Atmospheric classification as a cloud and precipitation evaluation tool in models and observations

Stuart Evans

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington
2014

Reading Committee:
Thomas Ackerman, Chair
Roger Marchand
Dargan Frierson

Program Authorized to Offer Degree:
Atmospheric Sciences
Abstract

Atmospheric classification as a cloud and precipitation evaluation tool in models and observations

Stuart Evans

Chair of the Supervisory committee:
Professor Thomas Ackerman
Department of Atmospheric Sciences

Atmospheric classifications are created for two regions using an automated clustering technique first described in Marchand et al. 2009. This method applies an iterative clustering algorithm to regional snapshots of dynamic and thermodynamic variables from the ERA-Interim reanalysis to define atmospheric states. An atmospheric state in this context is a frequently occurring weather pattern for the region. In creating the states, a time series of atmospheric state for the region is created. These states then serve as a basis for compositing other weather observations, creating distributions of associated weather variables such as cloud occurrence, precipitation, and radiative fluxes. The states are also suitable for sorting output from general circulation models, allowing the comparison of observed and modeled cloud variables on a state-by-state basis. Analysis of which atmospheric states lead to the most error in modeled values provides insight into the particular atmospheric conditions and processes that are problematic for models.
A classification for a region surrounding Darwin, Australia is used to define periods of monsoon activity and investigate the interannual and intraseasonal variability of the Australian monsoon, as well as long-term trends in precipitation at Darwin. The classification creates a time series of atmospheric states, two of which are identified as corresponding to the active monsoon and the monsoon break. Occurrence of these states is used to define onset, retreat, and intensity of the monsoon season and the timing of individual active periods. Previous studies investigated the role of the MJO during the monsoon season, but have differed on whether the MJO creates a characteristic period or duration of active monsoon periods. We use our classification-based metrics of monsoon activity to examine the timing of individual active periods each season relative to the phase of the MJO, showing that the passage of the convective anomaly over Darwin helps both trigger and end individual active periods. Lastly, we look at trends in the occurrence of the atmospheric states and find that the number of active monsoon days has increased over the previous 33 years. We show that this, rather than changes in the daily rainfall rate during active monsoon periods, is responsible for an increasing trend in annual precipitation at Darwin during that time.

A second classification for the region surrounding the Southern Great Plains (SGP) site in Oklahoma produced 21 atmospheric states. Analysis of these states showed that they represented different stages of passing synoptic systems, with some states representing warm fronts, others cold fronts, still others high pressure systems, and so on. Snapshots from a 2° run of the AM3 model from GFDL were sorted according to these states, producing a time series of state within the model. The time series of model state was used to composite ISCCP simulator output from the model according to each state.
These simulated ISCCP composites are compared to composites of observed ISCCP data. In all states the model does not produce nearly enough high thin cloud. The representation of other cloud types depends on the state. We show that for states which have large-scale ascent (warm fronts and cold fronts) the model does not produce enough deep thick cloud, while for states that have parameterized convection (e.g. high pressure systems) the model produces too much cloud. The former we interpret as the model struggling to resolve the dynamics of fronts, while the latter indicates that the deep convective parameterization triggers too often. We demonstrate that the overall cloud amount bias is driven primarily by within-state errors in cloud amount rather than by errors in the relative frequency of occurrence of states.

We compare the results of the $2^\circ$ run of the AM3 to a second run of the model at $0.5^\circ$ resolution. We find that thick cloud increases in states with large-scale ascent, and interpret this result as a result of fronts being closer to resolved in the higher resolution run of the model. We also find that high thick cloud decreases in states with parameterized convection. We find that this result is due to the distribution of CAPE values becoming more skewed at higher resolution, and thus triggering the deep convective parameterization less often. As a result of these improvements in within-state cloud amount error, the overall bias in cloud amount is the higher resolution run is as much due to errors in the frequency of occurrence of states as it is due to within-state errors. Interestingly, while individual state cloud amount biases decrease, the overall bias increases, due to changes in the distribution of states. Whether the higher resolution run constitutes a better simulation of clouds is thus an interesting question which we address.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Figures</td>
<td>i</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>Chapter 1: Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 2: Clustering method and data</td>
<td>5</td>
</tr>
<tr>
<td>Chapter 3: Monsoon variability analysis at Darwin</td>
<td>12</td>
</tr>
<tr>
<td>3.1 Description of the Darwin states</td>
<td>15</td>
</tr>
<tr>
<td>3.1.1 Dry season states</td>
<td>16</td>
</tr>
<tr>
<td>3.1.2 Transition season states</td>
<td>17</td>
</tr>
<tr>
<td>3.1.3 Monsoon season states</td>
<td>18</td>
</tr>
<tr>
<td>3.2 Defining monsoon metrics</td>
<td>19</td>
</tr>
<tr>
<td>3.3 Interannual variability</td>
<td>20</td>
</tr>
<tr>
<td>3.4 Intraseasonal variability</td>
<td>25</td>
</tr>
<tr>
<td>3.5 Trend Analysis</td>
<td>29</td>
</tr>
<tr>
<td>Chapter 4: Evaluating model cloud properties at Southern Great Plains</td>
<td>51</td>
</tr>
<tr>
<td>4.1 Description of the SGP states</td>
<td>52</td>
</tr>
<tr>
<td>4.1.1 Southerlies</td>
<td>53</td>
</tr>
<tr>
<td>4.1.2 Cold fronts</td>
<td>54</td>
</tr>
<tr>
<td>4.1.3 Northerlies</td>
<td>55</td>
</tr>
<tr>
<td>4.1.4 Anticyclones</td>
<td>56</td>
</tr>
<tr>
<td>4.1.5 Summer</td>
<td>57</td>
</tr>
<tr>
<td>4.2 Evaluation of cloud properties in the AM3</td>
<td>58</td>
</tr>
<tr>
<td>4.2.1 Classifying model output</td>
<td>58</td>
</tr>
<tr>
<td>4.2.2 Comparison of normalization techniques</td>
<td>61</td>
</tr>
<tr>
<td>4.2.3 Evaluation of atmospheric state cloud properties</td>
<td>63</td>
</tr>
<tr>
<td>4.2.4 Decomposition of errors</td>
<td>67</td>
</tr>
<tr>
<td>4.3 Evaluation of a high resolution run</td>
<td>69</td>
</tr>
<tr>
<td>Chapter 5: Discussion and conclusions</td>
<td>102</td>
</tr>
<tr>
<td>Bibliography</td>
<td>109</td>
</tr>
</tbody>
</table>
List of Figures

Figure Number  Page

2.1  Schematic flowchart of the classification process..................................................11

3.1  Monthly histograms of the occurrence of each state. The number in each panel’s title indicates the total number of occurrences of each state.............................................37

3.2  1000 mb dew point (shaded, °C) and 1000 mb wind vectors for the eight atmospheric states. Black lines show the land-sea boundaries. Latitude and longitude are labeled on the axes. Darwin is at the center of the domain..................38

3.3  Surface pressure anomaly (mb, shaded) and 750 mb wind vectors for the eight atmospheric states. Anomalies are calculated from the mean of the full input dataset. Black lines show the land-sea boundaries. Latitude and longitude are labeled on the axes. Darwin is at the center of the domain.................................39

3.4  500 mb relative humidity (fraction, shaded) and 500 mb wind vectors for the eight atmospheric states. Black lines show the land-sea boundaries. Latitude and longitude are labeled on the axes. Darwin is at the center of the domain..................40

3.5  Vertical profiles of hydrometeor occurrence by atmospheric state, defined as the fraction of the time clouds or precipitation are detected with a reflectivity of -40dBZ or larger by the millimeter wavelength cloud radar at the ARM site in Darwin. Dry season states are shown in shades of red, transition states in green, and monsoon states in blue. Shaded areas denote the 95% confidence limits, as determined by the bootstrap resampling method. The numbers in the legend indicate the number of 3 hour blocks with radar observations for each state.........................................................41
3.6 Time series of atmospheric state (upper panel) and daily precipitation (lower panel) for the 1981-1982 monsoon season. Upper panel shows our monsoon onset and retreat (vertical green lines) and active periods (blue bars) as defined in Section 3.2. Lower panel also shows onset defined by Hendon & Liebmann (1990, vertical red line), and onset and retreat defined by Drosdowsky (1996, vertical blue lines).....42

3.7 Time series of monsoon onset (lower solid lines) and retreat (upper solid lines) according to this study (black), Drosdowsky (1996, blue), and Hendon and Liebmann (1990, red). The dashed black line shows the average September-October-November Southern Oscillation Index. The year listed on the horizontal axis is the starting year of the monsoon season, i.e. 1960 is the 1960/1961 season. .................................................................43

3.8 Time evolution of the correlation coefficients between monthly average SOI and the number of active monsoon days (solid), onset date (dashed), and retreat date (dotted) for the years preceding and following the monsoon. Horizontal dashed lines indicate 95% significance levels, estimated from the Fisher-Z transformation. Vertical red lines indicates mean onset and retreat dates........................................44

3.9 The relative occurrence of different states for each MJO phase during the monsoon season. Monsoon season is defined as onset to retreat. Shaded regions indicate 95% confidence limits as estimated by a bootstrap algorithm operating on 33 years of data. The overall fraction of the monsoon season occupied by each state/category is shown in the legend. Days with weak MJO phase (36% of all monsoon days) are excluded from this figure. Weak days are 67% active, 28% break, 5% other, thus excluding these days does not notably alter the distribution of states. ..................45
3.10 Distribution of MJO phase for the first day of each active period (dark blue), and the last day of each active period (light blue). Shading represents 95% confidence limits, as calculated by a bootstrap algorithm. Days with weak MJO phase are excluded from this figure, eliminating 38% of first and last days. ........................................46

3.11 Domain mean values of zonal wind (upper panel, black, left axis), meridional wind (upper panel, red, left axis), and relative humidity (upper panel, blue, right axis) at 875 mb (solid lines) and 500 mb (dashed lines). Mean daily precipitation at Darwin Airport (lower panel). Values are composited relative to the first day of each active period. ........................................................................................................47

3.12 Frequency of occurrence of atmospheric states over the ERA-Interim period, calculated as the anomaly, in days, from the mean. Years are calculated from the midpoint of one calendar year to the midpoint of the next. Labeled years are the first calendar year of each year period, e.g., 1983 indicates the July 1983 to July 1984 period. Red lines indicate states with statistically significant trends. The mean number of occurrences of each state is shown in the panel titles. ........................................48

3.13 Annual rainfall at Darwin Airport. Years are calculated from the midpoint of one calendar year to the midpoint of the next. Labeled years are the first calendar year of each year period, e.g., 1980 indicates the 1980-81 period. The trend over the full period (6.5 mm/yr, p=0.004) is shown in red, while the trend over the period of this study (9.5 mm/yr, p=0.23) is shown in blue.................................................................49

3.14 Contributions to the 33 year trend in annual rainfall at Darwin Airport. Blue bars are the contribution due to trends in within-state mean precipitation and red bars are the contribution due to trends in the frequency of occurrence of the states.....50
4.1 Average meteorological conditions for four states from the SGP classification which often follow each other in sequence. The ARM SGP site is at the center of the domain in Oklahoma. Upper row shows 875 mb dew point (°C) and winds; lower row shows surface pressure anomaly (mb) and 750 mb winds. States 12 and 18 show a cold front and low pressure system sweeping across the region, while States 14 and 6 show cold air spilling in from the north as high pressure builds in the wake of the frontal passage. .................................................................78

4.2 Monthly histograms of occurrence for each atmospheric state. Panels with blue backgrounds occur primarily in winter, red background in summer, white backgrounds in spring/fall. The number in each panel title is the number of occurrences of each state. .................................................................79

4.3 Four measures of mean state temperature (°C). Blue curves are near-surface (875 mb) values at the SGP site (light blue) and across the domain (dark blue). Vertically averaged temperatures are shown at SGP (red) and averaged across the domain (black). .................................................................80

4.4 Four measures of mean state relative humidity (RH). Blue curves are near-surface (875 mb) values at the SGP site (light blue) and averaged across the domain (dark blue). Vertically averaged values are shown at SGP (red) and averaged across the domain (black). .................................................................81

4.5 Vertical profiles of cloud occurrence as observed by the vertically pointed millimeter cloud radar at the ARM SGP site for each atmospheric state. Cloud occurrence is defined as the fraction of time a particular altitude has reflectivity greater than -40 dBz. Solid black lines show the mean values, while grey shaded
areas represent the bootstrapped 95% confidence limits. Panels with blue backgrounds occur primarily in winter, red background in summer, white backgrounds in spring/fall. The number in each panel title is the number of observations of each state. ..........................................................82

4.6 Probability of transition from any state (rows) to any other state (columns). Each row sums to 1..................................................................................................................83

4.7 Conceptual flow chart of the progression of atmospheric states. The outer loop is comprised of states that occur primarily during winter. The inner loop represent states the occur primarily in spring and fall.................................................................84

4.8 Monthly domain mean temperature at different pressure levels (different colors) from Era-Interim (solid lines), the GFDL AM3 at 2° resolution (dashed lines), and at 0.5° resolution (dotted lines)..........................................................................................85

4.9 Distribution of occurrence of states in observations (black) and in the 2° run of the AM3 using various normalization techniques (colors). Normalization techniques shown are no bias correction (blue), annual bias correction (red), seasonal bias correction (green), and histogram renormalization (magenta). Different techniques are described in Section 4.2.1 and Table 4.1. Values in the legend are the RMS error between each normalization and the observed distribution.........................................................86

4.10 Vertical profiles of cloud occurrence for each of the 21 states in the 2° AM3 for different normalization techniques (see Section 4.2.1 and Table 4.1). Background shading for each panel indicates states that occur primarily in summer (red), winter (blue), or spring/fall (white). Shaded areas indicate 95% confidence interval based
on a bootstrap resampling test. Shaded areas overlap at essentially all levels, as
described in Section 4.2.2 and summarized in Table 4.2.................................87

4.11 Joint cloud top pressure (vertical axis) – optical depth (horizontal axis) histogram of
cloud occurrence at the SGP site, as observed by ISCCP. Values are in percentage
points, with the sum of the histogram (the total cloud fraction at the site) reported
at the top of the figure.................................................................88

4.12 Joint cloud top pressure (vertical axis) – optical depth (horizontal axis) histograms
of cloud occurrence at the SGP site, as observed by ISCCP and composited by
atmospheric state. Values are in percentage points, with the sum of the histogram
(the total cloud fraction at the site) reported at the top of each panel. Color of panel
title indicates seasonality of state, with red indicating summer, blue indicating
winter, and black indicating spring/fall........................................89

4.13 As in Figure 4.11, but for the 2° run of the AM3, using output from the ISCCP
simulator..............................................................90

4.14 As in Figure 4.12, but for the 2° run of the AM3, using output from the ISCCP
simulator......................................91

4.15 Model bias of cloud occurrence at the SGP site, displayed as a joint histogram of
cloud top pressure and optical depth. Values are the difference, in percentage
points, between ISCCP simulator output from the 2° run of the AM3 and the
observed ISCCP values; positive represent excess cloud in the model. Summing over
the bins of the histogram produces the total cloud occurrence bias of the model, -24%,
shown at top of figure. Asterisks indicate histogram bins whose bias values
are significant at 95% confidence, according to a bootstrap resampling test........92
4.16 Model bias of cloud occurrence at the SGP site for each atmospheric state, displayed as a joint histogram of cloud top pressure and optical depth. Values are the difference, in percentage points, between ISCCP simulator output from the 2° run of the AM3 and the observed ISCCP values; positive represent excess cloud in the model. Summing over the bins of the histogram produces the total cloud occurrence bias of the model for each state, shown at top of figure. Asterisks indicate histogram bins whose bias values are significant at 95% confidence, according to a bootstrap resampling test.

4.17 Distributions of outgoing longwave radiation as observed by CERES (black) and in the 2° run of the model using no bias correction (blue). Shaded regions indicate 95% confidence interval, as determined by a bootstrap resampling test. Blue shaded panels are wintertime states, red shaded panels summertime states, and white panels are spring/fall states. States whose model mean OLR is statistically different from the observed mean have that bias shown in the panel title.

4.18 Contributions to the total bias in high thin cloud occurrence in the 2° run of the AM3, broken down by state and source (see Equation 4.1). Blue bars indicate contributions from state mean properties, green bars from relative frequency of occurrence, and red bars the second order contribution. The numbers in the legend indicate the mean magnitude and mean value for each bar type.

4.19 As in Figure 4.18, but for high thick cloud occurrence.

4.20 Occurrence of the atmospheric states in ERA-Interim (black), the 2° run of the AM3 (blue), and the 0.5° run of the AM3 (purple). Shaded areas are 95% confidence intervals, calculated with bootstrap resampling of annual distributions. Values in
the legend are RMS errors between each model distribution and the reanalysis distribution.

4.21 Difference in cloud occurrence, as reported by the ISCCP simulator, between the high and low resolution runs of the AM3. Blues indicate greater cloud occurrence in the high resolution run; reds indicate less. The sum of each histogram is the total change in cloud occurrence for a given state, and is reported in the panel title. 

4.22 Cumulative distributions of CAPE at the SGP site for each atmospheric state from ERA-Interim (black), the low resolution AM3 run (blue), and the high resolution run (red). Dashed vertical black line marks 1000 J/kg, the cutoff threshold for the deep convective parameterization in the AM3. Values in the legend indicate the percentage of CAPE values greater than this cutoff value.

4.23 Change in the magnitude of the total cloud occurrence bias for each state in the 0.5° model run as compared to the 2° run. Negative values indicate improvement; positive values indicate worsening. States are grouped and ordered as in Tables 4.3-4.5.

4.24 As for 4.18, but for high thick cloud occurrence in the high resolution model run.
List of Tables

<table>
<thead>
<tr>
<th>Table Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Cumulative precipitation, as a percentage of the total July to July precipitation that has occurred at the time of monsoon onset and retreat, according to various definitions. E12 refers to onset and retreat dates defined in this study, D96 to dates from Drosdowsky (1996), and H90 to dates from Hendon and Liebmann (1990). The mean value and standard deviation is given for each definition.</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>Percent of the variance in annual (July-July) precipitation at Darwin over the 33 year record that is explained by (1) AnnO, the annual occurrence of each state, measured in days, and (2) AvgP, the annual mean precipitation associated with each state. The last three columns are combined values of $R^2$, <em>i.e.</em>, the combined annual occurrence and weighted mean precipitation for multiple states. Dry refers to the combined occurrence and precipitation of States 1 to 4, Trans. to the combination of States 5 and six, and Mons. to the combination of States 7 and 8.</td>
<td>36</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary of different model output normalization techniques.</td>
<td>73</td>
</tr>
<tr>
<td>4.2</td>
<td>Statistics regarding the similarity of state cloud profiles in the AM3 for different normalization techniques, relative to the no bias correction normalization. Mean $r$ refers to the average of 21 correlation coefficients between the state cloud profiles of the no correction normalization and each of the other techniques. Mean RMS is the mean of 21 RMS errors between the no correction cloud profiles and those of the other techniques. Percent different is the percentage of cloud levels for the 21 states where the profile's 95% confidence intervals (from bootstrap resampling) do not overlap with the same interval of the no correction normalization.</td>
<td>74</td>
</tr>
</tbody>
</table>
4.3 Observed ISCCP values of cloud occurrence by atmospheric state and state category (as described in Section 4.1). Values are summed over quadrants of the ISCCP joint histogram, where high and low cloud is defined as cloud tops above and below 560 mb respectively, and thick and thin cloud is defined as optical depths greater and less than 9.4. ...........................................75

4.4 As for Table 4.3, but modeled values from the 2° run of the AM3..............................76

4.5 Cloud occurrence bias by atmospheric state and state category (as described in Section 4.1). Values are the difference between the 2° model and observation, i.e negative indicates a lack of cloud in the model. Values are summed over quadrants of the ISCCP joint histogram, where high and low cloud is defined as cloud tops above and below 560 mb respectively, and thick and thin cloud is defined as optical depths greater and less than 9.4. Table entries in red are statistically significant at 95% confidence according to a bootstrap resampling test.................................77
Acknowledgements

I would like to thank my advisors, Tom Ackerman and Roger Marchand for all their help, guidance, and patience over the years. I’ve learned an incredible amount from them and feel very fortunate to have had two advisors invested in me. I was both pointed towards interesting questions and encouraged to explore topics that piqued my interest, and for that I am grateful. I thank my GSR, Marina Meila, for her thoughtful contributions, and teaching me more about clustering. I am grateful to Dargan Frierson and David Battisti, not only for serving on my committee, but also for engaging with me on a personal level my entire time in the department. I am very appreciative of Leo Donner for his interest in my work and my career, and for arranging the collaboration that made much of this dissertation possible. Special thanks are due to Kyle Armour, Ben Hillman, Sandra Penny, Steve Po-Chedley, Alyssa Atwood, Mike Warner, Kelly McCusker, and Zoe Harrold for countless conversations on science, work, and life. Lastly, I would like to thank my family for their love and support throughout my life.
Chapter 1

Introduction

General Circulation Models (GCMs) have difficulty representing clouds and cloud impacts on climate. This is due, in large part, to the fact that the processes that control the creation, dissipation, and radiative properties of clouds occur on scales much smaller than the typical resolution of a GCM. This inability to explicitly resolve cloud processes necessitates the use of parameterizations to make statistical predictions of cloud properties. Given the broad variety and complexity of clouds found in nature, it is not surprising that these parameterizations produce errors in the statistical distributions of cloud properties when compared to observations. Determining the sources of errors in the parameterizations can be challenging, especially in GCMs (Jakob 2010). Unlike numerical weather prediction models, GCMs do not predict the specific sequence of weather events that any given location experiences, which makes it difficult to directly compare model output to observations at a particular time. A common way of addressing this problem is to create long-term temporal averages of both model output and observations, which can then be compared to each other. Long-term temporal averages can frequently determine the presence of errors. Unfortunately, it conflates the errors that are produced when the model fails to represent the distribution of synoptic scale conditions correctly with those produced by a misrepresentative cloud parameterization. Temporal averages also do not provide much information regarding the meteorological conditions that lead to the occurrence of the errors. Information regarding the specific weather types that contribute most to the errors is lost to the averaging process.
An alternate approach that can address these issues is to composite data by atmospheric state, rather than by season. Comparing model output with observations by atmospheric state can detect compensating errors and allows for the separation of errors due to modeling of the synoptic conditions from errors due to the cloud parameterization. Comparison of the relative frequency of atmospheric states in models and in observations provides information regarding how well the model represents the state of the atmosphere. Comparing modeled cloud properties with observed cloud properties within a particular state can help detect parameterization errors; when an error is detected, the meteorological conditions under which the error occurred are to some degree known.

There is a variety of ways of defining atmospheric states. One way is to use a particular weather variable or variables as a basis for compositing observations and model output. Bony and Dufresne (2005) used 500 mb vertical velocity in their study of tropical cloud feedbacks, while others, such as Franklin et al. (2013) combine 500 mb vertical velocity with lower tropospheric stability or sea surface temperature to create joint histograms of cloud properties for models that can be compared to observations. Another method of defining atmospheric states involves applying a clustering algorithm or principal component analysis to a set of weather observations obtained, for example, from radiosondes (Pope et al. 2009, Zivkovic & Louis 1992) or reanalysis (Fereday et al. 2008, Marchand et al. 2006 and 2009). Studies such as these can ask questions regarding the effect of different atmospheric conditions on clouds, radiation, and precipitation.

Recently, clustering approaches have also been applied to cloud observations, such as the joint histograms of cloud top pressure and optical depth provided by the International Satellite Cloud Climatology Project (ISCCP), to create a set of cloud regimes which can be used in
variety of ways. In doing so, these studies answer questions regarding the effects of the presence or lack of different cloud types. For example, cloud clustering techniques have been used to characterize cloud and radiative properties in the tropics (Jakob et al. 2005, Rossow et al. 2005, Jakob & Schumacher 2007), and over the Southern Ocean (Haynes et al. 2011, Mason et al. 2014). By combining data from different satellites, the method can even be applied to the entire globe (Oreopolous et al. 2014). Observationally based cloud regimes have also been used to evaluate cloud properties in GCMs (Zhang & Lin 2005, Williams & Webb 2008).

In this thesis we apply a classification technique developed by Marchand et al. (2009), hereafter M09, to a pair of regions in tropical Australia and Oklahoma. The Australian region is centered on the Department of Energy’s (DOE) Atmospheric Radiation Measurement (ARM) program site in Darwin, while the Oklahoma region is centered on ARM’s Southern Great Plains (SGP) site. Our interest in this technique stems from its use of a combination of reanalysis and radar observations to effectively identify common recurring meteorological states. An unresolved issue for most clustering or classification techniques is trying to determine how many states exist. In the M09 approach, the classification is applied to the reanalysis data and the associated radar observations are used to ensure that the identified atmospheric states are statistically meaningful. This method which was previously applied at the SGP site in Oklahoma (Marchand et al. 2006, M09). We describe the method, and the input data, in Chapter 2.

Once defined, atmospheric states can also be applied to understanding large-scale climate features. Examples include the subtropical marine stratocumulus zone in the North Pacific (Norris and Iacobellis 2005), midlatitude storm tracks (Gordon et al. 2005), and the Australian monsoon (Catto et al. 2012). The Australian monsoon is dominant source of water to
northwestern Australia (McBride and Nichols, 1983) and as a major feature of the tropical circulation is both an important and interesting phenomenon to study. In Chapter 3 we use the classification at Darwin to examine the variability, on both interannual and intraseasonal time scales, the predictability, and trends in the Australian monsoon over the last 3 decades. Chapter 4 uses the SGP classification to evaluate the representation of clouds in one IPCC climate model (GFDL AM3) under a variety of weather conditions. We will identify the states under which the model struggles to represent clouds, likely physical processes involved, and evaluate how each state contributes to the model’s total (temporally averaged) error. We conclude in Chapter 5 with a summary of findings and discussion of further applications of classification techniques.
Chapter 2

Clustering Method & Data

Our classification approach is described in detail in Marchand et al. (2009, hereafter M09) and is briefly summarized here. The technique is essentially the same for both the classification at Darwin and at Southern Great Plains (SGP), though there are slight differences in the inputs and parameter values which will be made clear. The approach has two stages, which are shown schematically in Figure 2.1. In the first stage, meteorological variables from a numerical weather prediction analysis are input to a competitive neural network. In this study we use data from the ERA-Interim reanalysis project as the input. For the Darwin classification the reanalysis is sampled eight times daily, as this data was available to us at the time, while at SGP the classification uses the publicly available four times daily data. Following M09, we use relative humidity, temperature, and the horizontal winds at seven predetermined (sigma) pressure levels as well as the surface pressure for 81 horizontal grid points, arranged in a 9 x 9 grid centered on the ARM site, either Darwin or SGP. At Darwin each gridbox covers an area of 2° x 2.5°, while at SGP the data is on ERA-Interim’s standard 1.5° x 1.5° grid. After defining states at Darwin we converted the state definitions to the ERA-Interim grid and four times daily frequency in order to determine atmospheric state over the full ERA-Interim time period. We performed this conversion by using the temporal overlap to determine member snapshots and thus definitions of the states. Four three-dimensional variables on a 9 x 9 x 7 grid plus surface pressure makes for a snapshot consisting of 2349 input variables at each time step. For the Darwin study we
used approximately 4 years of snapshots, from June 2006 through April 2010, while at SGP we used over 13 years worth, from December 1996 through March 2012, matching the availability of cloud radar data from the ARM sites at these locations.

The first stage in state classification is to use a neural network to divide the data into a predefined number of atmospheric states, where an atmospheric state may be thought of as a frequently occurring dynamic and thermodynamic pattern. Each state represents a group, or cluster, of similar observations. Essentially, each atmospheric state describes a specific weather pattern that is representative of a large number of snapshots. At Darwin, we instruct the neural network to define 15 states, while at SGP we instruct it to define 40 states in anticipation of finishing with more states at the more synoptically variable mid-latitude site. Mathematically, each state is defined by a vector of meteorological values for each of the input variables. To the neural network, the best representation of the input space is given by the set of state definitions such that the sum of the distances between each input vector (i.e., the meteorological variables at each time step) and the closest state definition is minimized. Distance is defined as the sum of the absolute values of the input vector elements minus the same elements in the state definition relative to the standard deviation of each element (calculated from the entire input set). As the above definition suggests, each meteorological observation (the complete set of input variables at a particular timestep) is associated with the closest atmospheric state definition, which then determines the members of each cluster.

The second stage is the evaluation stage, when each of the atmospheric states are evaluated to determine if the state has cloud properties that are temporally stable and distinct from every other state. Cloud properties are not input to the neural network, and
allow for a test of state quality using data that is independent of the state creation process. In order for the identified atmospheric states to be useful in the analysis of climate model cloud properties, we require that the associated distribution of observed cloud properties be statistically stable. This implies that every time the state occurs the observed cloud properties may be thought of as a random realization taken from a fixed distribution. We evaluate temporal stability by comparing the mean profile of hydrometeor (i.e. cloud, rain, and ice) occurrence in the first half of the dataset with that from the second half of the dataset. Hydrometeor profiles are created by aggregating millimeter wavelength cloud radar observations of reflectivity (Clothiaux et al. 2000) from the ARM sites according to the atmospheric states. The profiles are compared using a statistical hypothesis test based on a moving-blocks bootstrap resampling technique (Marchand et al. 2006). If the hydrometeor occurrence profile in the first half of the dataset differs statistically from that of the second half of the dataset then that state is either divided into two parts (increasing the total number of states by one) or eliminated (reducing the total number of states by one). If all the states are found to be stable, then a second test is applied, requiring that each state’s hydrometeor occurrence profile be distinct from the others. Again, states that fail this test are either divided or eliminated. The process of dividing or eliminating states is repeated until all of the remaining states are temporally stable and distinct. The decision on whether to divide or eliminate a particular state is based on the size of the state. In M09 a threshold of 6% of the total number of observations was used. In the Darwin study, we found we had to increase the threshold to 7% in order for the process to consistently converge, while at SGP we were able to lower it to 3%. We consider a lower threshold
desirable so long as the process still converges to a final set of states, as it increases the number of states the classifier is able to discern.

At Darwin we examined the sensitivity of the results to the choice of inputs and algorithm parameters. We varied the size threshold for dividing/deleting bad states, the list of input variables and their horizontal resolution, and implemented a screen for outliers. All these sensitivity tests were performed on half the dataset.

The size threshold used in the iteration stage affects the number of states produced by the classifier because it decides whether a bad state is eliminated or divided. Lower values for this parameter force the classifier to try to find smaller clusters by more frequently choosing to subdivide rather than eliminate states. However, if one tries to find more states than can be supported by the data, then the classifier struggles to converge. For example, if a good state is artificially divided into two parts, both halves will likely have similar (i. e., non-distinct) cloud profiles causing the division to be rejected. Thus, the goal when setting the threshold value is to set it as low as possible (in order to obtain as many good states as possible) while still achieving convergence on a set of stable and distinct states.

Another design choice that might affect the results is the selection of the input data. To test this, we ran the classifier multiple times on datasets with a number of variations from the standard case described above. We tested for the importance of variable selection by adding either geopotential height or divergence to the list of variables taken from the ECMWF reanalysis and tested for the importance of domain size by expanding our grid of points from 9 x 9 to 9 x 18 and shrinking it to 5 x 5. We also tested the importance of horizontal resolution, by creating datasets that covered the same region as the base case,
but at $1^\circ \times 1^\circ$ resolution and $0.5^\circ \times 0.5^\circ$ resolution. All six of these variants produced very similar results to those described in 3.1. A similar insensitivity to domain size and horizontal resolution was found by Marchand et al. (2009) for the SGP site.

We also investigated the extent to which our states are affected by outlier observations. Presently, a state is defined as the mean of all its members, and all observations are assigned to a state. Some observations are poor matches for any of the states, and our methodology forces them into a state and to contribute to the overall statistics. We experimented with removing the 5% furthest outlying observations from each state and recalculating the statistics for each state. By doing so, we were essentially adding a requirement that members of each state be relatively tightly clustered near the centroid. We found that removing the outliers from each state did not notably alter the statistics for each state. As such, it appears that the state definitions are not strongly affected by outliers.

The results of the competitive neural network (stage 1) are slightly dependent on the order in which input is presented to the neural network as well as the iterative state refining process (stage 2). In order to test for robustness, we ran the classification process ten times with different random seed values and looked for states that occurred often. At Darwin, each of the ten runs produced eight to ten states, most of which had a close match in most of the runs. To objectively determine which states were robust, we performed a cluster analysis on all of the states produced by all 10 runs. We varied the number of clusters for this stage from 8 to 12 as we looked to maximize the number of clusters that are robust, defined to be a cluster that has members from at least eight of the ten runs. The largest number of robust clusters we were able to create is eight, and their means are the
state definitions described in Chapter 3. At SGP, each of the individual runs produced 25 to 33 states, with 21 robust states that will be presented in Chapter 4. Requiring that each state occur repeatedly means that while the states presented here are not identically reproduced in each run of the classifier, they comprise a good representation of the possible outcomes of the classifier.
Figure 2.1

Schematic flowchart of the classification process.
Chapter 3

Monsoon variability analysis at Darwin

Northern Australia experiences a strong monsoon cycle over the course of the year. During austral winter, dry continental air from the southeast makes rainfall both rare and weak. Beginning in September, a transition period gradually increases temperature and moisture in the region, leading to more clouds and precipitation. The monsoon is generally considered to begin at some point in December when the winds become steady westerlies, bringing large amounts of tropical moisture to northern Australia (Hung and Yanai 2004). The vast majority of northern Australia’s rainfall comes during the monsoon wet season, continuing from December into March or April (McBride and Nicholls 1983). During this period the monsoon occasionally undergoes break periods, during which winds are easterly and rainfall is less frequent (Holland 1986). As the major source of rainfall for the region, variability in the onset, retreat, and intensity of the monsoon is of substantial importance to northern Australia.

Predictability of various metrics of the Australian monsoon has been an active topic of research for decades. Onset date and its relationship to the El Nino – Southern Oscillation (ENSO) is a particularly well-studied feature of the monsoon. There are many different definitions of monsoon onset. Nicholls et al. (1982) created precipitation-based definitions of monsoon onset, defining onset as the date by which various cumulative precipitation totals have been reached at Darwin, Australia. Others, such as Holland (1986), Drosdowsky (1996), and Kajikawa et al. (2010), have used wind based definitions
to identify monsoon onset: for example, when the zonal winds, either at a particular level at Darwin (Holland 1986), vertically averaged (Drosdowsky 1996), or averaged across the region (Kajikawa et al. 2010), reverse direction and become westerly. Still others have created multivariate definitions to identify monsoon onset. Hendon and Liebmann (1990), for example, defined onset as the first simultaneous occurrence of westerly winds and widespread rainfall in northern Australia.

Regardless of the definition, it is universally agreed that there is a strong relationship between onset date and ENSO activity. Sea surface temperatures (SSTs) around northern Australia are affected by ENSO. During La Nina, SSTs are greater, leading to enhanced convection and precipitation (Catto et al., 2012). As a result, La Nina conditions in the months leading up to onset (September - November) are usually associated with an earlier monsoon onset. La Nina conditions also tend to favor a more intense monsoon season, whether the monsoon seasonal intensity is measured as total seasonal rainfall (e.g., McBride and Nicholls 1983), number of active days (Drosdowsky 1996), or strength of an index based on westerly wind speed in the region (Kajikawa et al., 2010). In contrast, the retreat date for the monsoon has not been found to be correlated with ENSO indices at any lead time (e.g. Drosdowsky 1996), and the causes of its variability are, to the best of our knowledge, unknown.

The Madden-Julian Oscillation (MJO) is the major source of variability in the tropics on intraseasonal timescales (Madden and Julian 1971, Zhang 2005). Previous studies have focused on two characteristics of the Australian monsoon that may be affected by the MJO: onset date and the frequency of active periods. Hendon and Liebmann (1990) composited zonal winds and outgoing longwave radiation (OLR) at various leads and lags of their
monsoon onset date to show that the monsoon begins with the passage of an MJO event that begins in the Indian Ocean. An analysis of the frequency of active periods is limited to those studies and techniques that define individual active periods within a season, not just an onset date. Of those who do define active periods, Holland (1986) found that the mean interval from the beginning of one active period to the next is 40 days, implying that the MJO plays a role in the timing and duration of active periods; however, Drosdowsky (1996) challenged this, showing that there is no preferred duration or interval time for active periods under his definition, and speculated that the low-pass filtering of winds in the Holland study contributed to the apparent presence of a 40 day period. Both of these studies predate the very useful definition of MJO phase by Wheeler and Hendon (2004), which provides a new way of analyzing the timing of active periods, complementing the previous methods that used spectral analysis to detect the affects of the MJO. To the best of our knowledge, the role of the MJO in regulating active periods of the monsoon remains an unresolved question, and an analysis of the timing of individual active and break periods relative to the phase of the MJO has not been undertaken.

Observations suggest that over recent decades precipitation has increased in northern Australia. Hennessey et al. (1999) found a significant increase in the number of days with rain over the 20th century, while Taschetto and England (2009) found that the positive trend in northern Australia precipitation from 1970-2006 is due primarily to an increase in the number of deep convection rainfall events. Whether these trends are due to changes in the frequency of occurrence of various circulation patterns or changes in the number of deep convection events associated with particular circulation patterns is an interesting question. Catto et al. (2012) use a regime classification to address this question
and conclude that an increase in the frequency of occurrence of monsoon regimes was responsible for the trend from 1957-2007. Evans et al. (2012) also define a classification for the northern Australia region that can be used for this purpose, and could provide additional insight into whether it is changes in the frequency of circulation patterns or the precipitation associated with them that is responsible.

3.1 Darwin States

In this section, we describe the dynamic and thermodynamic fields that define each of the eight states that we identified by our classification scheme (Chapter 2), as well as the associated cloud structures. We use these eight states as the basis for further analysis in later sections. We refer to the states by number, though these numbers are arbitrary, and have been assigned so as to present the states in a fluid way. As shown in Figure 3.1, States 1 through 4 are dry season states and occur primarily from April to September, States 5 and 6 are transition season states, occurring mostly from March to May and September to November, and States 7 and 8 are monsoon season states, occurring almost exclusively between December and April. The mean meteorological characteristics of each state are shown in Figure 3.2 for the surface dew points and winds, in Figure 3.3 for the surface pressure anomalies and 750 mb winds, and in Figure 3.4 for the 500 mb relative humidity and winds. Figure 3.5 shows the mean profile of hydrometeor occurrence (from the ARM millimeter cloud radar) for each of the eight states, clearly showing that the two monsoon season states have more clouds and precipitation than any of the dry or transitional states. Collectively, the figures show a clear delineation between the hot, moist, ascending monsoon season and the cooler, arid, subsiding dry season.
3.1.1 Dry Season States

States 1-4 occur nearly exclusively during the months of April to September. While all four of the states occur in all the months of the dry season, they peak in occurrence at different times during the season (Figure 3.1). States 1 and 2 occur most frequently toward the beginning of the dry season, peaking in occurrence during the months of May and June. State 3 is most common in the heart of the dry season during July and August. State 4 becomes increasingly common as the dry season ends and the monsoon begins to build, peaking in September and continuing to occur, albeit infrequently, throughout October and November. All four dry season states have a number of features in common. They all have positive surface pressure anomalies (Figure 3.3) throughout the region, with a gradient from highest pressure in the south to lower pressure in the north. In addition, they all exhibit southeasterly flow at the surface throughout most of the region (Figure 3.2), and anti-cyclonic flow in the upper atmosphere, but with winds that differ above the Darwin site.

State 1 is a suppressed convection state. It is by far the most humid of the dry season states (Figure 3.4, lower levels not shown), and is nearly as cloudy as the monsoon states at lower levels (Figure 3.5). While still positive, it has the weakest surface pressure anomaly of the dry season states. At 750 mb the flow is southeasterly at the Darwin site, while at 500 mb the flow has reversed and become northwesterly. This northwesterly flow brings moisture from the equator, resulting in upper level humidities that are much greater than those of the other dry season states (Figure 3.4). State 1 is the cloudiest of the dry
season states, with most of its cloud occurring below 7 km suggesting little deep convection.

States 2 and 3 are the driest states, with State 2 being merely dry and State 3 being very dry. Both have large positive surface pressure anomalies, large dew point depressions at the surface and low humidities in the upper atmosphere. States 2 and 3 are differentiated primarily by their 500 mb flow (Figure 3.4). At 500 mb, State 2 has almost no meridional component, with easterlies over the Indonesian seas, westerlies over the continent, and weak flow at Darwin. The 500 mb flow of State 3 is drier and has a much stronger meridional component at Darwin, with southwesterly flow over the continent and an anti-cyclonic circulation pattern over the region. State 2 produces a very small amount of cloud, primarily above 7 km. State 3 produces almost no cloud at all.

State 4 is an isolated convection state. The flow is easterly at Darwin, and there is little gradient in surface pressure. The surface flow diverges when it reaches the continent, producing northeasterly winds over the continental interior that are similar to the flow in the monsoon states. It has surface temperatures nearly as hot as those of the monsoon season states (not shown), making it by far the hottest of the dry season states. While not nearly as moist as State 1, it is still more humid than either State 2 or State 3, and features a bimodal cloud height distribution with distinct peaks near 1.5 and 12 km. As the changes in surface temperature and flow suggest, this state occurs most often toward the end of the dry season and during the build-up to the monsoon.

3.1.2 Transition Season States
States 5 and 6 both occur primarily during the months of March to May and September to November (Figure 3.1). State 5 occurs approximately equally in both shoulder seasons, while State 6 occurs almost entirely in the pre-monsoon transition, and continues to occur, somewhat less frequently, into December. Both states have high surface temperatures equal to those of the monsoon season (not shown) but dry humidity profiles that are similar to dry season states (see for example at 500 mb, Figure 3.4). State 5 is somewhat more humid, while State 6 is slightly hotter (not shown). Both states have close to neutral surface pressure anomalies with weak horizontal gradients. The two states are most readily distinguished by their flow patterns. At Darwin, State 5 is southeasterly at the surface and at 750 mb (Figures 3.2 and 3.3) before becoming southwesterly at 500 mb (Figure 3.4). State 6 has primarily easterly flow at the surface, though there is westerly onshore flow on the western coast of the continent. At 500mb, State 6 features strong southeasterly flow at Darwin with a strong anti-cyclone centered in the western part of the domain. States 5 and 6 have very similar hydrometeor occurrence profiles (Figure 3.5), though the statistical test indicates that the larger hydrometeor coverage around 10 km in State 5 is significant (at the 95% confidence interval) - consistent with higher upper-level humidity. In both states, the profiles are double-peaked at approximately 2 and 13 km, with greater occurrence than the dry season states and much less than the monsoon states.

3.1.3 Monsoon Season States

States 7 and 8 are monsoon season states, occurring almost entirely between December and April (Figure 3.1). State 7 occurs more frequently in the core of the monsoon season, in January and February, and represents the monsoon during its active
monsoon phase. State 8 occurs more often on the edges of the monsoon in December and March and represents a break monsoon condition. Both have high surface temperatures, high humidities and negative surface pressure anomalies. They differ in their flow patterns however. At Darwin, State 7 has westerly, slightly cyclonic flow at the surface, which becomes strongly cyclonic and slightly northwesterly at 500 mb. The strong and moist onshore flow makes State 7 the most humid of all the states. State 8 has easterly flow at the surface which becomes anti-cyclonic by 500 mb. In this way it resembles a more humid version of State 6, though the center of the anti-cyclone is further south. States 7 and 8 are by far the cloudiest and second-cloudiest states, respectively, with a primary peak in the occurrence profile at 12 km, and a smaller peak at approximately 2 km (Figure 3.5).

3.2 Defining monsoon metrics

We determine the timing of monsoon events using the occurrence of atmospheric states for the northern Australia region described above and in Evans et al (2012). Here we focus on States 7 and 8, which occur primarily during the monsoon season from December to April. Evans et al. (2012) identified State 7 as the active monsoon and State 8 as the monsoon break based on their composite meteorology, cloud occurrence profiles, and rain rates.

Because our state definitions are based on ERA-Interim data, we are able subsequently to classify every time step of the Interim period into one of the eight states by calculating the Pythagorean distance from each reanalysis snapshot to each of the eight state definitions. The snapshot is then classified as the state to which it is closest. In doing so, we create a 34-year time series from 1979-2012 of atmospheric state for the region
surrounding Darwin (containing 33 complete monsoon seasons). This time series allows us to create simple metrics describing the Australian monsoon. We define the onset of the monsoon as the first time each season four consecutive snapshots (24 hours) are classified as the active monsoon. Similarly, monsoon retreat is the last time each season that there are 24 continuous hours of active monsoon. During the monsoon season, the beginning and end of individual active periods are identified the same way. We describe the intensity of the monsoon season by two metrics: the number of active days each year, and the fraction of all monsoon days each year (State 7 and State 8) that are classified as active (State 7).

As an example, the upper panel of Figure 3.6 shows the time series of atmospheric state for the 1981-1982 season, with the onset, retreat, and active periods as defined by the metrics above. The atmospheric state begins the year varying amongst the dry season states, spends much of September and October in the transitional states, and then nearly exclusively inhabits one of the two monsoon season states from onset in late November until April. The lower panel shows the daily precipitation for the season, demonstrating how it increases in frequency and intensity during active periods of the monsoon. We discuss the daily precipitation further in Section 3.3.

3.3 Interannual variability

Our atmospheric classification technique provides rather different monsoon definitions than those used in previous studies, and is not specifically tailored to identifying monsoon activity. Rather, as described in the previous section, it is based on a set of objective tests (stability and distinctiveness of radar profiles) and was originally designed
and applied at mid-latitudes. The technique includes no expert knowledge on the meteorology of Darwin and was not, for example, specifically guided toward identifying active and break periods of the Monsoon. We therefore consider it important to verify that we are properly identifying periods of monsoon activity before using the classification to analyze research questions in Sections 4 and 5. Our intent is not to show that this approach is superior to other approaches for the identification of monsoon onset; rather the purpose of this section is to demonstrate that our method produces results consistent with previous studies by showing that it recreates some of the best known features of Australian monsoon variability. We begin with the date of monsoon onset, perhaps the most studied feature of the monsoon. In Figure 3.7 we show the time series of monsoon onset and retreat as defined by this study as well as those from two studies that published tables of dates: Drosdowsky (1996) and Hendon and Liebmann (1990). Drosdowsky defines onset as the first period of at least two days when the vertically averaged 1000 hPa – 500 hPa zonal wind exceeds a minimum westerly threshold while being overlaid by easterlies winds from 300 hPa – 100 hPa. Hendon and Liebmann’s definition of onset is the first simultaneous occurrence of westerly 850 hPa winds in the temporally smoothed times series at Darwin and average rainfall of 7.5 mm/day at several northern Australian weather stations. We also show the austral spring (Sept.-Oct.-Nov.) average Southern Oscillation Index (SOI) as provided by the Australian Bureau of Meteorology. As noted in the introduction, the date of monsoon onset has a well-established negative correlation with the Southern Oscillation in the preceding austral spring, i.e., the monsoon starts earlier in La Nina years and later in El Nino years, while no such relationship has been found for the monsoon retreat. Our classification captures the same relationship between
the Southern Oscillation and monsoon onset date as the previous studies, and actually has a stronger relationship ($r=-0.49$ for our definition, $r=-0.38$ and $r=-0.29$ for Hendon and Liebmann and Drosdowsky respectively), thought this may simply be due to covering different time periods. In overlapping years of study, our onset dates correlate with those of Drosdowsky with $r=0.54$ and with Hendon and Liebmann with $r=0.71$, while the two correlate with each other with $r=0.61$. Kajikawa, et al. (2009) also report similar magnitude correlations for their onset date ($r=0.54$ with Drosdowsky, and $r=0.59$ with Hendon and Liebmann).

Figure 3.7 also shows that our onset dates are earlier than those of the other studies. The mean onset date for Drosdowsky is December 28-29\textsuperscript{th}, for Hendon and Liebmann it is December 24\textsuperscript{th}, while by our definition it is December 13\textsuperscript{th}. Other published mean onset dates include December 24\textsuperscript{th} (Holland 1986) and December 16\textsuperscript{th} (Kajikawa et al. 2010). We think that this up to two-week shift relative to the other definitions is caused in part by how we determine the state to which each ERA-Interim timestep belongs, and in part due to analyzing differing time periods. Regarding our determination of state, while other studies establish a threshold and require some observed environmental conditions to exceed it, our definition requires that the large-scale pattern more closely resemble the active monsoon than the transitional or break monsoon patterns. We discuss these differences further at the end of Section 3.4. For similar reasons, our definition makes our mean retreat date (March 28\textsuperscript{th}) later than others (March 13\textsuperscript{th} for Drosdowsky). The limited number of overlapping years used in our study and those used in Drosdowsky and Hendon and Liebmann may also contribute to the differences in start date, as we cannot determine from the available data whether decadal variability or very long term trends might exist.
We note that our mean onset date closely agrees with that of Kajikawa, et al. (2009), which is the only one of the studies that calculates onset dates more recent than the 1991/1992 season. Visual inspection of the results in Kajikawa, et al. shows that their onset dates for recent decades are earlier in the season than their onset dates from the early part of their dataset (the 1950s-1960s). This suggests that there may in fact be a long term trend or variability at work which would contribute to our onset dates being earlier than those of previous studies.

We further investigate the differences between our dates and those of other studies by analyzing the daily precipitation recorded at Darwin Airport. Returning to Figure 3.6, we see an early period of precipitation in late November and early December. Both our study and Hendon and Liebmann (1990) find this active period to be the start of the monsoon, while Drosdowsky's definition (1996) identifies the monsoon onset to be associated with the subsequent period of monsoon activity in mid-January. At the time of our onset, the cumulative precipitation for the water-year (July-July) is 17% of the total precipitation that fell in the 1981-82 year. At the time of Drosdowsky's onset 48% of the year's precipitation had already fallen, indicating that a substantial portion of the monsoon season had already occurred. We chose to show the 1981-1982 season in Figure 3.6 both because it has large disagreement for the onset date, and because it is representative of how the disagreement occurs. Inspection of the other cases of large disagreement (1980-81 onset, 1982-83 retreat, 1983-84 onset and 1985-86 onset and retreat) shows that for onsets the disagreeing study has not included an early period of precipitation, and instead has assigned monsoon onset to the subsequent period of activity. Similarly, retreat date disagreements involve a study not including a late period of precipitation. In doing so,
these two definitions for the monsoon sometimes do not include a substantial portion of the year's precipitation within the monsoon season. In the extreme case, only 16% of the precipitation in the 1985-86 season falls during the Drosdowsky-defined monsoon. Table 3.1 summarizes the differences in cumulative precipitation at the time of onset and retreat for each study. It shows that by more regularly detecting these early and late periods of activity, this study has a more consistent fraction of the yearly precipitation falling at the time of monsoon onset and retreat, despite precipitation not being a part of our definitions. We take this as evidence that we are very reliably identifying the beginning and end of each monsoon season.

A second well-studied characteristic of the monsoon is the interannual variability of its seasonal intensity. As described in the introduction, La Nina conditions are correlated with more intense monsoon seasons by a number of metrics (Kajikawa et al. 2010; Drosdowsky 1996; McBride and Nicholls 1983). We find that the total number of days classified as active is significantly correlated with springtime (September-October-November) SOI (p<0.05), as is the fraction of the monsoon season classified as active, albeit at a weaker significance level (p=0.07). This weaker significance level is likely due to this metric being a function of monsoon season duration, which is in turn a function of retreat date. The retreat date is uncorrelated with the SOI, thus introducing some variability into this metric. The number of active days and the active fraction of observations are correlated at r=0.53, implying that it is not purely a trade-off between active and break periods, but also a lengthening of the monsoon season that contributes to the strengthening of the monsoon season during La Nina years.
A practical question is how far in advance ENSO indices provide information regarding the onset and intensity of the coming monsoon season. Nicholls et al. (1982) found significant correlations between austral winter (JJA) mean surface pressure at Darwin and a variety of onset dates based on different levels of cumulative precipitation. The onset dates of Nicholls et al. have a similar range as ours (late October to early January), and surface pressure at Darwin is strongly correlated with the SOI, so we would expect to see similar predictability using our methods. To determine how far back we have predictive skill, we calculate correlation coefficients for onset date and number of active days with monthly average SOI (Figure 3.8). We find that as early as the preceding July, the SOI index is significantly correlated with both onset date and the number of active days. While statistically significant, these correlations are rather weak for July at $r=0.48$ and $0.45$ for onset date and active days, respectively. The correlation with the SOI gradually grows stronger as monsoon season approaches, eventually reaching $r=0.67$ and $0.53$ for onset date and active days in November. This suggests that, while some degree of seasonal predictability is connected with ENSO, there are other important sources of variability, such as the timing of MJO events and mid-latitude troughs, as well as the land-sea thermal contrast (Hung and Yanai 2004).

3.4 Intraseasonal variability

Within a single season, the monsoon goes through active and inactive periods on the scale of weeks. On these timescales, the MJO is a major driver of tropical variability (Zhang 2005). We investigate the role of the MJO using the Wheeler and Hendon (2004) definition of MJO phase. This definition is such that, as the MJO propagates eastward from the Indian
Ocean out into the Pacific, the phase number steadily increases to represent the geographic location of enhanced convection. Darwin experiences enhanced convection during Phases 4-6 and suppressed convection during Phases 8, 1, and 2. In our analysis, we omit periods of weak MJO activity, as defined by the Wheeler and Hendon index having an amplitude less than one. Overall, this eliminates 36% of days. Both active and break periods of the monsoon are equally affected by this, as the distribution of atmospheric states during weak MJO periods (67% active, 28% break) is essentially the same as during strong MJO periods (65% active, 27% break). Wheeler et al. (2009) demonstrate that the phase of the MJO according to these definitions has significant relationships with summertime precipitation in northern Australia, suggesting that the phase definition should be suitable for investigating the timing of active periods of the monsoon.

Figure 3.9 shows the distribution of atmospheric states during the monsoon season for each phase of the MJO. Monsoon season here is defined as the period from onset until retreat, as defined in Section 3.2. Overall, 65% of the observations from the 33 monsoon seasons are classified as active, while 28% are classified as break periods. When the MJO is in a position to enhance convection at Darwin, the monsoon is much more likely to be active, reaching probabilities of nearly 90% during Phases 5 and 6. Similarly, when the MJO is in a location associated with suppressed convection at Darwin, break periods become more likely and are a majority of the observations during Phases 1 and 2. For both the active and break periods, the largest changes in probability of occurrence happen at the transition from Phase 3 to 4 and Phase 7 to 8. This suggests that active periods are initiating with the arrival of the MJO and break periods with the departure.
To investigate this further, we compute the distribution of MJO phase for the first day of each active period and the last day of each active period (Figure 3.10). We define an active period as at least 4 consecutive observations (24 hours) being classified as the active monsoon (State 7). We see that active periods most frequently begin during MJO Phases 3 and 4, earlier than the peak of overall active occurrence, Phases 5 and 6. This implies that by the time the MJO reaches Phases 5 and 6, most active periods have already begun. Similarly, the active monsoon periods end most often during Phases 7 and 8, which is earlier in the MJO cycle than the peak of break monsoon occurrence. This again suggests that, by the time the MJO reaches Phases 1 and 2, most of the active periods have already ended. We conclude that the MJO plays a strong role in both initiating and terminating the active periods of the monsoon.

That the plurality of active periods begin during Phase 4 is not surprising, as this is the first phase with enhanced convection at Darwin (Wheeler and Hendon, 2004). By the time the MJO reaches Phase 4 the wind anomalies become westerly and cyclonic, making the snapshots very much like our definition of the active monsoon and readily initiating an active period. It is interesting, however, that the second-most active periods begin during MJO Phase 3, before the enhanced convection reaches Darwin. Composites by MJO phase for our region of classification confirm that the westerly wind anomaly has, on average, not yet reached the region during Phase 3, so one might reasonably ask what prompts the classifier to identify these observations as the start of active periods. Compositing relative humidity for each MJO phase shows a moisture anomaly at both low and mid-levels that leads the change in winds, consistent with observations of MJO related anomalies (Benedict and Randall 2007). By Phase 3 of the MJO, this moisture anomaly has changed from the
deeply negative values of Phases 1 and 2 to near neutral values, making the moisture profile of Phase 3 more similar to the active monsoon than to any other atmospheric state. Similarly, we see a dry anomaly which precedes the change in winds helping to end the active periods during Phases 6 and 7.

Still, the zonal wind is as important to the classification as relative humidity, making it somewhat puzzling that a moisture anomaly could, on its own, initiate an active period. To explore this we composite the Phase 3 snapshots that are the first day of an active period, and find that these snapshots are both more humid and more westerly than the overall average Phase 3 values. This is to say that the Phase 3 snapshots that begin active periods are ones with the convective anomaly further east than average. The variability in atmospheric conditions that exists for a single MJO phase is enough to allow active periods of the monsoon to begin in Phase 3, even if the phase, on average, does not enhance convection. We confirm that active periods are being triggered by changes in both wind and humidity by compositing the horizontal winds, relative humidity, and daily precipitation at Darwin for 10 days before and after the initiation of each active period (Figure 3.11). As can be seen, the zonal wind, meridional wind, and the relative humidity are all rapidly changing at the start of active periods. These changes begin the day before the active period begins, grow strong enough to initiate the active period, and continue to grow for the next 1-3 days. The daily precipitation is low 2-4 days before the event begins and rises steadily for several days before peaking at 3 and 4 days after the active period begins. This is in keeping with the idea that the active period begins as the MJO approaches, strengthens as the convective anomaly reaches Darwin, and then ends as the anomaly departs. Collectively, we take this as evidence that the active periods identified
are not simply the result of an anomaly in a single variable, but rather are the coherent establishment of a monsoon circulation, and that we are identifying them from their beginning, rather than their peak.

3.5 Trend analysis

The length of our 33-year period of study allows us to investigate trends in the monsoon metrics. While we find trends of earlier onset, later retreat, and greater duration, none of these trends are significant at the 95% confidence limit (p= 0.35, 0.42, 0.25 respectively). We do, however, find a significant positive trend in the number of days classified as the active monsoon and a significant negative trend in the number of days classified as the break monsoon (Figure 3.12). The trend for the active monsoon is approximately 10 days (40 snapshots)/decade and -6 days/decade for the break monsoon, both of which are significant at the 95% confidence level. The discrepancy between these values is accounted for by a an approximately -5 days/decade trend that the classifier identifies as transition season states, however this trend is less significant (p=0.13). The increase in active monsoon days at the expense of break monsoon and transition days suggests that the monsoon has become both longer and more active in recent decades. This is consistent with the observed trend in SOI. However, when we tested the relationship between the number of active days and the SOI in Section 3, we found that the number of active days increases by approximately 1.2 for each unit of the SOI index and the trend in the SOI index during the period of study is 3.4 units/decade. Multiplying the two together suggests that the trend in the SOI during the period of study is only responsible for approximately 4 extra active days per decade, or 40% of the trend in the number of
active days. Of course this simple calculation assumes a linear relationship between the SOI and the number of active days is a good model for this relationship. Nonetheless, while a fraction of the increase in precipitation at Darwin is related to ENSO, it is apparent that factors other than ENSO are also important.

To further investigate the idea of an intensifying monsoon, we turn to daily precipitation observations at Darwin Airport. We use these data because so far as we know they are the only high quality observations of precipitation that span the full ERA-Interim period, which starts in 1979. These data are available through the Australian Bureau of Meteorology and extend back to 1941. We composite these data according to state and year by assigning each classification time the daily-averaged precipitation from the day on which it occurred. Days during which multiple states are identified assign the day’s precipitation proportionately to the different states. Again, our year periods begin and end in July. This preserves a continuous monsoon period and guarantees that the pre- and post-monsoon transition seasons are counted in the same “year” as the monsoon. We consider it important to include the transition seasons in the same water year as the monsoon in order to account for the effect of the apparent trade-off in days classified as the active monsoon and days classified as transition periods. The dry seasons at Darwin are divided in our July-to-July water years, but their contribution is negligibly small with regard to total annual rainfall.

Annual rainfall at Darwin Airport has increased by 6.5 mm/yr (p=0.004) since the start of the record in 1941, as shown in Figure 3.13. During the period of the Evans et al. (2012) classification (1979-2012), the trend has increased to 9.5 mm/yr, though it is no longer significant over this shorter period (p=0.23). Nonetheless, as the more recent
period represents nearly half of a longer, highly statistically significant trend, and as the recent trend is notably large, it is interesting to explore why Darwin has received more rain in recent years. The precipitation measured at a single station represents only a very small area, which inherently leads to a large annual variability. However, analyzing data over many years significantly reduces the impact of variability due to spatial sampling. The change in total rainfall can be caused by either a change in the frequency of occurrence of the atmospheric states (rainy states becoming more frequent) or by a change in the precipitation associated with each state (states becoming rainier). Among the individual atmospheric states, only State 3 has a trend, in this case a negative one, in mean precipitation that is significant at 95% confidence. State 3 is a very dry state that rains very rarely and then only weakly, so a trend in its precipitation is not important to the overall trend. Following the method of Catto et al. (2012, hereafter C12), we decompose the overall precipitation trend into the contributions from changes in frequency of occurrence and contributions from changes in mean precipitation. Mathematically, this can be described as follows:

$$\Delta R = \sum_i N_i \Delta P_i + \sum_i \Delta N_i P_i + \sum_i \Delta N_i \Delta P_i$$  \hspace{0.5cm} (Eq. 3.1)

where $\Delta R$ is the total rainfall trend (mm/yr), $N$ is the number of occurrences of state $i$ (i.e. the number of events), $P$ is the mean precipitation for the state (mm/event), and $\Delta N$ and $\Delta P$ are the linear trends in those values for each state.

Figure 3.14 shows the contributions to the overall trend from both changes in the mean precipitation of each atmospheric states (the first term of Eq. 1) and changes in their frequency of occurrence (the second term). We omit the third term of Eq. 1 because the second order contributions are much smaller than the other terms. The decomposition
demonstrates that the observed increase in annual precipitation at Darwin is almost entirely due to the increase in the number of active monsoon days since 1979. This increase in the frequency of occurrence of the rainiest state of the classification more than makes up for the reduction in both the occurrence and mean precipitation of State 8. This is broadly in keeping with the findings of C12, who find that the increase in wet season precipitation at Darwin is due to an increase in the occurrence of rainy regimes at the expense of fewer occurrences of dry regimes. However, the regimes used in C12 and the states we use do not perfectly match up.

C12 uses the classification described by Pope et al. (2009) to decompose the observed precipitation trend. In Evans et al. (2012), we compare our classification to that of Pope et al., and conclude that their Deep West regime is similar to our State 7, and that their Moist East regime is similar to our State 8. Directly comparing our time series of atmospheric state to the time series of regime used in C12 provides a more nuanced understanding of the comparison between the two classifications. We find that approximately 60% of the days in the Deep West regime fall within our State 7 (active monsoon), with 30% and 10% being classified as State 8 (break monsoon) and State 6 (pre-monsoon transition) respectively. Our State 7 comprises a much larger fraction of the monsoon season than the Deep West regime, and also contains significant amounts of the Pope et al. Shallow West and Moist East regimes. We conclude that our State 7 is a broader definition of the active monsoon than the Deep West regime. The Pope et al. Moist East regime, meanwhile, is a much larger fraction of the monsoon season than our State 8 and contains large fractions of our State 6 (which is drier) and State 7 (which is much wetter).
This leads us to believe that our State 8 is a more restrictive definition of a break period of the monsoon.

As described in Section 5, we find a statistically significant increase in the frequency of occurrence of the active monsoon at the cost of the break monsoon and transition states. C12, however, find only a weakly positive trend that is not statistically significant in the frequency of the Deep West regime. Taken together, this suggests that it is the non-Deep West components of State 7, i.e., the moderately active periods that have become more frequent. C12 find a positive trend in the occurrence of their Moist East regime, while we find a negative trend in the occurrence of our State 8. This suggests that it is the more active elements of the Pope et al. Moist East regime that are responsible for their positive trend, while it is the least active elements of State 8 which are responsible for that state's negative trend in our result. This supports the conclusion that periods of moderate active monsoon have been increasing in frequency at the expense of the weakest periods of the monsoon.

Interestingly, while the annual frequency of occurrence of the active monsoon dominates the 33-year trend in precipitation at Darwin, the interannual variability of total annual precipitation is better explained by interannual changes in the mean daily precipitation of each state. The percent of the variance in annual precipitation ($r^2$) over the 33-year record due to both the annual occurrence of each state and the annual mean precipitation of each state is provided in Table 3.2. Variations in the active monsoon state explain by far the largest amount of variability in total annual precipitation, both in terms of frequency of occurrence (38%) and mean precipitation (47%), with variations in mean precipitation being somewhat more important. Because the frequency of occurrence of one
state is strongly correlated with the frequency of occurrence of other states during each season (dry, transition, and monsoon), we also include three combinations of states. The collective total of States 1-4 is the dry season, States 5-6 the transition season, and States 7-8 the monsoon season. As expected, variations in mean daily precipitation during the monsoon season explain the majority of the variance in total rainfall ($r^2$ of 75%). However, changes in annual mean daily precipitation may well be due in part to the spatial heterogeneity of precipitation, given the high spatial variability of precipitation measurements. It is unclear how much of the within-state variability is simply due to spatial variability (which will tend to average out in the longer term trend analysis) rather than true dynamically driven variability.
<table>
<thead>
<tr>
<th>Cumulative seasonal precip. (%)</th>
<th>E12 onset</th>
<th>D96 onset</th>
<th>H90 onset</th>
<th>E12 retreat</th>
<th>D96 retreat</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>16</td>
<td>30</td>
<td>26</td>
<td>92</td>
<td>85</td>
</tr>
<tr>
<td>std. dev.</td>
<td>6</td>
<td>12</td>
<td>10</td>
<td>8</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 3.1

Cumulative precipitation, as a percentage of the total July to July precipitation that has occurred at the time of monsoon onset and retreat, according to various definitions. E12 refers to onset and retreat dates defined in this study, D96 to dates from Drosdowsky (1996), and H90 to dates from Hendon and Liebmann (1990). The mean value and standard deviation is given for each definition.
Table 3.2

Percent of the variance in annual (July-July) precipitation at Darwin over the 33 year record that is explained by (1) AnnO, the annual occurrence of each state, measured in days, and (2) AvgP, the annual mean precipitation associated with each state. The last three columns are combined values of $R^2$, i.e., the combined annual occurrence and weighted mean precipitation for multiple states. **Dry** refers to the combined occurrence and precipitation of States 1 to 4, **Trans.** to the combination of States 5 and six, and **Mons.** to the combination of States 7 and 8.

<table>
<thead>
<tr>
<th>$R^2$ (%)</th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
<th>State 4</th>
<th>State 5</th>
<th>State 6</th>
<th>State 7</th>
<th>State 8</th>
<th>Dry</th>
<th>Trans.</th>
<th>Monsoon</th>
</tr>
</thead>
<tbody>
<tr>
<td>AnnO</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>38</td>
<td>8</td>
<td>0</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>AvgP</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>47</td>
<td>9</td>
<td>3</td>
<td>16</td>
<td>75</td>
</tr>
</tbody>
</table>
Figure 3.1

Monthly histograms of the occurrence of each state. The number in each panel's title indicates the total number of occurrences of each state.
Figure 3.2

1000 mb dew point (shaded, °C) and 1000 mb wind vectors for the eight atmospheric states. Black lines show the land-sea boundaries. Latitude and longitude are labeled on the axes. Darwin is at the center of the domain.
Figure 3.3

Surface pressure anomaly (mb, shaded) and 750 mb wind vectors for the eight atmospheric states. Anomalies are calculated from the mean of the full input dataset. Black lines show the land-sea boundaries. Latitude and longitude are labeled on the axes. Darwin is at the center of the domain.
Figure 3.4

500 mb relative humidity (fraction, shaded) and 500 mb wind vectors for the eight atmospheric states. Black lines show the land-sea boundaries. Latitude and longitude are labeled on the axes. Darwin is at the center of the domain.
Figure 3.5

Vertical profiles of hydrometeor occurrence by atmospheric state, defined as the fraction of the time clouds or precipitation are detected with a reflectivity of -40dBZ or larger by the millimeter wavelength cloud radar at the ARM site in Darwin. Dry season states are shown in shades of red, transition states in green, and monsoon states in blue. Shaded areas denote the 95% confidence limits, as determined by the bootstrap resampling method. The numbers in the legend indicate the number of 3 hour blocks with radar observations for each state.
Figure 3.6

Time series of atmospheric state (upper panel) and daily precipitation (lower panel) for the 1981-1982 monsoon season. Upper panel shows our monsoon onset and retreat (vertical green lines) and active periods (blue bars) as defined in Section 3.2. Lower panel also shows onset defined by Hendon & Liebmann (1990, vertical red line), and onset and retreat defined by Drosdowsky (1996, vertical blue lines).
Figure 3.7

Time series of monsoon onset (lower solid lines) and retreat (upper solid lines) according to this study (black), Drosdowsky (1996, blue), and Hendon and Liebmann (1990, red). The dashed black line shows the average September-October-November Southern Oscillation Index. The year listed on the horizontal axis is the starting year of the monsoon season, i.e. 1960 is the 1960/1961 season.
Figure 3.8

Time evolution of the correlation coefficients between monthly average SOI and the number of active monsoon days (solid), onset date (dashed), and retreat date (dotted) for the years preceding and following the monsoon. Horizontal dashed lines indicate 95% significance levels, estimated from the Fisher-Z transformation. Vertical red lines indicate mean onset and retreat dates.
Figure 3.9

The relative occurrence of different states for each MJO phase during the monsoon season. Monsoon season is defined as onset to retreat. Shaded regions indicate 95% confidence limits as estimated by a bootstrap algorithm operating on 33 years of data. The overall fraction of the monsoon season occupied by each state/category is shown in the legend. Days with weak MJO phase (36% of all monsoon days) are excluded from this figure. Weak days are 67% active, 28% break, 5% other, and excluding these days does not notably alter the distribution of states.
Figure 3.10

Distribution of MJO phase for the first day of each active period (dark blue), and the last day of each active period (light blue). Shading represents 95% confidence limits, as calculated by a bootstrap algorithm. Days with weak MJO phase are excluded from this figure, eliminating 38% of first and last days.
**Figure 3.11**

Domain mean values of zonal wind (upper panel, black, left axis), meridional wind (upper panel, red, left axis), and relative humidity (upper panel, blue, right axis) at 875 mb (solid lines) and 500 mb (dashed lines). Mean daily precipitation at Darwin Airport (lower panel). Values are composited relative to the first day of each active period.
Figure 3.12

Frequency of occurrence of atmospheric states over the ERA-Interim period, calculated as the anomaly, in days, from the mean. Years are calculated from the midpoint of one calendar year to the midpoint of the next. Labeled years are the first calendar year of each year period, e.g., 1983 indicates the July 1983 to July 1984 period. Red lines indicate states with statistically significant trends. The mean number of occurrences of each state is shown in the panel titles.
Figure 3.13

Annual rainfall at Darwin Airport. Years are calculated from the midpoint of one calendar year to the midpoint of the next. Labeled years are the first calendar year of each year period, e.g., 1980 indicates the 1980-81 period. The trend over the full period (6.5 mm/yr, p=0.004) is shown in red, while the trend over the period of this study (9.5 mm/yr, p=0.23) is shown in blue.
Figure 3.14

Contributions to the 33 year trend in annual rainfall at Darwin Airport. Blue bars are the contribution due to trends in within-state mean precipitation and red bars are the contribution due to trends in the frequency of occurrence of the states.
Chapter 4

Evaluating model cloud properties at Southern Great Plains

A major application of the atmospheric classifications described in Section 2 is model evaluation. For this application we choose to focus on the classification performed at the SGP site. The SGP classification was initially created as a test of the improvements to the classification method and increased cloud-radar data availability since previous studies by Marchand et al. (2006, 2009). The greater number of states found at SGP compared to Darwin provides more detail in identifying the meteorological conditions that cause different kinds of model errors.

In evaluating a model run using atmospheric classification there are three questions that we are particularly interested in pursuing. We would like to know under what meteorological conditions does the model produces the largest errors, whether the errors associated with particular patterns are more or less important than the frequency of occurrence of the patterns, and whether the various sources of error compensate for each other. All three of these questions are difficult to answer when simply comparing annual or seasonal averages. With atmospheric states, however, we can identify which states are associated with particular types of model bias, how frequently the states occur in both model and observations, and whether these biases are of the same or different sign.

In Section 4.1 we describe the meteorology of the SGP states in order to better understand what physical processes are likely to be important for each state. Section 4.2
describes the processing of model output and the evaluation of model cloud properties from an approximately 2° run of the Geophysical Fluid Dynamics Laboratory’s (GFDL) Atmospheric Model 3 (AM3; Donner et al., 2011). In Section 4.3 we compare the results of Section 4.2 to those of a 0.5° resolution run of the same model in order to determine how model resolution affects cloud occurrence biases.

4.1 Description of the SGP states

The procedure described in Chapter 2 applied to the SGP region produces 21 atmospheric states. It can be challenging to keep a mental picture of what they represent, hence we group the states into broad categories and then discuss how the states within each category differ from each other. Our categories are Southerlies, Cold Fronts, Northerlies, Anticyclones, and Summer. The first four of our broad categories are based on the near surface winds in the region, as we find this to be a convenient way of separating the atmospheric states into different stages of the synoptic-scale weather systems that predominate in the Great Plains. The summer states receive their own category, as frontal systems become much less common in summer. All 21 states are described below and their differences are illustrated by Figures 4.1 – 4.7. Figure 4.1 shows the mean meteorology of States 12, 18, 14, and 6 as an example of the progression of a frontal system across the region. Figure 4.2 shows the monthly occurrence of the states. Figures 4.3 and 4.4 show the state mean temperature and relative humidity respectively. Figure 4.5 shows the observed cloud occurrence profiles associated with each state. Figure 4.6 shows the probability of transitioning from one state to another, and is conceptually summarized in Figure 4.7. We introduce the states in the sequence shown in Figure 4.7, beginning with the
Southerlies. This roughly approximates the sequence of weather patterns the SGP site would experience during the passage of a frontal system, beginning with a warm front. Figure 4.7 omits the summertime states, as they do not fit into a typical frontal sequence, and we describe those last.

4.1.1 Southerlies

States 4, 5, 8, and 17 all have southerly winds throughout the region with the ARM SGP site near a warm front or in the warm sector of a frontal system. States 4 and 17 both feature low pressure to the northwest, high pressure to the southeast and near zonal flow in the upper atmosphere. Both of these states feature warm fronts in the northern part of the region, identifiable by the rotation of winds and the gradients in temperature and relative humidity. While the two states have similar flow patterns, they are distinguished by their temperatures (Figure 4.3). State 4 is approximately 10 °C warmer than State 17 and occurs primarily during spring and fall, while State 17 occurs most frequently during midwinter. Thus, while these two states represent a similar large scale dynamical setting, there are important differences in their thermodynamics. For example, State 17 features a lower boundary layer and tropopause as evidenced by the height of the upper and lower cloud peaks, and greater relative humidity in the middle and upper troposphere, all of which affect cloud occurrence (Figures 4.4 and 4.5). We will see the same sort of distinction among some of the states in the other categories as well.

State 8 also features a northwest to southeast pressure gradient, albeit a weaker one than States 4 and 17. It is a spring/fall state (Figure 4.2), and has temperatures similar to State 4. In contrast, however, it appears that the warm front is just north of the analysis
region rather than within it. As there are warm southerlies throughout the region, we interpret State 8 as a warm sector with a warm front to the north, and the cold front yet to arrive from the west.

State 5 has low pressure to the southwest, high pressure to the northeast and a warm front in the southern part of the domain. Near surface winds are southerly and slightly easterly throughout the domain with the northern half having cool or cold surface temperatures. Mid and upper levels show convergent flow and high relative humidity consistent with deep ascending air (Figure 4.4).

**4.1.2 Cold fronts**

States 12, 14, 18, and 20 feature cold fronts within the region. States 12, 14, and 18 (Figure 4.1) all have relatively even rates of occurrence from November through April, with less common occurrences in October and May (Figure 4.2). These three states represent the progression of a cold front across the region from northwest to southeast. In State 12 the cold front, angled northeast to southwest, is in the northwestern part of the region and has not yet reached the SGP site. The front can be identified due to the strong gradients in temperature and relative humidity, as well as the rotation of the winds from warm southerlies in the southeast to cold northerlies in the northwest. The front is accompanied by a deep low to the west of the SGP site. State 18, which typically follows State 12 (Figure 4.6), has the low and the front just to the east of the SGP site. This state is cooler and more humid than the preceding state as a larger portion of the region is occupied by the cold northerlies behind the cold front. Following State 18 is State 14 (Figure 4.6), by which time the low has moved out of the region to the east, and the cold front just barely remains in
the southeast corner of the domain. While all of these states likely contain a mix of frontal and postfrontal conditions, State 14 is the most post frontal, making it the coldest of this category (Figure 4.3). As the states progress from one to another, the dropping temperatures result in increasing relative humidity, particularly in the northern part of the region. By State 14, that humidity has been advected to the SGP site, increasing the frequency of cloud occurrence (Figure 4.5).

In terms of dynamics, State 20 is quite similar to State 12, with a cold front just northwest of the SGP site. The key difference between the two states is that State 20 is approximately 5 °C warmer everywhere in the region (Figure 4.3). State 20 has significant peaks in occurrence in May and October with almost no occurrences in either midsummer or midwinter. As was the case for States 4 and 17, States 12 and 20 share similar large-scale dynamical settings, but different seasonality and different thermodynamics. State 20, for example, has lower relative humidity in the mid and upper troposphere, leading to less cloud at those levels (Figures 4.4 and 4.5).

4.1.3 Northerlies

States 2, 6, 16, 19 and 21 all feature strong northerly flow near the surface becoming more westerly with height and high surface pressures throughout the domain. These states typically follow the various cold front states and are variations of cold air spilling out of the north in the wake of a cold front. Their upper-level winds show them to occur near the bottom of a trough or the very leading edge of a ridge. Of the five, State 2 stands out as 10-15 °C warmer than any of the others (Figure 4.3). State 2 typically follows State 20 (Figure 4.6), the spring/fall version of the cold front, and is itself a spring/fall state as well.
The other 4 states all occur most frequently in winter and are quite cold, with domain mean near surface temperatures ranging from 0-5 °C.

State 19 frequently follows State 14 (Figure 4.6), the last in the sequence of frontal states. State 19 features a high-pressure anomaly to the west of SGP and strong northerly flow near the surface that make it the coldest of the classification. State 19 is typically followed by either State 6 or State 21 (Figure 4.6), both of which on the leading edge of the following ridge and feature high pressure across the region. State 6 (Figure 4.1) has a stronger high-pressure anomaly, helping to make the flow more northerly, colder, and substantially dryer than State 21. Accordingly, State 6 is the least cloudy of the classification, with almost no cloud at any level, while State 21 has modest amounts of cloud in the middle levels of the troposphere despite its high pressure (Figure 4.5).

State 16 does not have a clear position in the sequence of states. Near the surface it has cold northerly flow and high pressure, but in the mid and upper levels the flow is zonal, warm, and humid, leading to the most cloud among the northerly states. We believe this is most likely capturing cases where a relatively weak front has passed to the north of the site, allowing cold air to spill in near the surface, but without having a major impact on the upper level flow.

4.1.4 Anti-cyclones

States 1, 7, and 13 feature strong anti-cyclones over the domain and represent upper level ridges. As with the other categories, there is a spring/fall state here, in this case State 7, which typically follows State 2 of the northerlies (Figure 4.6). Much like the other shoulder season states, State 7 is notably warmer than the others in this category, in
this case by 5-10 °C near the surface (Figure 4.3). In all three cases the anti-cyclonic rotation produces southerly winds in the western part of the region and northerly in the east. States 1 and 7 have northwesterly flow in the upper levels and have high surface pressure throughout the region, but don’t fit amongst the northerly states on account of the near surface southerly flow in the western half of the region. State 13 has zonal flow beginning at 500 mb, and high upper level humidities (Figure 4.4). We interpret this as moisture riding up and over the ridge. As a result, State 13 is substantially cloudier than States 1 and 7 at all levels (Figure 4.5).

4.1.5 Summer

States 3, 9, 10, 11, and 15 occur from May to October (Figure 4.2) and deserve their own category as they lack the characteristics of the frontal systems which characterize the winter and shoulder seasons. They are the five warmest states of the classification (Figure 4.3) and feature near-surface southerlies over most of the domain. State 3 has anti-cyclonic winds at the low levels and northwesterly flow beginning at 500 mb. This gives it a circulation pattern very similar to State 7 of the anti-cyclones. State 15 also has anti-cyclonic flow, centered to the east of SGP at the surface and in the center of the domain at 500 mb. Despite being more humid than States 3 or 10, this subsidence leads to the least cloud among the summer states, which also makes it the hottest state of the classification. States 9, 10, and 11 all have low-level southerly flow throughout the region that becomes westerly with height. State 9 most readily stands out as being both 3-5 °C cooler than the other summer states and much more humid, making it very cloudy at all levels (Figures 4.3, 4.4, and 4.5). This most likely represents deep convection in the region. State 10 has a
negative pressure anomaly over much of the domain and convergence in the northern part of the domain, making this a “cool front” of the sort that relieve the summertime heat and humidity in the Great Plains. Lastly, State 11 has relatively weak flow over the SGP site and resembles a much dryer, and hence less cloudy version of State 9 (Figure 4.5).

4.2 Evaluation of cloud properties in the AM3

We evaluate here a run of the GFDL AM3 model, (Donner, et al. 2011). The model is run at approximately 2° horizontal resolution with 48 sigma levels. The model is forced by historical sea surface temperatures for the period 2000-2010, during which snapshots of atmospheric variables are output 4x daily (0, 6, 12, 18Z). Temperature, relative humidity and horizontal winds are interpolated onto the ERA-Interim grid (1.5° x 1.5°) and the pressure levels at which the reanalysis data was sampled. This is done in order to make the comparison of model snapshots to observed states a comparison of the same variables at the same locations.

4.2.1 Classifying model output

Model snapshots are classified as belonging to one of the observed states by calculating the Euclidean distance from the snapshot to each of the observed state centroids and assigning the snapshot to the closest state. The centroids of each state are defined as the number of standard deviations each state variable is from the observed mean, i.e. each variable’s z-score. By using z-scores, all input variables are weighted equally in the sorting, regardless of their units. In order to calculate the distance a model snapshot is from a centroid, the model snapshot must be normalized in a way to make it
comparable. We have tested a number of ways of performing this normalization, which we describe below and summarize in Table 4.1.

The simplest method of normalizing a model snapshot is to subtract the observed mean values for each variable and divide by the observed standard deviations. This creates a description of the model snapshot that is directly comparable to the observed snapshots, i.e. a model snapshot that has the same z-score for a particular variable as an observed snapshot would also have the same raw value for that variable. The downside to this method is that the distribution of a variable in the model is not necessarily the same as in observations. As a result, it means that a bias in a particular model variable would make it more or less likely to produce certain z-scores, and thus more or less likely for model snapshots to be classified as belonging to particular states. While this could be argued as the most accurate description of state occurrence in the model, it does mean that, for example, a slight cold bias in the model could shift occurrence of warmer states to cooler states even if the frequency of circulation patterns were the same in the model as in the observations.

A simple way to account for bias in the model is to subtract it out before normalizing the snapshots. Mathematically this is equivalent to subtracting the model mean for a variable before dividing by the observed standard deviation. Doing so means that biases in the model will not produce shifts in the frequency of occurrence of states in the model. The potential need for this can be seen in Figure 4.8, which shows that this model run has a cold bias in the upper atmosphere, and a seasonally varying bias in the lower levels of the atmosphere. We can also account for changing biases like these by performing a seasonally varying bias correction. In this case we divide the year into four seasons – March to May,
June to August, September to November, and December to February – and calculate modeled means for each season. We then subtract out these seasonal means before dividing by the observed standard deviation. In doing so, we account for any seasonally varying biases.

The last normalization technique we test is histogram renormalization. While the bias correction techniques account for shifts in the distribution of variables in the model, they do not account for differences in the shape of the distribution. Histogram renormalization does not produce z-scores for a snapshot, but rather assigns each variable in a snapshot to percentile bin. In this case we use 20 bins for each variable, meaning each bin represents 5% of the variable’s range. Each snapshot is thus reduced to a list of bin numbers for each variable. This is done separately for both the observations and the model output, and model snapshots are then assigned to the state they most closely match in bin-space. With this method, a modeled variable in the 5th percentile of model output would be most closely matched to an observed state whose variable was in the 5th percentile of observed data.

A potential drawback to the bias correction and histogram renormalization techniques is that they make comparison of within-state properties more difficult to understand. The cloud and radiative properties associated with a model snapshot are a function of the raw values of the meteorological variables of that state. When no correction is done in normalizing the data, the raw values of the snapshot are required to be close to the raw observed values. This means that the various parameterizations in the model will be operating on values similar to those experienced in nature. In contrast, when a bias correction or histogram renormalization is performed, the model parameterizations are no
longer necessarily operating on values that are the same as those in nature. This makes it more difficult to interpret the origin of differences between the observed properties of a state and the modeled properties of the state. We analyze the effects of this in the following section.

### 4.2.2 Comparison of normalization techniques

Before beginning the evaluation of model cloud performance we first test the importance of normalization technique to the frequency of occurrence and cloud profiles of the states. The frequency of occurrence of the atmospheric states in both the reanalysis and the model is shown in Figure 4.9. As can be seen, all the states occur in relatively similar proportions in the reanalysis (4-8%), a consequence of dividing populous states and deleting sparse states during the clustering process (Chapter 2). In the model, the states occur with a greater range of frequencies, approximately 1-9% of the time, which is to say that the model struggles to create certain weather patterns, while creating others too frequently. When performing no bias correction during normalization (blue dots in Figure 4.9) we see the effect of model bias like that shown in Figure 4.8. The three notably underrepresented states in the model, States 9, 11, and 15, are all summertime states, with State 15 being both the hottest state of the classification and the most underrepresented. That they occur less frequently in the model is due to the model's overall cold bias (Figure 4.8) making it harder for snapshots to be classified as a summer state. The most overrepresented state, State 7, is a warm anticyclonic state. Many of the missing summertime snapshots are being classified as the slightly cooler State 7, along with the other two summer states, 3 and 10.
Applying a bias correction improves the distribution of states in the model for all cases. Removing the annual bias and performing a histogram renormalization produce similar magnitude improvements in the distribution of states, as measured by the RMS difference with the distribution in ERA-Interim. Among the four normalization techniques, these two produce the most similar distributions, suggesting that the shape of the distribution of the variables comprising a snapshot is similar in the model and in the reanalysis or at least that differences in the shape are of secondary importance. The seasonally varying correction makes for the biggest improvement in distribution, reducing the RMS error by more than half. It makes particularly strong improvements to the frequency of the states which were most under and overrepresented, States 7, 9, 11, and 15. This is not surprising, as states that occur in only one season, summer, are the most susceptible to a seasonally varying bias.

While the normalization technique has an important effect on the frequency of states, we find that it is not important to the within-state cloud properties. Figure 4.10 shows the state mean vertical profiles of cloud occurrence in the model for each of the four normalization techniques. There are shaded lines plotted here for the uncorrected and various bias correction approaches, where the shading is the 95% bootstrap confidence limits. The shaded lines are rarely distinct. This can be expressed in a variety of statistical ways as well, which are summarized in Table 4.2. The mean cloud profiles for the various corrected normalizations are extremely correlated with those of the no correction normalization ($r > 0.99$), the deviations between the two are very small (RMS < 1% point), and the fraction of those deviations that is significant at the 95% confidence level is small (5% or less). While not shown here, the distributions of top-of-atmosphere radiative fluxes
associated with each state are similarly insensitive to normalization technique. All of this implies that the snapshots which are being classified differently depending on the normalization do not have a significant impact on the cloud properties of the states to which they can be assigned. This is because states that are similar enough in meteorology for a snapshot to be assigned to either one, depending on bias correction, have relatively similar distributions of cloud properties. Given the insensitivity of cloud properties to the normalization technique, we will only use the no correction technique in the following sections of model evaluation, as the state cloud properties of other normalizations are unlikely to be notably different. Were we to use one of the bias corrections the within-state errors would not change in a meaningful way, while the bias in state occurrence would appear artificially small due to having corrected the mean state to match observations.

4.2.3 Evaluation of atmospheric state cloud properties

We evaluate the within state cloud properties of the model using data from the International Satellite Cloud Climatology Project (ISCCP). We use the ISCCP D1 product (Rossow and Schiffer, 1999) which includes joint cloud top pressure (CTP) – optical depth (τ) histograms of cloud occurrence. The histograms are stored on an approximately 2.5° grid, making them of comparable size to a model gridbox. Histograms are calculated at all timesteps when there is sufficient sunlight, and are then assigned to the atmospheric state occurring at that time. We average these joint histograms by state to produce an observed joint histogram of mean cloud occurrence for each atmospheric state. Summing the values of the joint histogram provides the total cloud occurrence for the gridbox. As shown in Figure 4.11, the total cloud occurrence observed by ISCCP at the SGP site is 60%, with much
of that occurring as high thin cloud (the upper left of the joint histogram). Figure 4.12 shows that same data compositied by atmospheric state. Here the difference in cloud types associated with the states becomes apparent. States with fronts near the ARM site, e.g. States 5, 12, 14, and 18 show large amounts of thick cloud, consistent with deep convection produced by those conditions. The height of the thick cloud varies somewhat between states, with the warm front (State 5), and arriving cold front (State 12) having the higher thick cloud and states in which the cold front has already passed the SGP site (States 14 and 18) having somewhat lower thick cloud. In contrast, fair-weather states such as States 1, 3, and 7 have less cloud overall, most of which is high and thin. These values are summarized in Table 4.3, which shows the cloud occurrence values for each state summed over different regions of the ISCCP joint histogram. Summing over regions of the histogram provides a simplified way of interpreting the type of clouds represented by the histogram. For the remainder of the discussion here, we refer to high and low cloud as above and below 560 mb, and thick and thin cloud as more and less than 9.4 optical depths. Occurrence in the high thick portion of the histogram is primarily physically thick clouds produced by deep convection or sharp fronts, in the high thin portion it is primarily cirrus clouds, low thick represents stratus clouds, and low thin primarily represents broken cloud such as fairweather cumulus.

We choose to use ISCCP as the measure of cloud occurrence for the states in order to make use of the ISCCP simulator (Bodas-Salcedo et al., 2011) output produced by the model run. The simulator uses the satellite retrieval algorithm to estimate the values the satellite would report for the cloud field simulated by the model. In doing so, it creates a model output field that is directly comparable to the observed data, and which attempts to
account for the sensitivity and biases of the satellite product. Figure 4.13 shows the mean model cloud occurrence for the SGP region. It shows a total cloud fraction of 36%, with a peak in the high thin cloud portion of the histogram, and a secondary peak in the high thick cloud portion. As with the observed ISCCP data, we composite the model joint histograms by atmospheric state, as shown in Figure 4.14 and summarized in Table 4.4.

The differences in cloud occurrence between the model and observations are shown in Figure 4.15 (total bias), 4.16 (state by state bias), and Table 4.5 (summary of biases), from which a number of features can be identified. Overall, the model has a large negative bias in thin cloud, and a relatively small positive bias in high thick cloud (Figure 4.15). This is perhaps to be expected, as GCMs are typically tuned to produce a reasonable top of atmosphere radiative balance, which is much more sensitive to thick clouds than thin ones, and on a global basis most (if not all) GCMs overestimate the amount of optically thick cloud underestimate the occurrence of optically thin cloud. Figure 4.16 and Table 4.5 show that the lack of high thin clouds is universal across the atmospheric states, however some categories of states have larger biases than others. This indicates that the bias is likely due to a variety of problems in the representation of cirrus clouds. We discuss this idea further in Chapter 5.

As Figure 4.16 shows, however, the relatively accurate simulation of thick clouds by the model is the product of counterbalancing biases among the states. States that have a front within the region (12, 14, 18, 20, and State 5 from the Southerly category, which features a warm front in close proximity to the SGP site) show large negative biases in thick cloud. These are all conditions in which the atmosphere has large-scale ascent at the front and deep ascending air. Fronts are relatively small features compared to a model gridbox
at 2° resolution, and the model will struggle to produce a front as sharp as exists in nature. With weaker fronts, the model cannot generate as much deep convection, thus limiting the amount of thick cloud in frontal conditions. In Section 4.3 we test this hypothesis that low model resolution is a significant factor causing a lack of thick cloud in frontal conditions.

In contrast to the frontal conditions, the anticyclonic, northerly, and summer categories show an excess of high thick cloud (Table 4.5). This excess is driven by the drier and higher pressure states within these categories – States 1 and 7 among the anticyclones, States 3 and 15 among the summer states, and State 6 of the northerlies. All of these states represent conditions not associated with large-scale ascent, though with some potential for isolated convection to produce high thick cloud. As a result, all of these states have very little high thick cloud in observations (Figure 4.12). That the model produces too much high thick cloud in these conditions suggests that the deep convective parameterization is triggering too often.

In the AM3, deep convection is triggered when the convective available potential energy (CAPE) exceeds a critical threshold of 1000 J/kg, and the difference in pressure between the level of free convection and the level of neutral buoyancy is at least 500 mb (Donner et al., 2011; Benedict et al., 2013). The CAPE calculation assumes that no entrainment occurs as the parcel rises. This has been shown to produce a relationship between CAPE and upper level temperature and humidity that is weaker in the model than in reality (Donner and Phillips, 2003). This may explain how the model can exceed the CAPE threshold as often as it does for these states, despite the relatively dry and stable conditions. Interestingly, this would mean that while the lack of deep cloud in frontal conditions appears to be a result of model resolution, the compensating excess of deep
cloud in these fair-weather conditions may be the result of the model parameterization. We explore this possibility further in Section 4.4.

The effects of these thick cloud biases can be readily seen in the resulting differences in outgoing longwave radiation (OLR) between the model and observations. For observations we use the CERES (Wielecki et al., 1996) SYN product (Doelling et al. 2013), which provides 3-hourly observed values of OLR on a 1° grid. We composite these data by atmospheric state and compare their distributions to those of the model in Figure 4.17. The states with fronts in them which lacked thick cloud (State 5 for example) show a lack of low OLR values, while the states that had too much thick cloud (State 7 for example) have an excess of low OLR values. The result is that the mean OLR values for states with fronts in them are significantly biased positive, while the values for states with isolated convection are biased negative. The exception to this are the summertime states, when higher surface temperatures in the model create a positive OLR bias.

### 4.2.4 Decomposition of errors

One benefit of state-based classification is that it allows the overall bias in a variable of interest to be decomposed in contributions from errors in the frequency of occurrence of states and contributions from errors in the within-state properties. To do so we adapt an equation from Williams and Webb (2008) they use to calculate a “cloud regime error metric”. They define the error in a variable $X$ for a particular state $n$, that has a frequency of occurrence $f$, as

$$
\varepsilon_n = \sqrt{(\Delta f_n(X_{obs,n} - \langle X_{obs} \rangle))^2 + (f_{obs,n} \Delta X_n)^2 + (\Delta f_n \Delta X_n)^2} \quad \text{(Eq. 4.1)}
$$
where \( \Delta f \) and \( \Delta X \) are the state mean differences between the model and observations of frequency of occurrence and the variable in question respectively, and \( <X_{\text{obs}}> \) is the mean value of \( X \) across all states. The three terms of this equation represents the error due to the difference in frequency of occurrence of the state, the error due to the difference in the state-mean values, and a second order term. The mean value is subtracted from the state means in the first term because the importance of the frequency of occurrence error depends on whether the state in question was above or below average in variable \( X \), e.g. underprediction of a state with below average cloud occurrence constitutes a positive bias contribution. Comparison of these first two terms for each state shows which source of error dominates for each state, and when summed across all states, which source of error is most important to the overall model error.

Figure 4.18 shows each state's contributions to the model's total bias in high thin cloud occurrence (above 560 mb, less than 9.4 optical depths). It clearly shows that within-state biases are the dominant source of the overall bias. As discussed in the previous section, the model has a large negative bias in high thin cloud that is present in all of the states (Table 4.5). As a result, the distribution of states matters very little to the overall bias. Essentially, if all states produce similar errors, it does not matter which state occurs.

In contrast to this is the cause of errors in high thick cloud occurrence (above 560 mb, more than 9.4 optical depths), shown in Figure 4.19. In this case, both the distribution of states and the within-state biases are important contributors to the overall bias. Because some states have positive biases in high thick cloud while other states have negative biases, it is important to the overall bias how frequently each state occurs. The absolute error contributions from state cloud properties are larger than those from the
state distribution (the mean absolute errors in Fig 4.19), but there is substantial cancellation among them due to the differences in thick cloud biases in frontal and fair-weather states described in the previous section. As a result, the total error from state cloud properties is actually smaller than that from the distribution (the mean errors in Fig 4.19). The two mean errors oppose each other in sign as well, indicating that not only do the state cloud biases counterbalance each other, but the occurrence of states also counterbalances the state cloud properties. The existence of compensating errors in both the state properties and the relative frequency of states for a variable that is important to the radiative balance of the atmosphere might also be a result of tuning to ensure the model has the correct mean global average.

4.3 Evaluation of high resolution run

We test the role of model resolution by repeating the analysis from the previous section on a version of the AM3 run at 0.5° resolution. As with the 2° run, the model is forced by historical sea surface temperatures for the period 2000-2010, and we extract snapshots of the relevant atmospheric variables are 4x daily (0, 6, 12, 18Z). These variables are then interpolated onto the ERA-Interim grid and pressure levels to enable them to be sorted into the atmospheric states. We normalize the snapshots using the no bias correction technique described in Section 4.2.1. This allows for direct comparison to the results in Section 4.2.3, and a straightforward interpretation of the relative importance of different sources of bias. The only difference in the processing of this run is that
snapshots covering a reduced portion of the domain, the central 3x3 grid of points, are the data that was available. As a result, all sorting is done based only on this smaller area. Testing the affects of using this smaller domain on the 2° run showed that the effects on the state distribution and state properties were small, so we do not anticipate this being an issue in the analysis here.

We begin the comparison of the two runs by looking at the frequency of occurrence of the states. Figure 4.20 shows that high resolution run simulates a more accurate distribution of states than the low resolution run. Notably, States 7, 11, and 15, which were three of the states with the largest under- and overpredictions are all significantly improved. This is due to the model producing a more accurate mean state, as the biases in state variables such as temperature are generally smaller in the high resolution run (Figure 4.8).

The response of clouds to the increased resolution is more complex. Figure 4.21 shows the difference in cloud occurrence between the high and low resolution runs. In Section 4.2.3 we hypothesized that the lack of thick cloud in frontal states was likely a result of poor model resolution being unable to simulate fronts well. This is supported by the increase in thick cloud in four of the five states discussed in that section, States 5, 14, 18, and 20, with State 12 having little change. While 0.5° resolution still does not truly resolve fronts, it can certainly come closer to doing so than a 2° model. A better formed front would generate more large-scale ascent and thick cloud as a result. More frequent thick cloud in states with fronts is indeed what the high resolution model produces, and in doing so, the high resolution run improves the simulation of cloud occurrence in these states. The only state without a front to experience significant increases in thick cloud is
State 9, a summertime state representing deep convection across the region. State 9 is by far the cloudiest of the summer states, and this is likely another case of the improved model resolution being better able to capture large-scale ascent over the region.

The warm and dry fair-weather states which had identified as having excess high thick cloud also experience consistent changes in the 0.5° run. States 1, 3, 6, 7, and 15 all experience a loss of high thick cloud. Previously we hypothesized that the excess high thick cloud was due to the deep convection parameterization triggering too often. A comparison of CAPE from reanalysis and the model for each atmospheric state are shown in Figure 4.22. At low resolution, high values of CAPE occur too frequently in the model and trigger deep convection too often. This is especially true for the states that have an excess of deep thick cloud. When the model is run at higher resolution, high values of CAPE become rarer in nearly all states, reducing the frequency with which the deep convective parameterization triggers. In conditions with only isolated convection, one would expect to find large areas of subsiding air and small regions of rising air, producing a distribution of vertical velocities skewed towards negative velocities with a long tail towards positive velocities. As the model goes to higher resolution, this distribution becomes more skewed, producing smaller, stronger regions of rising air. This makes for fewer instances of deep convection, and thus a decrease in high thick cloud in these conditions. As with the frontal states, this improves the simulation of high thick cloud for these states.

The same mechanism may be happening for states other than the ones discussed so far, as a reduction in mid-level, mid-thickness cloud is common to nearly all the non-frontal states. Unlike the states discussed so far, however, reductions in cloud for these other states, which generally did not have excess of cloud in the low resolution run, generally
makes the simulation worse. This is summarized in Figure 4.23, which shows the change in total cloud occurrence bias for each state, where negative values indicate the bias decreased and the simulation got better. We can see the improvement of the frontal states owing to increased thick cloud, and the improvement of anticyclonic and summertime states due to a decrease in thick cloud, but we also see the worsening of the simulation of some northerly and southerly states which did not have enough cloud to begin with, and which do not feature large scale ascent or deep convection.

The overall improvements to both thick clouds in the model and the frequency of occurrence of states have an interesting effect on the contribution to the overall error from different sources. Figure 4.24 shows the contributions to the total bias in high thick cloud in the high-resolution run, using the same equation as in the previous section. In contrast to the low-resolution run (shown in Figure 4.19), the relative frequency of occurrence is now more important than the within-state biases. The improvements in thick cloud already discussed have substantially reduced the magnitudes of error from the state mean properties, but the improved distribution of states has not helped, actually causing a slight increase in the error due to frequency of occurrence. This is possible, as in the low-resolution run the frequency of occurrence errors counterbalance each other, but in the high-resolution run they do so less. As an example, the under-prediction of State 15, which has very little high thick cloud in observations (Table 4.3), contributed a positive bias to help counteract the negative biases from other states. By improving the distribution of states in the high-resolution run, the model has essentially removed one side of the counterbalancing errors. As a result, despite an improved distribution of states, and, in most cases, improved state properties, the overall bias of the model has become worse.
Table 4.1
Summary of different model output normalization techniques.

<table>
<thead>
<tr>
<th></th>
<th>Variables normalized to</th>
<th>Subtract from snapshot</th>
<th>Divide snapshot by</th>
</tr>
</thead>
<tbody>
<tr>
<td>No correction</td>
<td>z-scores</td>
<td>Observed mean</td>
<td>Observed sigma</td>
</tr>
<tr>
<td>Annual bias removed</td>
<td>z-scores</td>
<td>Model annual mean</td>
<td>Observed sigma</td>
</tr>
<tr>
<td>Seasonal bias removed</td>
<td>z-scores</td>
<td>Model seasonal mean</td>
<td>Observed sigma</td>
</tr>
<tr>
<td>Histogram renormalization</td>
<td>Percentile bins</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>
### Table 4.2

Statistics regarding the similarity of state cloud profiles in the AM3 for different normalization techniques, relative to the no bias correction normalization. Mean r refers to the average of 21 correlation coefficients between the state cloud profiles of the no correction normalization and each of the other techniques. Mean RMS is the mean of 21 RMS errors between the no correction cloud profiles and those of the other techniques. Percent different is the percentage of cloud levels for the 21 states where the profile's 95% confidence intervals (from bootstrap resampling) do not overlap with the same interval of the no correction normalization.

<table>
<thead>
<tr>
<th></th>
<th>Mean r</th>
<th>Mean RMS (% pts)</th>
<th>% different</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ann corr</td>
<td>0.9954</td>
<td>0.60</td>
<td>2.18</td>
</tr>
<tr>
<td>Sea corr</td>
<td>0.9912</td>
<td>0.85</td>
<td>5.06</td>
</tr>
<tr>
<td>Hist corr</td>
<td>0.9962</td>
<td>0.59</td>
<td>1.49</td>
</tr>
<tr>
<td>Category</td>
<td>State</td>
<td>Hi Thin</td>
<td>Hi Thick</td>
</tr>
<tr>
<td>------------</td>
<td>-------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>Southerly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>State 4</td>
<td>34</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>State 5</td>
<td>24</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>State 8</td>
<td>36</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>State 17</td>
<td>42</td>
<td>4</td>
</tr>
<tr>
<td>Frontal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>State 12</td>
<td>31</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>State 14</td>
<td>29</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>State 18</td>
<td>27</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>State 20</td>
<td>32</td>
<td>14</td>
</tr>
<tr>
<td>Northerly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>State 2</td>
<td>42</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>State 6</td>
<td>37</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>State 16</td>
<td>39</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>State 19</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>State 21</td>
<td>45</td>
<td>3</td>
</tr>
<tr>
<td>Anticyclone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>State 1</td>
<td>42</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>State 7</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>State 13</td>
<td>32</td>
<td>1</td>
</tr>
<tr>
<td>Summer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>State 3</td>
<td>22</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>State 9</td>
<td>38</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>State 10</td>
<td>39</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>State 11</td>
<td>30</td>
<td>2</td>
</tr>
</tbody>
</table>

**Table 4.3**

Observed ISCCP values of cloud occurrence by atmospheric state and state category (as described in Section 4.1). Values are summed over quadrants of the ISCCP joint histogram, where high and low cloud is defined as cloud tops above and below 560 mb respectively, and thick and thin cloud is defined as optical depths greater and less than 9.4.
<table>
<thead>
<tr>
<th>Category</th>
<th>State</th>
<th>Hi Thin</th>
<th>Hi Thick</th>
<th>Lo Thin</th>
<th>Lo Thick</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southerly</td>
<td>4</td>
<td>20</td>
<td>12</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>22</td>
<td>9</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>20</td>
<td>31</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>18</td>
<td>8</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>8</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Frontal</td>
<td>12</td>
<td>19</td>
<td>19</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>27</td>
<td>20</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>12</td>
<td>14</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>14</td>
<td>25</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>18</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Northerly</td>
<td>2</td>
<td>10</td>
<td>6</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>12</td>
<td>7</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>8</td>
<td>4</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>15</td>
<td>11</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>9</td>
<td>4</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Anticyclone</td>
<td>1</td>
<td>11</td>
<td>7</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>12</td>
<td>7</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>14</td>
<td>8</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Summer</td>
<td>3</td>
<td>15</td>
<td>10</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>14</td>
<td>9</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>27</td>
<td>23</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>13</td>
<td>8</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>16</td>
<td>7</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

**Table 4.4**

As for Table 4.3, but modeled values from the 2° run of the AM3.
<table>
<thead>
<tr>
<th>Category</th>
<th>State</th>
<th>Hi Thin</th>
<th>Hi Thick</th>
<th>Lo Thin</th>
<th>Lo Thick</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southerly</td>
<td>State 4</td>
<td>-14</td>
<td>-3</td>
<td>-5</td>
<td>-4</td>
</tr>
<tr>
<td></td>
<td>State 5</td>
<td>-11</td>
<td>5</td>
<td>-7</td>
<td>-2</td>
</tr>
<tr>
<td></td>
<td>State 8</td>
<td>-4</td>
<td>-20</td>
<td>0</td>
<td>-10</td>
</tr>
<tr>
<td></td>
<td>State 17</td>
<td>-18</td>
<td>-6</td>
<td>-11</td>
<td>-5</td>
</tr>
<tr>
<td></td>
<td>State 17</td>
<td>-21</td>
<td>4</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>Frontal</td>
<td>State 12</td>
<td>-12</td>
<td>0</td>
<td>-5</td>
<td>-8</td>
</tr>
<tr>
<td></td>
<td>State 14</td>
<td>-7</td>
<td>-5</td>
<td>-2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>State 18</td>
<td>-18</td>
<td>-7</td>
<td>-3</td>
<td>-17</td>
</tr>
<tr>
<td></td>
<td>State 20</td>
<td>-13</td>
<td>6</td>
<td>-3</td>
<td>-10</td>
</tr>
<tr>
<td></td>
<td>State 20</td>
<td>-11</td>
<td>4</td>
<td>-12</td>
<td>-9</td>
</tr>
<tr>
<td>Northerly</td>
<td>State 2</td>
<td>-32</td>
<td>2</td>
<td>-2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>State 6</td>
<td>-25</td>
<td>3</td>
<td>-9</td>
<td>-2</td>
</tr>
<tr>
<td></td>
<td>State 16</td>
<td>-31</td>
<td>3</td>
<td>-2</td>
<td>-3</td>
</tr>
<tr>
<td></td>
<td>State 19</td>
<td>-35</td>
<td>1</td>
<td>0</td>
<td>-3</td>
</tr>
<tr>
<td></td>
<td>State 21</td>
<td>-35</td>
<td>1</td>
<td>-2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>State 21</td>
<td>-35</td>
<td>1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Anticyclone</td>
<td>State 1</td>
<td>-21</td>
<td>4</td>
<td>-2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>State 7</td>
<td>-24</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>State 13</td>
<td>-10</td>
<td>6</td>
<td>-5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>State 13</td>
<td>-34</td>
<td>1</td>
<td>-2</td>
<td>-2</td>
</tr>
<tr>
<td>Summer</td>
<td>State 3</td>
<td>-15</td>
<td>2</td>
<td>-10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>State 9</td>
<td>-15</td>
<td>6</td>
<td>-7</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>State 10</td>
<td>-12</td>
<td>-5</td>
<td>-11</td>
<td>-4</td>
</tr>
<tr>
<td></td>
<td>State 10</td>
<td>-16</td>
<td>1</td>
<td>-11</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>State 11</td>
<td>-22</td>
<td>-3</td>
<td>-17</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>State 15</td>
<td>-13</td>
<td>2</td>
<td>-8</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 4.5**

Cloud occurrence bias by atmospheric state and state category (as described in Section 4.1). Values are the difference between the 2° model and observation, i.e., negative indicates a lack of cloud in the model. Values are summed over quadrants of the ISCCP joint histogram, where high and low cloud is defined as cloud tops above and below 560 mb respectively, and thick and thin cloud is defined as optical depths greater and less than 9.4. Table entries in red are statistically significant at 95% confidence according to a bootstrap resampling test.
Figure 4.1

Average meteorological conditions for four states from the SGP classification which often follow each other in sequence. The ARM SGP site is at the center of the domain in Oklahoma. Upper row shows 875 mb dew point (°C) and winds; lower row shows surface pressure anomaly (mb) and 750 mb winds. States 12 and 18 show a cold front and low pressure system sweeping across the region, while States 14 and 6 show cold air spilling in from the north as high pressure builds in the wake of the frontal passage.
Figure 4.2

Monthly histograms of occurrence for each atmospheric state. Panels with blue backgrounds occur primarily in winter, red background in summer, white backgrounds in spring/fall. The number in each panel title is the number of occurrences of each state.
Figure 4.3

Four measures of mean state temperature (°C). Blue curves are near-surface (875 mb) values at the SGP site (light blue) and averaged across the domain (dark blue). Average temperatures for the seven sigma levels of the classification are shown at SGP (red) and averaged across the domain (black).
**Figure 4.4**

Four measures of mean state relative humidity (RH). Blue curves are near-surface (875 mb) values at the SGP site (light blue) and averaged across the domain (dark blue).

Average values for the seven sigma levels of the classification are shown at SGP (red) and averaged across the domain (black).
Figure 4.5

Vertical profiles of cloud occurrence as observed by the vertically pointed millimeter cloud radar at the ARM SGP site for each atmospheric state. Cloud occurrence is defined as the fraction of time a particular altitude has reflectivity greater than -40 dBz. Solid black lines show the mean values, while grey shaded areas represent the bootstrapped 95% confidence limits. Panels with blue backgrounds occur primarily in winter, red background in summer, white backgrounds in spring/fall. The number in each panel title is the number of observations of each state.
Figure 4.6

Probability of transition from any state (rows) to any other state (columns). Each row sums to 1.
Figure 4.7

Conceptual flow chart of the progression of atmospheric states. The outer loop is comprised of states that occur primarily during winter. The inner loop represent states the occur primarily in spring and fall.
Figure 4.8

Monthly domain mean temperature at different pressure levels (different colors) from Era-Interim (solid lines), the GFDL AM3 at 2° resolution (dashed lines), and at 0.5° resolution (dotted lines).
Figure 4.9

Distribution of occurrence of states in observations (black) and in the $2^\circ$ run of the AM3 using various normalization techniques (colors). Normalization techniques shown are no bias correction (blue), annual bias correction (red), seasonal bias correction (green), and histogram renormalization (magenta). Different techniques are described in Section 4.2.1 and Table 4.1. Values in the legend are the RMS error between each normalization and the observed distribution.
Figure 4.10

Vertical profiles of cloud occurrence for each of the 21 states in the $2^\circ$ AM3 for different normalization techniques (see Section 4.2.1 and Table 4.1). Note these are model diagnosed cloud fraction profiles, which are not expected to be equivalent to the observed radar profiles. Background shading for each panel indicates states that occur primarily in summer (red), winter (blue), or spring/fall (white). Shaded areas indicate 95% confidence interval based on a bootstrap resampling test. Shaded areas overlap at essentially all levels, as described in Section 4.2.2 and summarized in Table 4.2.
Figure 4.11

Joint cloud top pressure (vertical axis) – optical depth (horizontal axis) histogram of cloud occurrence at the SGP site, as observed by ISCCP. Values are in percentage points, with the sum of the histogram (the total cloud fraction at the site) reported at the top of the figure.
**Figure 4.12**

Joint cloud top pressure (vertical axis) – optical depth (horizontal axis) histograms of cloud occurrence at the SGP site, as observed by ISCCP and composited by atmospheric state. Values are in percentage points, with the sum of the histogram (the total cloud fraction at the site) reported at the top of each panel. Color of panel title indicates seasonality of state, with red indicating summer, blue indicating winter, and black indicating spring/fall.
Figure 4.13

As in Figure 4.11, but for the $2^\circ$ run of the AM3, using output from the ISCCP simulator.
Figure 4.14

As in Figure 4.12, but for the $2^\circ$ run of the AM3, using output from the ISCCP simulator.
Figure 4.15

Model bias of cloud occurrence at the SGP site, displayed as a joint histogram of cloud top pressure and optical depth. Values are the difference, in percentage points, between ISCCP simulator output from the 2° run of the AM3 and the observed ISCCP values; positive represent excess cloud in the model. Summing over the bins of the histogram produces the total cloud occurrence bias of the model, -24%, shown at top of figure. Asterisks indicate histogram bins whose bias values are significant at 95% confidence, according to a bootstrap resampling test.
Figure 4.16

Model bias of cloud occurrence at the SGP site for each atmospheric state, displayed as a joint histogram of cloud top pressure and optical depth. Values are the difference, in percentage points, between ISCCP simulator output from the 2° run of the AM3 and the observed ISCCP values; positive represent excess cloud in the model. Summing over the bins of the histogram produces the total cloud occurrence bias of the model for each state, shown at top of figure. Asterisks indicate histogram bins whose bias values are significant at 95% confidence, according to a bootstrap resampling test.
Figure 4.17

Distributions of outgoing longwave radiation as observed by CERES (black) and in the 2° run of the model using no bias correction (blue). Shaded regions indicate 95% confidence interval, as determined by a bootstrap resampling test. Blue shaded panels are wintertime states, red shaded panels summertime states, and white panels are spring/fall states. States whose model mean OLR is statistically different from the observed mean have that bias shown in the panel title.
Figure 4.18

Contributions to the total bias in high thin cloud occurrence in the 2° run of the AM3, broken down by state and source (see Equation 4.1). Blue bars indicate contributions from state mean properties, green bars from relative frequency of occurrence, and red bars the second order contribution. The numbers in the legend indicate the mean magnitude and mean value for each bar type.
Figure 4.19

As in Figure 4.18, but for high thick cloud occurrence.
Figure 4.20

Occurrence of the atmospheric states in ERA-Interim (black), the 2° run of the AM3 (blue), and the 0.5° run of the AM3 (purple). Shaded areas are 95% confidence intervals, calculated with bootstrap resampling of annual distributions. Values in the legend are RMS errors between each model distribution and the reanalysis distribution.
**Figure 4.21**

Difference in cloud occurrence, as reported by the ISCCP simulator, between the high and low resolution runs of the AM3. Blues indicate greater cloud occurrence in the high resolution run; reds indicate less. The sum of each histogram is the total change in cloud occurrence for a given state, and is reported in the panel title.
Figure 4.22

Cumulative distributions of CAPE at the SGP site for each atmospheric state from ERA-Interim (black), the low resolution AM3 run (blue), and the high resolution run (red). Dashed vertical black line marks 1000 J/kg, the cutoff threshold for the deep convective parameterization in the AM3. Values in the legend indicate the percentage of CAPE values greater than this cutoff value.
Figure 4.23

Change in the magnitude of the total cloud occurrence bias for each state in the 0.5° model run as compared to the 2° run. Negative values indicate improvement; positive values indicate worsening. States are grouped and ordered as in Tables 4.3-4.5.
**Figure 4.24**

As for 4.18, but for high thick cloud occurrence in the high resolution model run.
Chapter 5
Discussion and conclusions

Our clustering method uses reanalysis data to define sets of atmospheric states for the regions surrounding the ARM program sites at Darwin, Australia and Southern Great Plains in Oklahoma. Cloud occurrence data from the vertically pointed cloud radars at the sites is used to test the quality of the states in an iterative process. The number of states and the state definitions are refined until all states have vertical profiles of cloud occurrence that are both temporally stable and distinct from one another. The result is eight states for the Darwin region, and 21 states for the SGP region.

The Darwin states are used to objectively identify dates of monsoon onset and retreat, periods during which the monsoon is active, and seasonal measures of monsoon intensity. We use these data to demonstrate that our method is capable of reproducing previously identified relationships between ENSO and monsoon onset and seasonal intensity. This, along with consistently capturing most of the cumulative precipitation within our monsoon periods, gives us confidence that our method is accurately identifying periods of monsoon activity. We then use our classification-based metrics to show that active periods of the monsoon begin as the MJO approaches the region and end as the MJO leaves the region. Finally, we use this classification to calculate the contributions to the observed precipitation trend at Darwin Airport from changes in large-scale circulation patterns (state occurrence) and changes in local-scale thermodynamics (mean precipitation within each state) and find that the positive trend in precipitation is entirely
due to an increase in the number of active monsoon days (large-scale circulation patterns) with no significant change in the mean daily precipitation during the active monsoon.

Our finding that the passage of the MJO initiates active and break periods of the monsoon may help to understand the difference in findings between Holland (1986) and Drosdowsky (1996). As in the Drosdowsky study, and in contrast to that of Holland, when we calculate the period from one active period to the next we find a broad, flat distribution with no preferred timescale. This can be reconciled with our findings from Section 4 by recalling that there is a large range in the lifetime of an MJO event as it crosses the Indian and Pacific Oceans. For example, Kim et al. (2013) found that MJO events with a strong dry anomaly leading the enhanced convection last, on average, 40% longer than those that do not. Such a range of MJO lifetimes makes an analysis of the return time or duration of active periods less likely to show the influence of the MJO. That we were able to find a clear influence of the MJO (Figures 5 and 6) shows the value of analyzing the role of the MJO using an index like that of Wheeler and Hendon (2004) which allows for the discovery of MJO-induced effects that do not have a clear 40-50 day period.

The trend toward greater occurrence of the active monsoon is a topic that merits further study. The trend in ENSO does not appear to be completely responsible for the trend in the number of active days. We cannot determine from the available data if the trend in active days is due to decadal variability or the result of anthropogenic forcing. It is possible that this question can be addressed with the use of general circulation models (GCMs). For example, Rotstayn et al. (2007) found that including anthropogenic aerosols in runs of the CSIRO GCM created circulation and precipitation changes over northern Australia that helped to explain the observed trends. The model in question has been
shown to have an inaccurate relationship between ENSO and precipitation in northwest Australia (Shi et al. 2008), but the possibility of an aerosol – precipitation relationship is worth investigating in other GCMs. We demonstrated that the MJO is important to the occurrence of active monsoon periods, so it may be that a GCM needs a good representation of the MJO as well in order to simulate the observed trend in monsoon activity.

GCM projections of precipitation over northern Australia for the 21st century disagree whether precipitation will increase or decrease. This difference can be understood in part by changes to the frequency of dynamic regimes and the precipitation associated with them (Moise et al. 2012). If GCM output were classified according to the Darwin atmospheric states presented here, it would be straightforward to carry out the type of analysis performed here on different GCM runs. The calculation we performed to identify the contributions to the observed precipitation trend can be repeated using 21st century trends in precipitation. In doing so, we can identify whether the precipitation trend in a particular GCM run is due to changes in the frequency of occurrence of certain atmospheric states, or due to changes in the associated properties of those states. An interesting question would be whether the important contributors to the future precipitation trend in different models varied depending on the model’s representation of ENSO, the MJO, or aerosols.

The SGP states primarily identify different stages of synoptic systems crossing the Great Plains. We use the time series of state to composite ISCCP joint histograms of observed cloud top height and optical depth for each state. We also sort snapshots of model output from two different runs of the AM3 at low and high horizontal resolution according to these states. Using output from the model’s ISCCP simulator, we composite
modeled joint histograms of cloud top height and optical depth for the states. Comparing the state joint histograms in observations and in the model allows us to identify which states, and thus what types of weather are associated with the model’s bias in the occurrence of different types of clouds.

We compare the low resolution run of the model to observations and find that the model is significantly lacking high thin clouds for all atmospheric states, while errors in the occurrence of high thick clouds varies greatly (including in sign) depending on the atmospheric state. In particular, states featuring large-scale ascent (i.e. fronts) lack high thick cloud in the model, while states that should only have isolated convection (i.e. fairweather states) have an excess of high thick cloud. We attribute the former to the model’s difficulty resolving fronts, and the latter to a deep convection parameterization that is triggered too often by excessive values of CAPE. In the high-resolution run of the model, the representation of high thick clouds is better for both these group of states, i.e. cloud occurrence increases in the frontal states and decreases in the fairweather ones. The former is likely due to improved resolution of atmospheric fronts, while the latter may be due to a better distribution of CAPE. Lastly, we partition the overall model bias in cloud occurrence into contributions from the distribution of states in the model, and contributions from the mean cloud occurrence of the states (within-state error). We find that in the low-resolution run the state mean cloud occurrence contributes approximately twice as much to the total bias as the distribution of the states, while in the high-resolution run the two are approximately equal sources of bias.

A logical next step after testing the importance of model resolution is to evaluate the model using different convective parameterizations. Benedict et al. (2013) evaluated
tropical cloud properties in the AM3 while modifying the deep convective parameterization in a variety of ways. The modifications they implemented included an activation trigger by Zhang (2002) that requires that the time-integrated low-level ascent be sufficient to lift a parcel to the level of free convection. They found that compared to a control run, deep convection in the model was strongly suppressed in the simulations with this alternative trigger. Potentially, a trigger like this one could improve the excess of high thick clouds in fair-weather states in the low-resolution model by making deep convection less frequent. Whether such an improvement would come at the cost of a reduction of high thick cloud in frontal states as well (of which the model does not produce enough), is an interesting question. Frontal states should have the low-level ascent necessary to trigger the convection, so it may be that they would not experience the suppression of deep convection that other states would. Benedict et al. (2013) also noted that the modified deep convective parameterizations also degrade the mean state of the model. As a result, it may be that a model run using the Zhang (2002) trigger could be an example where the frequency of occurrence of states becomes more important than the within-state errors.

The finding that high thin clouds are lacking in the model under all circumstances requires more analysis to fully understand. High thin clouds can be generated or maintained in a number of ways in the AM3, making it challenging to identify the cause of the bias in their occurrence with the model runs we have analyzed here. For example, advection, stratiform lifting, shallow convection, and deep convection (through detrainment of condensate into the mesoscale updraft), and parameterizations of particle sedimentation can all influence high thin cloud amounts in a gridbox. The relative importance of each of these sources or sinks of high thin cloud in different atmospheric
conditions may help explain why some states are missing much more thin cloud than others. For example, the two groups of states that are have the greatest lack of high thin cloud, the northerlies and anti-cyclones, are also the two groups that have the least high thick cloud in the model (Tables 4.4 and 4.5). High thick cloud is primarily created by the deep convective parameterization, which transfers 90% of non-precipitated condensate into anvil clouds (Donner et al. 2011). That the largest biases occur when this process is rare suggests that without detrainment from deep convection it may be difficult for the model to have enough upper level moisture to sustain high thin clouds. This could be tested by repeating this analysis on a series of runs that alter the condensate partitioning in the deep convection parameterization. Alternately, these states are also ones which occur during upper-level ridges, when advection of thin cloud should be occurring. The lack of high thin cloud in these states may thus indicate that cirrus lifetime in the model is too short, preventing it from being advected into the region. Again, this is testable through the manipulation of various model parameters, such as ice fall speed.

This study raises an interesting question of what it means for one model configuration to represent clouds better or worse than another one. The ISCCP observations for the SGP site have a total cloud occurrence of 60%, the low-resolution run of the AM3 has 34%, and the high-resolution has only 28%. In this sense, increasing the resolution of the model has degraded the simulation of clouds at the SGP site. However, as Figure 4.16 makes clear, cloud occurrence for most states improves at higher resolution, and overall the within-state biases shrink. It is only through changes in the distribution of states and the removal of compensating errors that the total cloud occurrence becomes worse. This makes it a question of semantics whether increasing model resolution
improves the representation of clouds in the AM3. A similar problem arises in studies of the effect of different parameterizations on model performance. It is a common feature of GCMs for improvements in tropical intraseasonal variability to come at the cost of larger biases in the mean state, and that as a result most GCMs have weak intraseasonal variability (Kim et al. 2011). That accuracy in the simulation of atmospheric processes has been sacrificed to improve the mean state is understandable, as most GCMs are designed for long-term climate change scenarios, for which an accurate mean state is more important. For GCM studies designed to test and understand particular processes however, non-standard GCM configurations may be the better choice of tool.

The ability of atmospheric classification to discover and characterize compensating errors in a GCM run is a potentially valuable tool in parameterization development. Most directly, it can help to identify the physical situations which are most problematic and require further attention to represent accurately. More broadly though, classification can help to evaluate whether a particular parameterization or change to the model configuration is an improvement or not. If the evaluation metric for a particular model run is how well it reproduces the mean state it would be easy for actual improvements to a parameterization to be wrongly discarded when the mean state fails to improve. Continuing to pursue these advances in parameterization may produce eventual improvements in the mean state that might not be realizable without making it worse first. Given the challenge of producing an accurate parameterization, following such a lead may eventually be very important.
Bibliography


