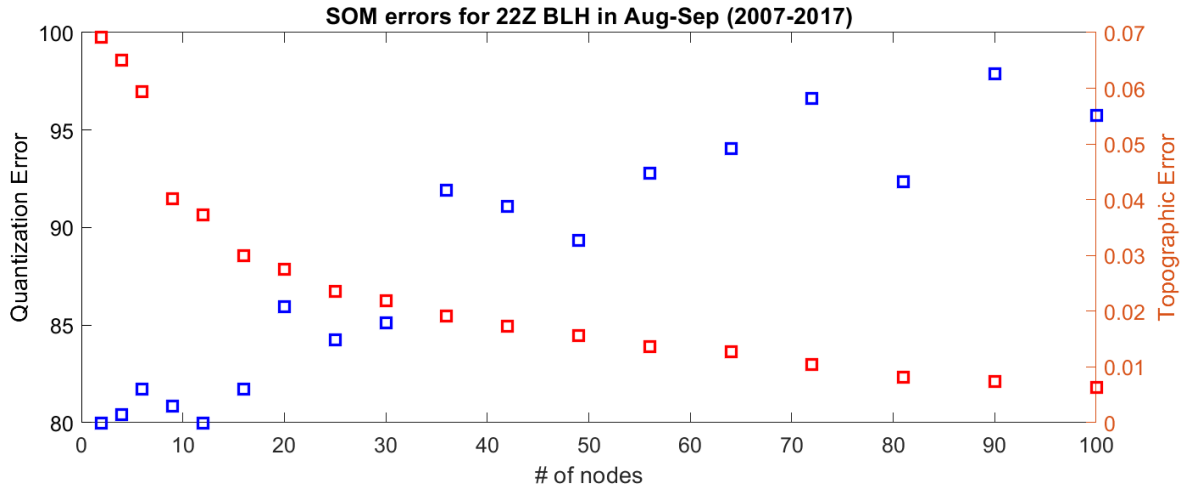


Figure E2: SOM errors for 22 UTC Aug-Sep boundary layer height.

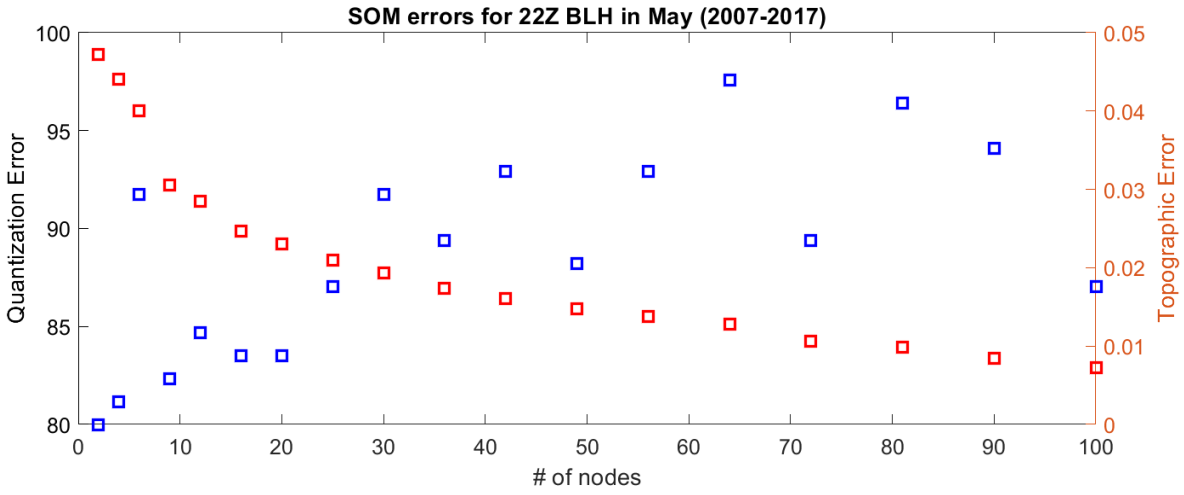


Figure E3: SOM errors for 22 UTC May boundary layer height.

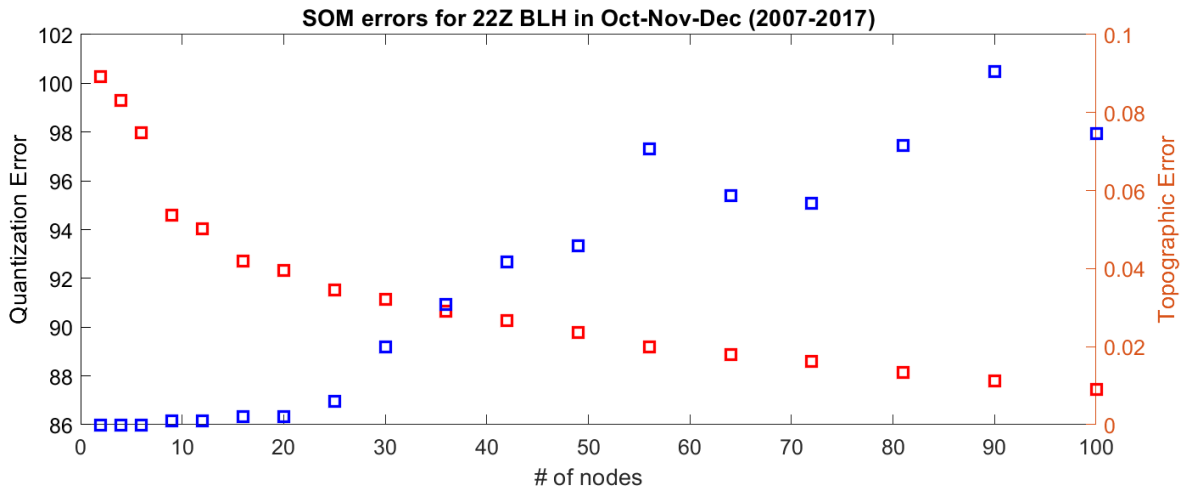


Figure E4: SOM errors for 22 UTC Oct-Nov-Dec boundary layer height.

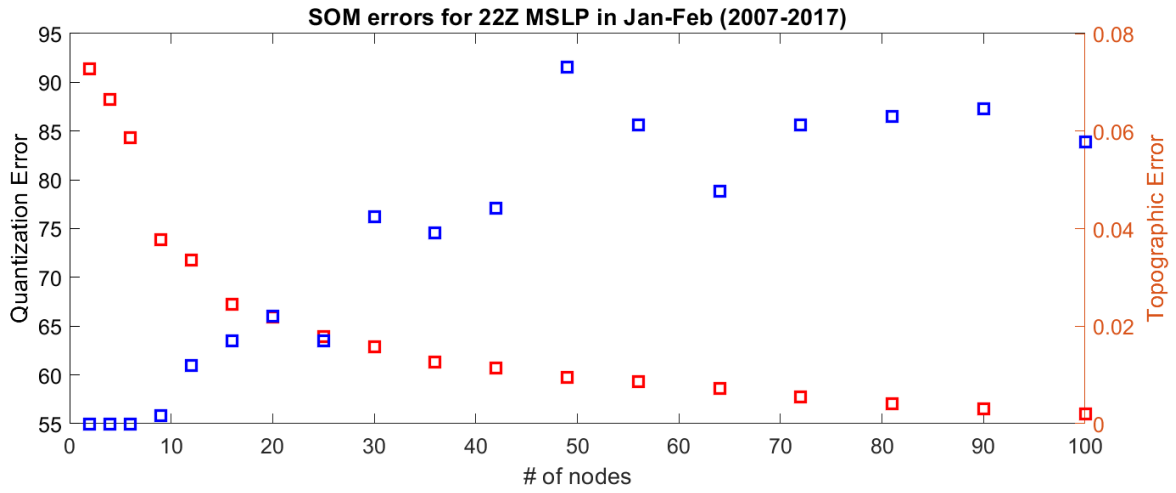


Figure E5: SOM errors for 22 UTC Jan-Feb mean sea level pressure.

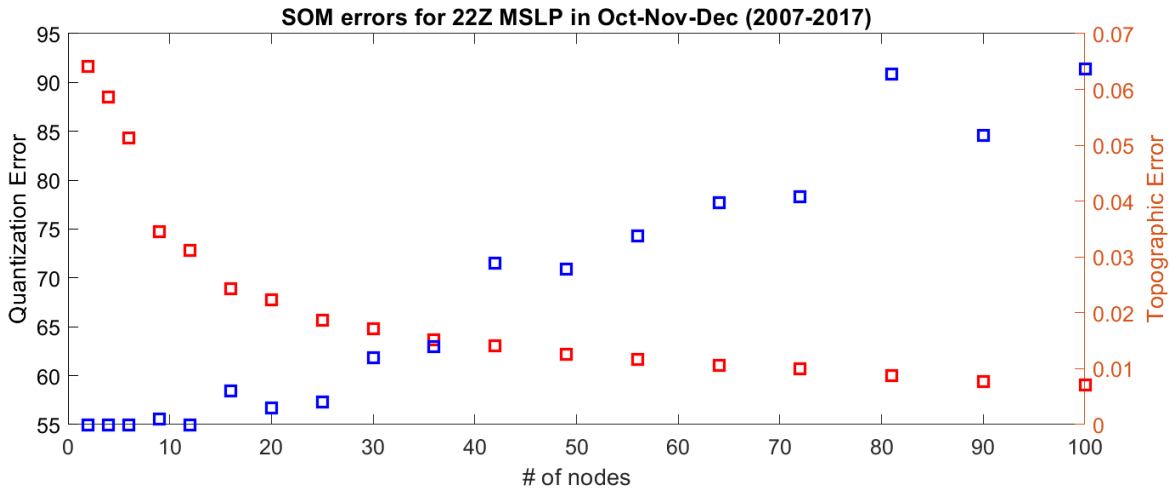


Figure E6: SOM errors for 22 UTC Oct-Nov-Dec mean sea level pressure.

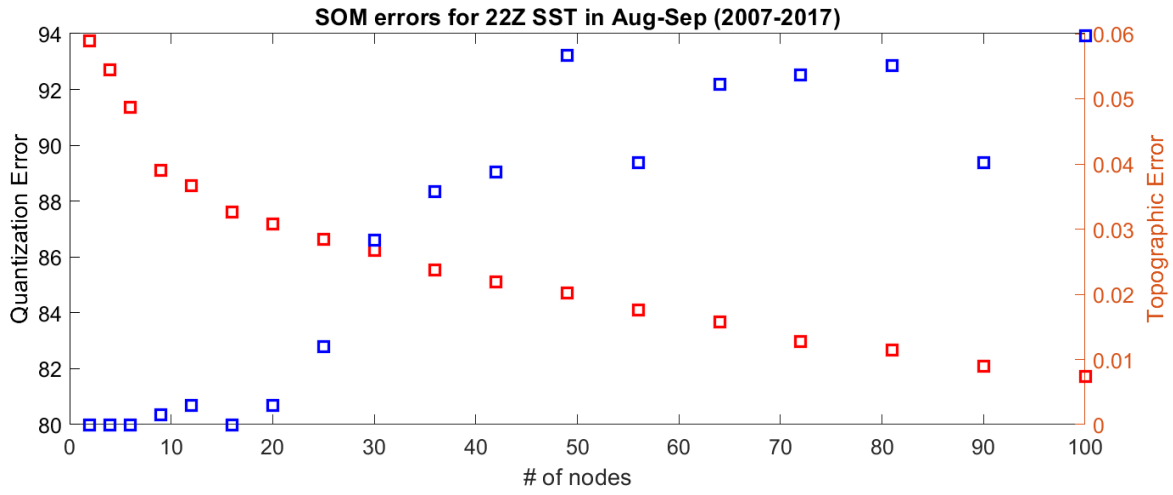


Figure E7: SOM errors for 22 UTC Aug-Sep sea surface temperature.

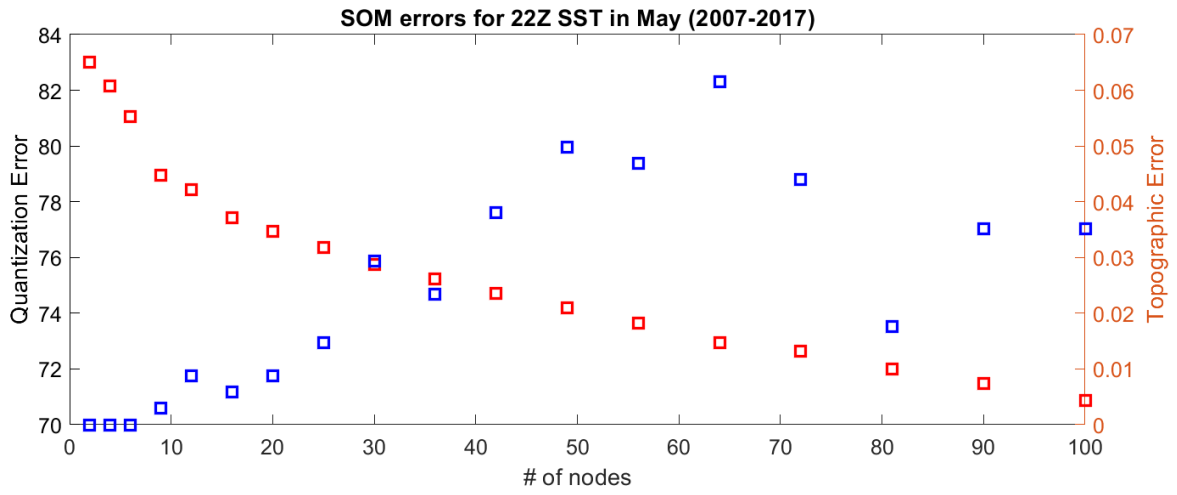


Figure E8: SOM errors for 22 UTC May sea surface temperature.

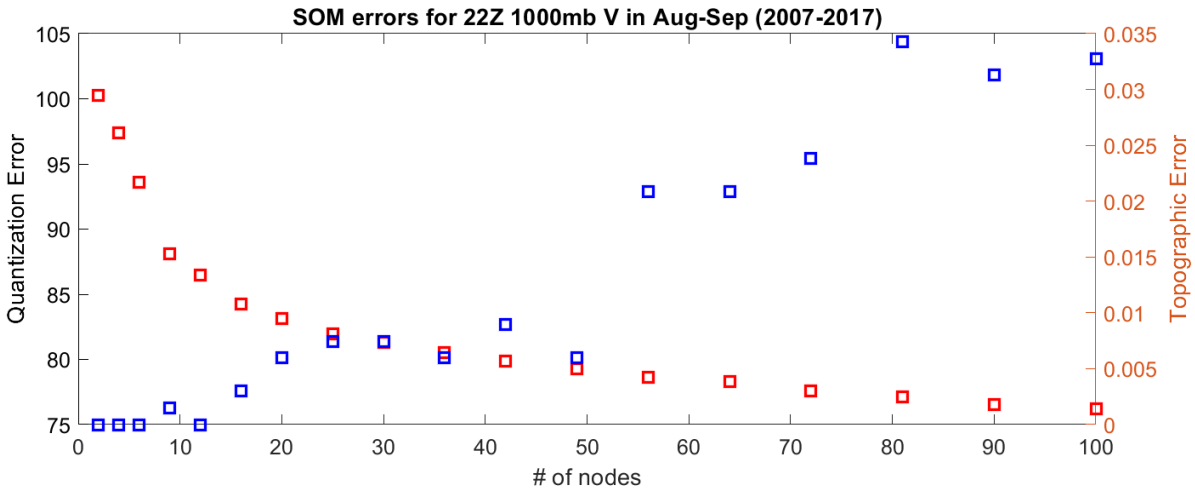


Figure E9: SOM errors for 22 UTC Aug-Sep 1000-hPa *v*-winds.

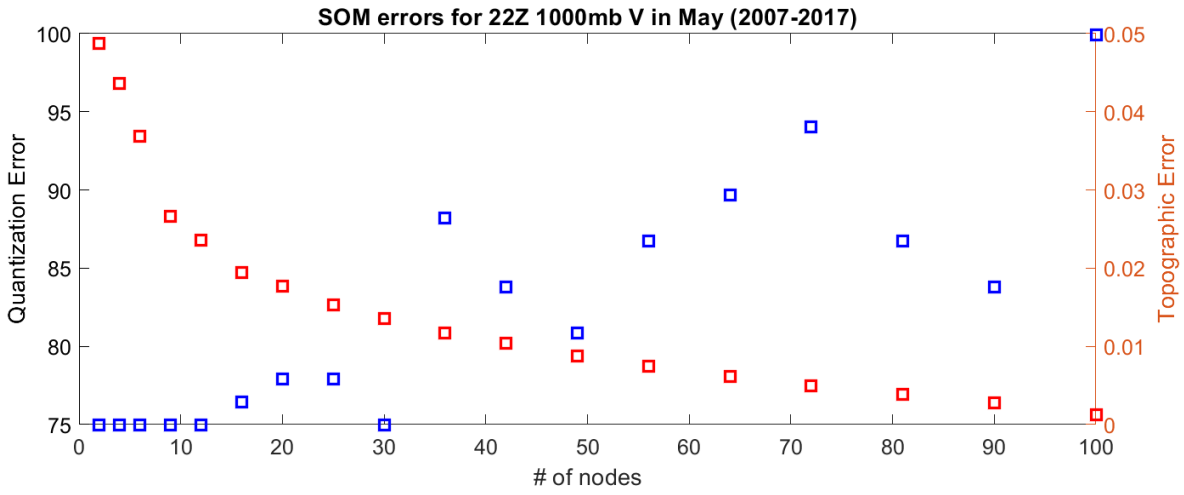


Figure E10: SOM errors for 22 UTC May 1000-hPa *v*-winds.

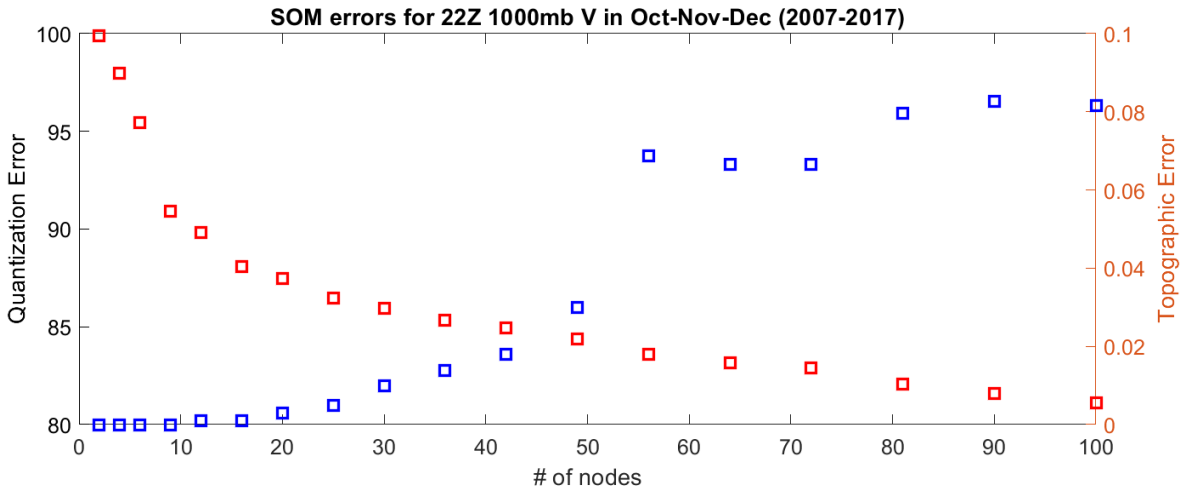


Figure E11: SOM errors for 22 UTC Oct-Nov-Dec 1000-hPa *v*-winds.

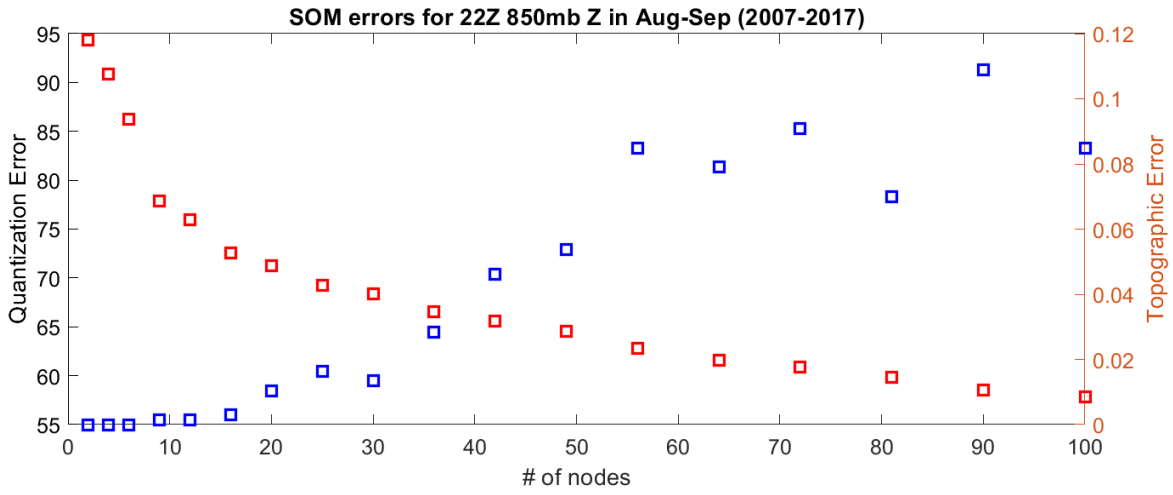


Figure E12: SOM errors for 22 UTC Aug-Sep 850-hPa geopotential height.

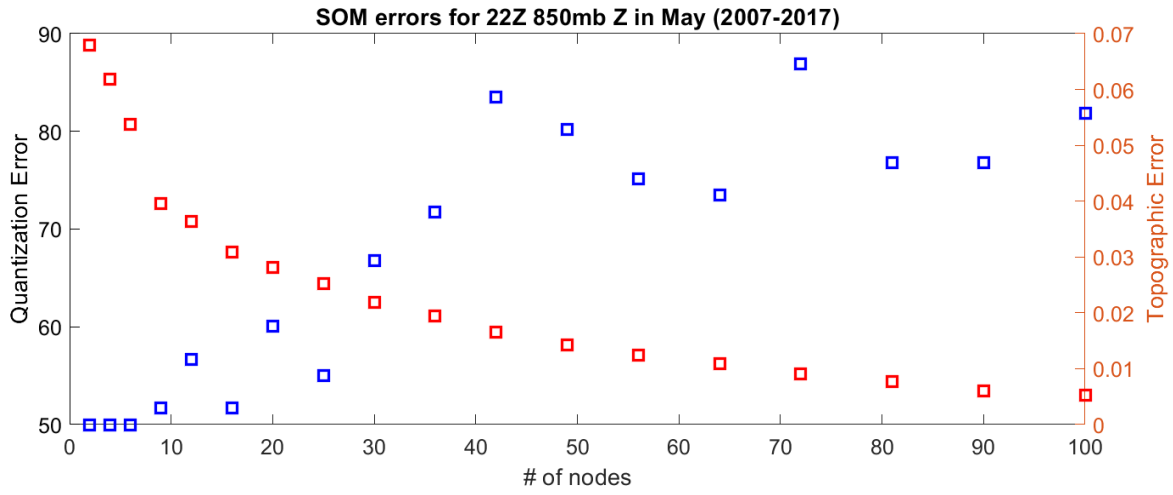


Figure E13: SOM errors for 22 UTC May 850-hPa geopotential height.

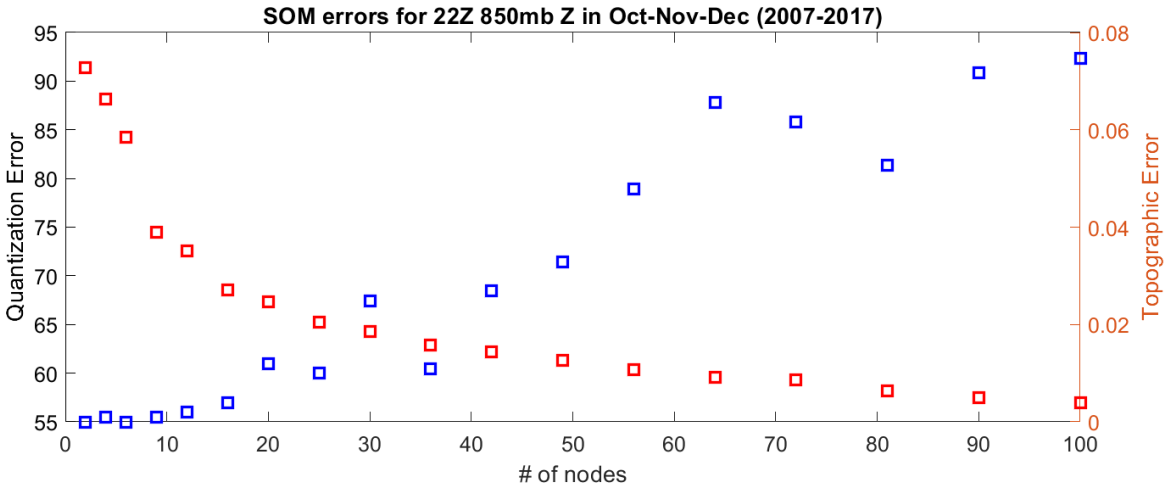


Figure E14: SOM errors for 22 UTC Oct-Nov-Dec 850-hPa geopotential height.

APPENDIX F – Supplemental MODIS Terra Figures

This appendix includes a collection of figures from analysis of satellite data to supplement the results in Section 3.1. The main results are highlighted well by the figures in the thesis, while these serve to paint the complete picture and fully document the investigation. All figures herein were computed using MODIS Terra (10:30 am LT) Level-3 Collection 6.1 1° × 1° data from 2000 to 2018 (see Section 2.6 for details). In figures showing data on Yes and/or No days in the southeast Atlantic, data are taken from 2007 to 2017 as this is the current extent of the cloud-eroding boundary dataset (see Section 2.1 for details).

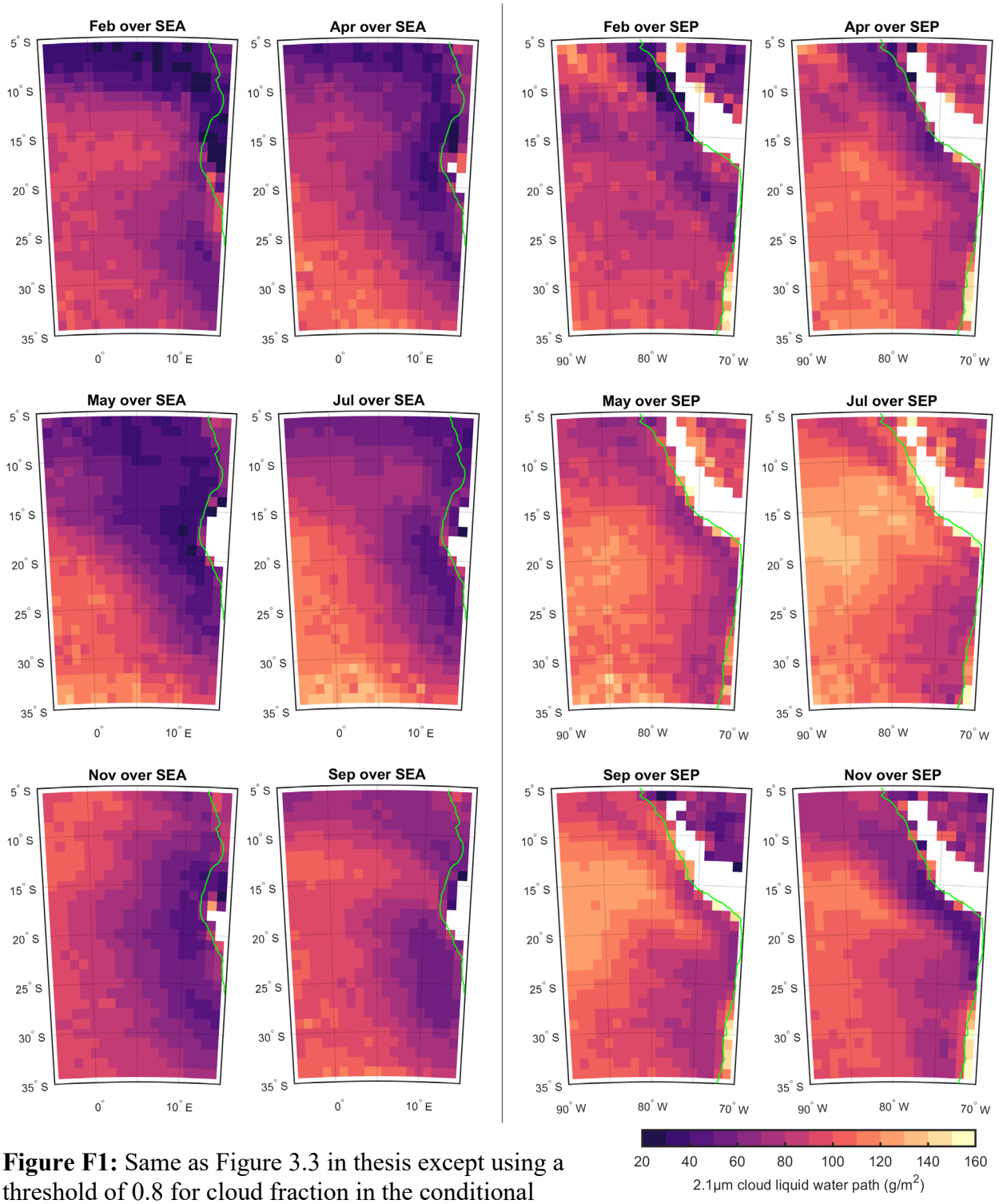


Figure F1: Same as Figure 3.3 in thesis except using a threshold of 0.8 for cloud fraction in the conditional averaging. Panels show conditional monthly mean liquid water path in the (left) southeast Atlantic and (right) southeast Pacific. Coastline is indicated by the green line.

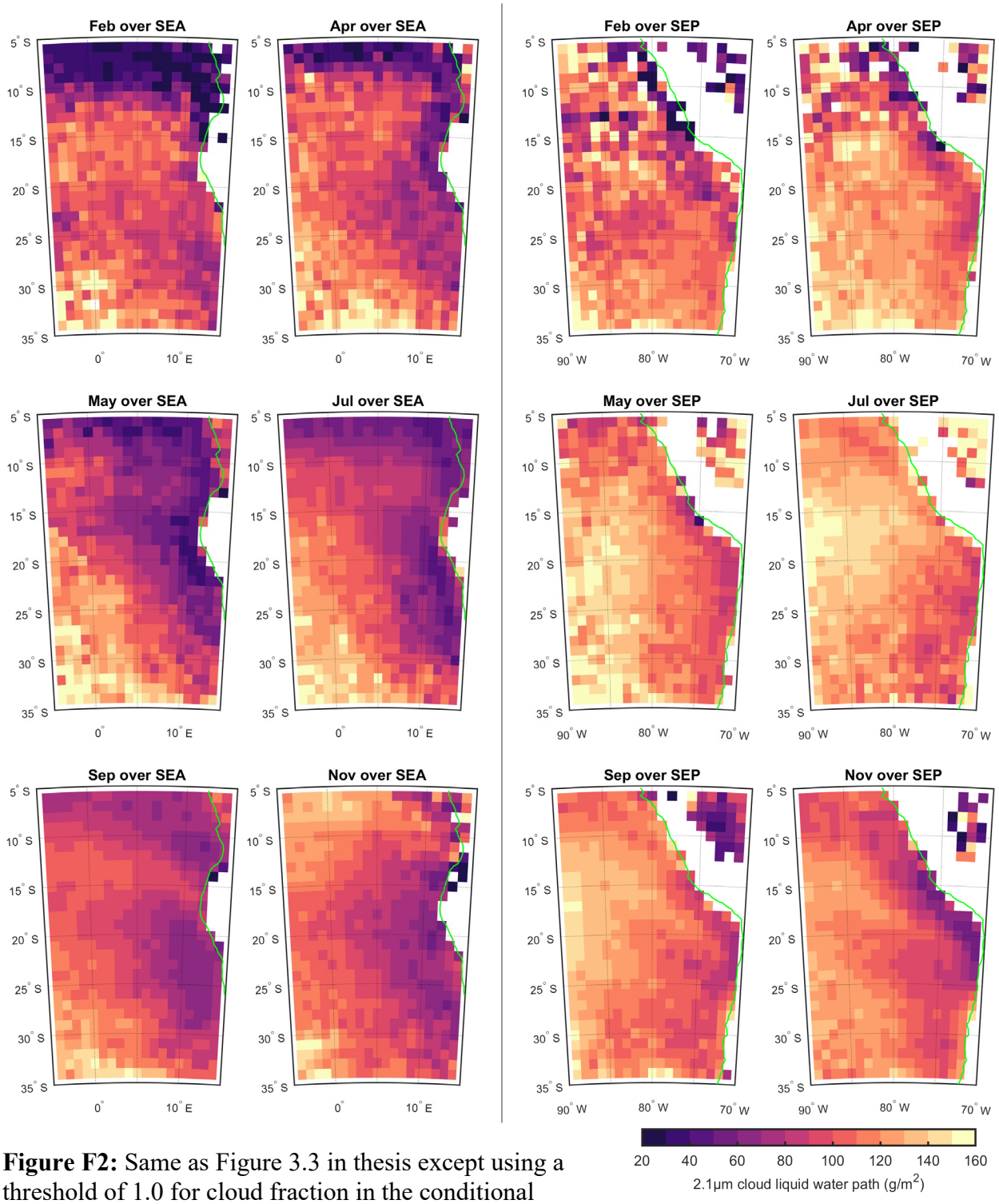


Figure F2: Same as Figure 3.3 in thesis except using a threshold of 1.0 for cloud fraction in the conditional averaging. Panels show conditional monthly mean liquid water path in the (left) southeast Atlantic and (right) southeast Pacific. Coastline is indicated by the green line.

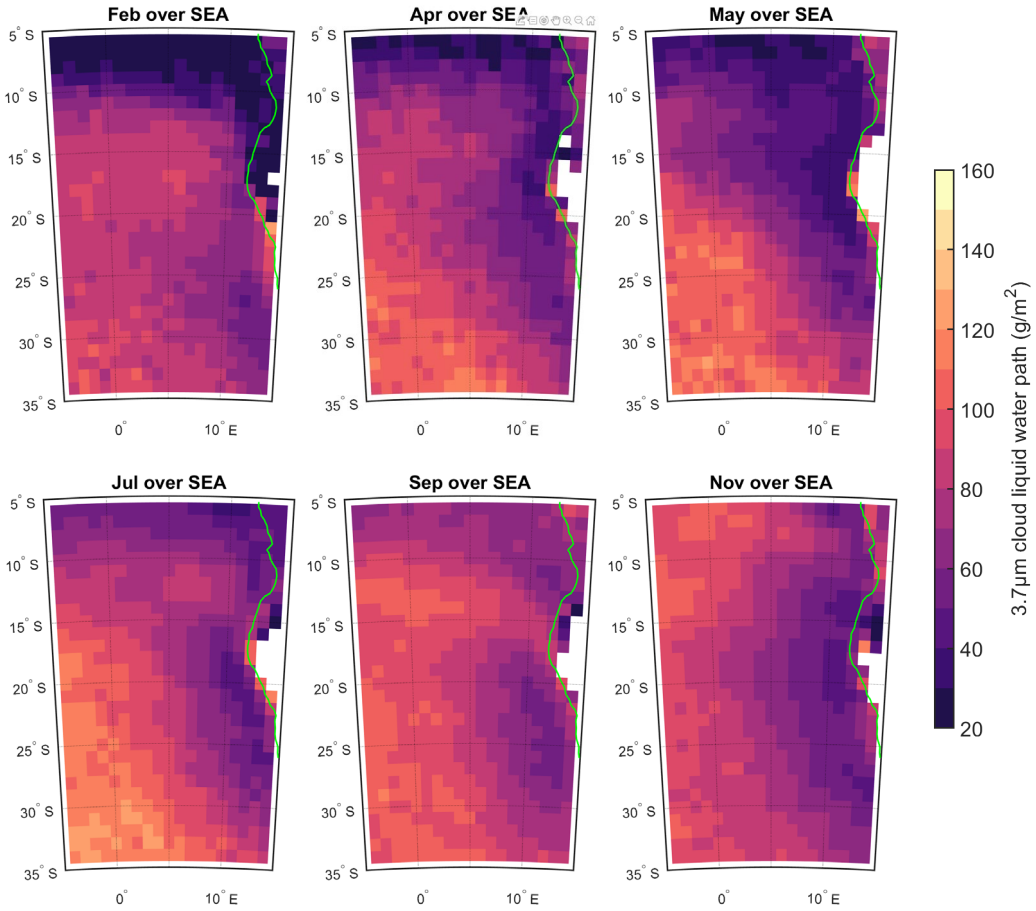
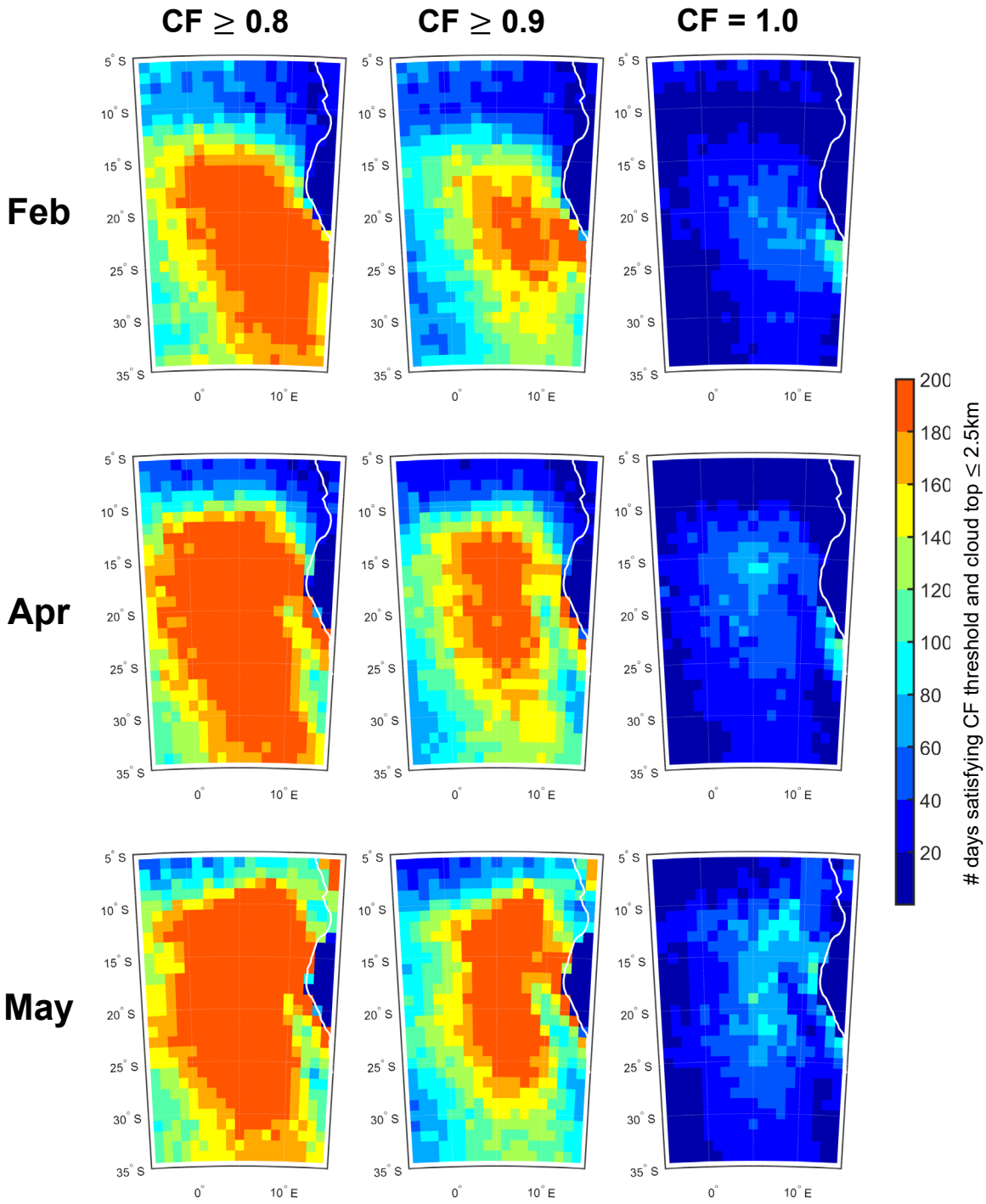
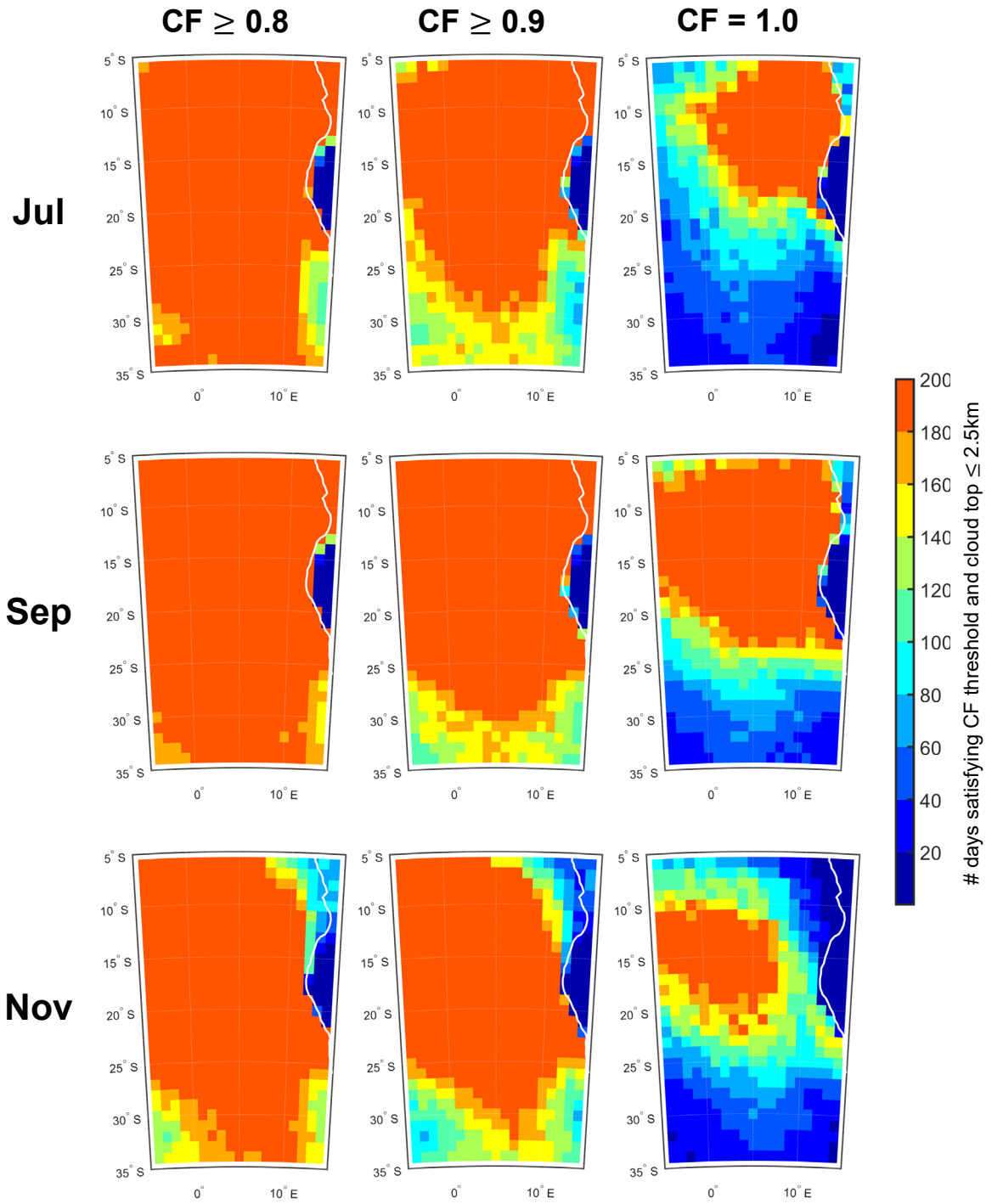


Figure F3: Monthly mean liquid water path in the southeast Atlantic from the 3.7 µm channel conditioned upon cloud fraction ≥ 0.9 and cloud top ≤ 2.5km. Coastline is indicated by the green line.

Figure F4: Monthly mean number of days satisfying the conditions for averaging of cloud top ≤ 2.5km and a certain cloud fraction threshold in the southeast Atlantic. Each column denotes a different cloud fraction threshold (0.8, 0.9, and 1) while each row is a different month. Coastline is indicated by the white line.


CF ≥ 0.8
CF ≥ 0.9
CF = 1.0
Feb
Apr
May
5° S
10° S
15° S
20° S
25° S
30° S
35° S
0°
10° E
200
180
160
140
120
100
80
60
40
20
days satisfying CF threshold and cloud top ≤ 2.5km


CF ≥ 0.8
CF ≥ 0.9
CF = 1.0
Jul
Sep
Nov
5° S
10° S
15° S
20° S
25° S
30° S
35° S
0°
10° E
200
180
160
140
120
100
80
60
40
20
days satisfying CF threshold and cloud top ≤ 2.5km

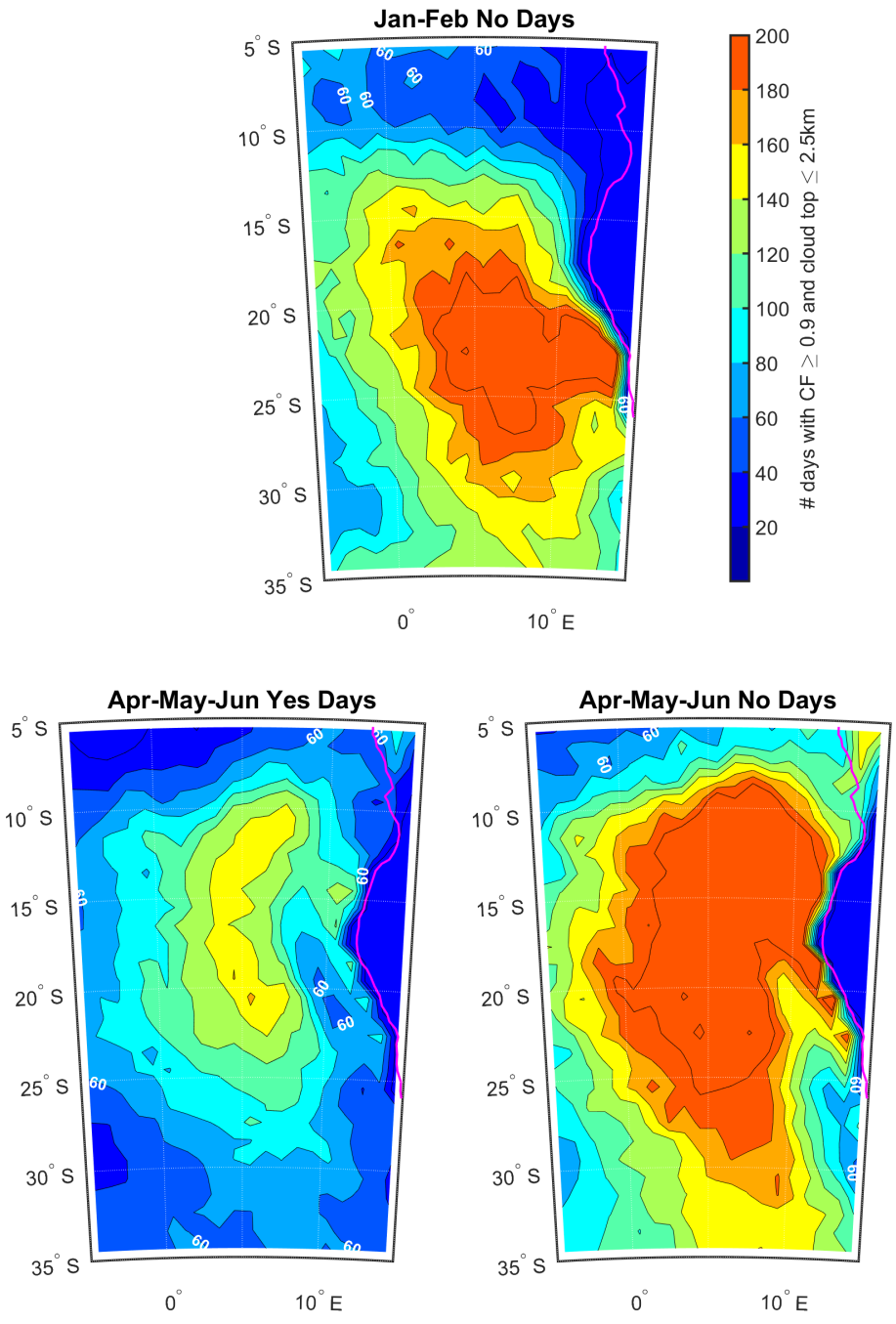


Figure F5: Mean number of days satisfying the conditions for averaging of cloud top ≤ 2.5km and cloud fraction ≥ 0.9 in the southeast Atlantic on No days in Jan-Feb and Apr-May-Jun and Yes days in Apr-May-Jun. The contour of 60 days satisfying the conditions is annotated, with contours every 20 days. Coastline is indicated by the pink line.

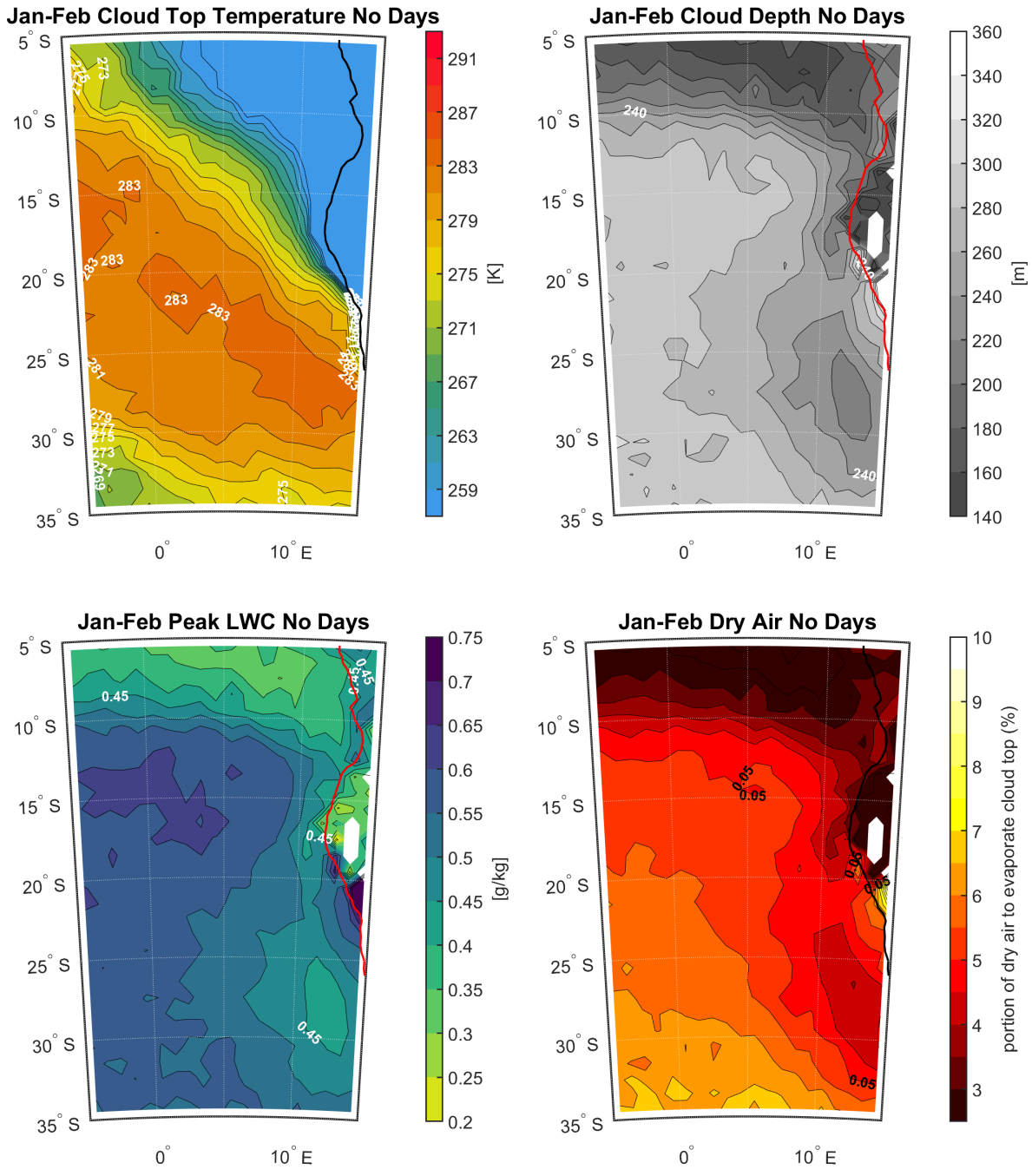


Figure F6: MODIS cloud top temperature and MODIS-derived cloud thickness, peak liquid water content (LWC), and fractional entrainment required to evaporate the peak LWC on No days in Jan-Feb. Contours are every 2 K for cloud top temperature, every 20 meters for cloud depth, every 0.05 g/kg for peak LWC, and every 0.5% for fractional entrainment. Coastline is indicated by the black or red line.

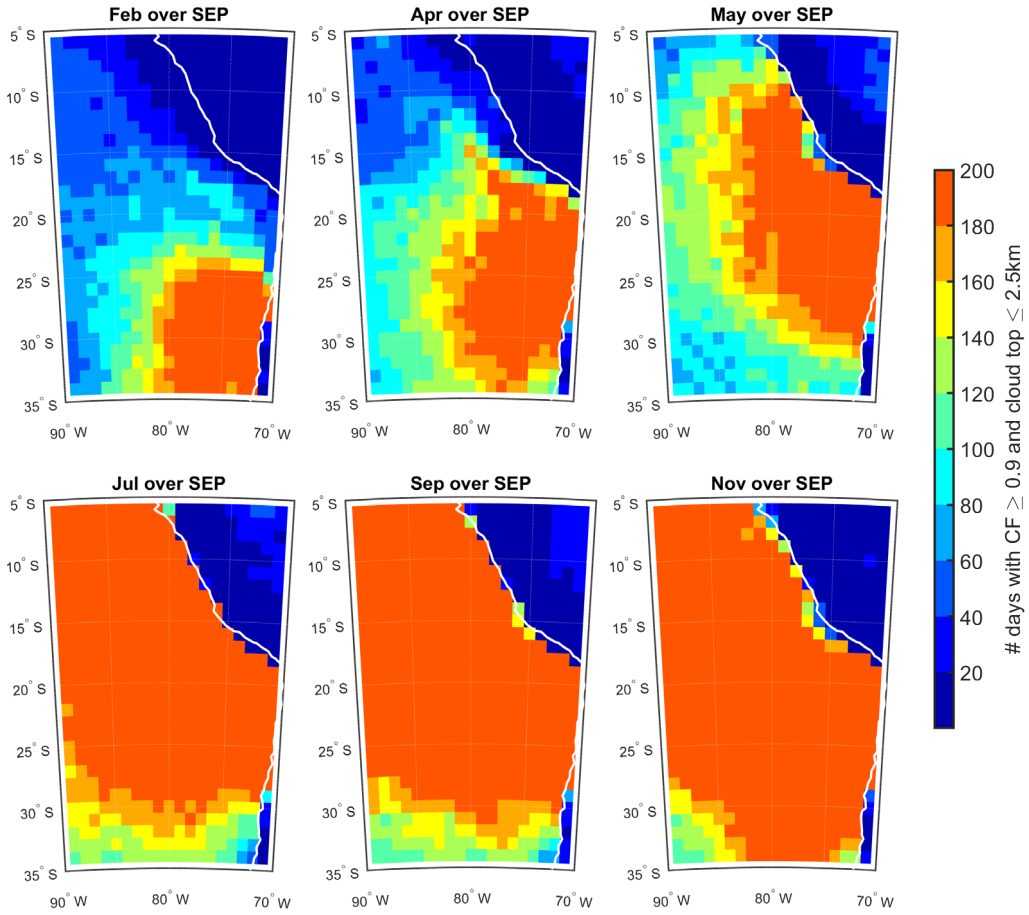


Figure F7: Monthly mean number of days satisfying the conditions for averaging of cloud top ≤ 2.5km and cloud fraction ≥ 0.9 in the southeast Pacific. Coastline is indicated by the white line.

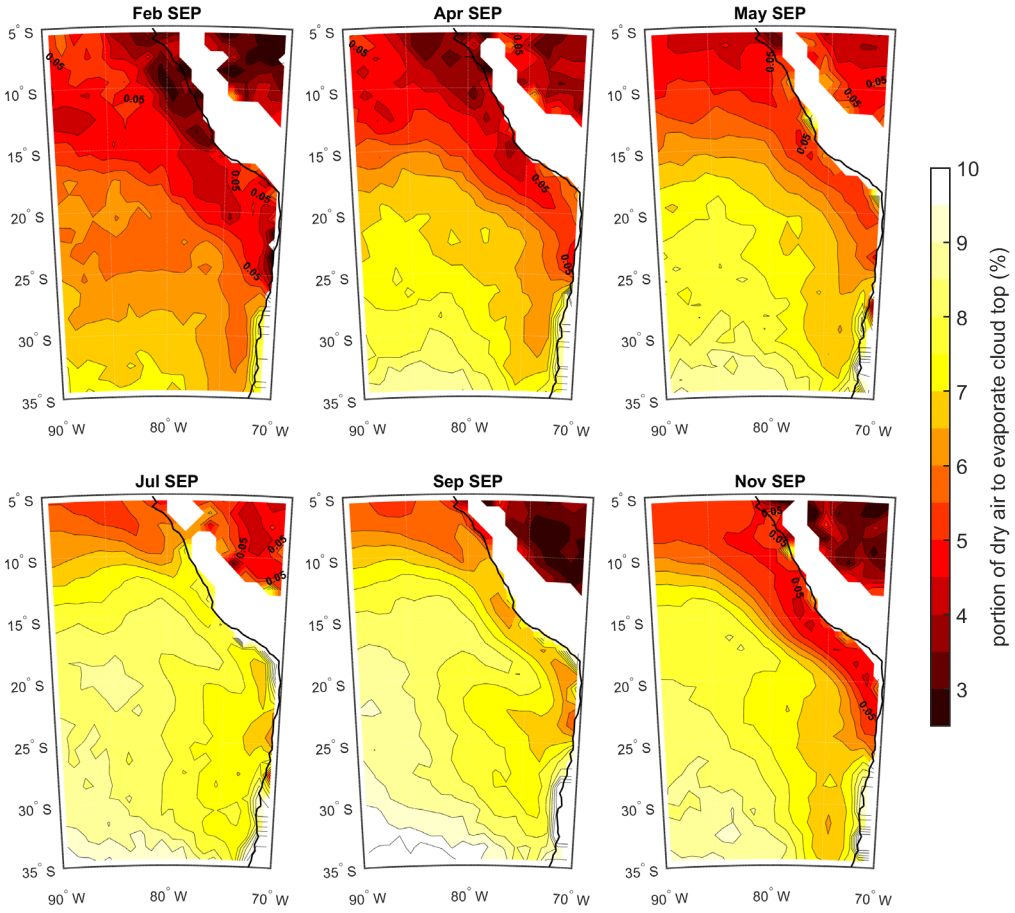


Figure F8: Monthly mean MODIS-derived fractional entrainment required to evaporate liquid water at cloud top in the southeast Pacific. Contours are every 0.5% with the 5% fractional entrainment contour annotated. Coastline is indicated by the black line.

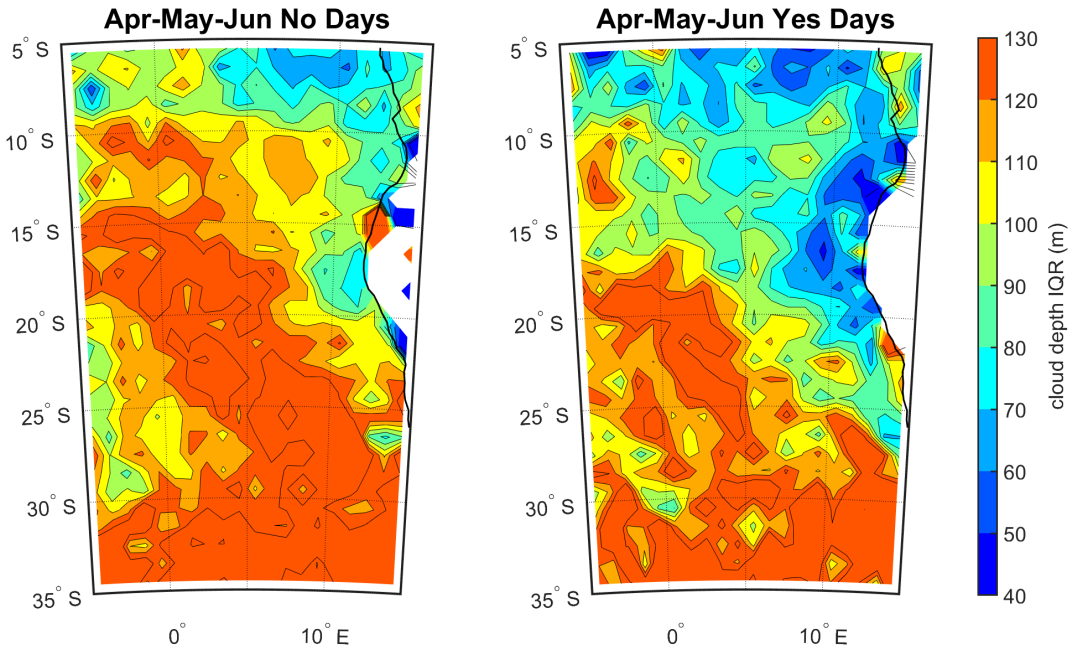


Figure F9: Inner-quartile range of MODIS-derived cloud thickness over the southeast Atlantic separated by Yes and No days in Apr-May-Jun. Contours are every 10 meters.

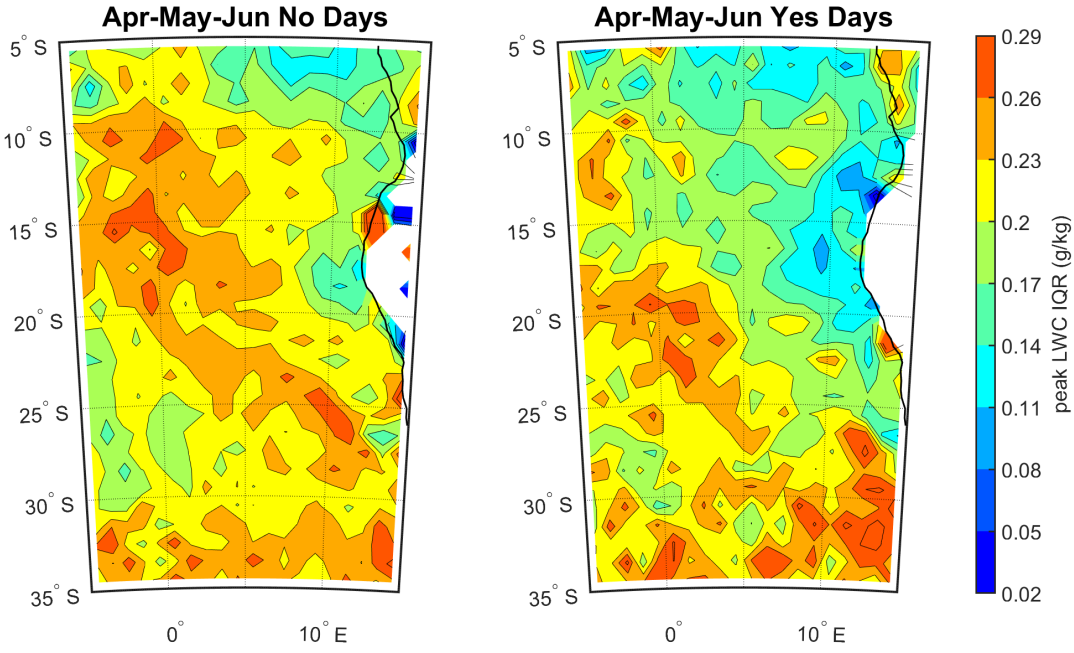


Figure F10: Inner-quartile range of MODIS-derived peak liquid water content over the southeast Atlantic separated by Yes and No days in Apr-May-Jun. Contours are every 0.03 g/kg

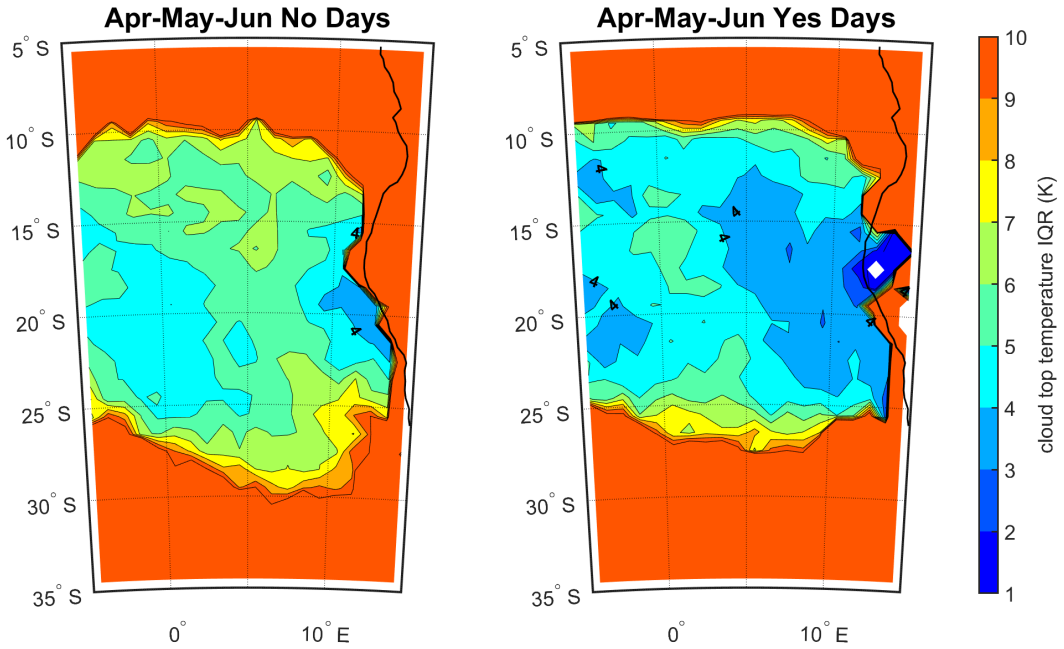


Figure F11: Inner-quartile range of MODIS cloud top temperature over the southeast Atlantic separated by Yes and No days in Apr-May-Jun. Contours are every 10 meters.

APPENDIX G – Supplemental SOM Figures

A selection of SOM figures is provided in this appendix for perusal. Specifically, we include figures showing the resulting nodes after training on normalized anomalies (G1–2), SOMs trained on boundary layer height during other time periods in the southeast Atlantic (G3–5), variables projected onto the Apr-May-Jun SOM of boundary layer height anomalies (G6–9) and the full boundary layer height field (G10–13), and an assortment of the other variables we tested with SOMs that did not exhibit relevance for cloud-eroding boundaries as defined by our criteria (Section 2.5). We also include projections of boundary layer height on the SOMs of sea surface temperature and v-wind at 975 hPa for comparison (Figures G15/16 and G18/19).

As in Section 3, projections onto the SOMs are averages of the best-matching fields for each node. All figures shown in this appendix are from SOMs trained on data at 22 UTC in the southeast Atlantic. Counts of matching Yes, No, and Maybe days using the cloud-eroding boundary dataset (Section 2.1) are annotated to the right of each node. Percentages represent the ratio of Yes, No, or Maybe days matching that node out of the total number of Yes, No, or Maybe days in the time period. Asterisks (*) mark unlikely random high counts and minus signs (-) mark unlikely random low counts (see Section 2.5 for details).

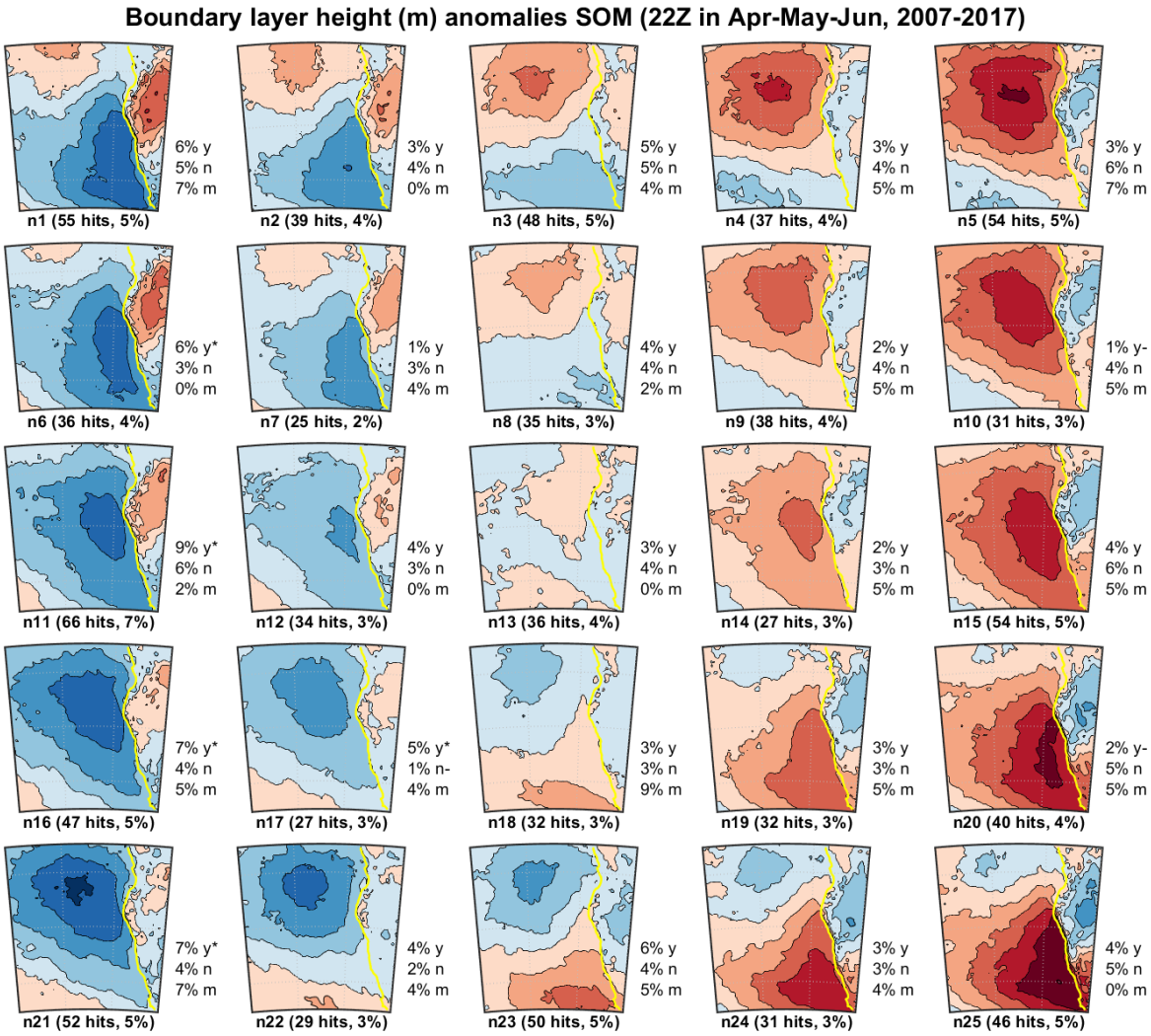


Figure G1: SOM of boundary layer height normalized anomalies in Apr-May-Jun, contoured every 0.2 units ranging from -1 to 1. Coastline is indicated by the yellow line.

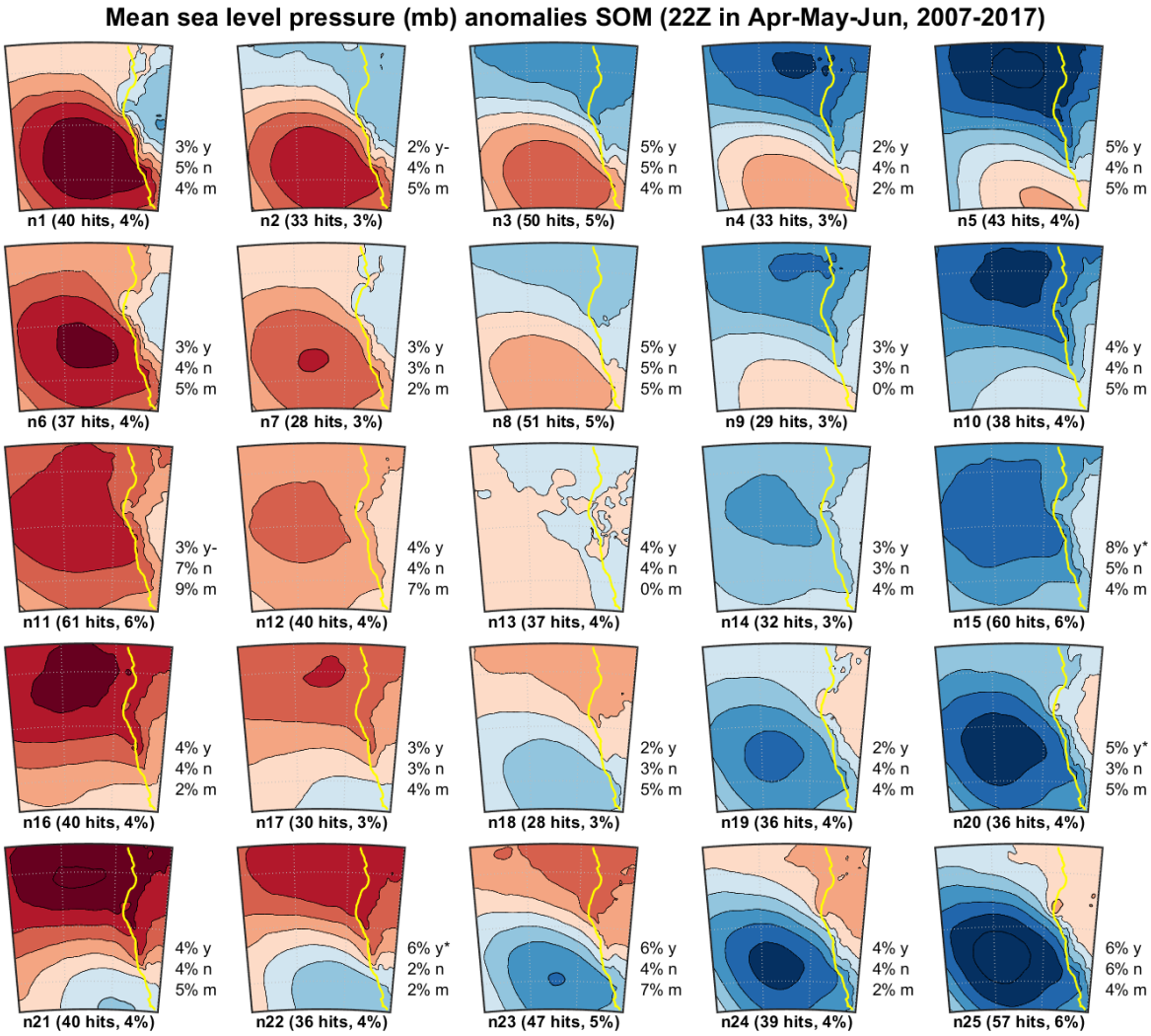


Figure G2: SOM of mean sea level pressure normalized anomalies in Apr-May-Jun, contoured every 0.2 units ranging from -1 to 1. Coastline is indicated by the yellow line.

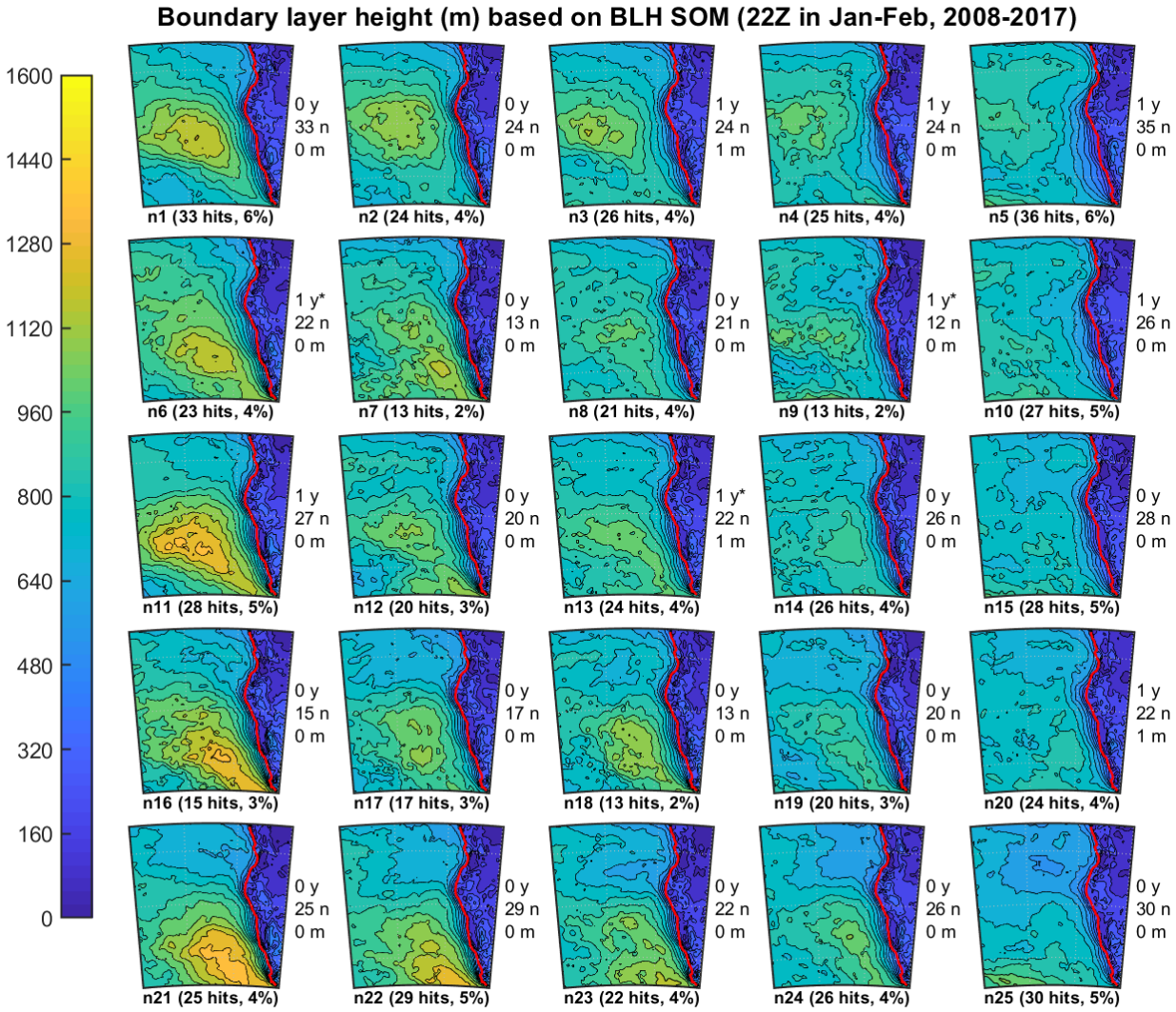


Figure G3: Boundary layer height projected onto the SOM of boundary layer height normalized anomalies in Jan-Feb, contoured every 80 meters. Coastline is indicated by the red line.

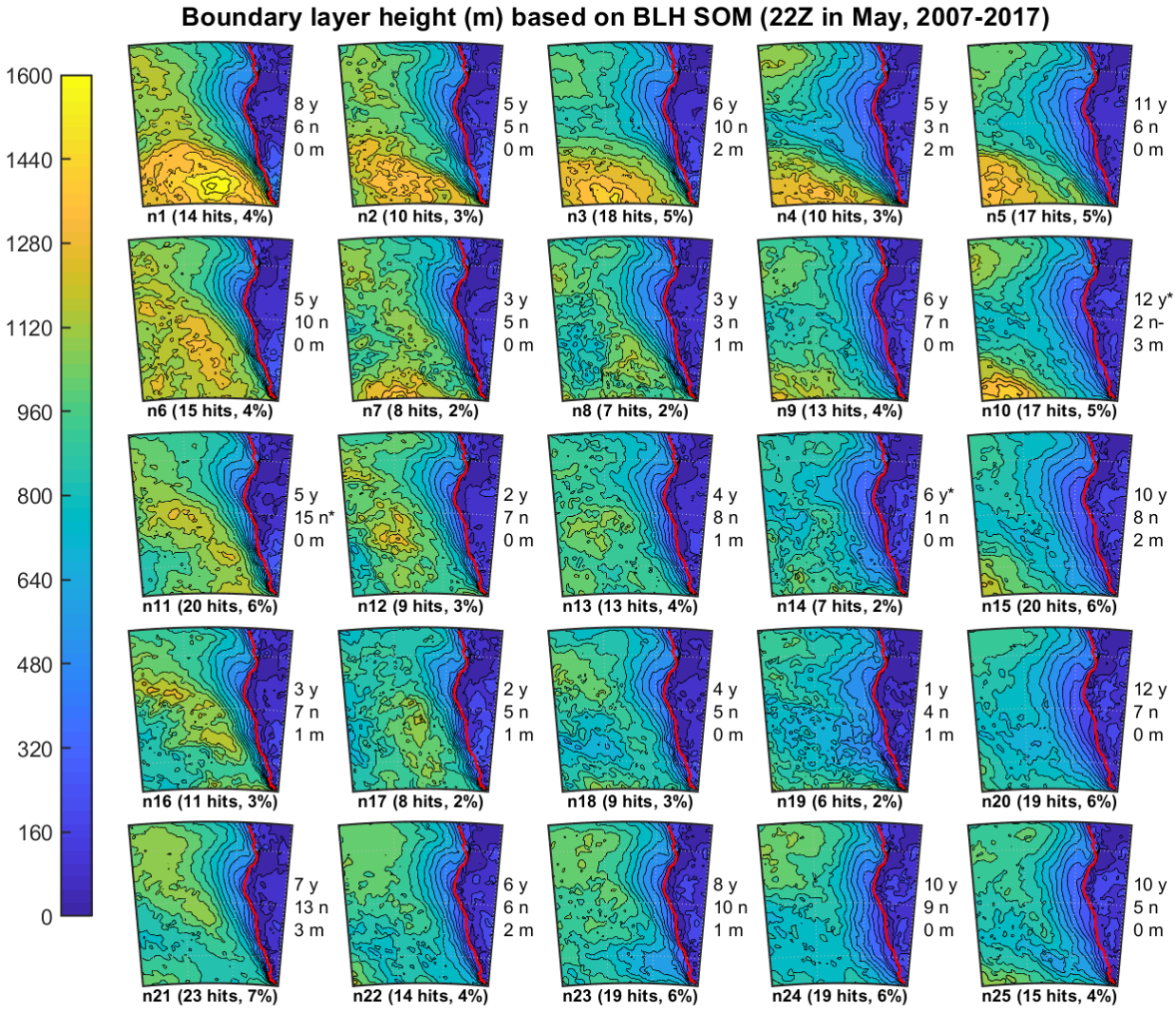


Figure G4: Boundary layer height projected onto the SOM of boundary layer height normalized anomalies in May, contoured every 80 meters. Coastline is indicated by the red line.

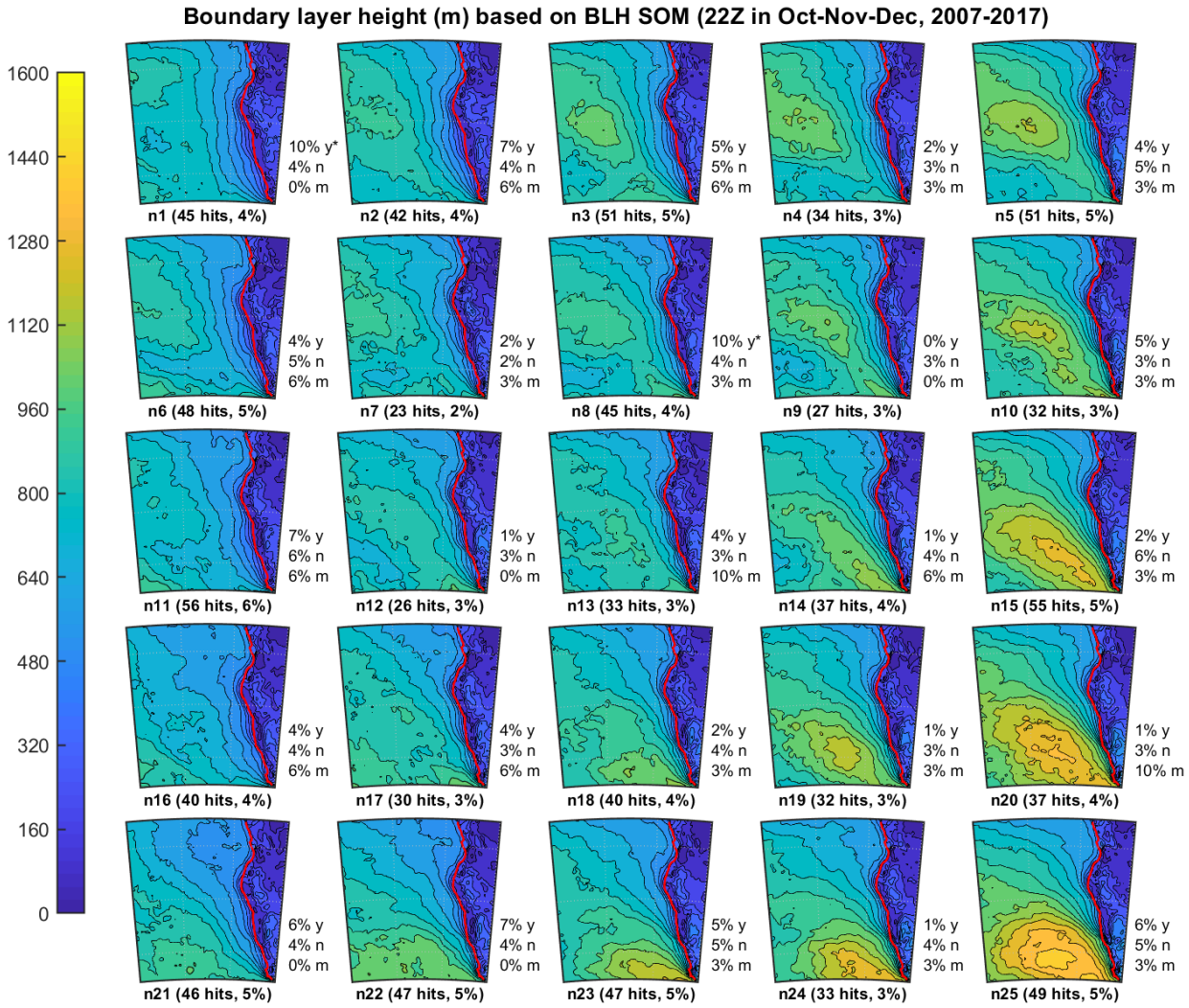


Figure G5: Boundary layer height projected onto the SOM of boundary layer height normalized anomalies in Oct-Nov-Dec, contoured every 80 meters. Coastline is indicated by the red line.

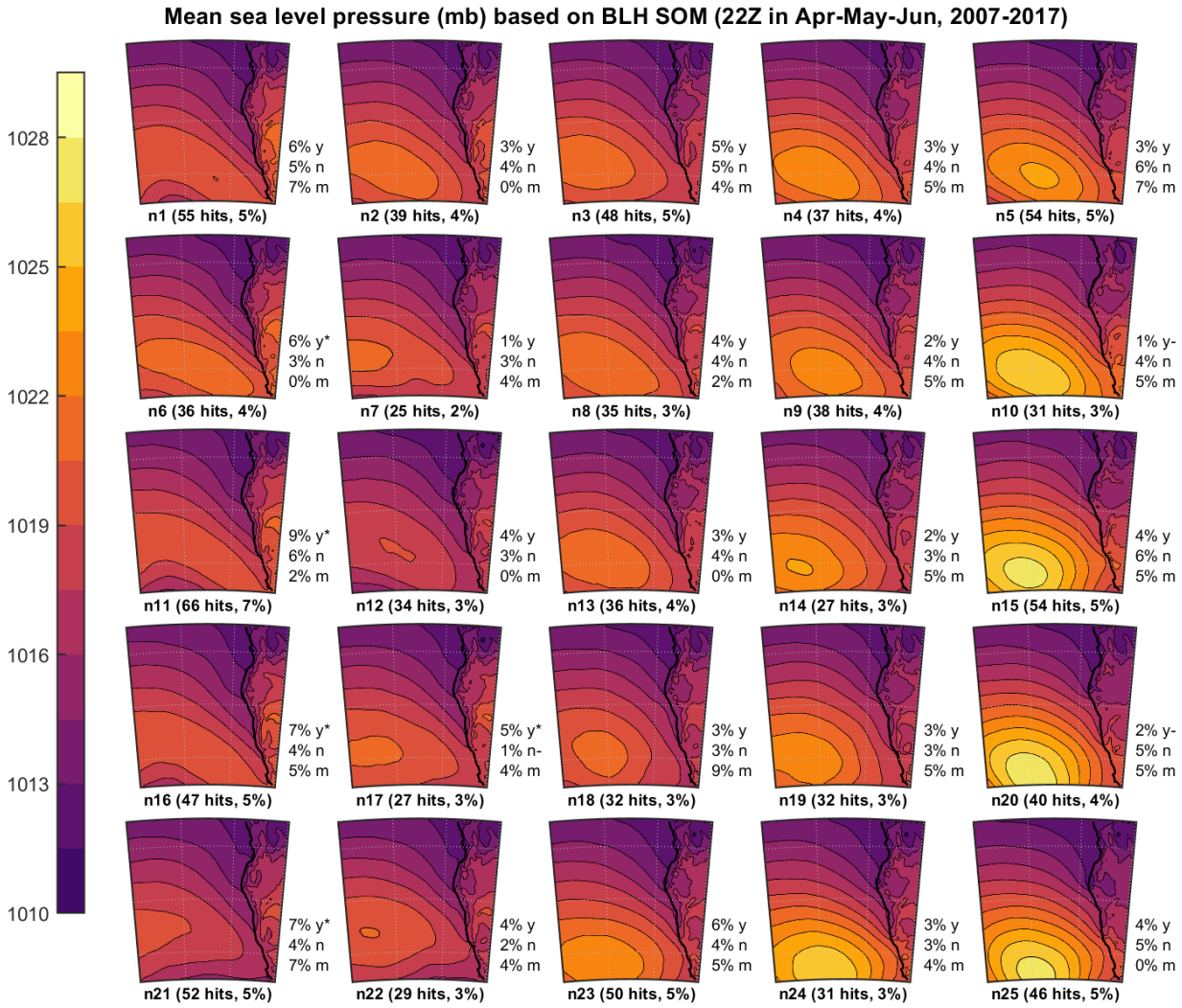


Figure G6: Mean sea level pressure projected onto the SOM of boundary layer height normalized anomalies in Apr-May-Jun, contoured every 1.5 millibars. Coastline is indicated by the black line.

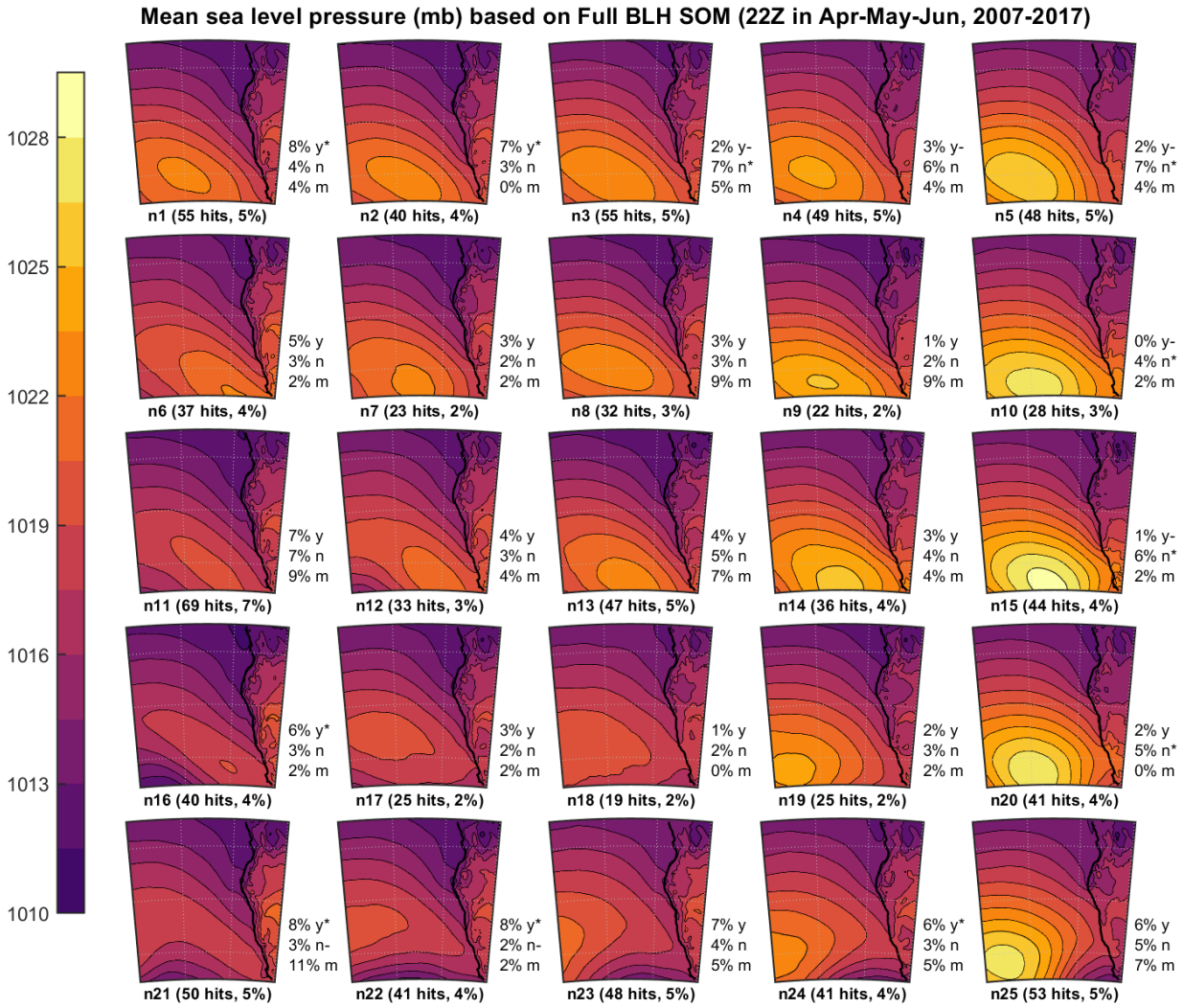


Figure G7: Mean sea level pressure projected onto the SOM of full boundary layer height in Apr-May-Jun, contoured every 1.5 millibars. Coastline is indicated by the black line.

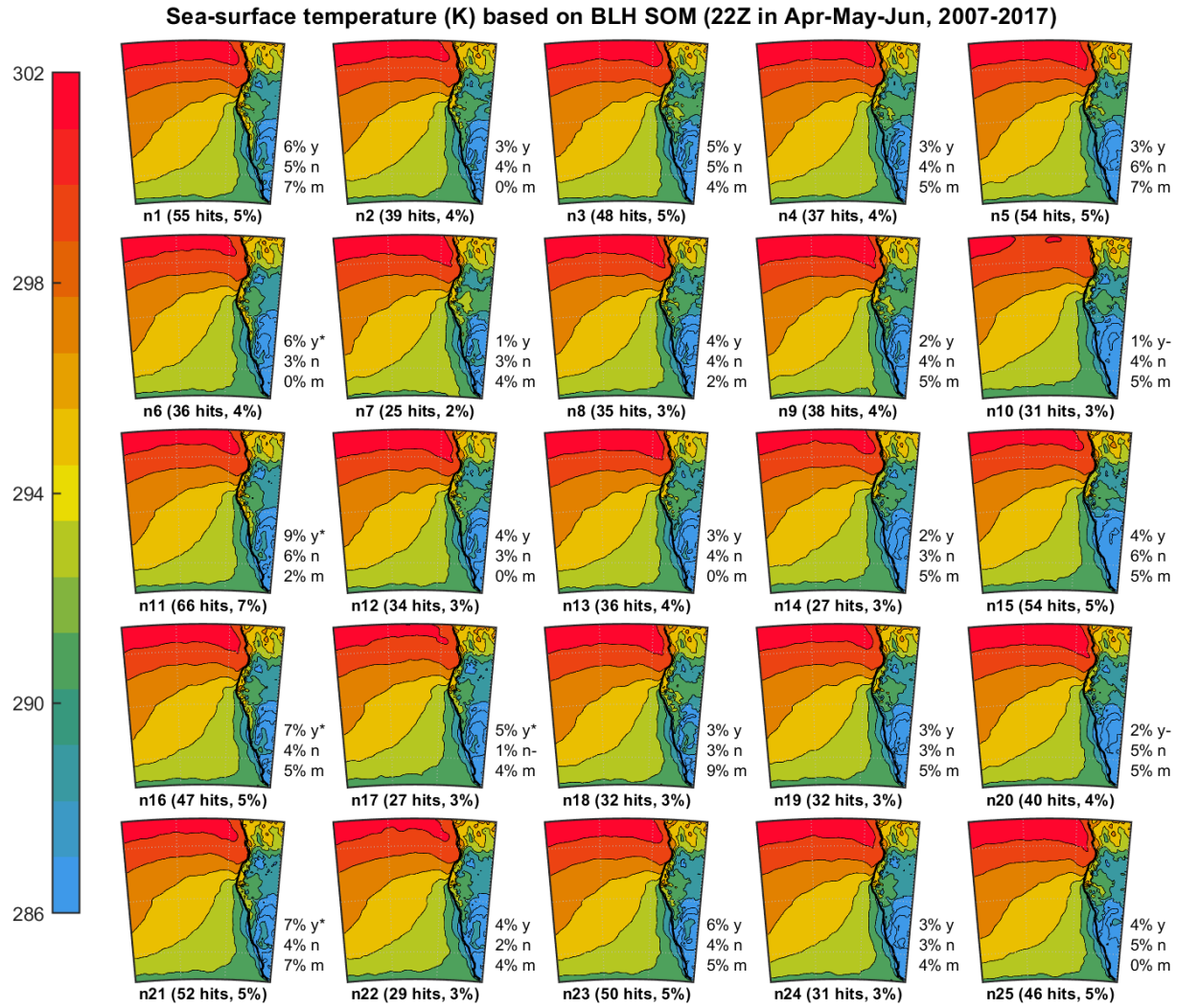


Figure G8: Sea surface temperature projected onto the SOM of boundary layer height normalized anomalies in Apr-May-Jun, contoured every 2 Kelvin. Coastline is indicated by the black line.

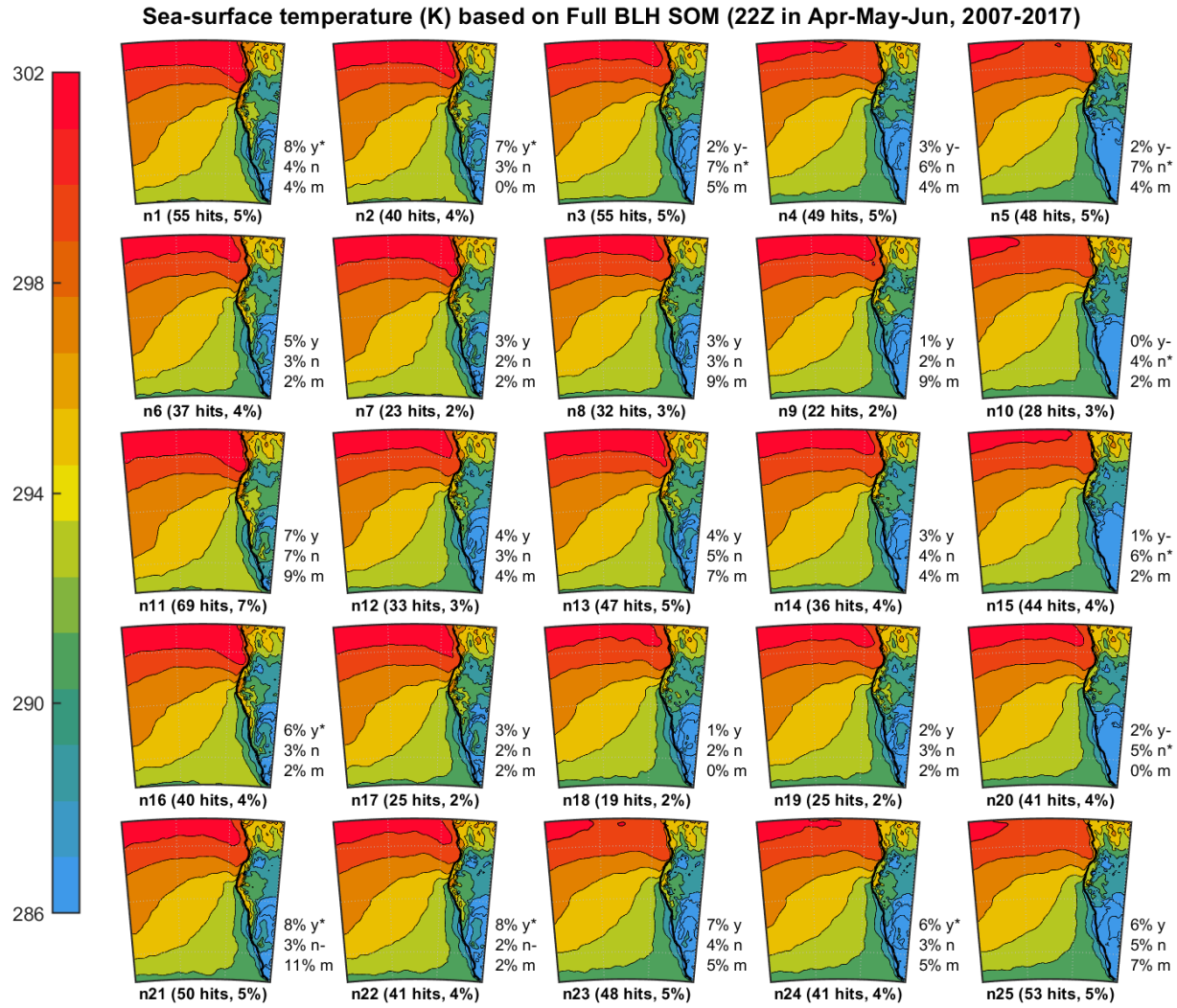


Figure G9: Sea surface temperature projected onto the SOM of full boundary layer height in Apr-May-Jun, contoured every 2 Kelvin. Coastline is indicated by the black line.

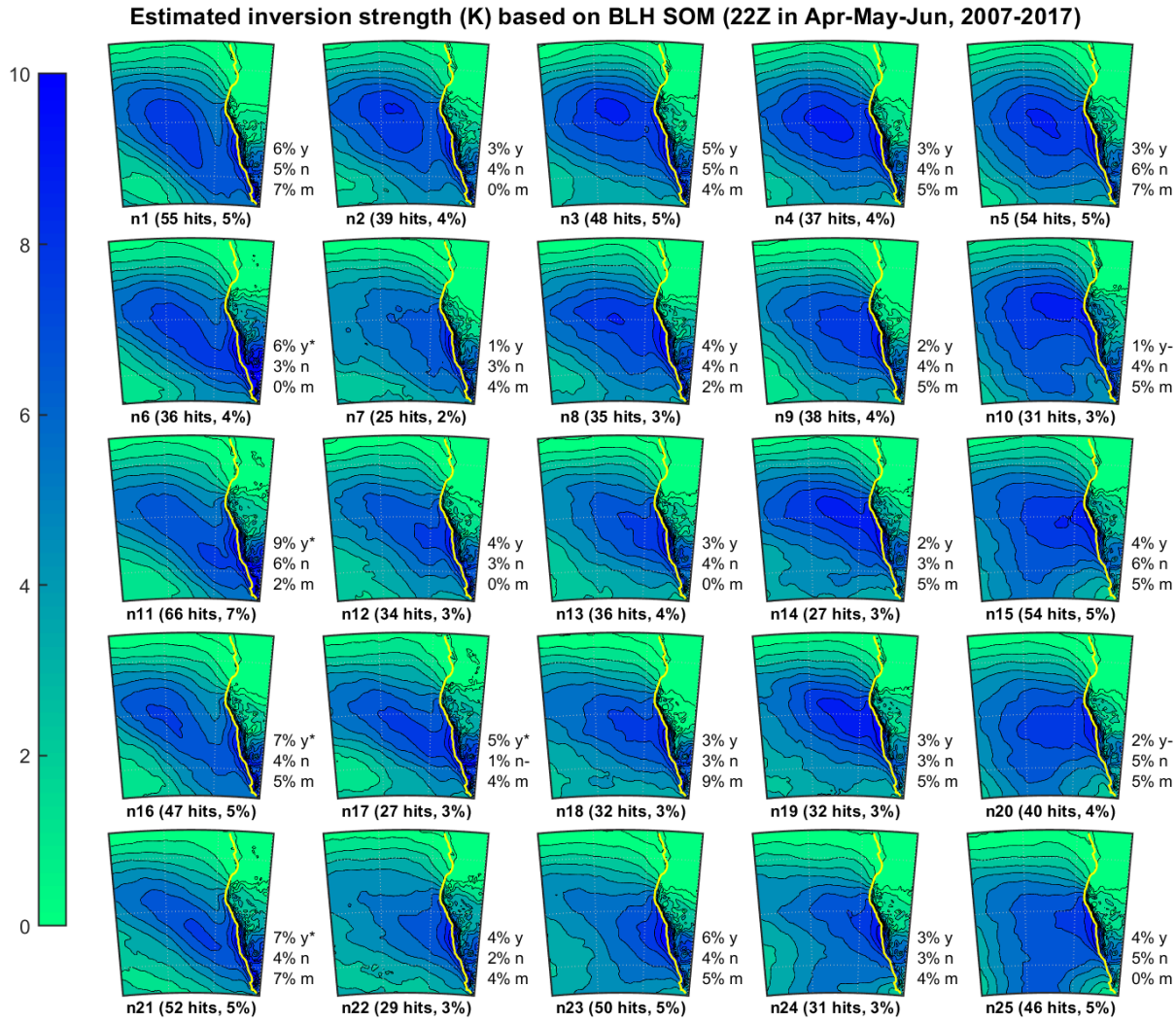


Figure G10: Estimated inversion strength projected onto the SOM of boundary layer height normalized anomalies in Apr-May-Jun, contoured every 1 Kelvin. Coastline is indicated by the yellow line.

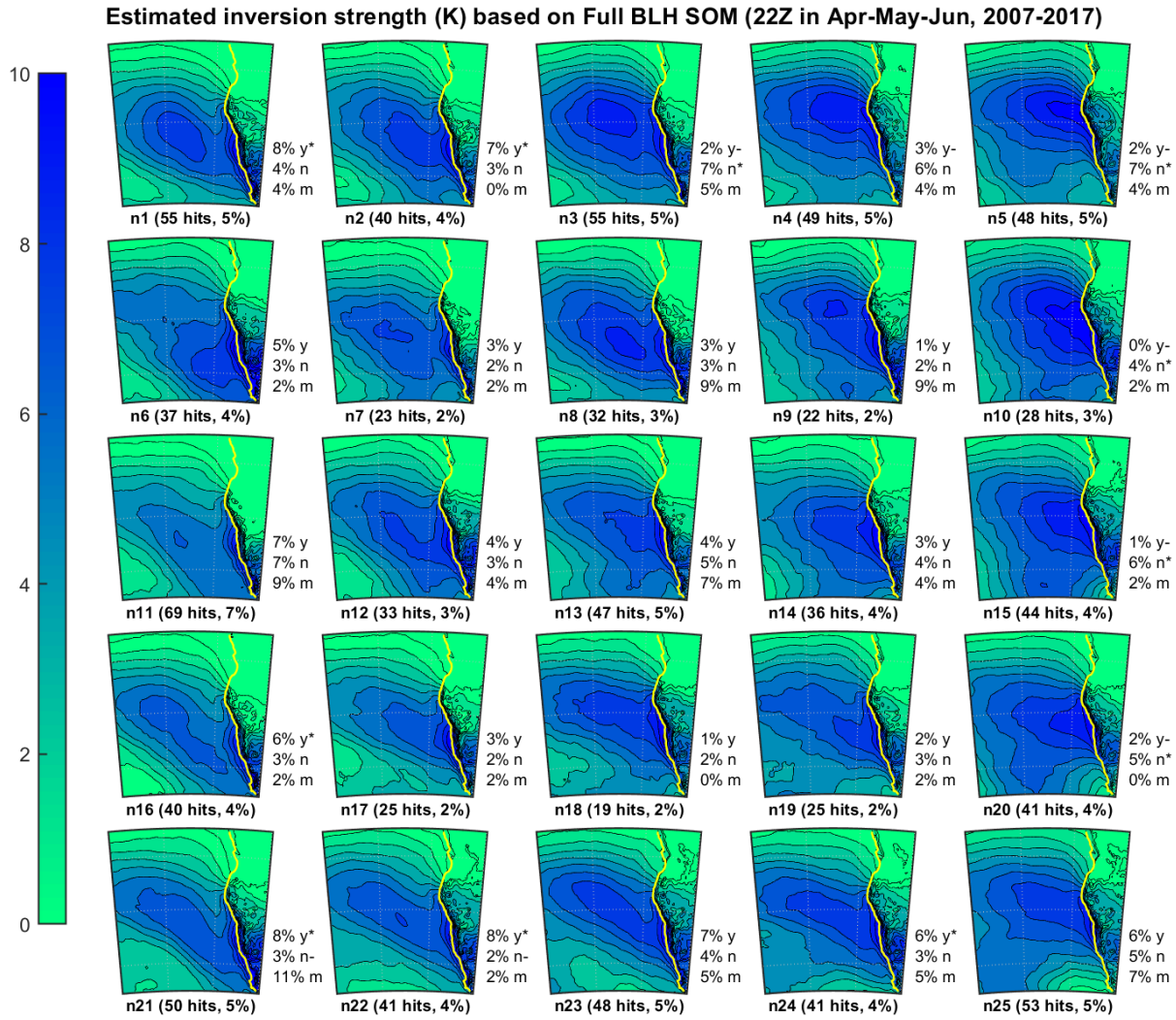


Figure G11: Estimated inversion strength projected onto the SOM of full boundary layer height in Apr-May-Jun, contoured every 1 Kelvin. Coastline is indicated by the yellow line.

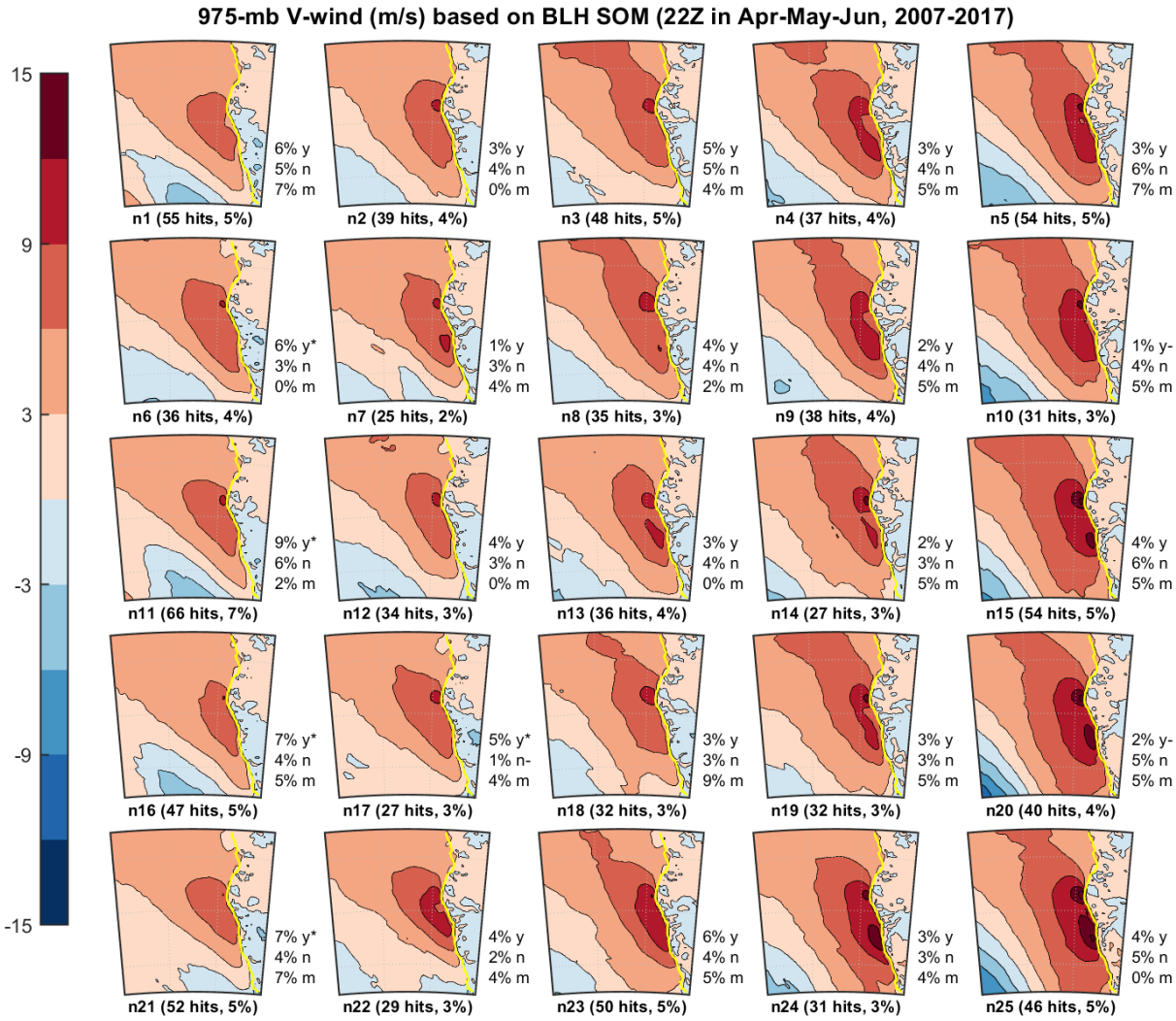


Figure G12: Meridional (*v*) wind at 975 hPa projected onto the SOM of boundary layer height normalized anomalies in Apr-May-Jun, contoured every 3 m/s. Coastline is indicated by the yellow line.

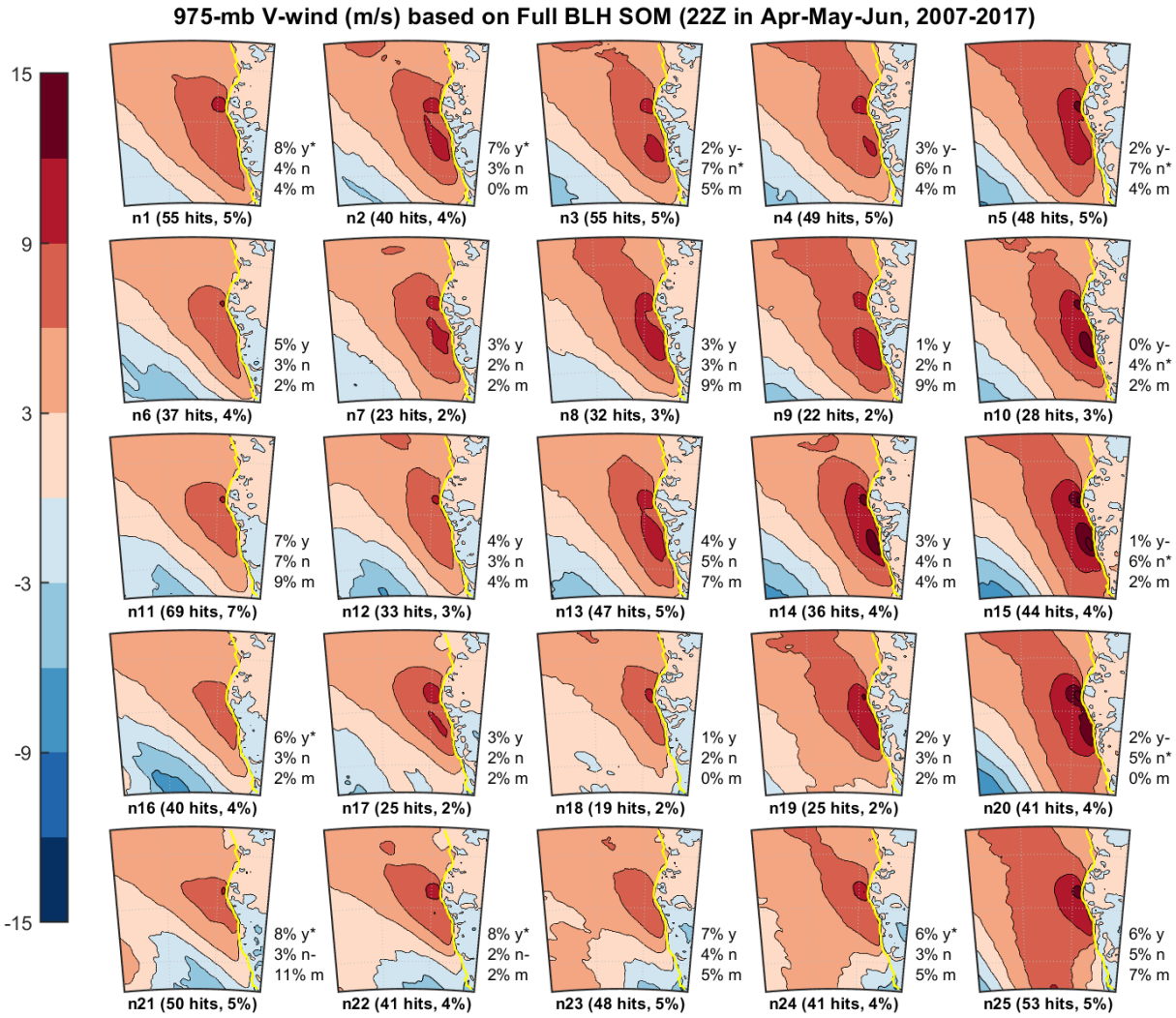


Figure G13: Meridional (*v*) wind at 975 hPa projected onto the SOM of full boundary layer height in Apr-May-Jun, contoured every 3 m/s. Coastline is indicated by the yellow line.

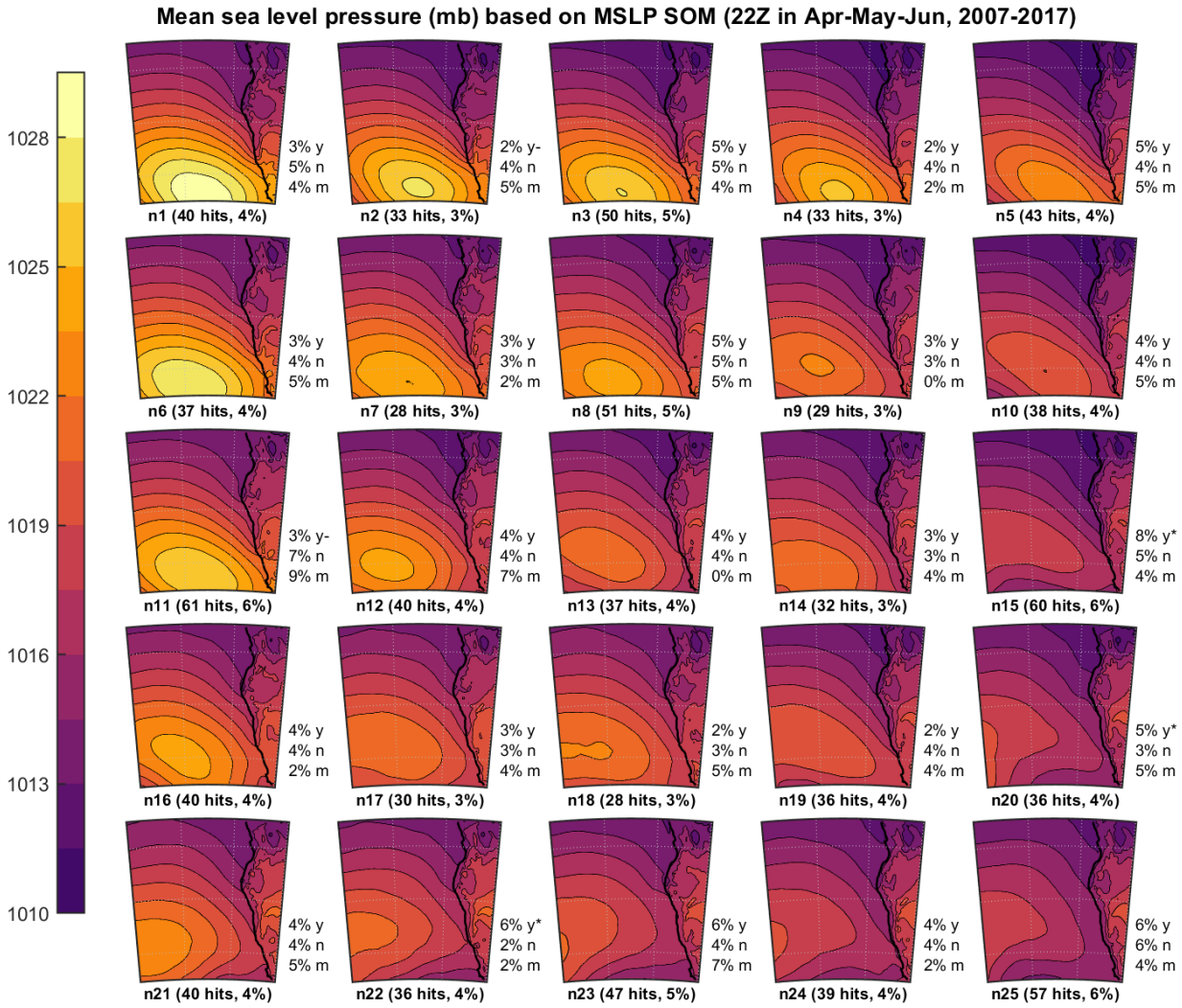


Figure G14: Mean sea level pressure projected onto the SOM of mean sea level pressure normalized anomalies in Apr-May-Jun, contoured every 1.5 millibars. Coastline is indicated by the black line.

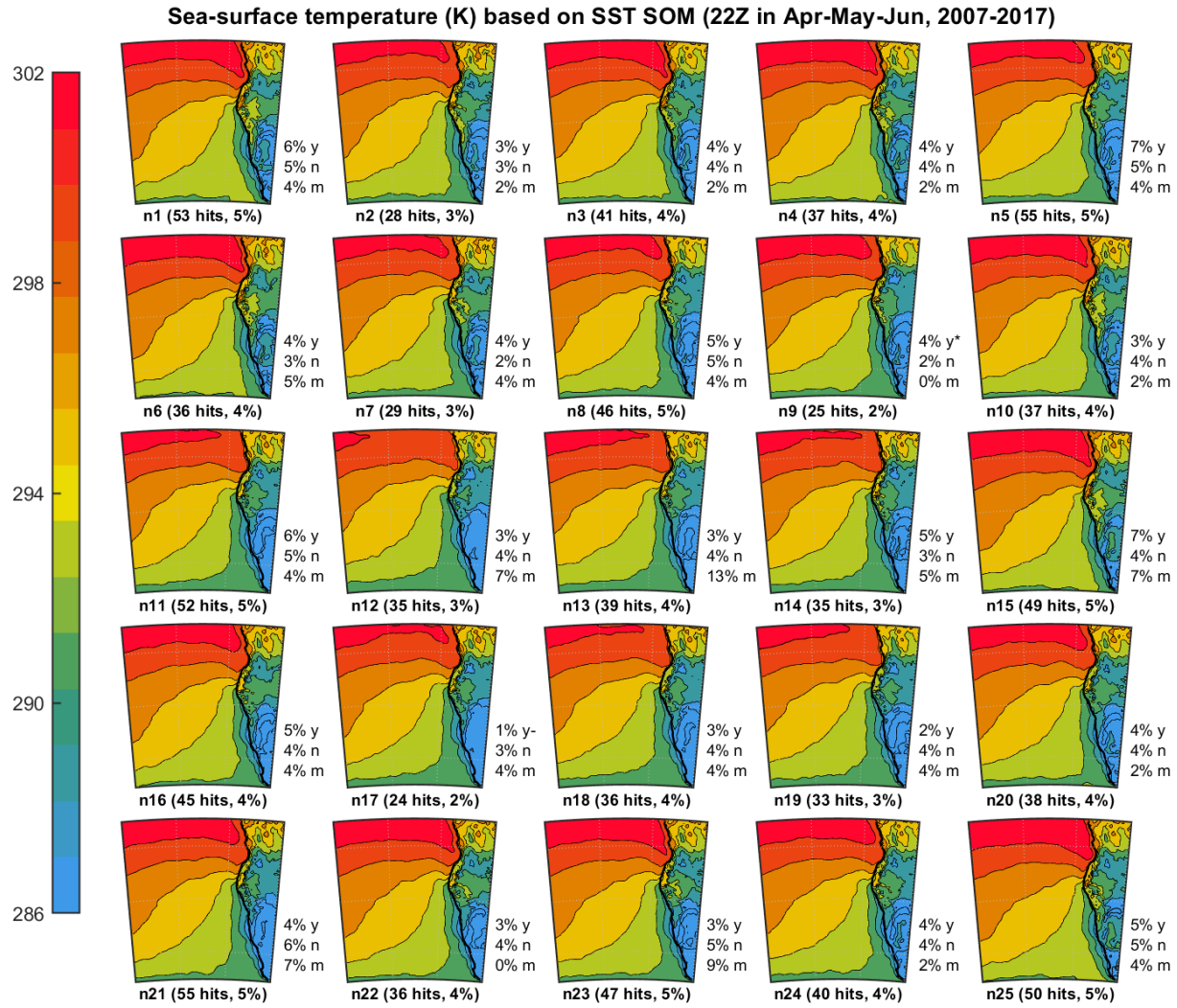


Figure G15: Sea surface temperature (SST) projected onto the SOM of SST normalized anomalies in Apr-May-Jun, contoured every 2 Kelvin. Coastline is indicated by the black line.

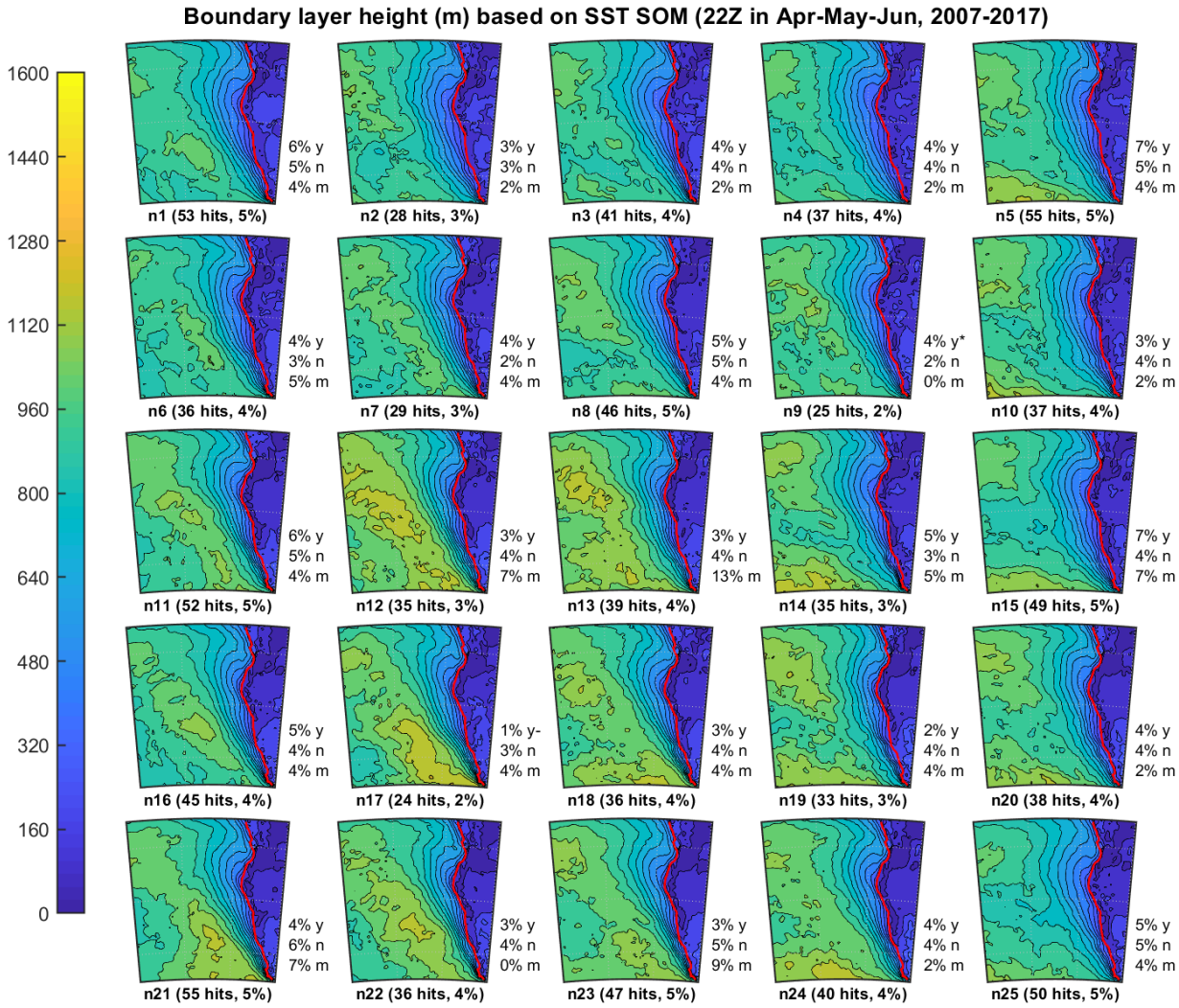


Figure G16: Boundary layer height projected onto the SOM of SST normalized anomalies in Apr-May-Jun, contoured every 80 meters. Coastline is indicated by the red line.

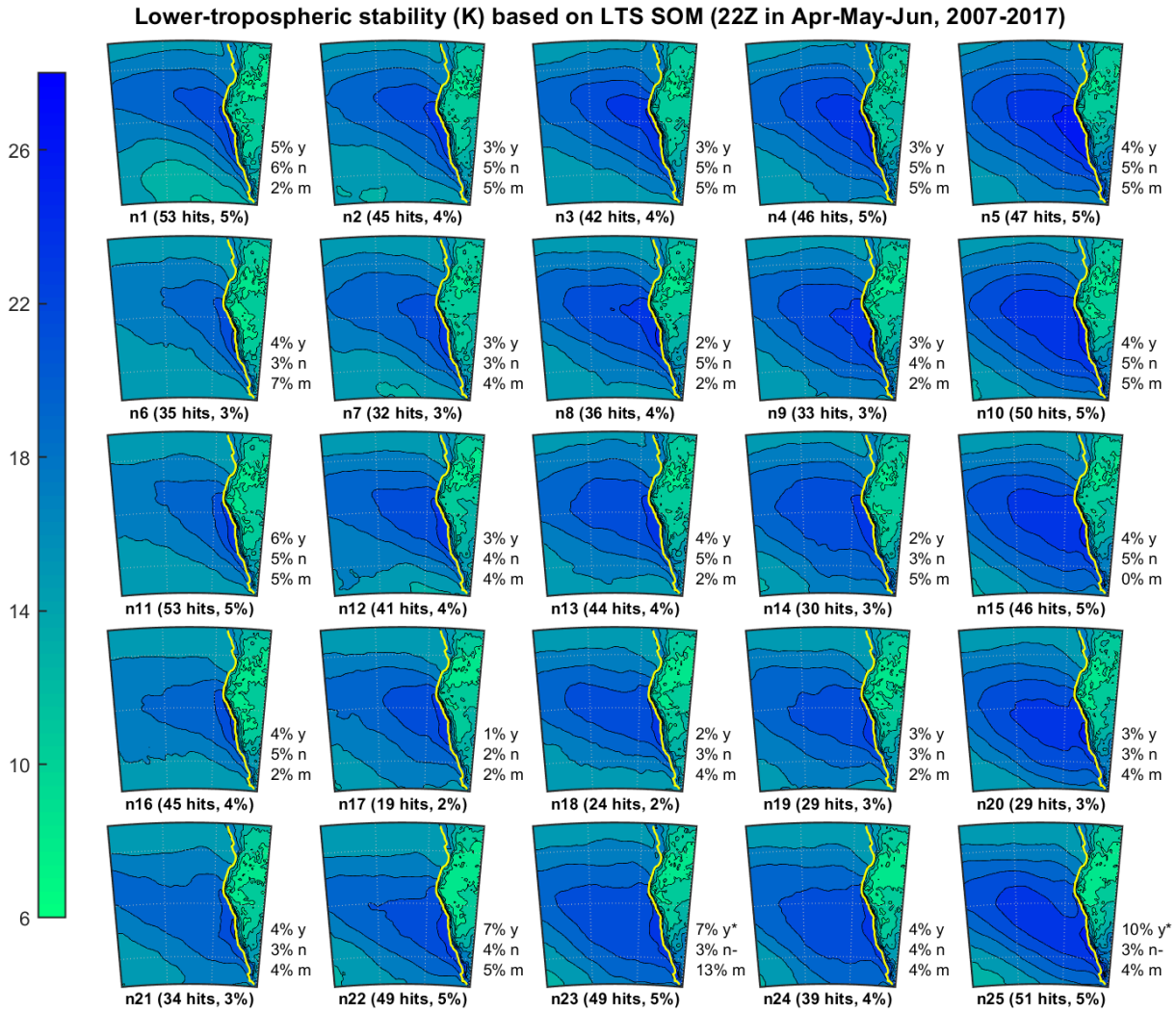


Figure G17: Lower-tropospheric stability (LTS) projected onto the SOM of LTS normalized anomalies in Apr-May-Jun, contoured every 1 Kelvin. Coastline is indicated by the yellow line.

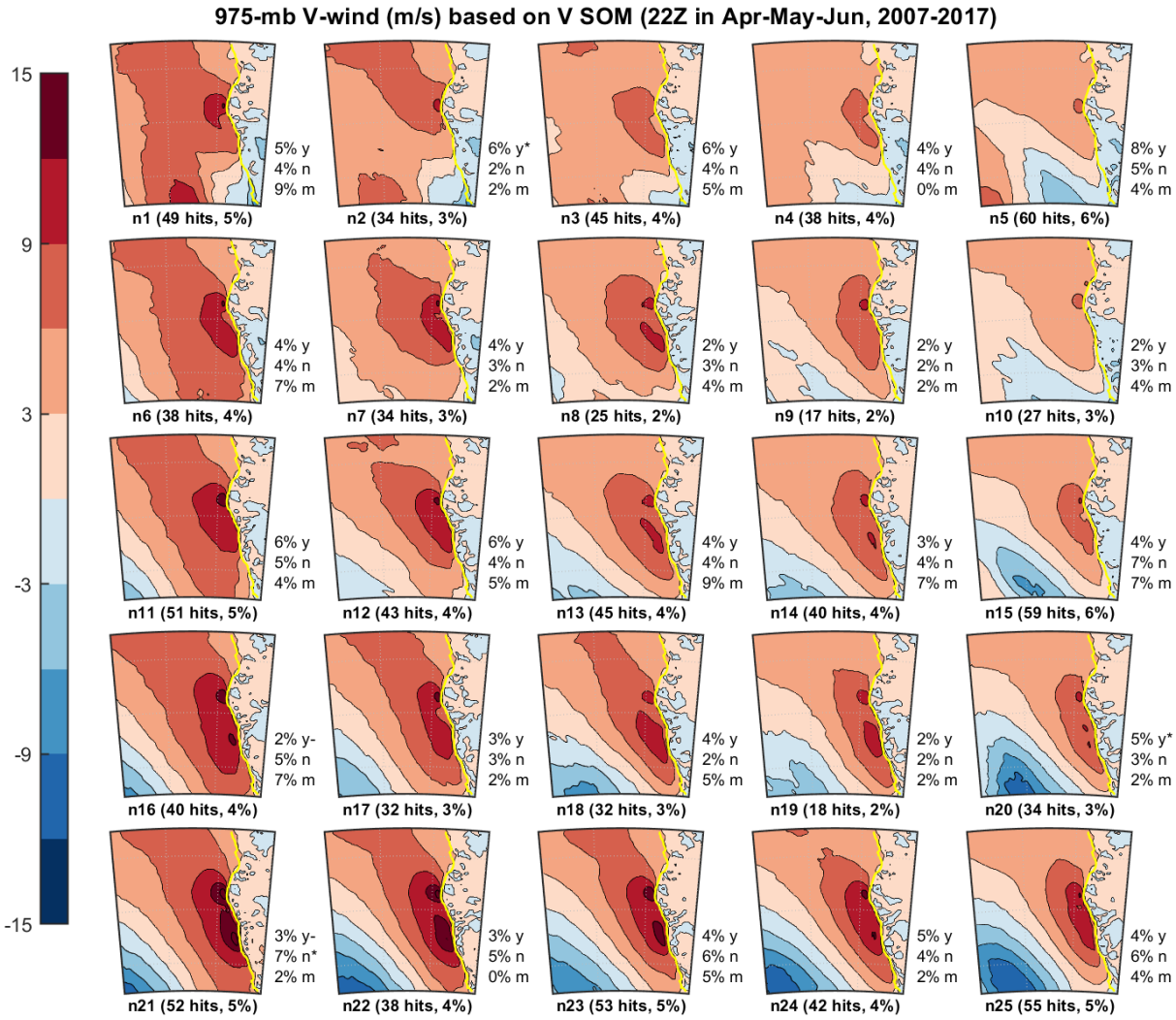


Figure G18: Meridional (*v*) wind at 975 hPa projected onto the SOM of 975-hPa *v*-wind normalized anomalies in Apr-May-Jun, contoured every 3 m/s. Coastline is indicated by the yellow line.

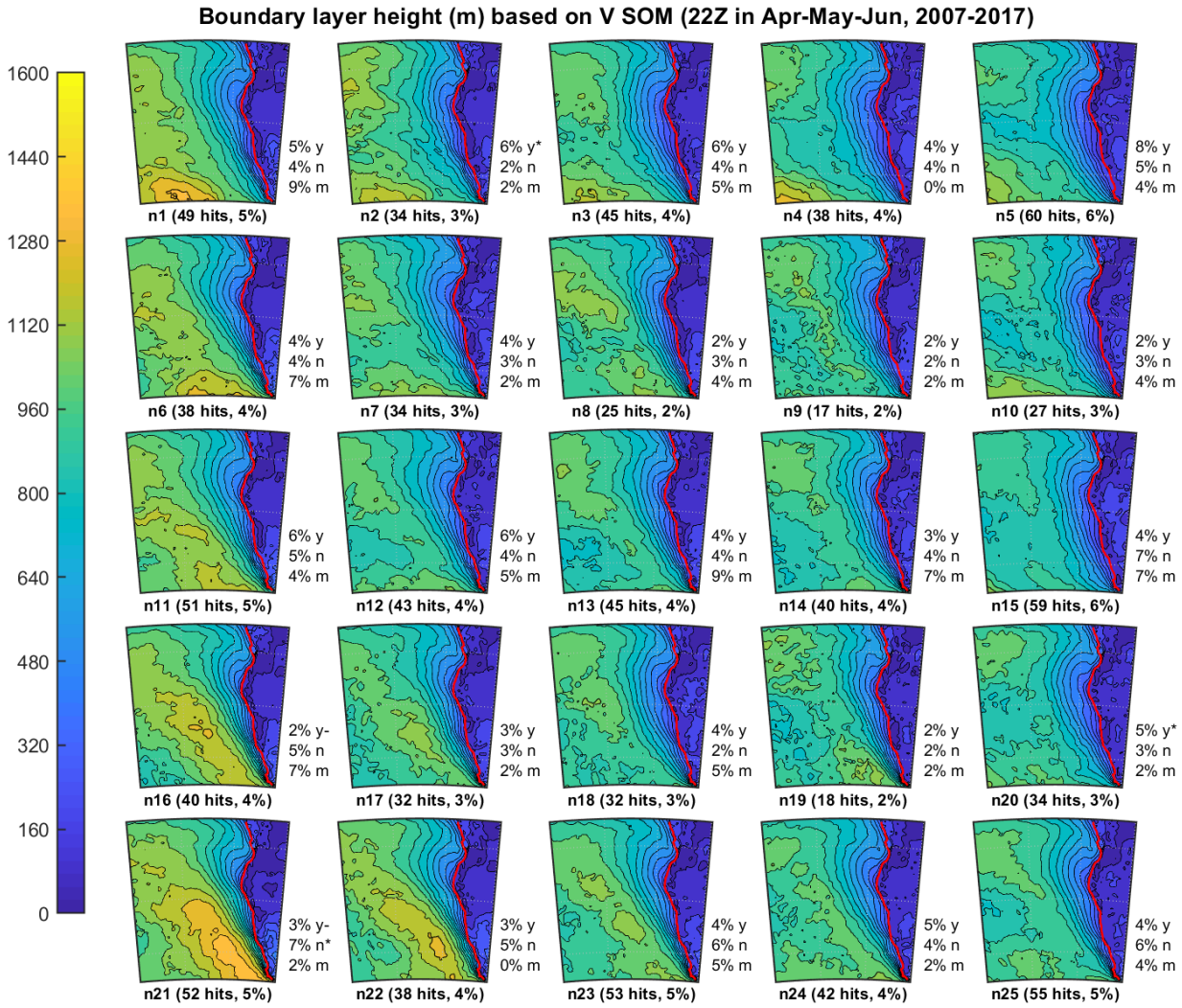


Figure G19: Boundary layer height projected onto the SOM of 975-hPa *v*-wind normalized anomalies in Apr-May-Jun, contoured every 80 meters. Coastline is indicated by the red line.

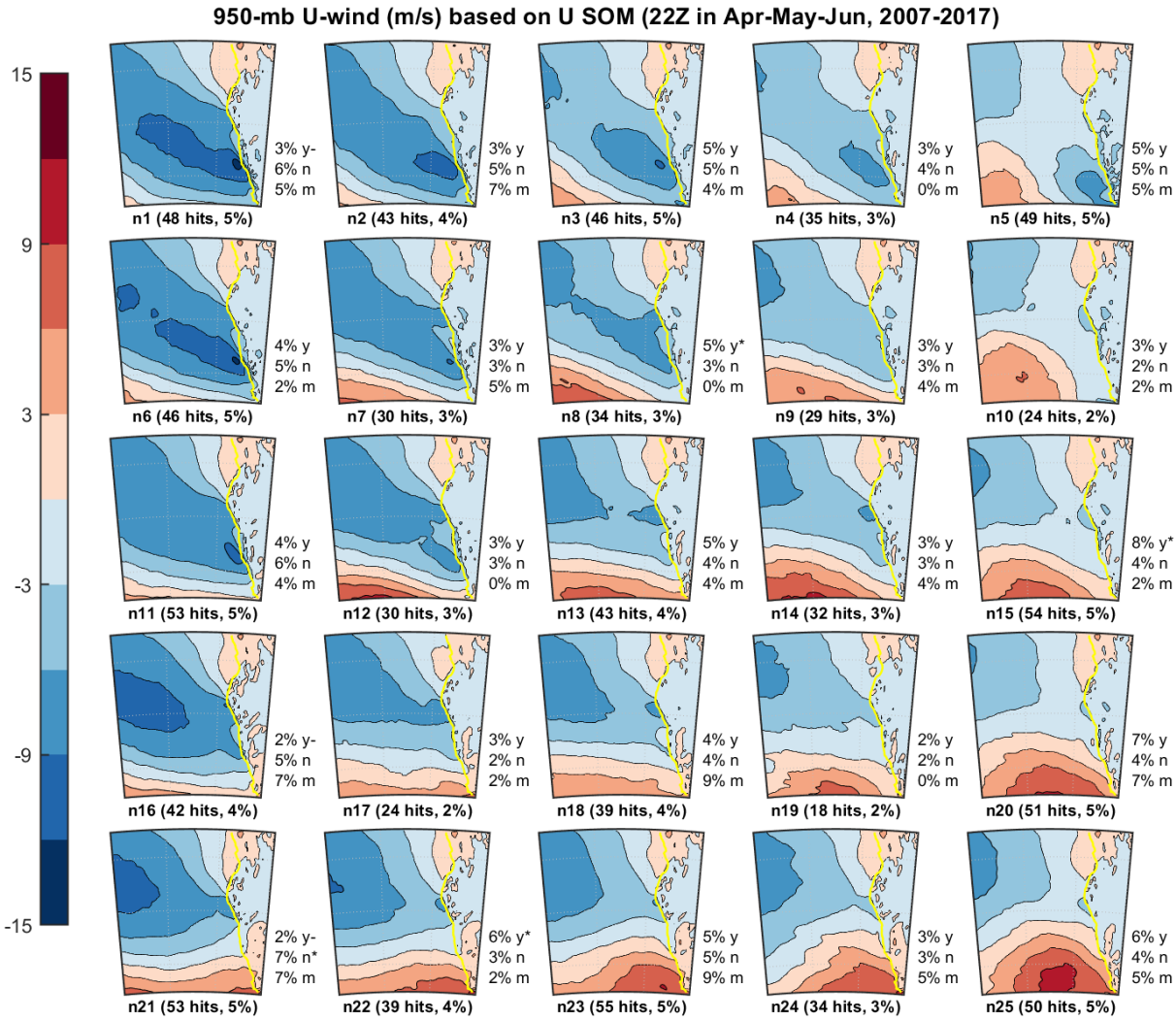


Figure G20: Zonal (*u*) wind at 950 hPa projected onto the SOM of 950-hPa *u*-wind normalized anomalies in Apr-May-Jun, contoured every 3 m/s. Coastline is indicated by the yellow line.

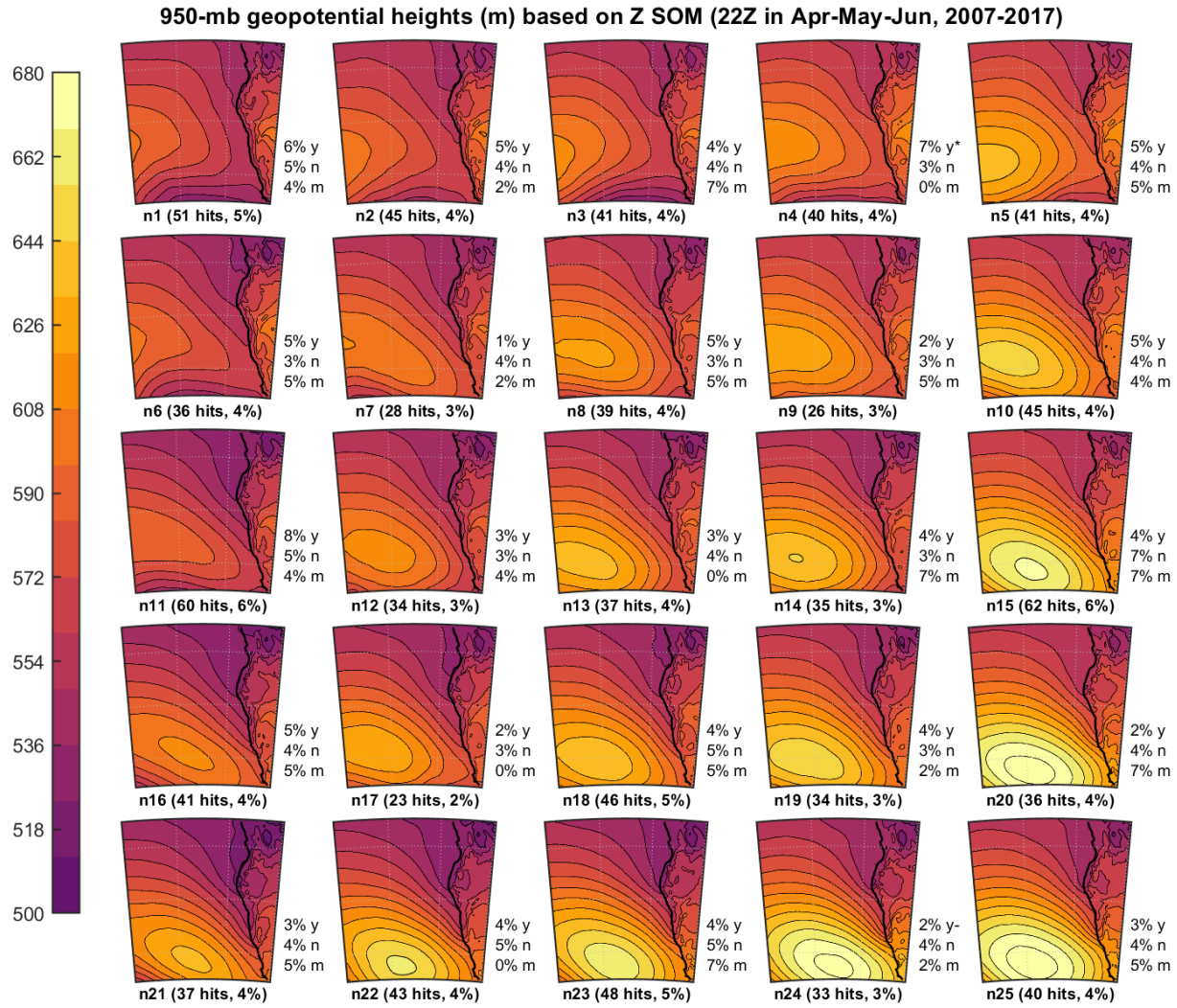


Figure G21: 950-hPa geopotential height (*Z*) projected onto the SOM of 950-hPa *Z* normalized anomalies in Apr-May-Jun, contoured every 9 meters. Coastline is indicated by the black line.

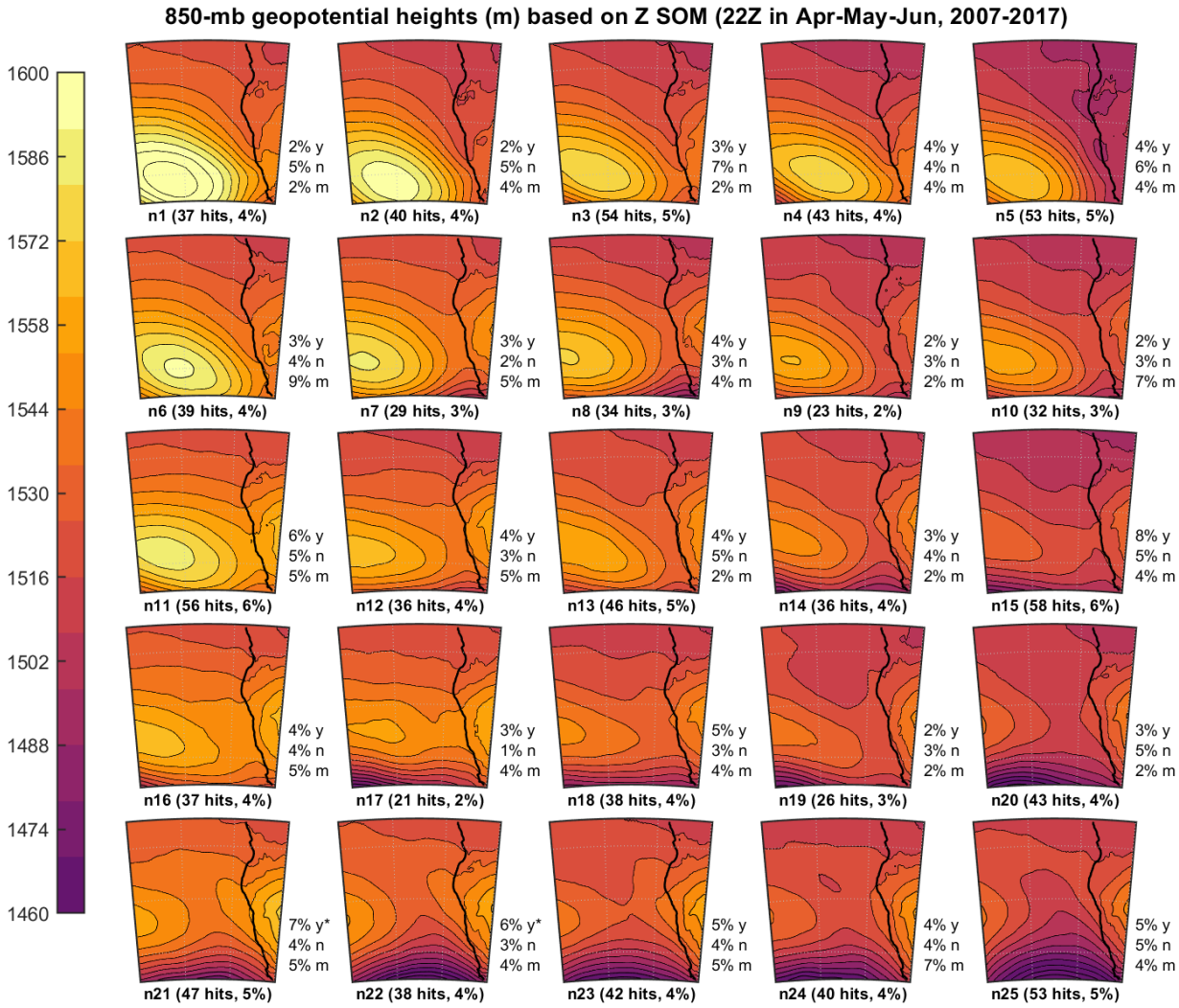


Figure G22: 850-hPa geopotential height (*Z*) projected onto the SOM of 850-hPa *Z* normalized anomalies in Apr-May-Jun, contoured every 7 meters. Coastline is indicated by the black line.

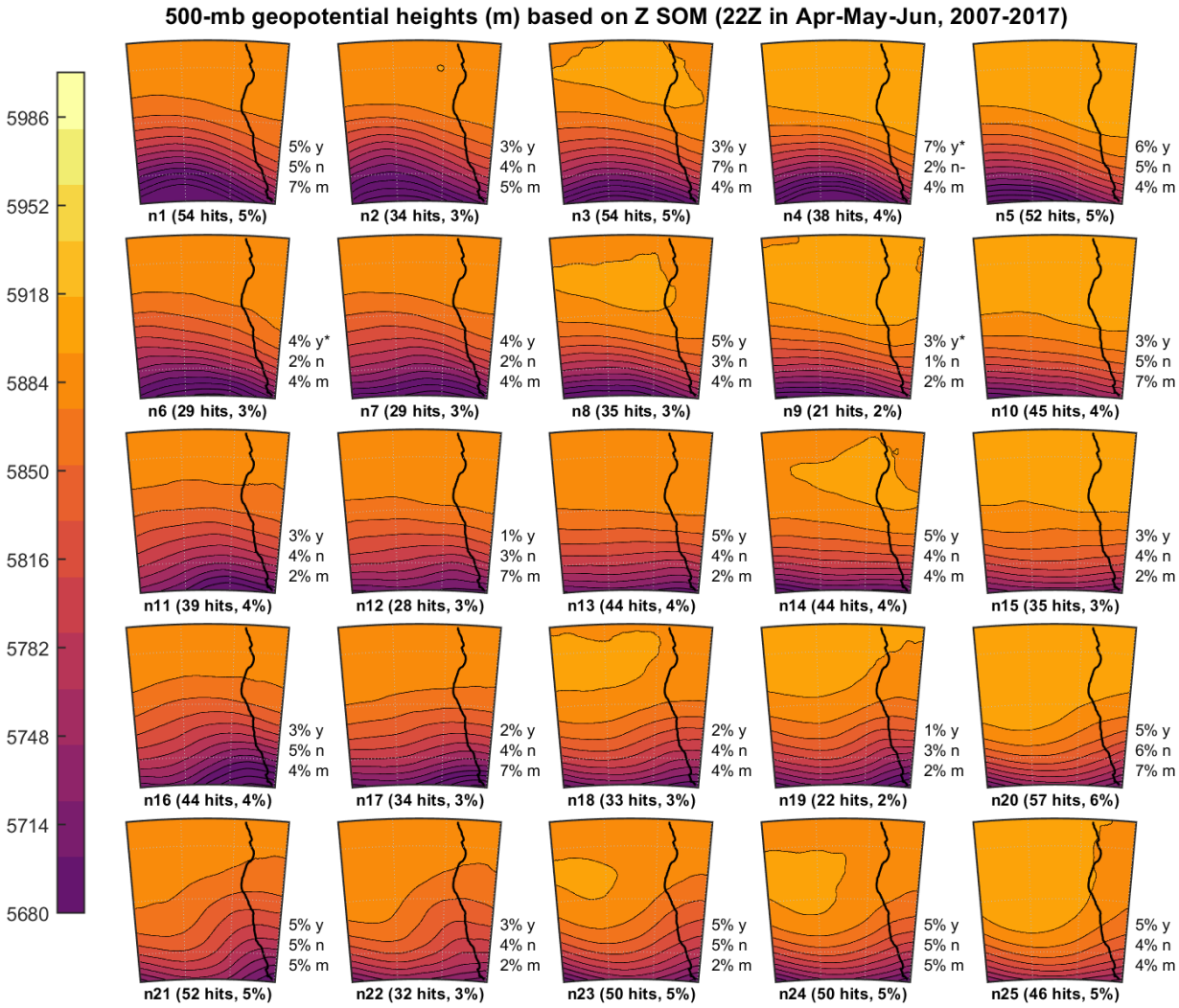


Figure G23: 500-hPa geopotential height (*Z*) projected onto the SOM of 500-hPa *Z* normalized anomalies in Apr-May-Jun, contoured every 17 meters. Coastline is indicated by the black line.

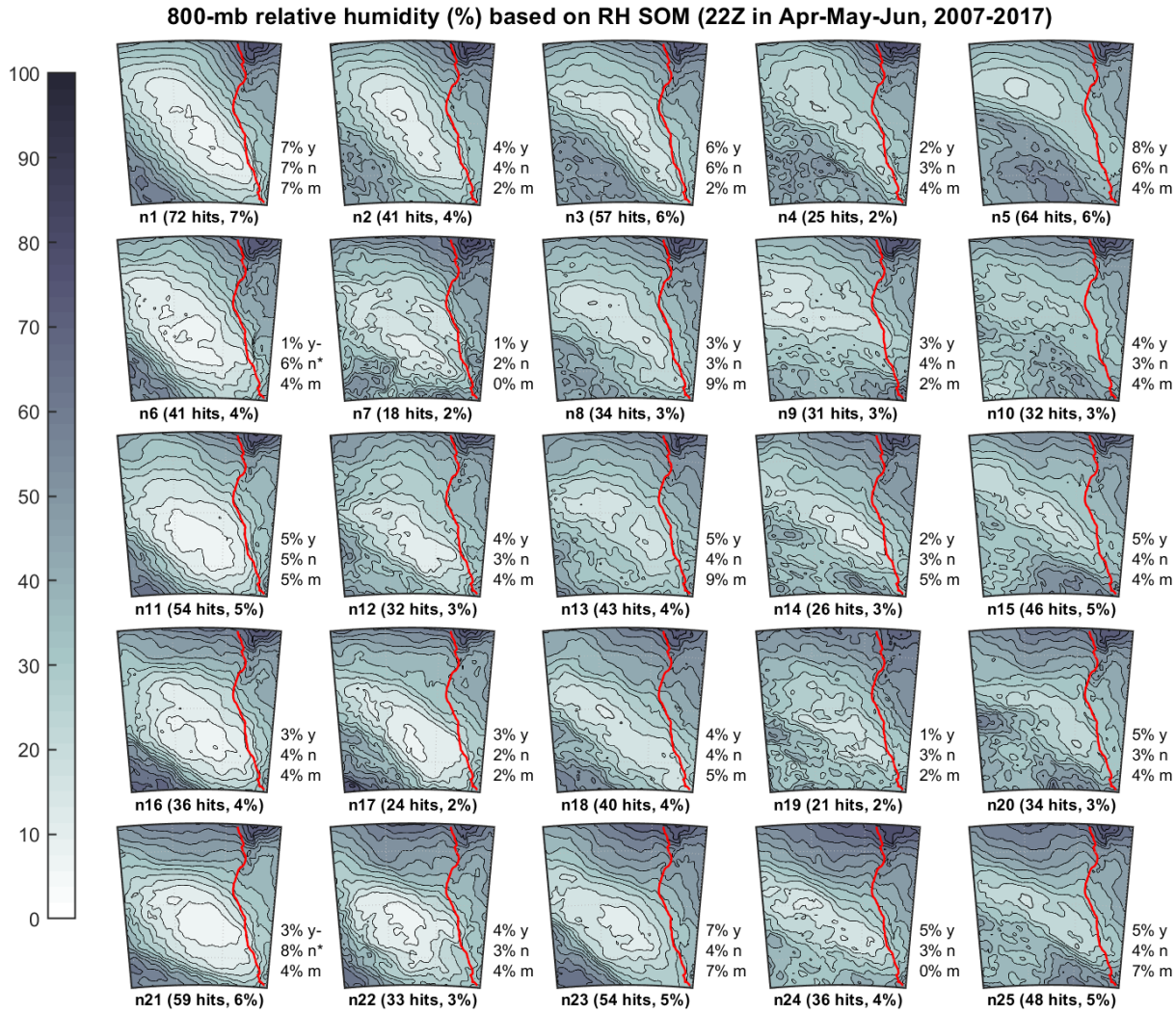


Figure G24: 800-hPa relative humidity (RH) projected onto the SOM of 800-hPa RH normalized anomalies in Apr-May-Jun, contoured every 5%. Coastline is indicated by the red line.

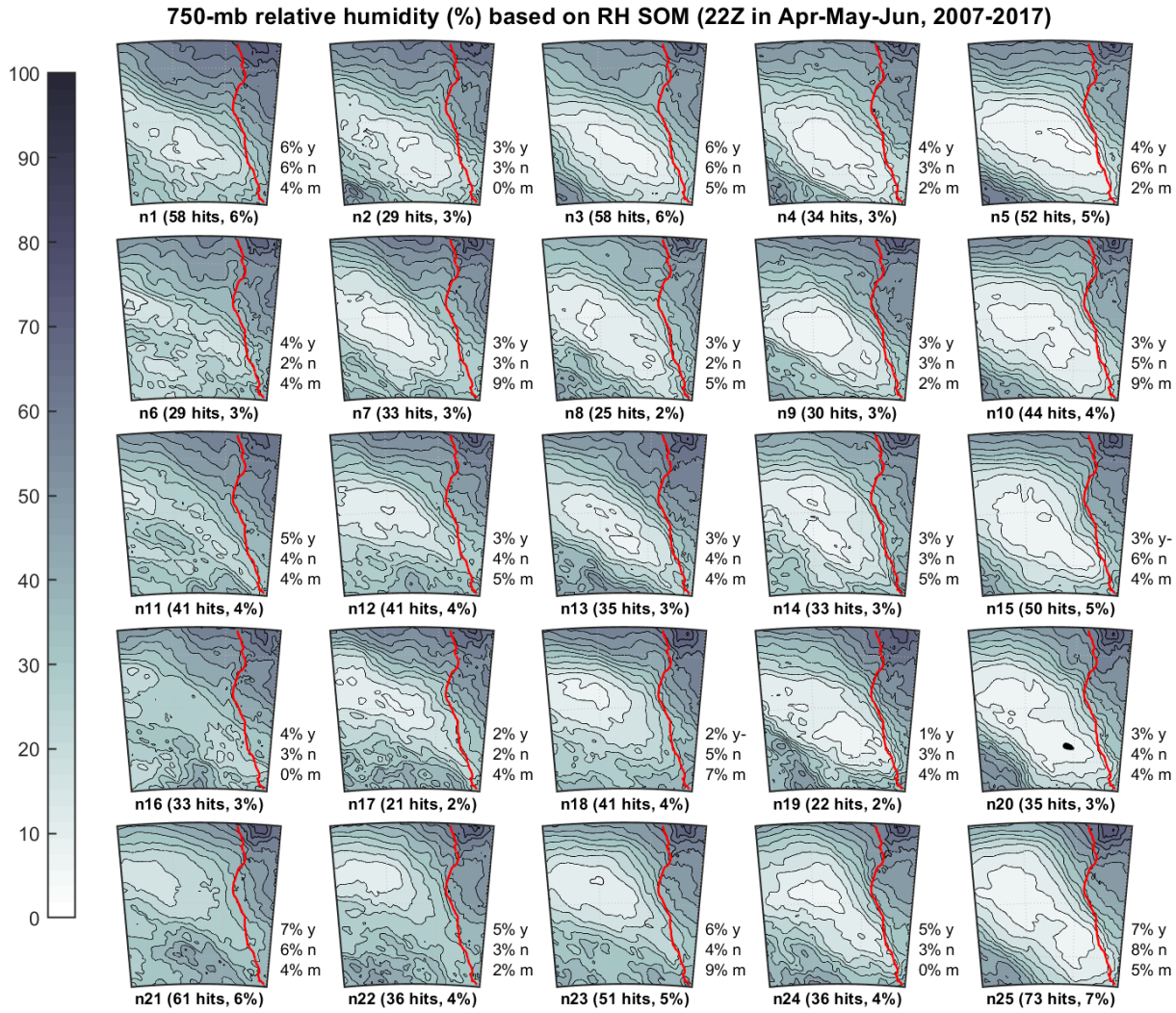


Figure G25: 750-hPa relative humidity (RH) projected onto the SOM of 750-hPa RH normalized anomalies in Apr-May-Jun, contoured every 5%. Coastline is indicated by the red line.