



January 2017

Assessment Of WRF Simulated Precipitation By Meteorological Regimes

Brooke Anne Hagenhoff

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: <https://commons.und.edu/theses>

Recommended Citation

Hagenhoff, Brooke Anne, "Assessment Of WRF Simulated Precipitation By Meteorological Regimes" (2017). *Theses and Dissertations*. 2224.
<https://commons.und.edu/theses/2224>

This Thesis is brought to you for free and open access by the Theses, Dissertations, and Senior Projects at UND Scholarly Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UND Scholarly Commons. For more information, please contact und.common@library.und.edu.

Assessment of WRF Simulated Precipitation by Meteorological Regimes

by

Brooke Anne Hagenhoff

Bachelor of Science, University of Oklahoma, 2015

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science


Grand Forks, North Dakota

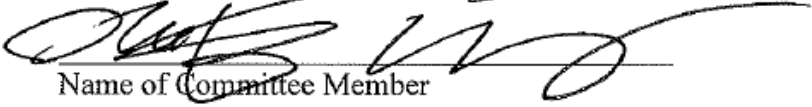
August


2017

Copyright 2017 Brooke Hagenhoff

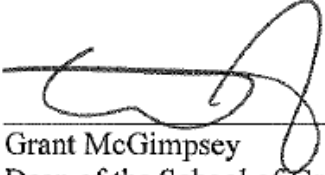
This thesis, submitted by Brooke Hagenhoff in partial fulfillment of the requirements for the Degree of Master of Science from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

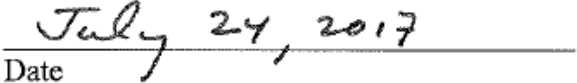

Name of Chairperson


Name of Committee Member


Name of Committee Member

This thesis is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.


Grant McGimpsey
Dean of the School of Graduate Studies


Date

PERMISSION

Title Assessment of WRF Simulated Precipitation by Meteorological Regimes

Department Atmospheric Science

Degree Master of Science

Department Atmospheric

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in her absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Brooke Hagenhoff

7/25/2017

TABLE OF CONTENTS

LIST OF FIGURES	viii
LIST OF TABLES	xiv
ACKNOWLEDGMENTS	xv
ABSTRACT	xvi
CHAPTER	
I. INTRODUCTION	1
Background	1
Purpose of the Study	4
Impacts of the Study	5
II. METHODOLOGY	6
Datasets and Techniques	6
NCEP NARR	6
NCEP Stage IV	7
NSSL-WRF	9
Self-Organizing Maps	10

Methods.....	11
Case Selection.....	11
SOM Generation.....	13
Precipitation Evaluation.....	15
Statistical Analysis.....	15
III. RESULTS	19
General Characteristics	19
Example SOM.....	21
Southern Great Plains (SGP)	23
CDF90.....	29
Statistical Analysis.....	35
CDF50.....	38
Statistical Analysis.....	43
Discussion.....	46
Northern Plains (NP).....	48

CDF90.....	54
Statistical Analysis.....	59
CDF50.....	62
Statistical Analysis.....	68
Discussion.....	71
Case Study	73
IV. CONCLUSIONS	75
Future Work.....	77
REFEFERENCES.....	79

LIST OF FIGURES

Figure	Page
1. The Northern Plains and Southern Great Plains regions of interest in this study.....	5
2. Domains for the 12 RFCs	8
3. Euclidian distance of each class from class 1 (upper left) for the SGP CDF90 SOM.....	17
4. Euclidian distance of each class from class 1 (upper left) for the SGP CDF50 SOM.....	17
5. The 500 hPa analysis from the NARR for 0000 UTC on 20 May 2013.....	21
6. The 500 hPa height anomalies for a 7x4 (28-class) SOM centered on SGP	22
7. Euclidian distance for 20 May 2013 on the SGP CDF90 SOM	22
8. Near-surface analyses for the 9x6 (54-class) SGP climatology SOM.....	24
9. Climatology SOM for SGP 500 hPa geopotential height anomalies.....	24
10. Daily average precipitation for the SGP climatology SOM	25
11. Average precipitation in six-hour increments for the SGP climatology SOM.....	25
12. Daily average precipitation bias for the SGP climatology SOM.....	26
13. Average precipitation bias in six-hour increments for the SGP climatology SOM.....	27

14. Distribution of CDF90 classes within the SGP climatology SOM.....	28
15. Distribution of CDF50 classes within the SGP climatology SOM.....	28
16. Percent of days in each class with precipitation	29
17. Near-surface analyses for the SGP CDF90 SOM.....	30
18. The 500 hPa geopotential height anomalies for the SGP CDF90 SOM.....	30
19. Number of cases in each class of the SGP CDF90 SOM	31
20. Daily average precipitation for the SGP CDF90 SOM.....	32
21. Average precipitation in six-hour increments for the SGP CDF90 SOM	32
22. Daily average precipitation bias for the SGP CDF90 SOM	34
23. Average precipitation bias in six-hour increments for the SGP CDF90 SOM.....	34
24. SGP CDF90 percent of cases in each class with maximum six-hour accumulated precipitation occurring for 0600-1100 UTC	35
25. Near-surface analyses for the SGP CDF90 averaged SOM.....	36
26. The 500 hPa geopotential height anomalies for the SGP CDF90 averaged SOM.....	36
27. Daily average precipitation bias for the SGP CDF90 averaged SOM.....	37
28. Number of cases in each class of the SGP CDF90 averaged SOM	38

29. Plot depicting where the SGP CDF50 cases are located on the SGP CDF90 SOM.....	39
30. Near-surface analyses for the SGP CDF50 SOM.....	40
31. The 500 hPa geopotential height anomalies for the SGP CDF50 SOM.....	40
32. Daily average precipitation for the SGP CDF50 SOM.....	41
33. Daily average precipitation bias for the SGP CDF50 SOM	41
34. Average precipitation in six-hour increments for the SGP CDF50 SOM	42
35. Average precipitation bias in six-hour increments for the SGP CDF50 SOM.....	42
36. Number of cases in each class of the SGP CDF50 SOM	43
37. Near-surface analyses for the SGP CDF50 averaged SOM.....	44
38. The 500 hPa geopotential height anomalies for the SGP CDF50 averaged SOM.....	44
39. Daily average precipitation bias for the SGP CDF50 averaged SOM.....	45
40. Number of cases in each class of the SGP CDF50 averaged SOM	46
41. Distribution of SGP CDF90 cases that occur for each pattern by month	48
42. Near-surface analyses for the 9x6 (54-class) NP climatology SOM	49
43. Climatology SOM for NP 500 hPa geopotential height anomalies	49

44. Daily average precipitation for the NP climatology SOM.....	50
45. Average precipitation in six-hour increments for the NP climatology SOM	50
46. Daily average precipitation bias for the NP climatology SOM	51
47. Average precipitation bias in six-hour increments for the NP climatology SOM.....	51
48. Distribution of CDF90 classes within the NP climatology SOM.....	52
49. Distribution of CDF50 classes within the NP climatology SOM.....	53
50. Percent of days in each class with precipitation	53
51. Near-surface analyses for the NP CDF90 SOM	54
52. The 500 hPa geopotential height anomalies for the NP CDF90 SOM	55
53. Daily average precipitation for the NP CDF90 SOM.....	55
54. Daily average precipitation bias for the NP CDF90 SOM	56
55. Average precipitation in six-hour increments for the NP CDF90 SOM.....	57
56. Average precipitation bias in six-hour increments for the NP CDF90 SOM	58
57. Number of cases in each class of the NP CDF90 SOM.....	58
58. Near-surface analyses for the NP CDF90 averaged SOM.....	59

59. The 500 hPa geopotential height anomalies for the NP CDF90 averaged SOM.....	60
60. Daily average precipitation bias for the NP CDF90 averaged SOM	60
61. Number of cases in each class of the NP CDF90 averaged SOM	62
62. Plot depicting where the NP CDF50 cases are located on the NP CDF90 SOM	63
63. Near-surface analyses for the NP CDF50 SOM	63
64. The 500 hPa geopotential height anomalies for the NP CDF50 SOM	63
65. Daily average precipitation for the NP CDF50 SOM.....	66
66. Daily average precipitation bias for the NP CDF50 SOM	66
67. Average precipitation in six-hour increments for the NP CDF50 SOM.....	67
68. Average precipitation bias in six-hour increments for the NP CDF50 SOM	67
69. Number of cases in each class of the NP CDF50 SOM.....	68
70. Near-surface analyses for the NP CDF50 averaged SOM.....	69
71. The 500 hPa geopotential height anomalies for the NP CDF50 averaged SOM.....	69
72. Daily average precipitation bias for the NP CDF50 averaged SOM	70
73. Number of cases in each class of the NP CDF50 averaged SOM	70

74. Distribution of NP CDF90 cases that occur for each pattern by month	72
75. NP CDF90 percent of cases in each class with maximum six-hour accumulated precipitation occurring for 0600-1100 UTC	72
76. Surface (left) and 500 hPa (right) analysis for the 13 June 2017 case study	73
77. Accumulated precipitation amounts for NSSL WRF (left) and Stage IV (right) for the 13 June 2017 case study	74

LIST OF TABLES

Table	Page
1. List and description of the SOMs used in this study	13
2. SOM settings used in this study.....	15
3. Average precipitation bias for the SOMs.....	20
4. Z-Test statistic results for the SGP CDF90 SOM.....	37
5. Z-Test statistic results for the SGP CDF50 SOM.....	46
6. Z-Test statistic results for the NP CDF90 SOM	61
7. Z-Test statistic results for the NP CDF50 SOM.....	71

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to the members of my advisory committee, Drs. Xiquan Dong and Matthew Gilmore, for their assistance and useful comments throughout the entire thesis process.

A special thank you to my advisor, Dr. Aaron Kennedy, for his guidance and insight during my time at the University of North Dakota. Thanks to his encouragement I attempted, and was even quite successful, at things I previously accepted as characteristic shortcomings in myself. These successes have not only increased my self-confidence, but have also helped propel me forward towards career goals which to me at one time seemed a distant dream. For the encouragement given and confidence gained I will be forever thankful.

To Dr. DaNa Carlis, who first encouraged to me attend graduate school and gave me the confidence to take such a step, and to Dr. Fred Carr, my University of Oklahoma advisor who offered support throughout my time as an undergraduate and encouraged me to consider the University of North Dakota. Without them I would not be writing this at all.

And last, to my fellow graduate student friends for the support as we muddled through this thing called graduate school together, and to my family for supporting me as I chose yet again to move across the country to chase a dream. Thank you so much.

ABSTRACT

This study evaluated warm-season precipitation events in a multi-year (2007-2014) database of Weather Research and Forecasting (WRF) simulations over the Northern Plains and Southern Great Plains. These WRF simulations were run daily in support of the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) by the National Severe Storms Laboratory (NSSL) for operational forecasts.

Evaluating model skill by synoptic pattern allows for an understanding of how model performance varies with particular atmospheric states and will aid forecasters with pattern recognition. To conduct this analysis, a competitive neural network known as the Self-Organizing Map (SOM) was used. SOMs allow the user to represent atmospheric patterns in an array of nodes that represent a continuum of synoptic categorizations. North American Regional Reanalysis (NARR) data during the warm season (April-September) was used to perform the synoptic typing over the study domains. Simulated precipitation was evaluated against observations provided by the National Centers for Environmental Prediction (NCEP) Stage IV precipitation analysis.

CHAPTER I
INTRODUCTION

Background

To generate better model forecasts of convective storms and resultant precipitation, it is important to continually improve the models that are used to explicitly simulate these storms. Output from so-called Convection Allowing Models (CAMs) such as the Weather Research and Forecasting Model (WRF) has the potential to improve forecasts by providing greater detail to operational forecasters (Kain et al. 2006). CAMs have a horizontal grid spacing of ≤ 4 km, as this is the coarsest spacing to sufficiently represent the evolution and structure within Mesoscale Convective Systems (MCS, Weisman et al. 1997). Smaller grid spacing is computationally time consuming, so this grid spacing is a compromise between being coarse enough to generate forecasts quickly enough for operational use and yet fine enough to resolve smaller scale features. Done et al. (2004) showed that WRF with 4 km grid spacing could depict realistic MCSs and improve the forecasts of convective mode and daily frequency of convection when compared to coarser simulations. Similarly, during the 2004 Storm Prediction Center (SPC) – National Severe Storms Laboratory (NSSL) Spring Program, convection-allowing WRF simulations were found to outperform the lower resolution Eta model for convective initiation, storm evolution, and convective mode (Kain et al. 2006). These benefits have pushed CAMs to the frontier of forecasting.

Despite these advantages, the high resolution of CAMs can make it more difficult to evaluate the accuracy of the model's forecast. Using traditional forecast verification statistics that only match spatial placement of storms, a forecast model that generates a storm with fine-scale (more realistic) but displaced precipitation maxima scores worse when compared to observations than a coarse resolution forecast (less realistic) lacking small-scale features (Baldwin et al. 2001). Mass et al. (2002) found that by using traditional verification methods such as mean absolute deviation, bias, and root-mean square error, there was an increase in accuracy from a model with 36 km to 12 km grid spacing, but less of a change when moving from 12 km to 4 km grid spacing. It was suggested that the lack of improvement was due to scores which fail to account for fine-scale features or spatial displacement of convection, a claim further supported by Clark et al. (2007).

There are alternatives to traditional statistical measures which may be useful in evaluating CAM performance. Studies such as Carbone et al. (2002) and Goines and Kennedy (2017) utilized Hovmöller diagrams and found that the models have varying skill in predicting propagating precipitation and scattered, diurnally forced precipitation. Alternatively, performance can be separated by varying meteorological conditions. Because precipitation and storm mode is influenced by synoptic patterns (Parker and Johnson 2000), a regime based analysis of model performance would help to identify specific synoptic patterns that models struggle with, and potentially hint at ways to improve model performance. A regime based analysis can range from simpler approaches, such as defining the dynamic regime by 500 hPa vertical motion as in Tselioudis and Jakob (2002), to more complicated approaches such as k-means

clustering. For example, Theobald et al. (2015) utilized the latter method to classify synoptic patterns to understand high precipitation events in the Snowy Mountains of Australia.

Self-Organizing Maps (SOMs, Kohonen et al. 1996), a type of competitive neural network, have been utilized to discriminate meteorological patterns. In Cavazos (1999), SOMs were used to classify winter circulation and humidity patterns for extreme precipitation events in Northeastern Mexico and Southeastern Texas. The results helped identify differences in atmospheric patterns which were more likely to produce extreme precipitation events. Hewitson and Crane (2002) applied SOMs to average precipitation associated with synoptic circulations in an effort to show how SOMs can be used to evaluate trends in data, and found that over a 40 year period the same synoptic circulation results in more precipitation now than it did 40 years ago, illustrating the usefulness of SOMs in climate studies. Cassano et al. (2006) used SOMs to analyze synoptic patterns in the Arctic during the summer and winter seasons using sea level pressure data from ten global climate models (GCMs) and reanalysis data. The study found that SOMs were useful in evaluating differences between ensemble simulations and reanalysis data, as well as differences between ensemble members. More recently, SOMs were used to identify cloud biases in climate models under different synoptic patterns (Kennedy 2011) and improve the climatology of clouds (Kennedy et al. 2016). Because of the range of patterns generated in SOMs based on the input cases, they are useful to evaluate model performance and to determine trends in data when cases are grouped and analyzed based on the atmospheric pattern.

Overall, the use of SOMs has largely been confined to long term datasets and climate studies. Historically, studies of CAMs such as WRF studies have been limited to shorter time periods or specific case studies. The few number of cases in past studies has prohibited the use of clustering techniques. In recent years, however, CAMs have been operated daily. For example, the National Severe Storms Laboratory (NSSL) has run daily 4 km WRF simulations in support of the Hazardous Weather Testbed (HWT). Making use of eight years of model runs from 2007-2014 allows for a long enough duration that SOMs could be applied. This provides the opportunity for a unique look at model performance for varying synoptic regimes over an extended time period.

Purpose of the Study

This study will evaluate how the performance of the NSSL WRF in simulating precipitation amounts varies for different synoptic patterns. These patterns will be classified using SOMs. The average precipitation bias for each pattern will be calculated to determine how, if at all, the precipitation bias varies. Because CAMs have predominately been used for convective forecasts, only the warm season (April-September, 2007-2014) will be investigated. Considering the entire country does not experience the same frequency of synoptic patterns and that SOMs offer a Euclidean look at these regimes, the United States has been sectioned into five regions to be analyzed separately. These regions include the Southern Great Plains (SGP), Northern Plains (NP), Midwest (MW), Gulf Coast (GC), and Northeast (NE), with a focus on the SGP and NP regions (Figure 1).

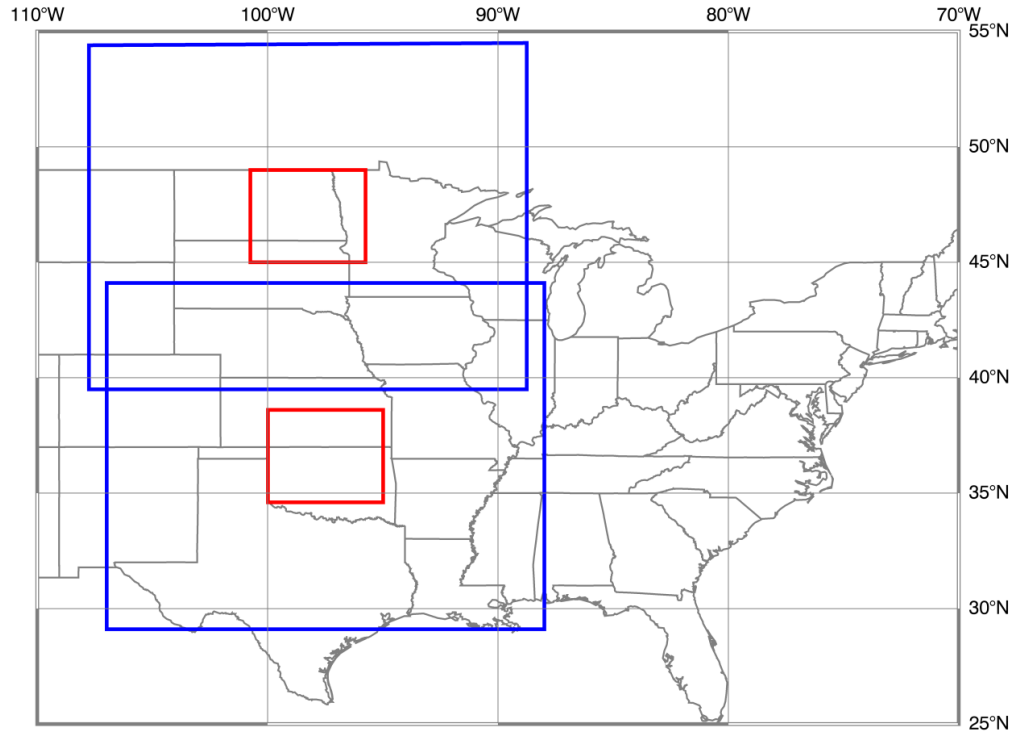


Figure 1. The Northern Plains and Southern Great Plains regions of interest in this study. Blue boxes represent the domains for the classification of meteorological patterns and the smaller red boxes represent the areas where precipitation was analyzed.

Impacts of the Study

Evaluating model biases by synoptic pattern has a number of direct impacts on the meteorological community. First, this type of study will aid forecasters in using CAM data in operational forecasting by recognizing model trends for particular synoptic patterns (e.g. pattern recognition). Second, this work has the potential to shed light on the physical reasons when the model is struggling and provide insight on how to improve the model. Finally, as climate change occurs, synoptic patterns may become more or less frequent, rendering a model more or less useful based on how well precipitation is simulated under a particular pattern.

CHAPTER II

METHODOLOGY

Datasets and Techniques

This section provides details regarding the datasets and techniques used to complete this study. To classify synoptic patterns the National Center for Environmental Protection (NCEP) North American Regional Reanalysis (NARR) is used. Rainfall measurements come from the NCEP Stage IV precipitation dataset. The evaluated model is the National Severe Storms Lab (NSSL) version of the WRF model run in support of the NOAA HWT. The SOM technique is also described.

NCEP NARR

The NCEP NARR is a reanalysis dataset with a variety of output variables useful for synoptic typing across the Continental United States (CONUS). NARR uses the NCEP Eta model and its 3D-VAR Data Assimilation System (EDAS) to assimilate observational data from radiosondes, satellites, aircraft, and surface observations. Unique to this reanalysis, assimilation of precipitation is performed to insure that model precipitation is close to observed values (Mesinger et al. 2006). The improved precipitation in NARR has been verified in studies by Becker et al. (2009) and Bukovsky and Karoly (2007). The data available dates from 1979 to present with 32 km horizontal

grid spacing, 45 layer vertical levels, and 3 hourly output (Mesinger et al. 2006). The output variables make NARR useful in representing the atmospheric state and it is significantly improved over previous NCEP reanalysis datasets (Mesinger et. al 2006). Variables used in this study include mean sea level pressure (MSLP), 900 hPa relative humidity (RH) and winds along with 500 hPa RH, winds (u and v components), and geopotential heights. These variables were selected because the two levels are frequently utilized in an operation setting. Additionally, these variables provide insight into moisture and flow at lower and upper levels. The 500 hPa geopotential heights were used to calculate 500 hPa geopotential height anomalies across domains of interest in order to mitigate seasonal variations which can bias SOMs (Kennedy et al. 2016). Additionally, 900 hPa RH and winds were used rather than surface values because there is less confidence in the accuracy of surface variables in NARR (Mesinger et al. 2006). Overall, Kennedy et al. (2011) has shown that NARR compares well with observed soundings over the Southern Plains, making it a useful dataset in representing the synoptic patterns in this study.

NCEP Stage IV

Stage IV data mosaics the regional hourly, six-hourly, and 24-hourly precipitation analyses produced by the 12 River Forecast Centers (RFCs) over the CONUS (Figure 2). The process of producing Stage II and III products, which are then used for Stage IV, is described in Briendenbach et al. (1998) and is summarized below.

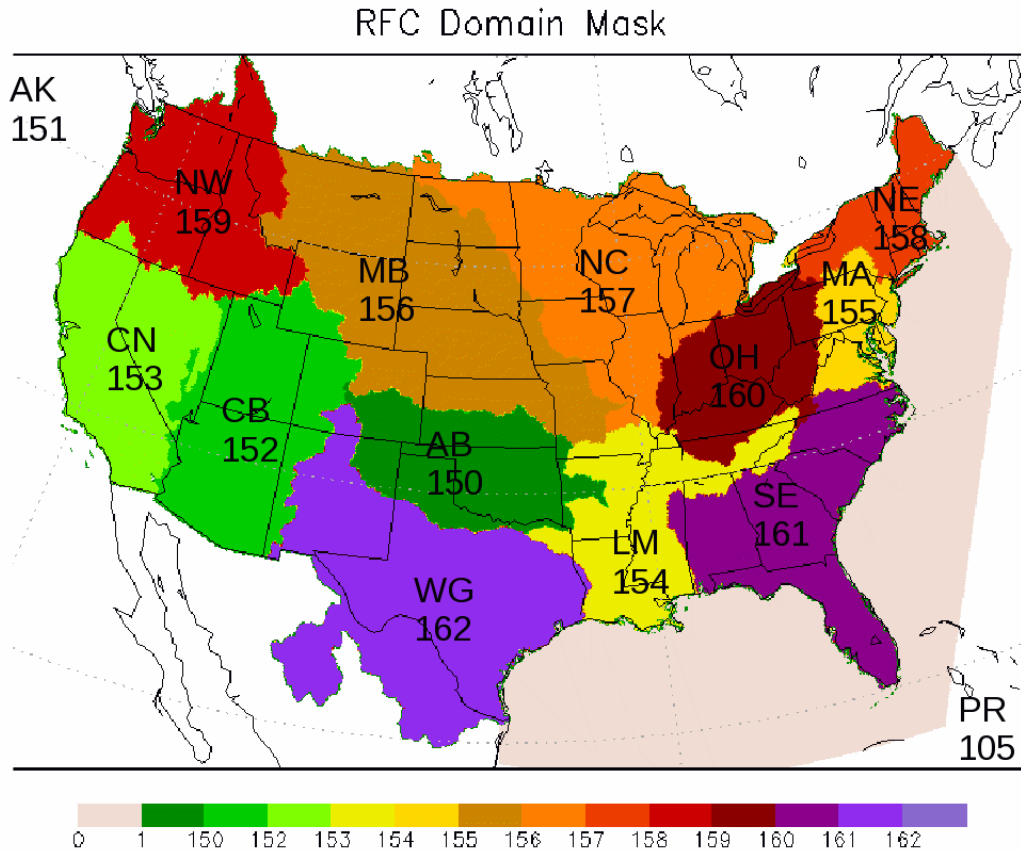


Figure 2. Domains for the 12 RFCs. From <http://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/stage4/>.

Generation of Stage IV data begins with estimated hourly precipitation totals from the National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS) Next-Generation Radars (NEXRAD). This product is known as Stage I precipitation data. RFCs adjust this data using precipitation gauges to create Stage II data. This Stage II merged dataset is computed out to 230 km from the radar using the hourly digital precipitation (HDP) product from the Stage I Post Processing System. This is done by computing the mean bias between the gauges and radar using a Kalman filter approach (Smith and Krajewski 1991) and is then used to adjust the HDP rainfall estimate. To account for inconsistent radar biases due to range or precipitation type (e.g. convective or non-convective), the importance of the radar estimate increases with increasing distance from the gauge. The Stage II estimates for multiple radars within the

RFC domain are then mosaicked to create the Stage III product. Where two or more radars overlap, the office has the choice of using either the mean or maximum value. Stage III also allows for interactive quality control of the radar and rain gauge precipitation amounts.

The Stage III products from each RFC are mosaicked by NCEP into a national product, retaining the manual quality control performed by the RFCs (Lin and Mitchell 2005). Because not all RFCs use the same precipitation quality control algorithms, there can be discontinuities between offices (Prat and Nelson 2015, Goines and Kennedy 2017). However, Wescott et al. (2008) and Gourley et al. (2010) compared Stage IV data with other products which do not use manual quality control and found that Stage IV data outperforms those products. In addition, Wu et al. (2012) illustrated the quality of Stage IV data by using it to evaluate other precipitation products.

NSSL-WRF

This study utilizes NSSL-WRF simulations from 2007-2014, which were run in part to support the Hazardous Weather Testbed (HWT) Spring Forecasting Experiment for operational forecasts. An advantage over short term verification studies is that the eight years' worth of simulations allow for analysis of the model's performance under different synoptic regimes without the daily or small-scale variations in performance which can skew results when only one or a few cases are analyzed. The longer dataset is beneficial from a statistical perspective of including more cases from each synoptic regime. The NSSL-WRF model configuration is version 3.4.1 and has 4-km horizontal grid spacing with 35 vertical levels and a 24 second time step. The initial and boundary

conditions are obtained from interpolation of the 40 km North American Mesoscale Forecast System (NAM) and the microphysics parameterization scheme used is the WRF Single Moment microphysics scheme with 6 water classes (WSM 6): water vapor, cloud water, rain, graupel, cloud ice, and snow (Hong and Lin 2006). Radiation schemes include the Dudhia shortwave scheme (Dudhia 1989) and RRTM longwave radiation (Mlawer et al. 1997). Land-atmosphere interactions are simulated using the Noah land-surface model (Mitchell et al. 2005). Boundary layer and turbulence are parameterized by the Mellor–Yamada–Janjic (MYJ) scheme (Janjic 1994). The model uses the positive definite advection (PDA) scheme for moisture variables, which has been shown to improve precipitation verification (Hahn and Mass 2009). It is initialized at 0000 and 1200 UTC with forecasts out to 36 hours. This study utilizes the 0000 UTC run and only the 12-36 hour (1200-1100 UTC) time period each day to allow for model spin-up in the first 12 hours. While newer versions of the WRF model and individual parametrizations are available (e.g. double moment vs. single moment microphysics schemes), this configuration has been kept because of the familiarity that Storm Prediction Center (SPC) forecasters have with the model (personal communication, Greg Carbine, SPC).

Self-Organizing Maps

Self-Organizing Maps (SOMs) allow the user to represent the data in an array of classes known as nodes that self-organize and represent an entire continuum of the dataset (Kohonen et al. 1996). This is similar to k-means clustering; however, a neighborhood function smooths the data and allows for classes that span the dimensions of the dataset. The nodes are initialized off of random values and as training vectors are added to the SOM, the node with the minimum Euclidian distance from the input vector

is selected and the case is classified to the winning node (Hewitson and Crane 2002). Through an iterative training process, each node is modified towards the input vector while surrounding nodes are partially modified (Kohonen et al. 1996). The training occurs in two steps, first by iterating over each case once with a large learning rate and neighborhood radius to orient the SOM. In the second step, the input cases are iterated over multiple times with a lower learning rate and neighborhood radius to converge to a final solution. This process allows the nodes to self-organize where nodes are related to one another via a two-dimensional feature map with more similar nodes closer together and less similar nodes farther apart, such that the corners of the SOM tend to represent the extreme cases with a smooth continuum in between (Sheridan and Lee 2011).

Methods

Case Selection

Precipitation bias in a model can occur for several reasons. The model may produce precipitation for the wrong atmospheric state or produce too much or too little precipitation for an atmospheric state. To understand these types of biases, two types of SOMs were developed. First, climatology SOMs were created using every day during the time period (2007-2014, April - September), regardless of precipitation, to determine whether the model has precipitation biases (positive or negative) across any of the synoptic patterns. Precipitation analysis was done for a $5^{\circ} \times 4^{\circ}$ (longitude by latitude) area within a larger $19^{\circ} \times 15^{\circ}$ region for synoptic patterns to focus on precipitation occurring at the center of the pattern. A sensitivity test using $2.5^{\circ} \times 2^{\circ}$ and $1^{\circ} \times 1^{\circ}$ areas showed that the 5×4 area exhibited the best correlation of model precipitation with observations.

The second type of SOM was developed from only precipitating cases. Utilizing a cumulative distribution function (CDF) of daily precipitation, cases were selected by thresholds that determined a specified percentage of rain within the CDF. While precipitation occurs over many different patterns, this type of SOM allowed for an analysis of patterns that have a significant contribution to annual rainfall. The cases examined using this technique contributed to the upper 50% and 90% of daily precipitation totals in each of the five regions, as those are potentially high-impact events. The cases in the CDF90 SOM were responsible for 87.7% of the seasonal precipitation in SGP and 86.9% of the seasonal precipitation in NP, while CDF50 SOM cases were responsible for 46.7% of seasonal precipitation in SGP and 45.8% of seasonal precipitation in NP. These more detailed SOMs provide better insight into mechanisms that may be responsible for any model precipitation bias. In summary, three sets of SOMs were produced for each region:

- Climatological: Produced from all available warm season days (precipitating and non-precipitating).
- CDF90: Produced from precipitation days within the upper 90% of the CDF
- CDF50: Produced from precipitation days within the upper 50% of the CDF

A full list and description of the SOMs created is shown in Table 1.

Table 1. List and description of the SOMs used in this study.

SOM	Region	Number of Cases	SOM Dimensions (x,y)	Description
climo	SGP	1366	9x6	all days
cdf90	SGP	387	7x4	at least 3 mm of total precip
cdf50	SGP	106	6x3	at least 12 mm of total precip
climo	NP	1374	9x6	all days
cdf90	NP	421	7x4	at least 2 mm of total precip
cdf50	NP	111	6x3	at least 8 mm of total precip
climo	MW	1367	9x6	all days
cdf90	MW	516	7x4	at least 2 mm of total precip
cdf50	MW	216	7x4	at least 8 mm of total precip
climo	GC	1369	9x6	all days
cdf90	GC	505	7x4	at least 3 mm of total precip
cdf50	GC	156	6x3	at least 9 mm of total precip
climo	NE	1335	9x6	all days
cdf90	NE	436	7x4	at least 3 mm of total precip
cdf50	NE	131	6x3	at least 10 mm of total precip

SOM Generation

To create the SOMs, the variables selected from NARR data were averaged to a 1° x 1° longitude by latitude, grid. This grid-spacing provides synoptic to meso-alpha scale detail in the SOM-generated synoptic patterns. This grid was then utilized for the previously mentioned 19 ° x 15 ° area for each domain. In total this yielded a vector of 285 (19 x 15) elements for any given variable. As shown in Figure 1, the area over which the synoptic patterns are made is significantly larger than the area where the precipitation is evaluated. This is done to ensure precipitation events lie well within the large-scale synoptic pattern and forcing.

SOMs were trained using all of the atmospheric variables previously described; near-surface and 500 hPa levels were chosen as these levels are often used in an

operational setting and therefore useful for pattern recognition purposes. The variables were normalized to a common range such that each variable provided equal weight to the classification of the SOM.

Common to many clustering algorithms, the user must determine the appropriate number of classes (or nodes; “class” may be used interchangeably with “node” hereafter). Too few classes can cause important differences in atmospheric states to be smoothed out while too many classes can result in classes that never occur and have no cases classified to a specific node. The number of classes selected was based on work by Kennedy et al. (2016). Climatology SOMs were created with 54 classes (9x6) while SOMs made for the precipitating cases, discussed below, had 28 (7x4) or 18 (6x3) classes. These numbers ensured that cases were classified into each and every node of the SOMs. Typically a 28-class scheme was used unless the number of cases fell below 150, and then an 18-class scheme was used. This helped ensure every class would have at least one case classified to it because there would be fewer classes when fewer cases were to be classified.

The SOMs were generated using freely available SOM_PAK software (Kohonen et al. 1996) and generation of the SOMs followed the methodology of Kennedy et al. (2016). The selected and normalized variables from NARR, at 0000 UTC to capture environment when convection is more likely to occur, were put into input vectors to train the SOM. With 285 points the each region (19x15) and 8 variables, vectors were 2280 elements long. Each SOM was initialized ten times and the SOM with the smallest average error (based off classified Euclidean distance) was saved and used in this study. The SOM settings used in this study are summarized in Table 2. An example of the classification technique is provided in the Results section.

Table 2. SOM settings used in this study.

SOM settings	Value
Trials	10
Training Length (stage 1)	# of cases
Training Length (stage 2)	# of cases *100
Learning Rate (stage 1)	.05
Learning Rate (stage 2)	.01
Neighborhood Radius (stage 1)	Xdim-1
Neighborhood Radius (stage 2)	1

Precipitation Evaluation

Precipitation for the SOM was plotted as the average precipitation total for the cases in each class, averaged over the 5° x 4° domain. This size was chosen to be large enough to account for precipitation that may be spatially displaced but still small enough to focus on precipitation that was well within the forcing of the synoptic pattern. This study analyzed not only the total precipitation and respective model bias, but also the bias in six-hour increments throughout the case day. This was done to evaluate how the precipitation bias evolved as the day progressed. The days were sectioned off as follows: 1200-1700 UTC, 1800-2300 UTC, 0000-0500 UTC, and 0600-1100 UTC.

Statistical Analysis

To determine the statistical significance of biases, SOMs were reduced to six patterns by grouping similar atmospheric patterns together. This aided the analysis by increasing the number of cases classified to each pattern, which made statistical testing more robust. In addition, this is useful in an applied sense because forecasters can quickly identify a pattern from six predominant patterns vs. trying to distinguish patterns in the larger SOMs.

To objectively group classes, patterns were grouped together based on the mean Euclidian distance between classes. This was done by selecting the vector of data representing one class and comparing it to all other classes and calculating the difference, much in the same way as one would measure the distance between two points. This process was performed for each class within the larger SOM. A set pattern was established to group similar classes. For the 7x4 (28 class) SOMs, four classes were grouped together in each of the corners while six classes were grouped together in the center (Figure 3). This configuration resulted because corner classes were spread further apart (in distance) than those in the center. Less classes were grouped at the corners because the spread in classes is larger, resulting in higher error with surrounding classes. The same method was used to group cases for the 6x3 (18 class) SOMs, shown in Figure 4. From there the average synoptic patterns could be plotted and the average bias for the new patterns was recalculated. This method of grouping together patterns from the original SOM rather than creating a new SOM with fewer classes was used to allow for the defining features of the SOM to still be distinguishable rather than being smoothed out by the SOM generation process.

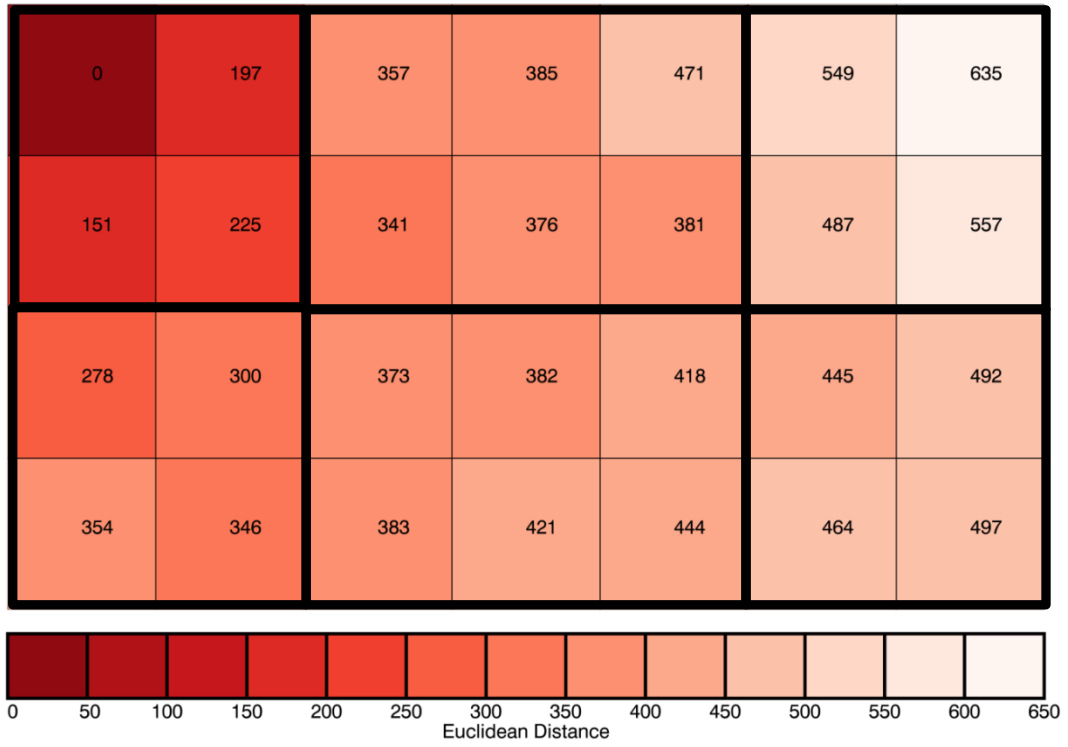


Figure 3. Euclidian distance of each class from class 1 (upper left) for the SGP CDF90 SOM. Bold black boxes indicated which classes where averaged together for statistical analysis of 7x4 SOMs.

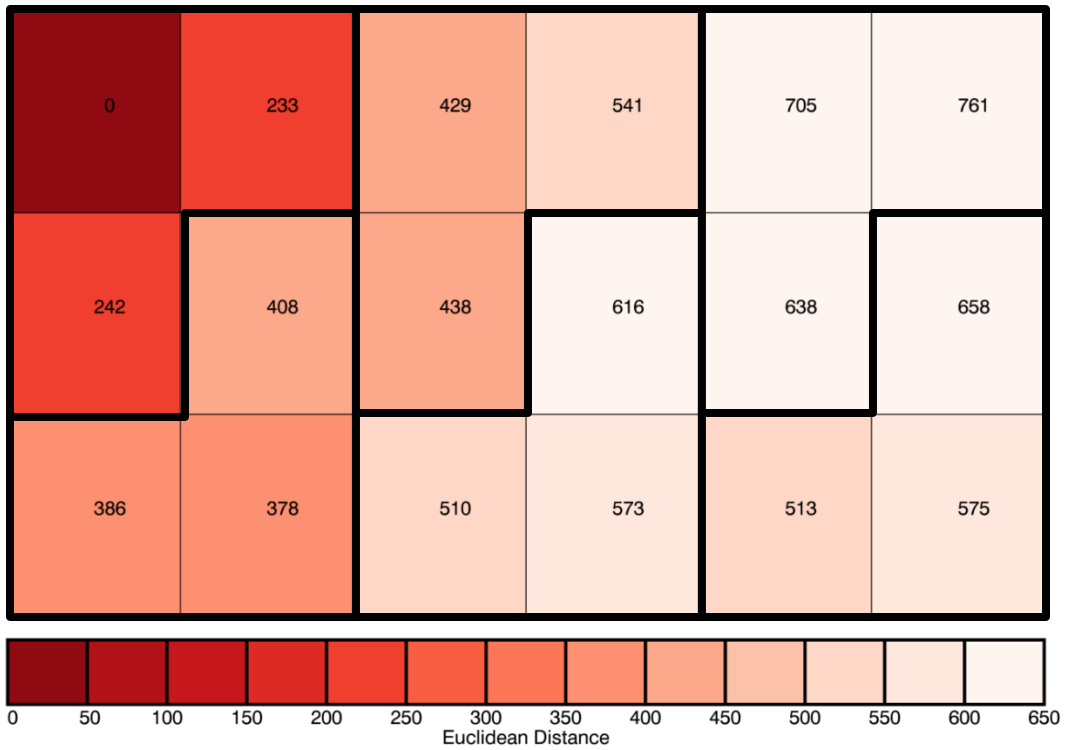


Figure 4. Euclidian distance of each class from class 1 (upper left) for the SGP CDF50 SOM. Bold black boxes indicated which classes where averaged together for statistical analysis of 6x3 SOMs.

To calculate whether the differences in the biases were statistically significant between grouped SOM classes, a z-test was used (Equation 1).

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{(\hat{var}[\bar{x}_1 - \bar{x}_2])^{1/2}} \quad (1)$$

For this study, the null hypothesis represents the situation where the mean precipitation bias being compared between classes is equal. To reject this hypothesis at the 95% significance level, the numerator must be approximately double the denominator ($z \geq 1.96$). Results from these tests are discussed in Chapter III.

CHAPTER III

RESULTS

General Characteristics

Prior to investigating the SOMs, average biases were calculated for the three sets of cases, and these biases were calculated for the entire forecast period of each case (1200-1100 UTC). These biases were further broken down into 6-hr increments throughout the day to evaluate how the precipitation bias evolved (Table 3). A few notable properties stand out and these results are consistent with work from Goines and Kennedy (2017). First, NSSL WRF simulations generally over predict total daily precipitation, regardless of region. Second, the positive bias is associated with a strong diurnal signal in convection, as seen by larger positive biases between 1800-2300 UTC. This positive bias is offset by negative biases during the overnight hours (0600-1100 UTC) for nearly every region except the NE. The CDF50 sets of cases for the MW, GC, and SGP regions show that the negative bias for the daily average is due to a strong, negative nocturnal precipitation bias which overcomes the positive bias from earlier in the day. In the remaining regions for the CDF50 cases, positive biases are lower than the more comprehensive lists of cases, again because of the strong negative bias occurring overnight for these higher precipitation cases. For these regions, the overnight negative

bias is not high enough to balance out the daytime positive bias. This suggests that overall positive biases are the result of weaker precipitating cases, because the negative nocturnal bias is dominant for the heavier precipitation cases. Further, the significant overnight negative biases predominately occur in regions where propagating, nocturnal convection occurs and contributes significantly to observed precipitation (Carbone et al. 2002). This result is consistent with Goines and Kennedy (2017) who found shorter precipitation streaks within Hovmöller diagrams across this region for the NSSL WRF model.

Table 3. Average precipitation bias for the SOMs. Biases are separated by time period; total refers to the entire day (1200-1100 UTC). Color shading, ranging from dark blue (strongly negative) to white (near zero) to dark red (strongly positive), is used to more easily see the patterns in this table.

Region	SOM	Cases	Average Bias (mm)					Description
			Total	1200 to 1700 UTC	1800 to 2300 UTC	0000 to 0500 UTC	0600 to 1100 UTC	
SGP	climo	1366	0.24	0.13	0.23	0.13	-0.25	all days
	cdf90	387	-0.38	0.10	0.39	0.15	-1.02	at least 3mm of precip
	cdf50	106	-2.90	-0.03	0.20	-0.39	-2.68	at least 12mm of precip
NP	climo	1374	0.57	0.13	0.28	0.20	-0.02	all days
	cdf90	421	0.99	0.21	0.53	0.43	-0.19	at least 2mm of precip
	cdf50	111	0.25	0.20	0.72	0.09	-0.76	at least 8mm of precip
MW	climo	1367	0.46	0.02	0.39	0.15	-0.10	all days
	cdf90	516	0.49	-0.07	0.67	0.28	-0.39	at least 2 mm of precip
	cdf50	216	-0.31	-0.15	0.88	-0.02	-1.03	at least 8 mm of precip
GC	climo	1369	0.74	0.20	0.50	0.05	-0.01	all days
	cdf90	505	0.90	0.29	0.65	0.04	-0.08	at least 3 mm of precip
	cdf50	156	-0.12	-0.18	0.41	-0.04	-0.30	at least 9 mm of precip
NE	climo	1335	0.64	0.11	0.41	0.03	0.10	all days
	cdf90	436	0.93	0.09	0.70	-0.06	0.19	at least 3 mm of precip
	cdf50	131	0.31	-0.19	0.46	-0.34	0.39	at least 10 mm of precip

Example SOM

To illustrate how the SOM technique works, Figure 5 shows a plot of the 500 hPa winds and geopotential heights for 20 May 2013, one of the precipitation cases selected for SGP. The SOM here is for the SGP CDF90 cases (which include 20 May 2013) at 500 hPa, and shows patterns ranging from troughs on the left side of the SOM to flow out of the southwest or northwest in other areas of the SOM (Figure 6). This case had strong upper-level flow out of the southwest associated with an upper-level trough. The SOM precipitation domain is outlined in blue on the figure. On the SOM this is a class that also has strong flow out of the southwest. Figure 7 shows the mean Euclidian distance - or average error - between that individual case and each class, showing how the case is classified to the class that is the close match (has the lowest error). This is a way to illustrate how the SOM technique matches individual cases to classes.

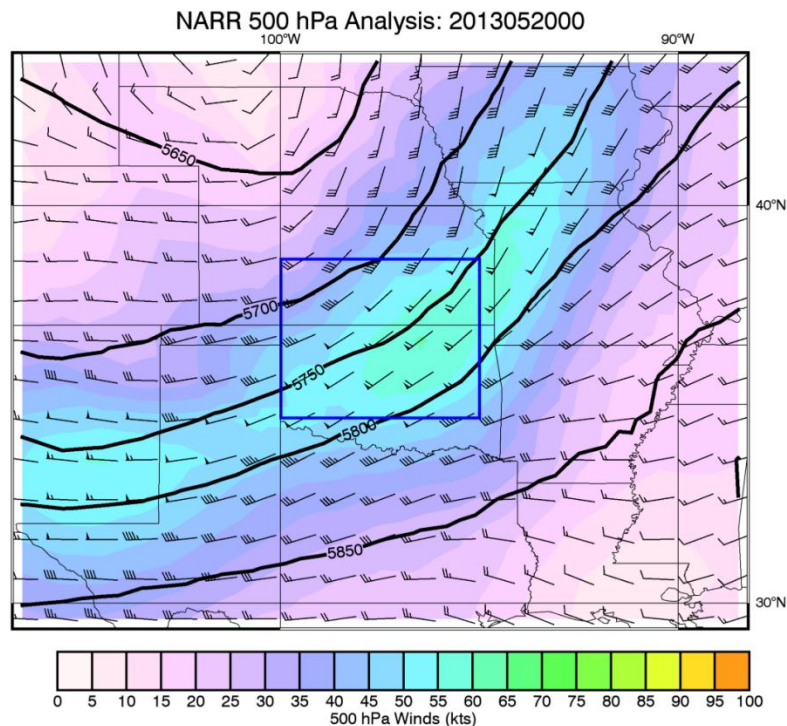


Figure 5. The 500 hPa analysis from the NARR for 0000 UTC on 20 May 2013. Solid black lines are geopotential heights (meters) while wind magnitudes (knots) are plotted using the traditional station model (wind speed and direction) as well as shaded.

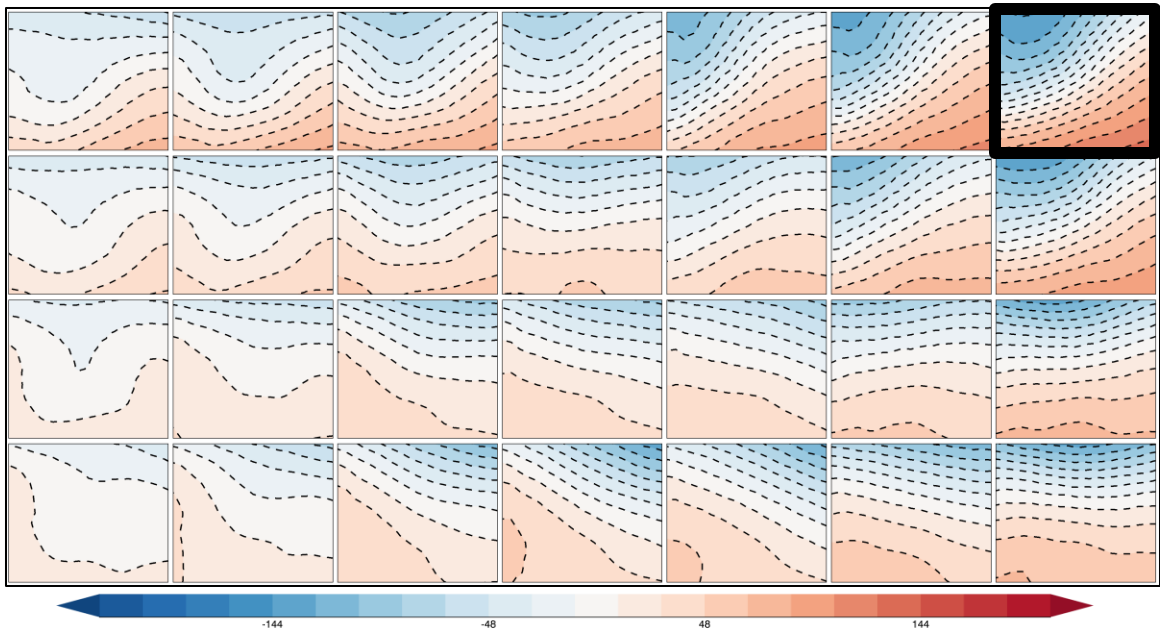


Figure 6. The 500 hPa height anomalies for a 7x4 (28-class) SOM centered on SGP. Positive (negative) height anomalies are represented by red (blue) shades. The black box identifies the class in which the 20 May 2013 case was matched to.

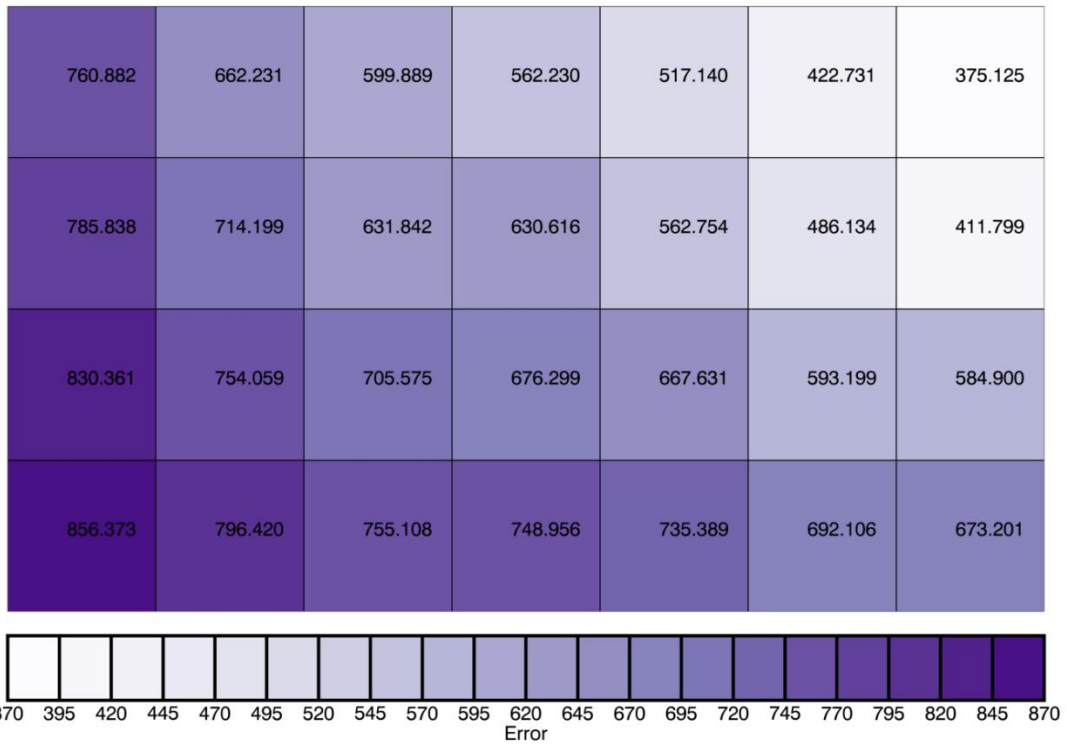


Figure 7. Euclidian distance for 20 May 2013 on the SGP CDF90 SOM. Error values are unitless.

Southern Great Plains (SGP)

The climatology SOM, which includes precipitating and non-precipitating days, for the SGP region is shown in Figure 8 (near-surface analyses), and Figure 9 (500 hPa analyses). SGP is characterized by synoptic patterns ranging from mid-latitude cyclones under southwest flow aloft (upper-left side of the SOM) to upper-level ridging and northwest flow aloft associated with surface high pressure (bottom-right side of the SOM). Nested within these climatological patterns are regimes conducive for precipitation which are of interest for this study (Figure 10). Not surprisingly, classes with higher precipitation totals are generally found on the left-hand side of the SOM. While surface patterns vary across these classes, the broad categorizing feature of these states is southwesterly flow at 500 hPa (Figure 9). A few precipitating classes are found outside of this region. Along the top-right side of the SOM, there is a region with higher precipitation totals (Figure 10). Under these states, there exists weak low pressure (Figure 8; top right) and flow out of the northwest. These patterns also tend to see higher precipitation amounts overnight (0600-1100 UTC) than during the rest of the day (Figure 11).

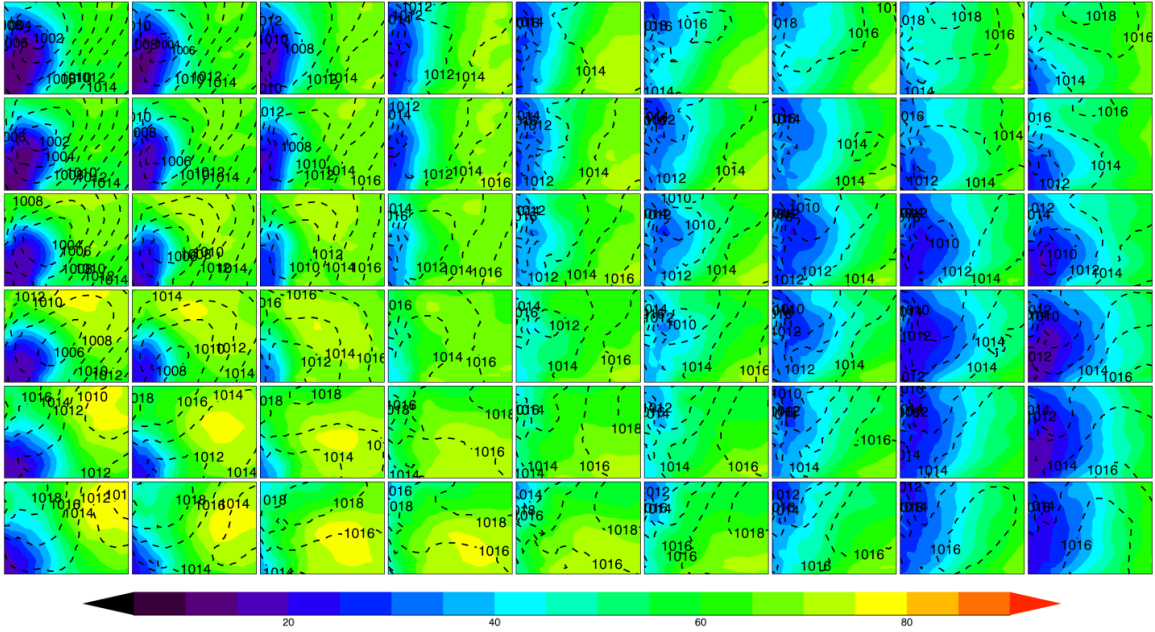


Figure 8. Near-surface analyses for the 9x6 (54-class) SGP climatology SOM. MSLP is contoured with dashed lines while filled contours represent 900 hPa RH. Cool (warm) colors represent drier (moister) air.

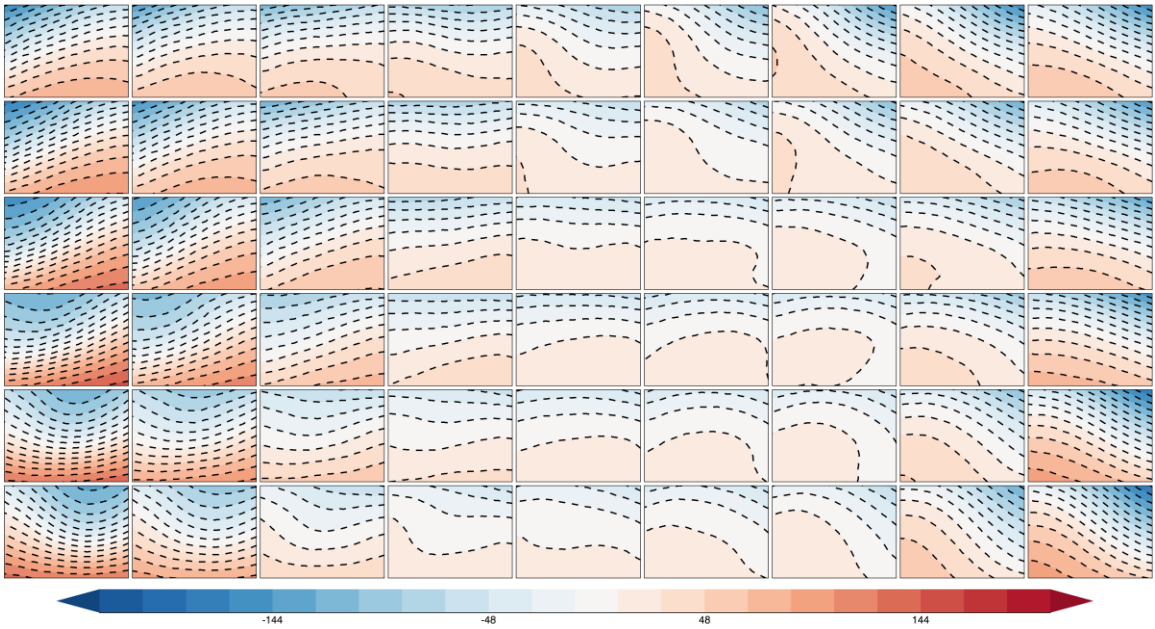


Figure 9. Climatology SOM for SGP 500 hPa geopotential height anomalies. Blues indicated negative height anomaly and red indicates positive height anomaly.

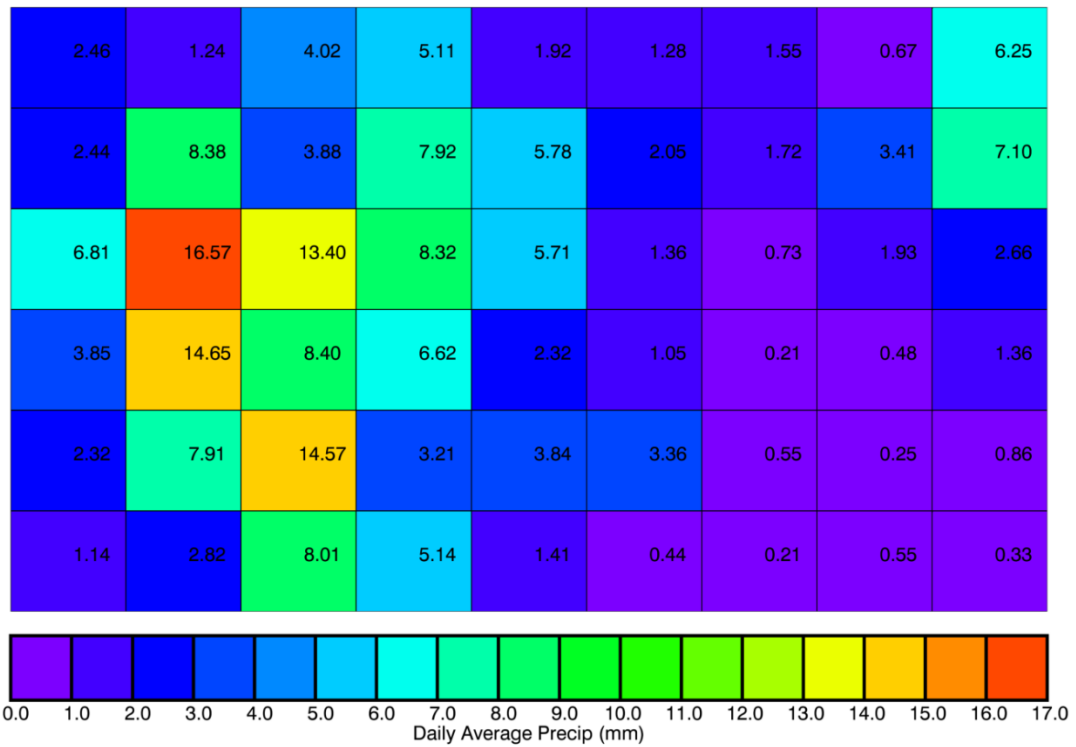


Figure 10. Daily average precipitation for the SGP climatology SOM.

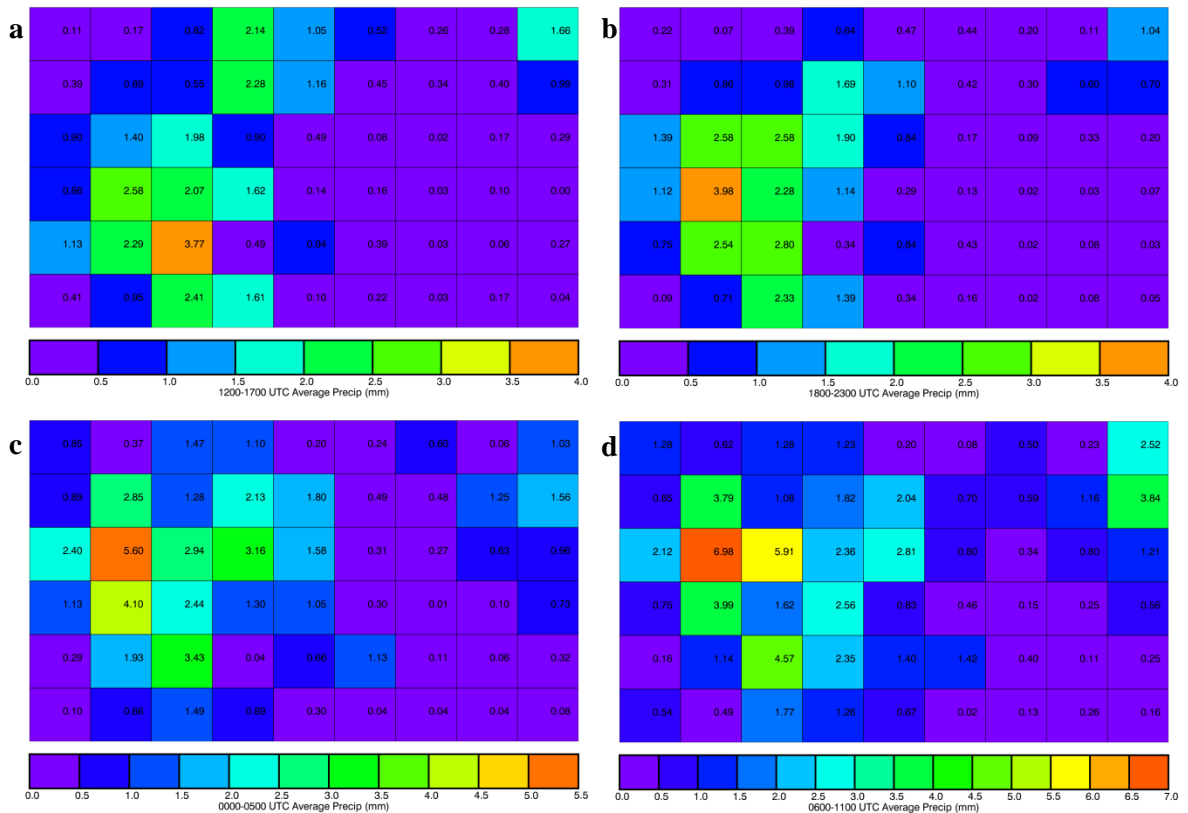


Figure 11. Average precipitation in six-hour increments for the SGP climatology SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), 0600-1100 UTC (d).

Biases for the NSSL WRF vary across the climatological SOM (Figure 12). Higher bias amounts, whether positive or negative, tend to occur for the classes on the left side that see higher precipitation amounts. However, as the day progresses a pattern begins to develop in the bias. While daytime (1200-1700 UTC and 1800-2300 UTC) precipitation bias is mostly positive, precipitation bias overnight (0000-0500 UTC and 0600-1100 UTC) becomes negative (Figure 13). This is seen across most of the SOM and also with the patterns in the upper-right where there is the smaller region of higher precipitation amounts. It is also worth noting that there is more variability in precipitation amounts overnight and the highest precipitation amounts occur during this time.

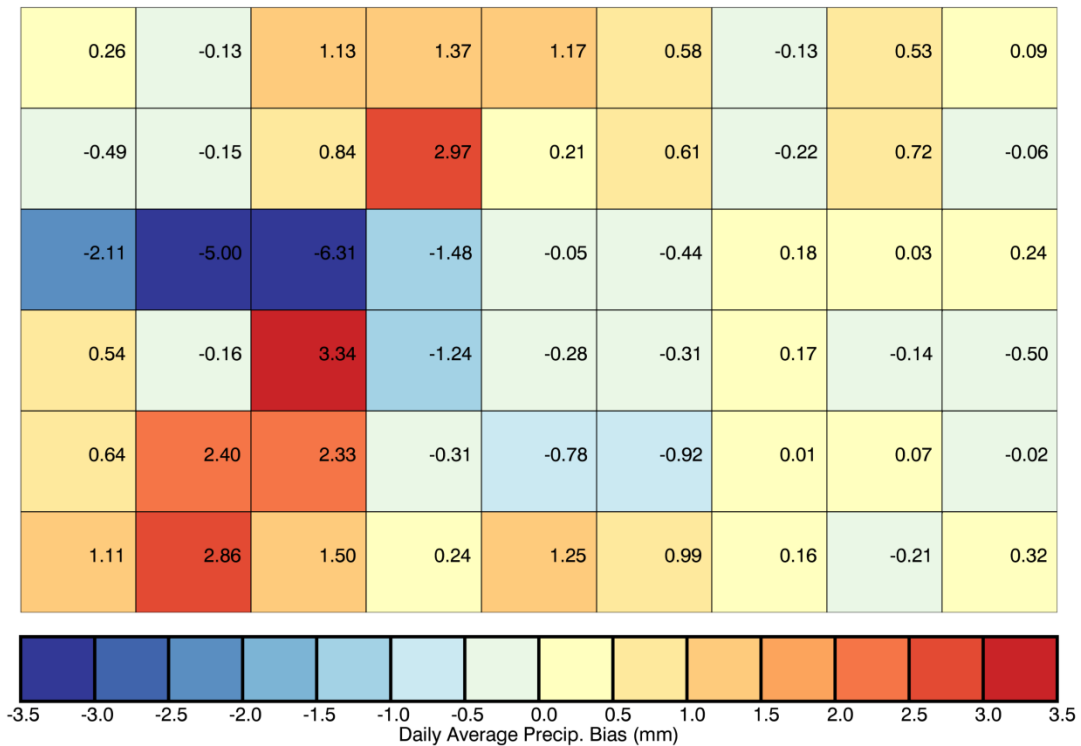


Figure 12. Daily average precipitation bias for the SGP climatology SOM.

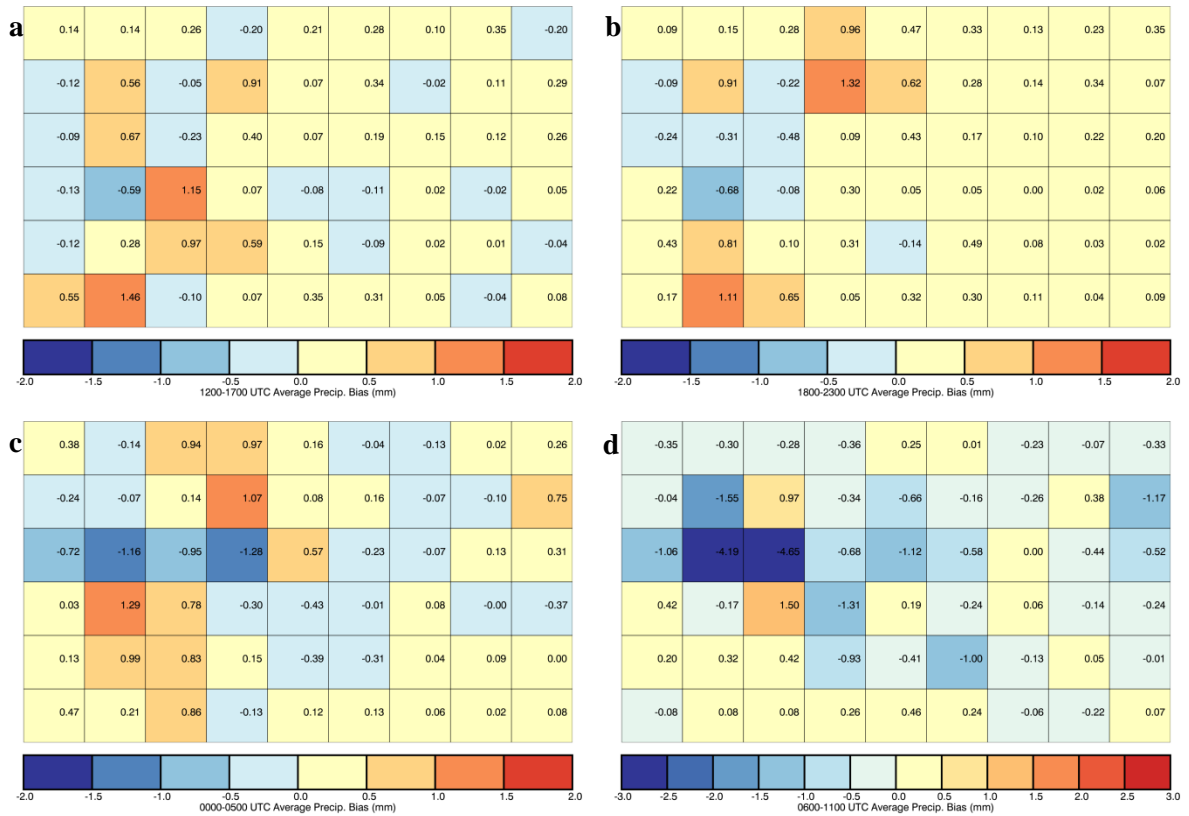


Figure 13. Average precipitation bias in six-hour increments for the SGP climatology SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), 0600-1100 UTC (d).

The selection of CDF90 and CDF50 cases largely falls within the higher precipitation classes (Figures 14 and 15), as would be expected. Figure 16 shows the percent of cases in each class that had precipitation, and many of the classes with a high percentage of precipitation cases were among those selected. However, some classes that had CDF cases selected have lower percentages of days with precipitation occurring (precipitation amounts greater than .1 mm), indicating that the class has high variability in rainfall meaning either a lot of precipitation occurs or none at all.

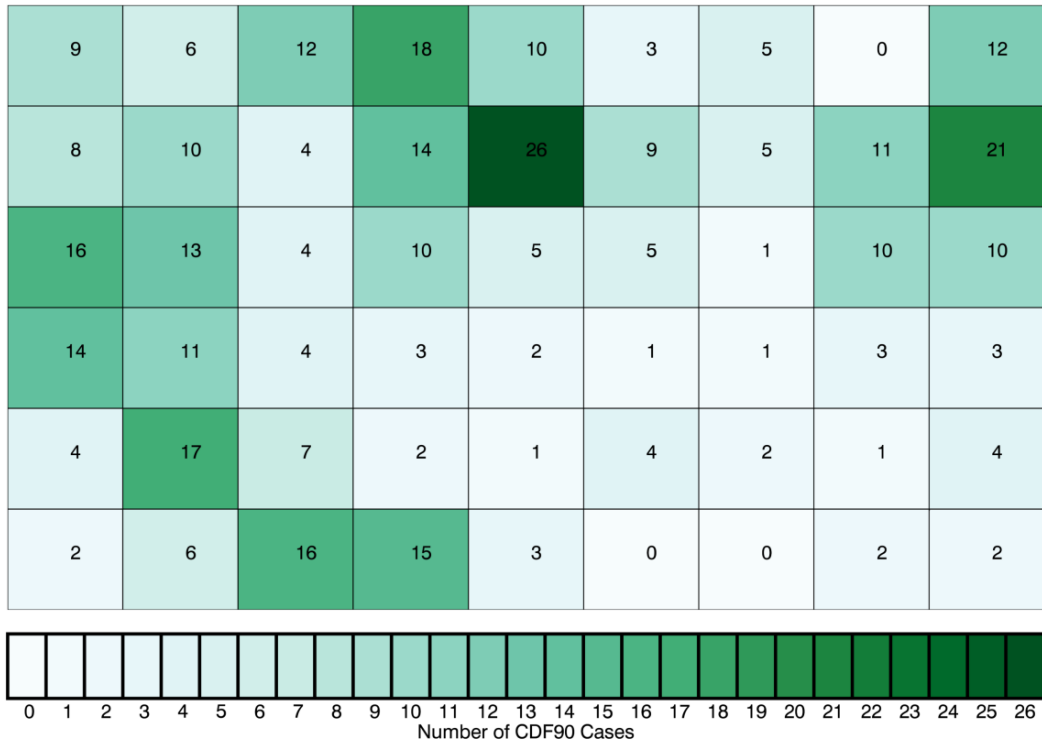


Figure 14. Distribution of CDF90 classes within the SGP climatology SOM.

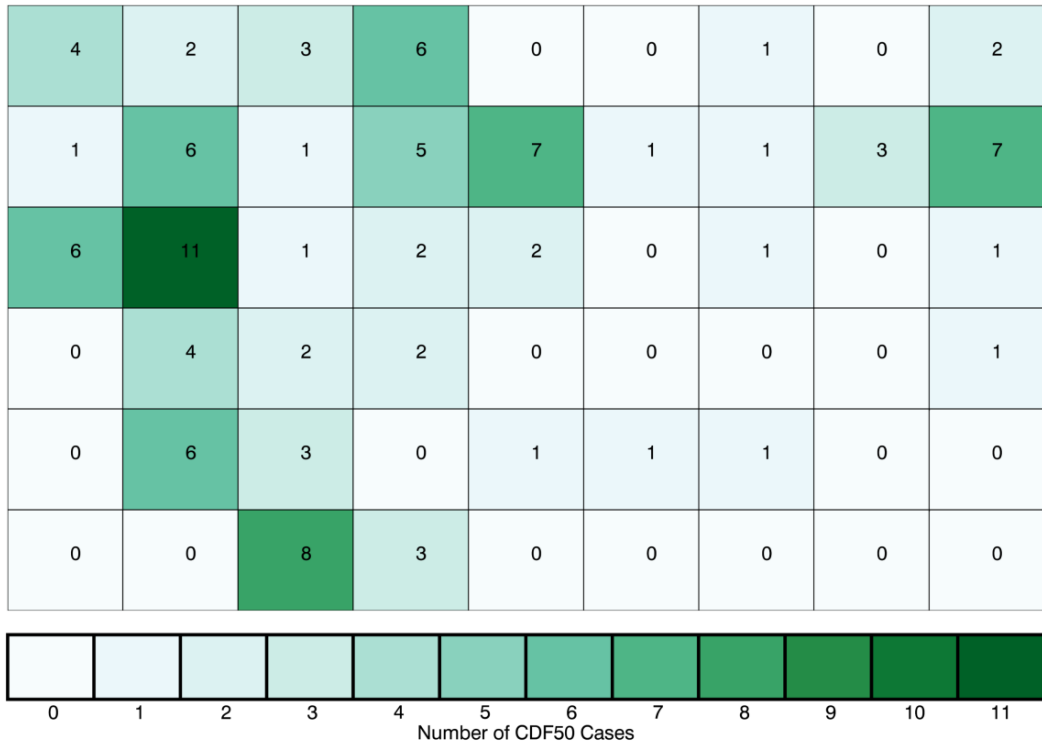


Figure 15. Distribution of CDF50 classes within the SGP climatology SOM.

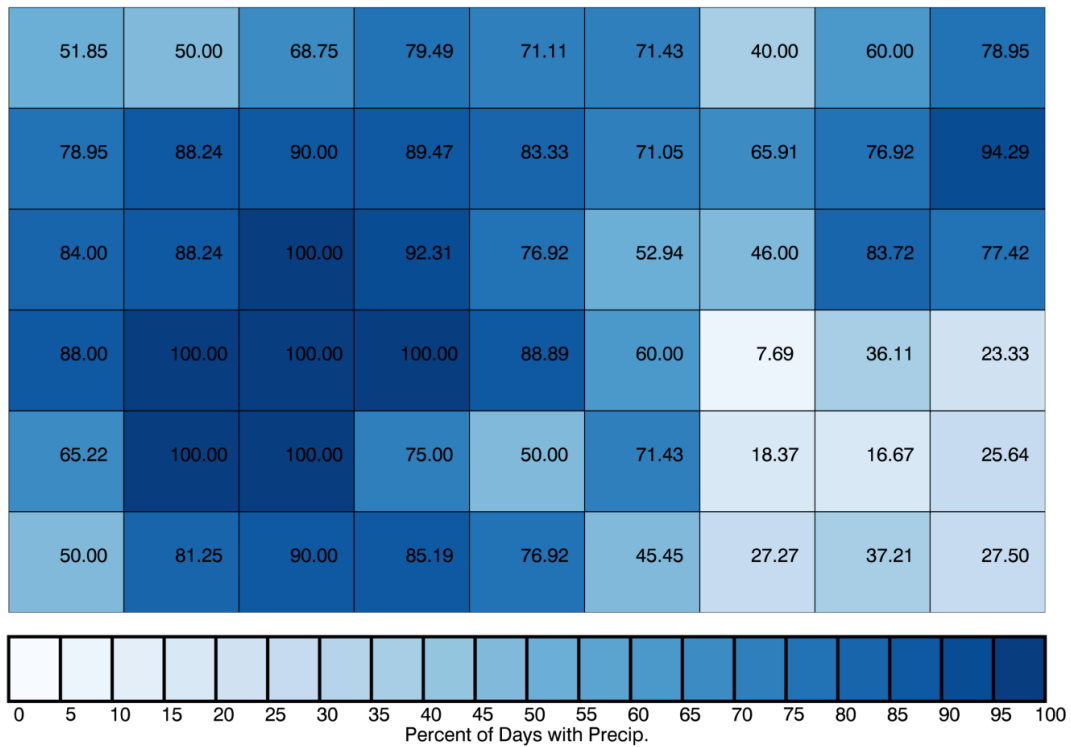


Figure 16. Percent of days in each class with precipitation.

CDF90

Generation of SOMs based on only precipitating days yields more variation in patterns responsible for these events. Figures 17 and 18 depict the near-surface and upper-level analyses, respectively, for the CDF90 SOM. SGP is characterized by a variety of patterns responsible for precipitation events. In lee of the Rocky Mountains, the majority of cases have a dryline in the west. Stronger forced events are found on the right hand side (RHS) of the SOM, as evident by the stronger surface lows (Figure 17) and coincident with the faster upper level flow implied by upper level troughs with tightly packed height anomalies (Figure 18). Additional patterns of note from Figures 17 and 18, respectively, include warm fronts under shortwave troughs (upper left of SOM) and relatively weak surface forcing underneath northwesterly flow (bottom center of

SOM). While cases are relatively well distributed across the classes of the SOM, some variability does exist in the number of cases classified to each class, particularly in the upper-center of the SOM (Figure 19).

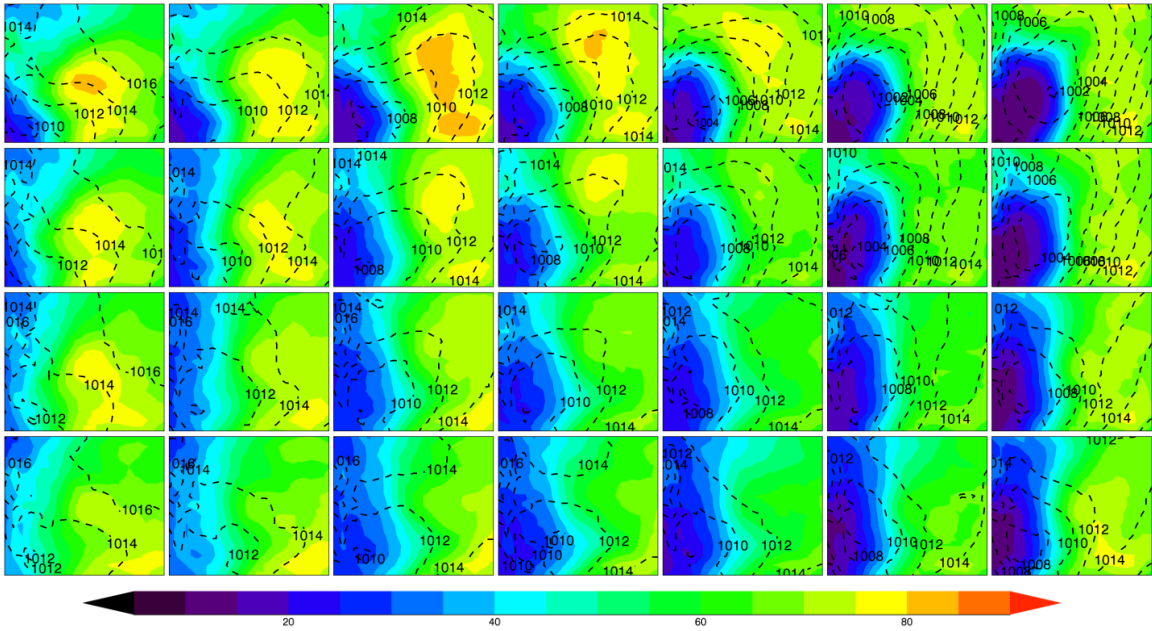


Figure 17. Near-surface analyses for the SGP CDF90 SOM.

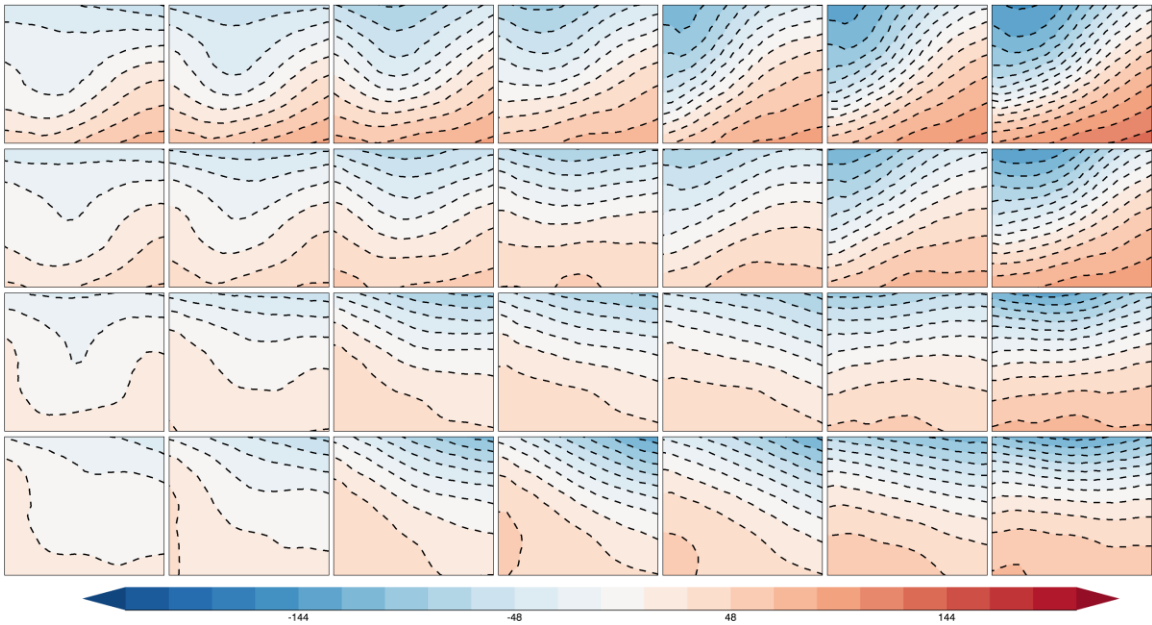


Figure 18. The 500 hPa geopotential height anomalies for the SGP CDF90 SOM.

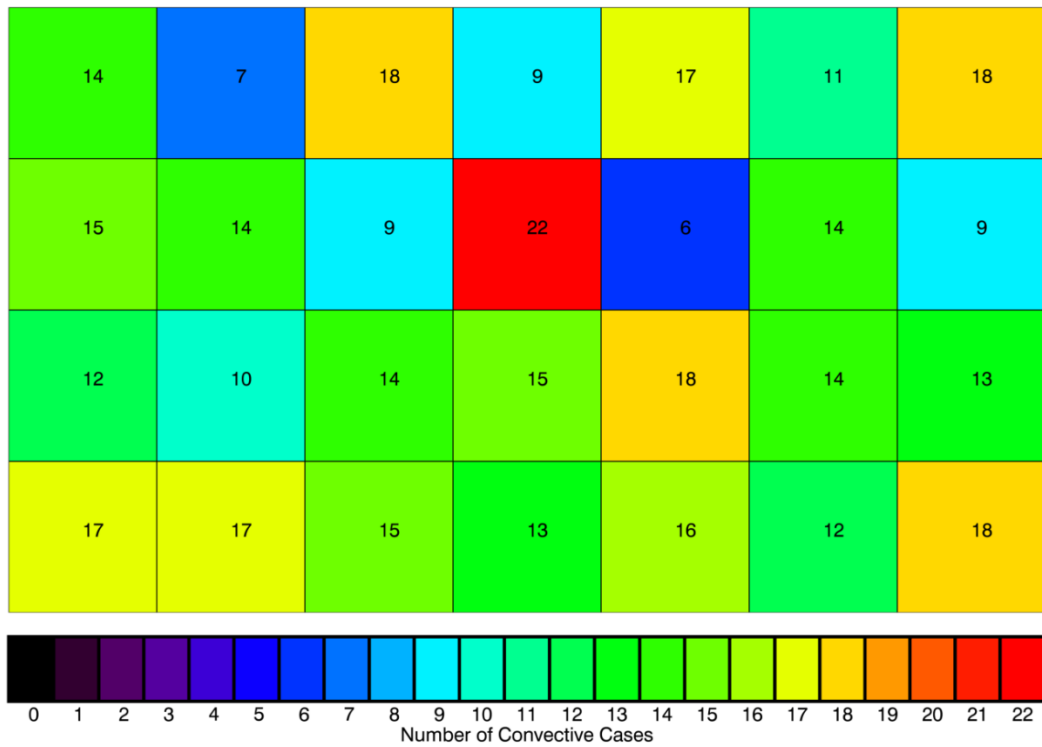


Figure 19. Number of cases in each class of the SGP CDF90 SOM.

Average daily precipitation amounts for the SOM range from less than 7 mm to nearly 19 mm, with higher amounts (>13 mm) occurring along the edges (Figure 20). Because the patterns on the edges of the SOM tend to represent those that are less similar to other patterns, this suggests that these higher precipitation patterns are not as well represented. As precipitation evolves throughout the day, Figure 21 shows that the patterns in the upper-left associated with surface warm fronts and upper-level troughs are more likely to have precipitation in the morning (1200-1700 UTC) and daytime (1800-2300 UTC). Meanwhile, patterns on the right with stronger cyclones and upper-level forcing or patterns with weaker surface patterns and northwest flow aloft have more precipitation during the late afternoon (0000-0500 UTC) and overnight (0600-1100 UTC).

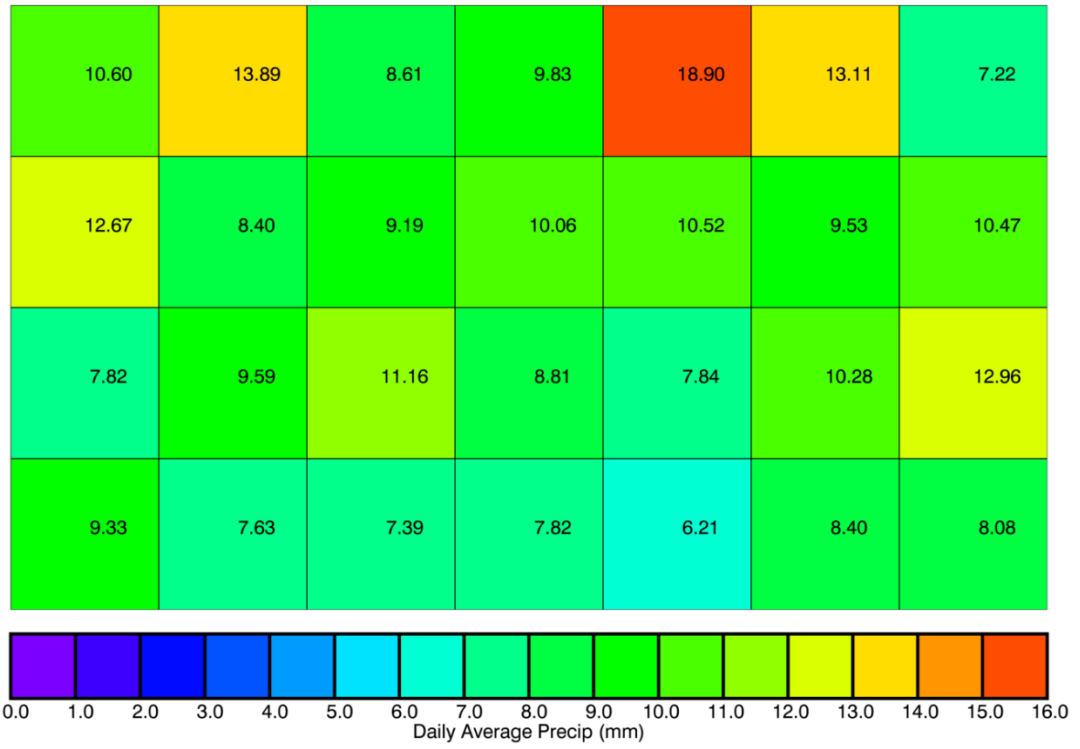


Figure 20. Daily average precipitation for the SGP CDF90 SOM.

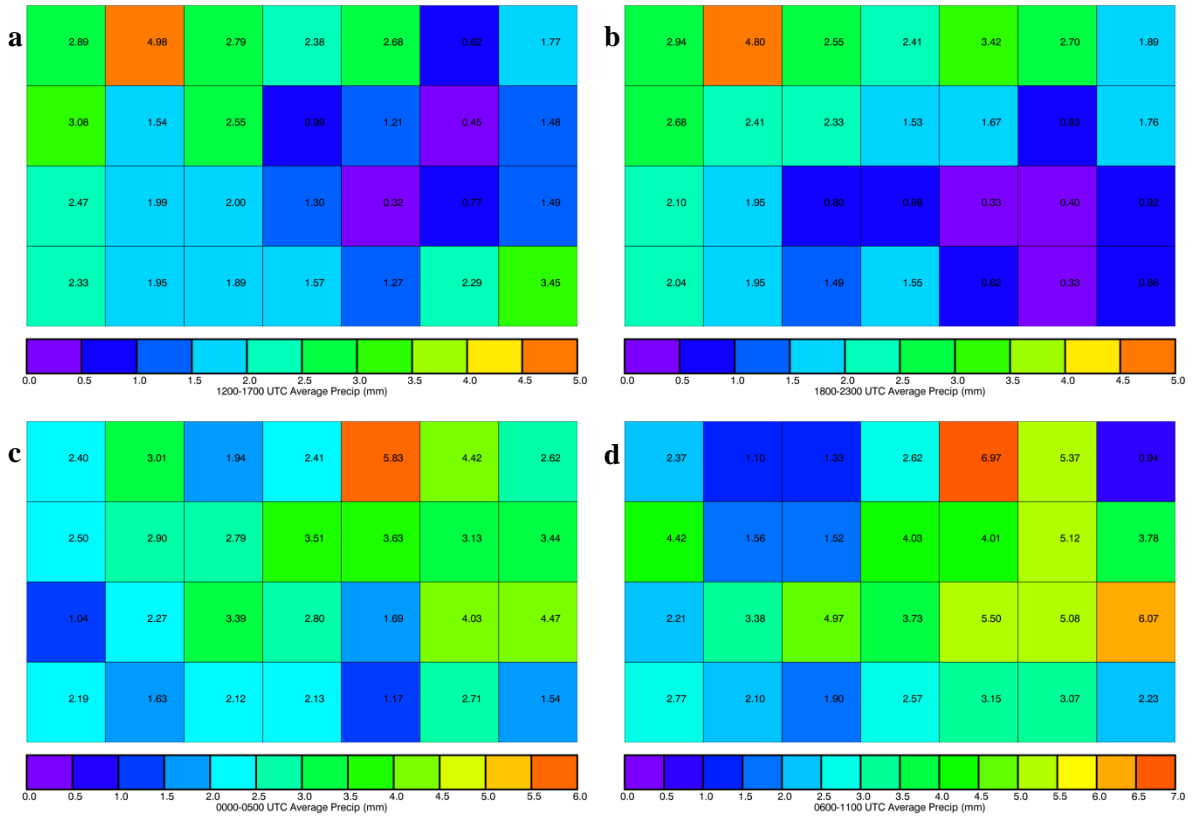


Figure 21. Average precipitation in six-hour increments for the SGP CDF90 SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), 0600-1100 UTC (d).

WRF precipitation biases for the SGP CDF90 SOM are provided in Figure 22. Despite class-to-class variability, patterns of bias are present across the SOM. In particular, the highest positive biases are found in the upper-left corner associated with high RH and surface warm fronts beneath an upper-level trough. The strongest negative biases are found in the upper-right with the stronger surface and upper-level forcing. Finally, an area of weaker negative biases (Figure 22) is found for the weakly forced events (Figure 17) under northwest upper level flow (Figure 18) in the bottom-center of the SOM. The latter cases and the stronger cyclone cases are responsible for the significant negative bias found for SGP during the overnight hours (0600-1100 UTC, Table 3), and these results are supported by 6-hr precipitation biases (Figure 23). While morning and early afternoon biases are predominately positive, biases during the overnight period are largely negative. The few cases that have positive biases during this latter period are confined to classes that have warm fronts. Overall, patterns responsible for significant nocturnal precipitation are confined to the right side of the SOM. These patterns have a high percentage of cases with maximum precipitation occurring overnight for each class, indicating that these patterns are responsible for significant, nocturnal precipitation events (Figure 24).

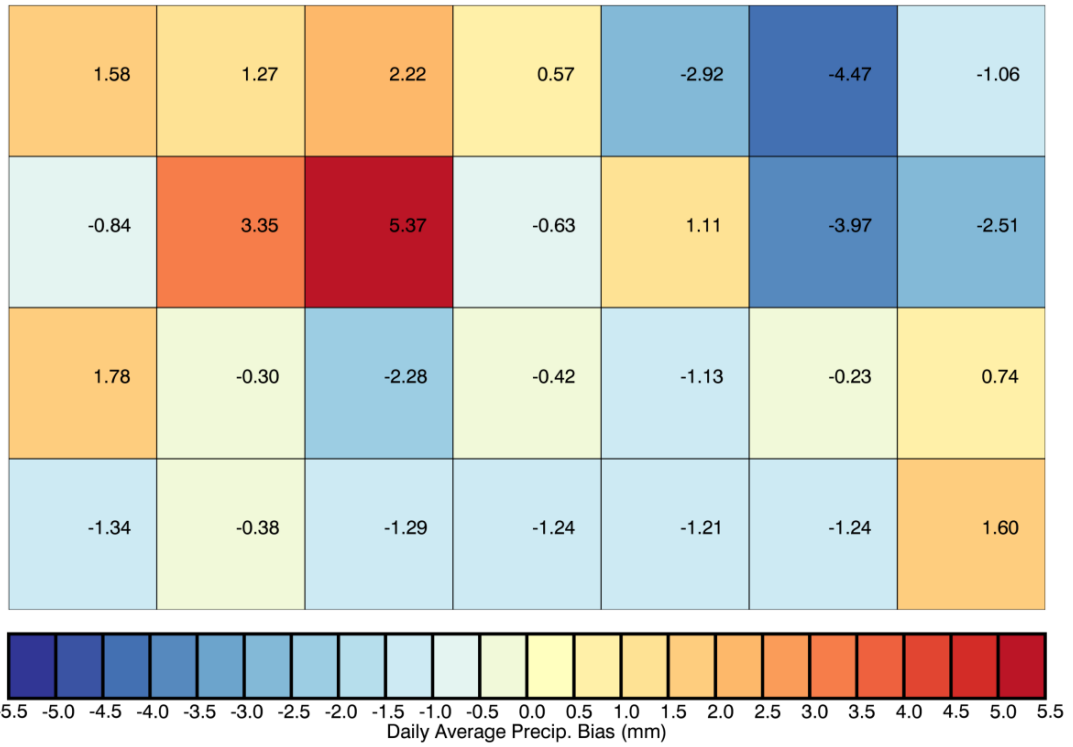


Figure 22. Daily average precipitation bias for the SGP CDF90 SOM.

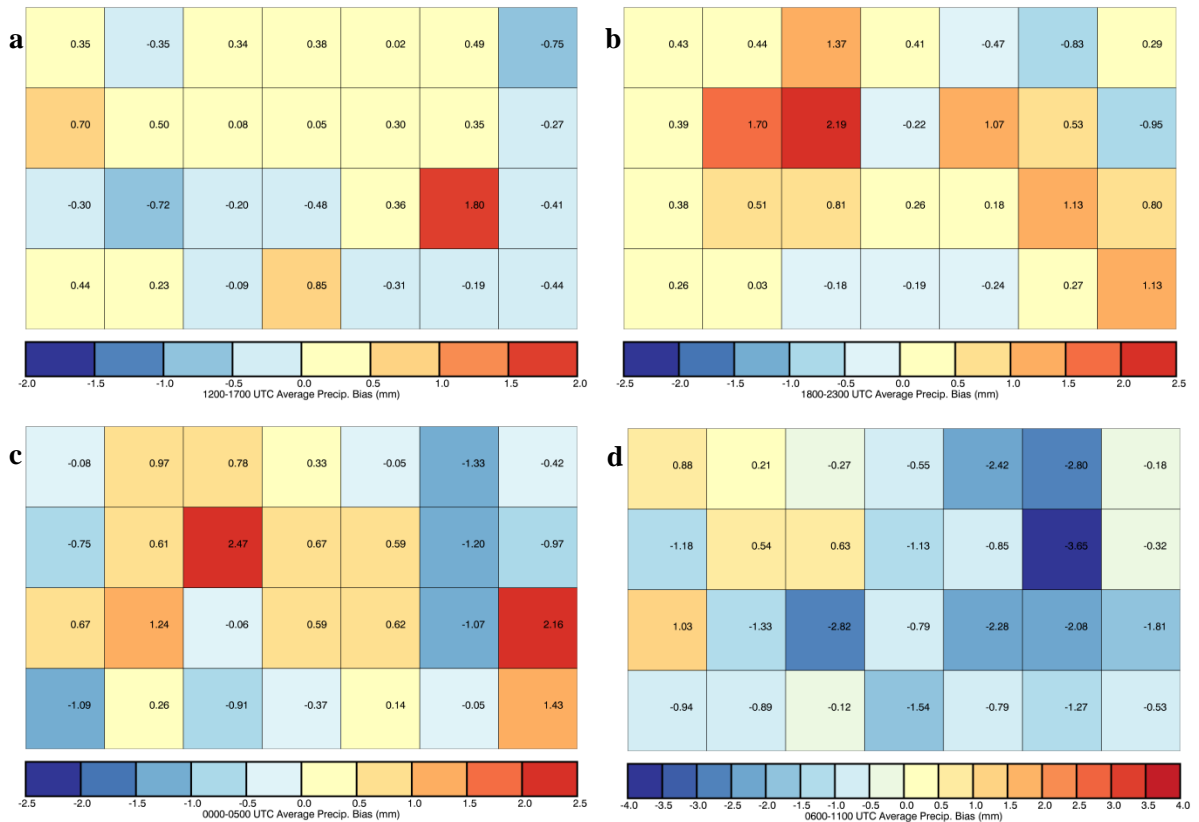


Figure 23. Average precipitation bias in six-hour increments for the SGP CDF90 SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), 0600-1100 UTC (d).

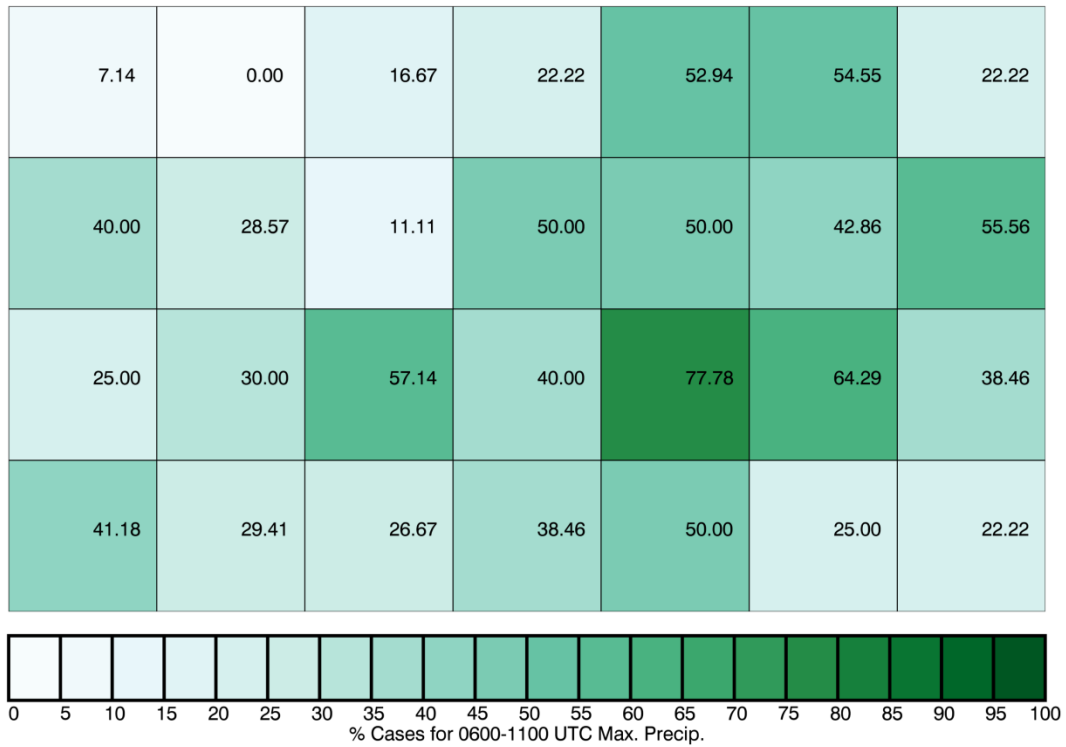


Figure 24. SGP CDF90 percent of cases in each class with maximum six-hour accumulated precipitation occurring for 0600-1100 UTC.

Statistical Analysis

After grouping the SGP CDF90 patterns as described previously (Chapter II), the results are Figures 25, 26, and 27. The key synoptic pattern features of the original SOM are retained even with the additional smoothing due to averaging, with the stronger drylines on the right of the SOM and weaker surface patterns with upper-level flow from the northwest on the left. Statistical significance for the averaged SOM is shown in Table 4. Recall that the z-test is being used to show how different the resulting averaged groups are from each other. Statistical significance varies across the averaged SOM. For example, classes 1 and 2 are not statistically different from each other, while classes 1 and 3 are. This makes sense because the first two classes have similar positive biases whereas class 3 is the strongest negative class in the averaged SOM. Along with class 5

which also has a negative bias, these classes have the strongest statistical separation from the other patterns. The number of cases contributing to each class is shown in Figure 28, and center classes have more cases because there are more original CDF90 classes contributing based on the previously determined grouping.

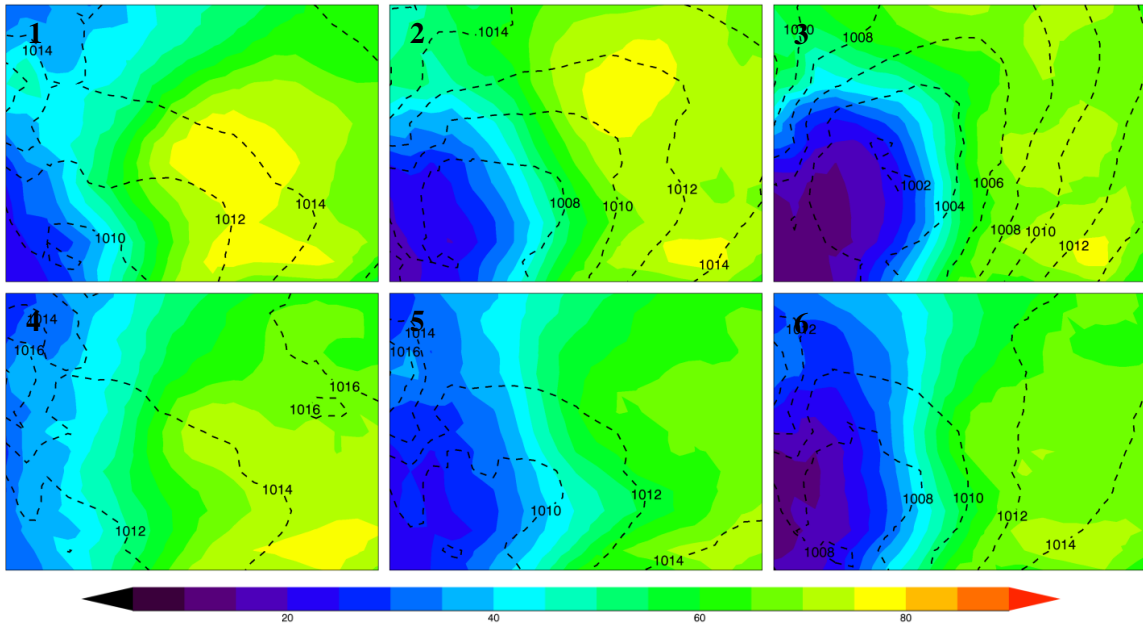


Figure 25. Near-surface analyses for the SGP CDF90 averaged SOM.

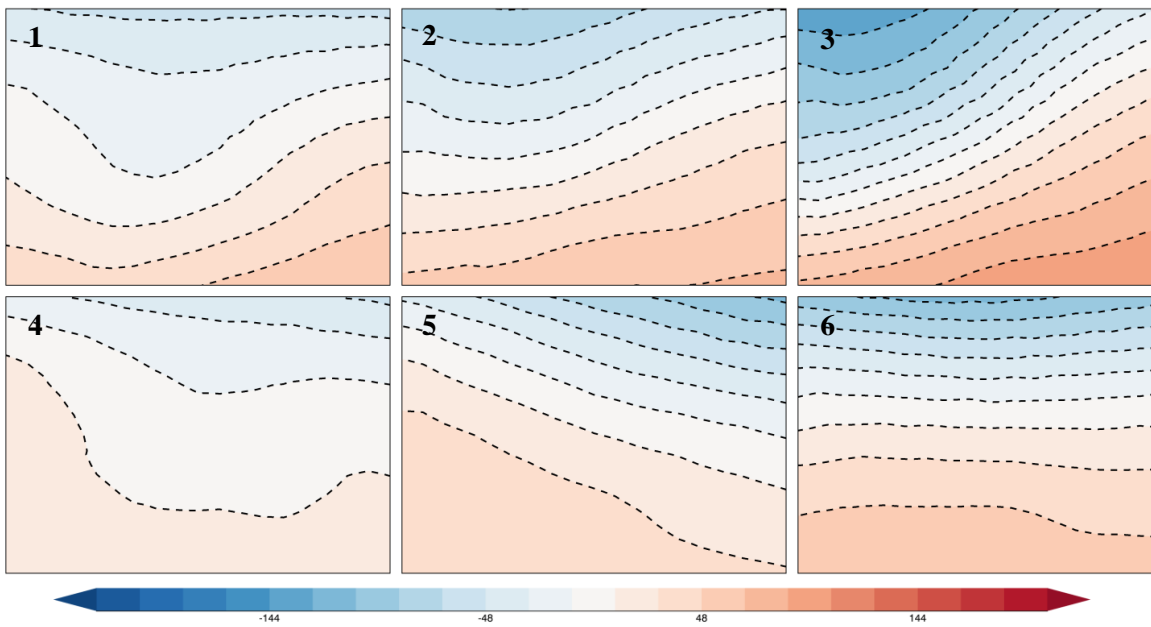


Figure 26. The 500 hPa geopotential height anomalies for the SGP CDF90 averaged SOM.

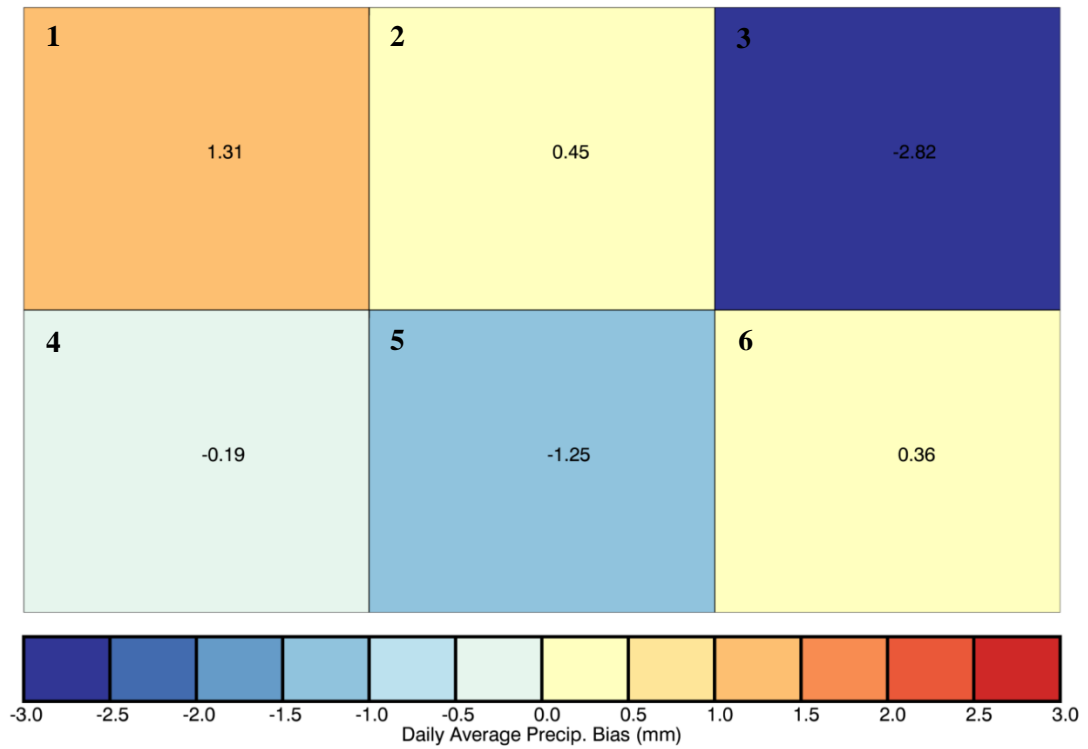


Figure 27. Daily average precipitation bias for the SGP CDF90 averaged SOM.

Table 4. Z-Test statistic results for the SGP CDF90 SOM. Statistically significant results in **bold**.

Comparison Class	1	2	3	4	5	6
1	0	-.797	-4.228	-1.571	-3.041	-.878
2	.797	0	-3.311	-.668	-1.991	-.089
3	4.228	3.311	0	3.055	2.149	3.169
4	1.571	.668	-3.055	0	-1.504	.559
5	3.041	1.991	-2.149	1.504	0	1.843
6	.878	.089	-3.169	-.559	-1.843	0

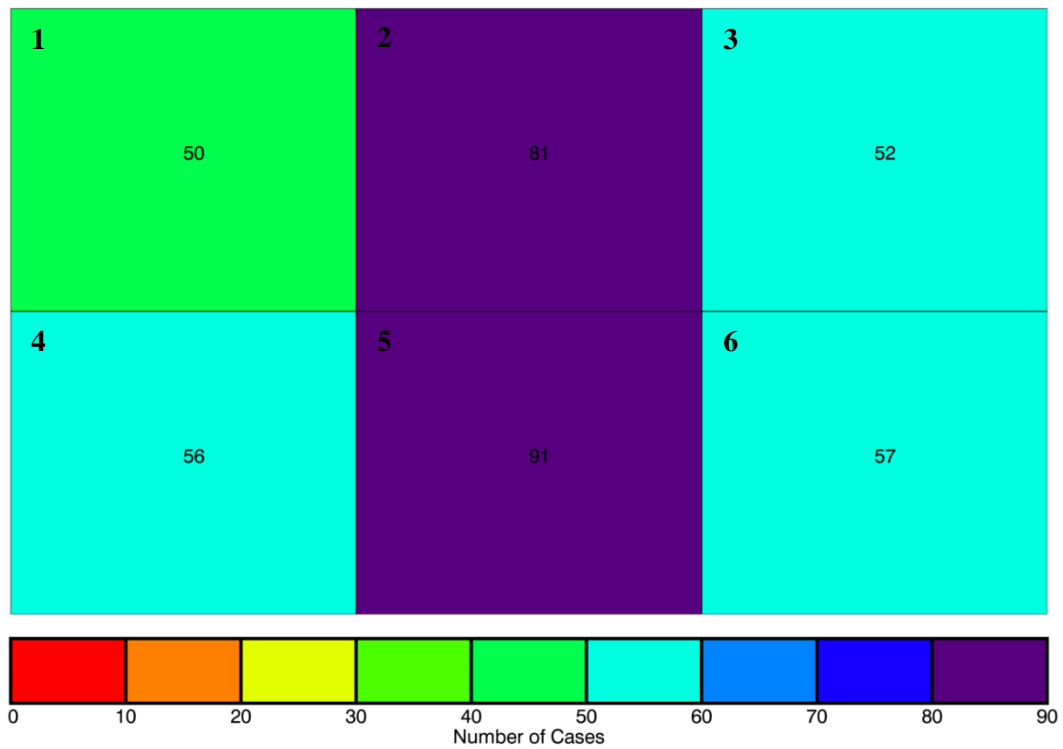


Figure 28. Number of cases in each class of the SGP CDF90 averaged SOM.

CDF50

Many of the CDF50 cases came from the strongest surface cyclone classes which were located in the upper-right of the CDF90 SOM (Figure 29). Others come from CDF90 classes with upper-level troughing or northwest flow. While similar to CDF90, the negative overnight bias signal for the CDF50 cases is much stronger, likely because they are higher precipitation events which were previously shown to have a negative nocturnal bias in the CDF90 SOM (Table 3).

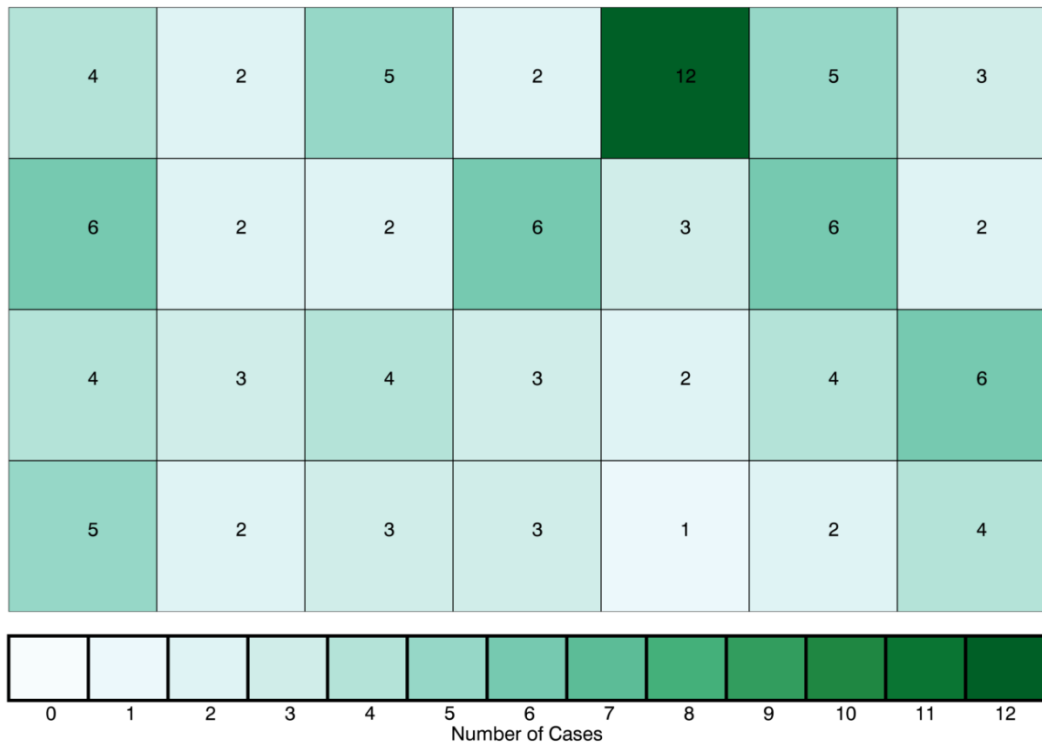


Figure 29. Plot depicting where the SGP CDF50 cases are located on the SGP CDF90 SOM.

Overall, the atmospheric patterns have higher RH values for many of the classes (Figure 30) and upper-level forcing is stronger based on the gradient in geopotential height anomaly (Figure 31). As would be expected based on the selection of the highest 50% of precipitation days, precipitation amounts are higher than for the CDF90 cases, particularly for classes with higher humidity (Figure 32). The bias is negative for nearly every pattern (Figure 33), with the exception being the far bottom right class in the SOM, which is positive. This class resembles the warm front classes from the CDF90 SOM with strong forcing aloft, which also had positive precipitation biases. While most of the precipitation for the SOM occurs from 0000-0500 UTC and 0600-1100 UTC, the class with the highest positive bias has the most precipitation in the morning and daytime hours (Figure 34). This class also has a positive bias in the model as the day progresses, as

opposed to the other classes which have an overwhelmingly negative bias, especially overnight (Figure 35).

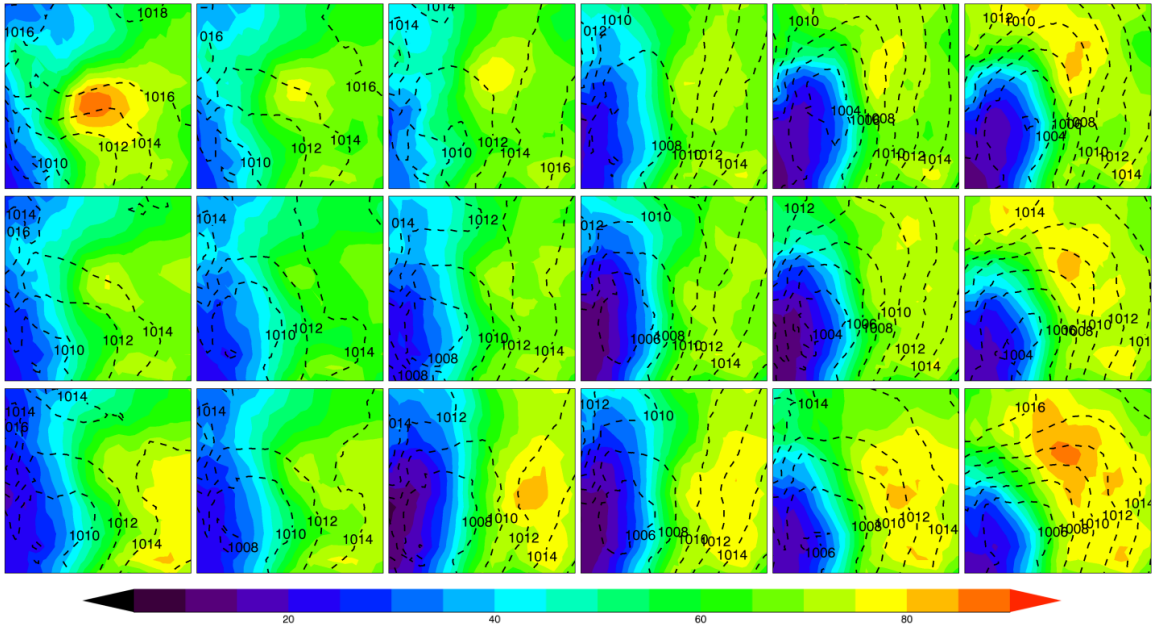


Figure 30. Near-surface analyses for the SGP CDF50 SOM.

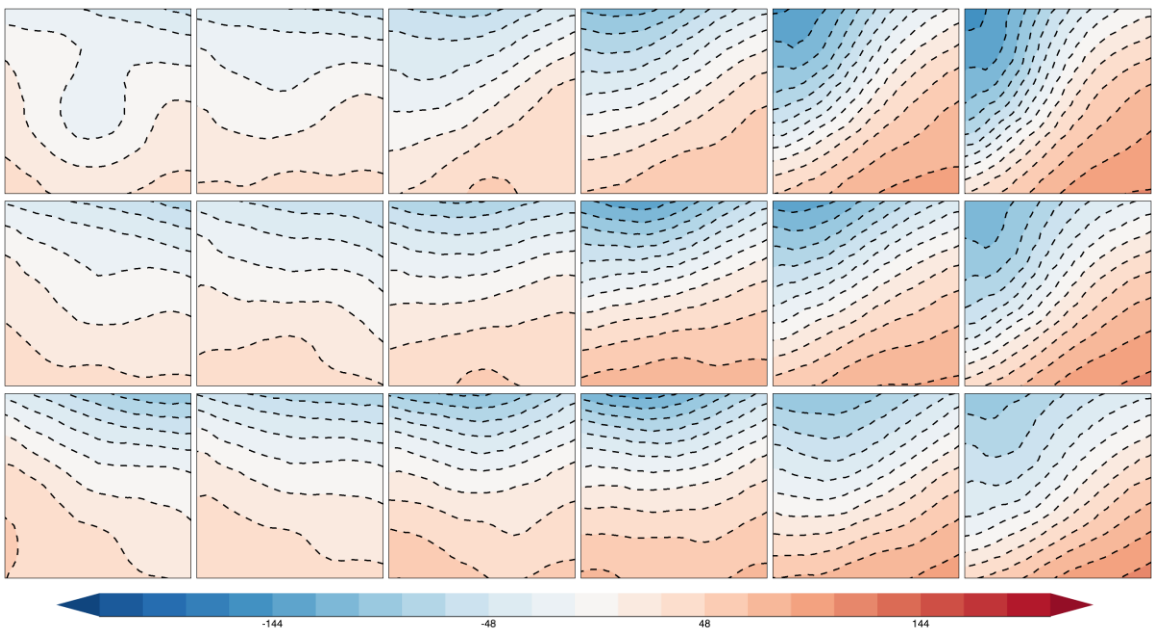


Figure 31. The 500 hPa geopotential height anomalies for the SGP CDF50 SOM.

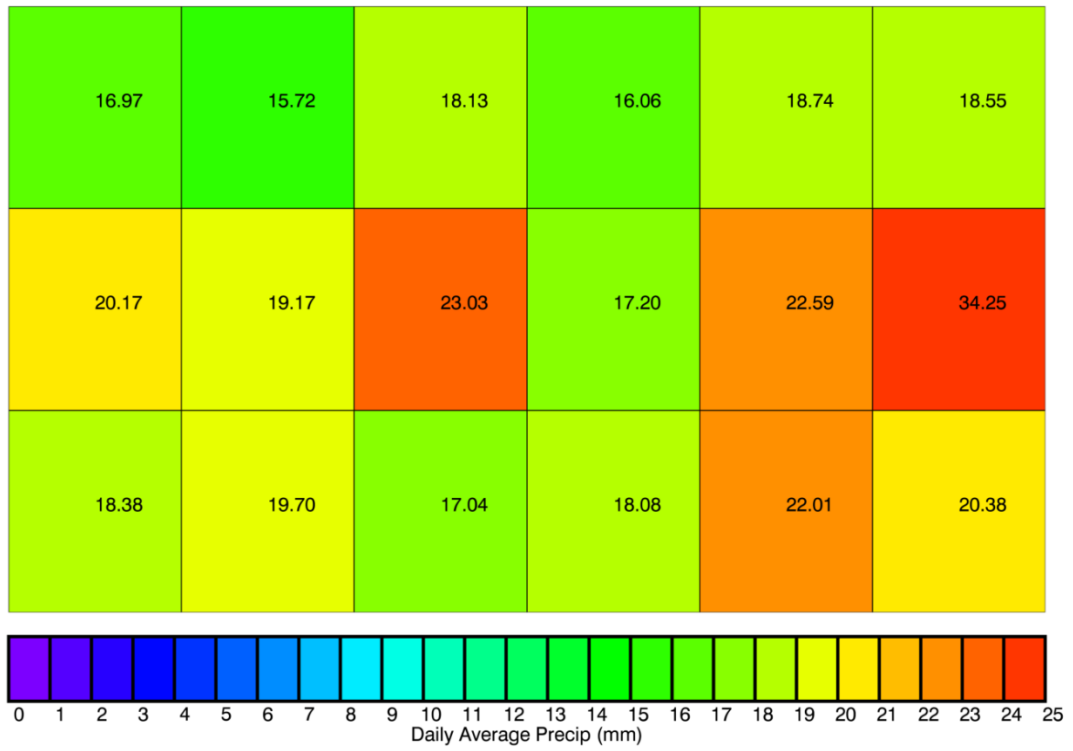


Figure 32. Daily average precipitation for the SGP CDF50 SOM.

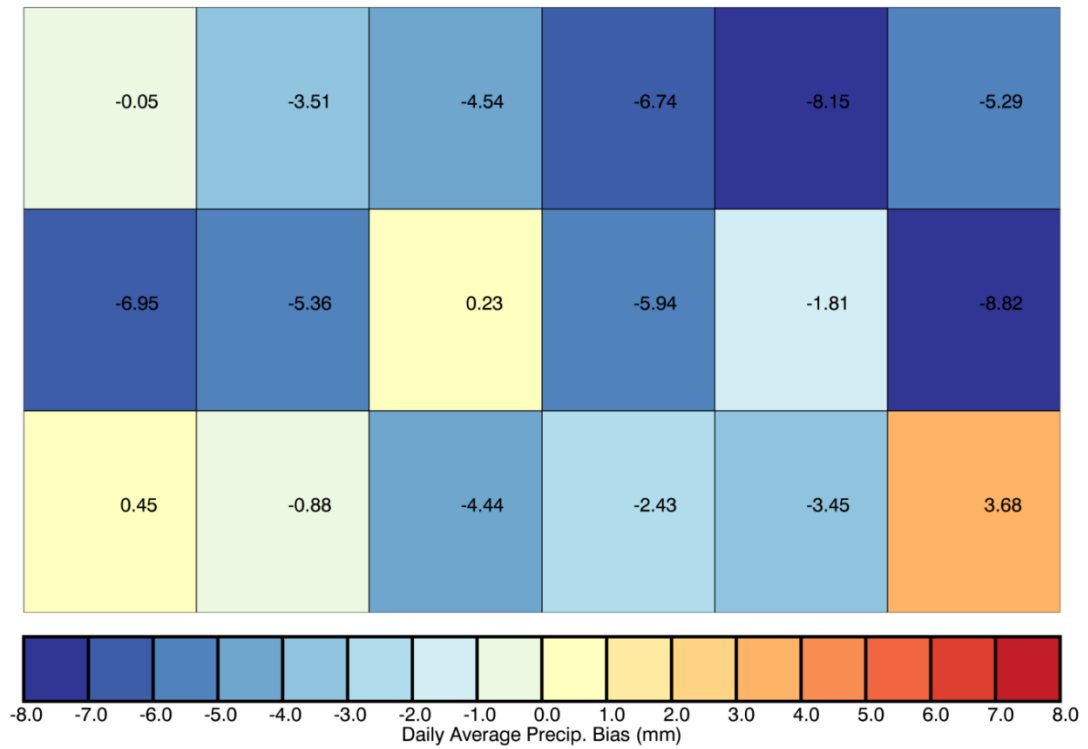


Figure 33. Daily average precipitation bias for the SGP CDF50 SOM.

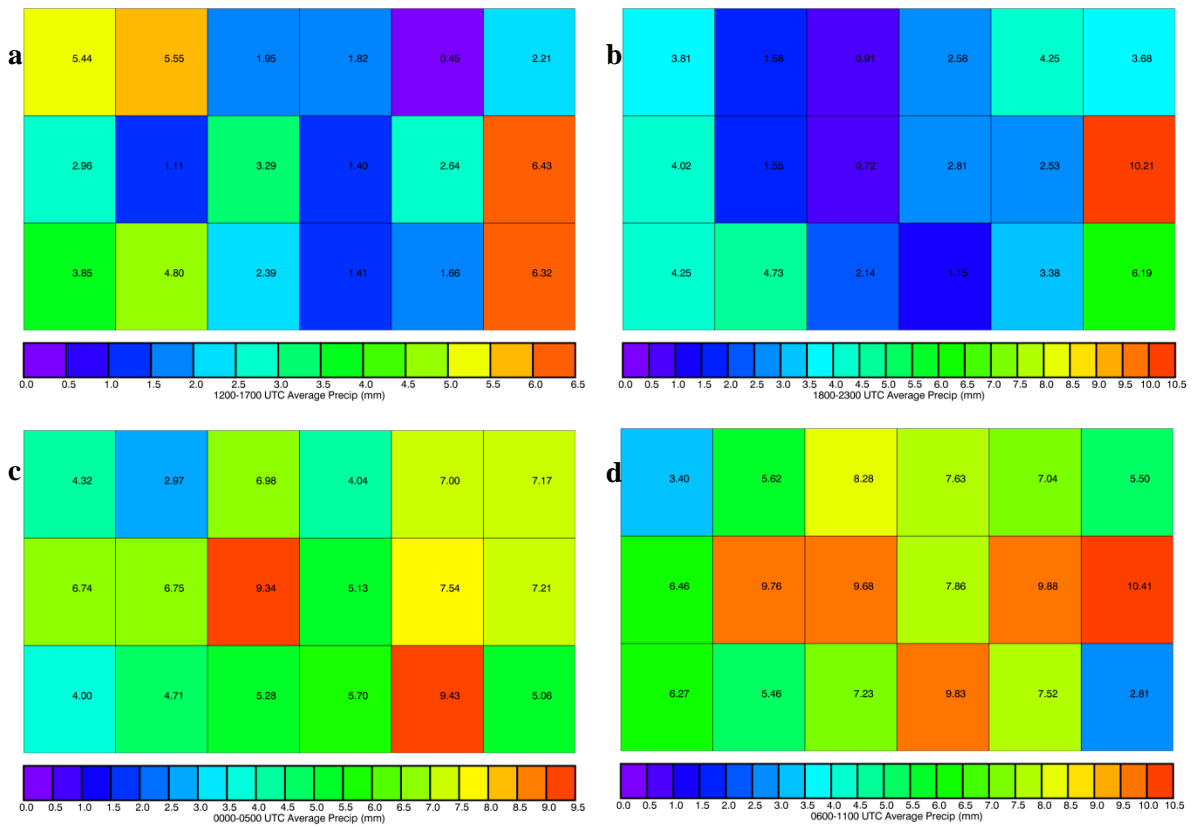


Figure 34. Average precipitation in six-hour increments for the SGP CDF50 SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), 0600-1100 UTC (d).

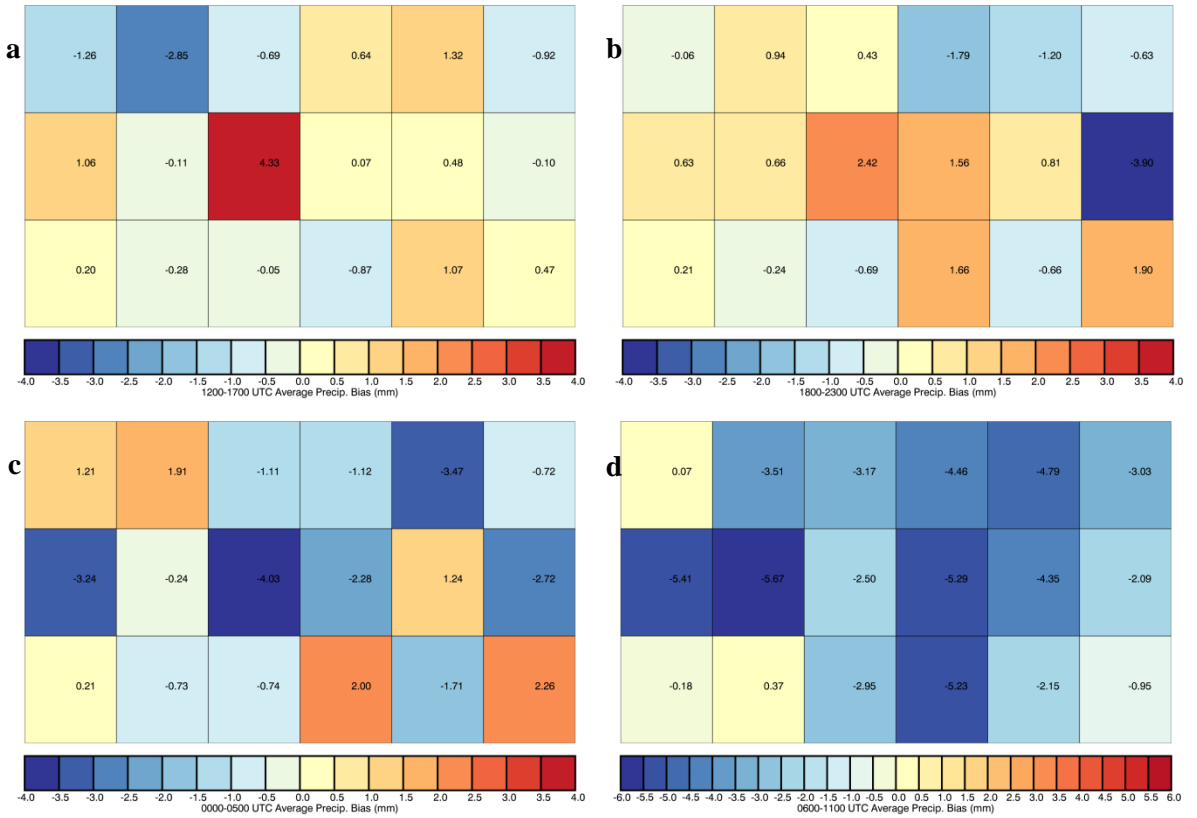


Figure 35. Average precipitation bias in six-hour increments for the SGP CDF50 SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), and 0600-1100 UTC (d).

Cases are not as well distributed for this SOM, with higher numbers of cases occurring in patterns along the edge of the SOM (Figure 36). This suggests that there is variability in the synoptic patterns here and the more extreme, or outlier, patterns are not as well represented by the SOM.

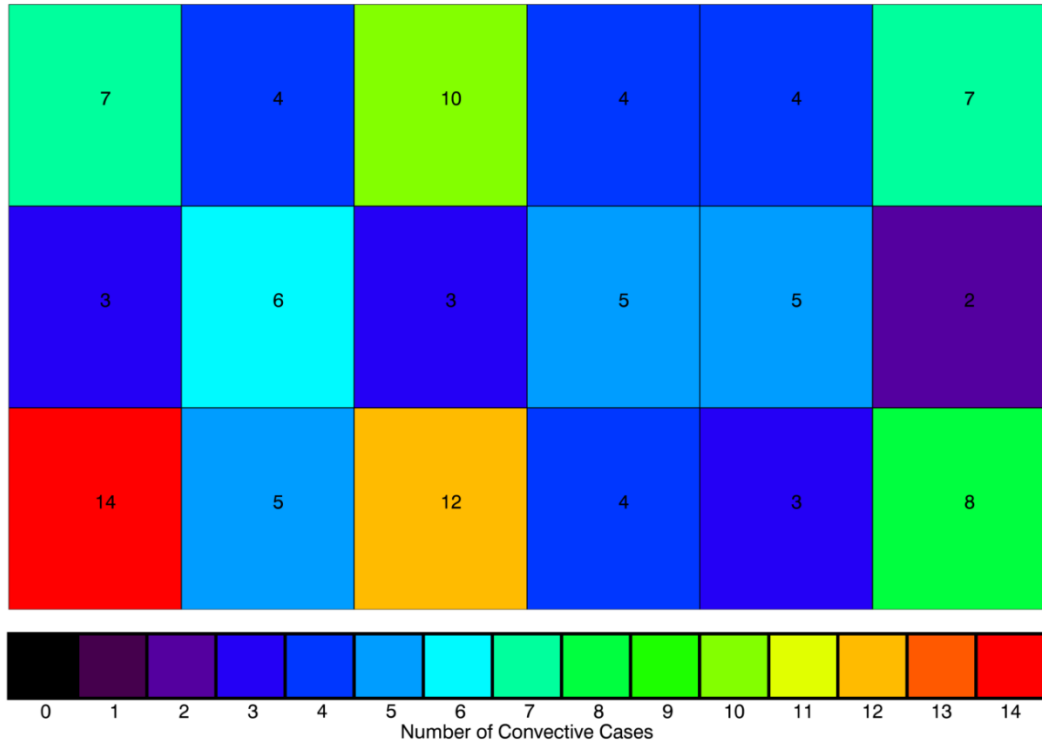


Figure 36. Number of cases in each class of the SGP CDF50 SOM.

Statistical Analysis

Grouping the SGP CDF50 classes together as described previously results in Figures 37, 38, and 39. As before, the key pattern features are retained in this averaged SOM. The only positive class is in the lower right of the averaged SOM, where the high positive bias associated with a strong warm front in the full CDF50 SOM was located. Statistical significance for the averaged SOM are shown in Table 5. Here the only statistical significance comes from class 4 being different from classes 3 and 6. Class 4 is

a weaker surface pattern under northwest flow aloft, while classes 3 and 6 are stronger dryline classes with either zonal or southwest flow aloft. The results of the z-test show that there is less statistical difference between patterns for the grouped CDF50 SOM in comparison to the grouped CDF90 SOM. This is likely caused by less cases contributing to the SOM and weakening the statistics (Figure 39). Another contributing factor to less significance is that the bias values are more similar to each other in comparison to that seen in the SGP CDF90 SOM.

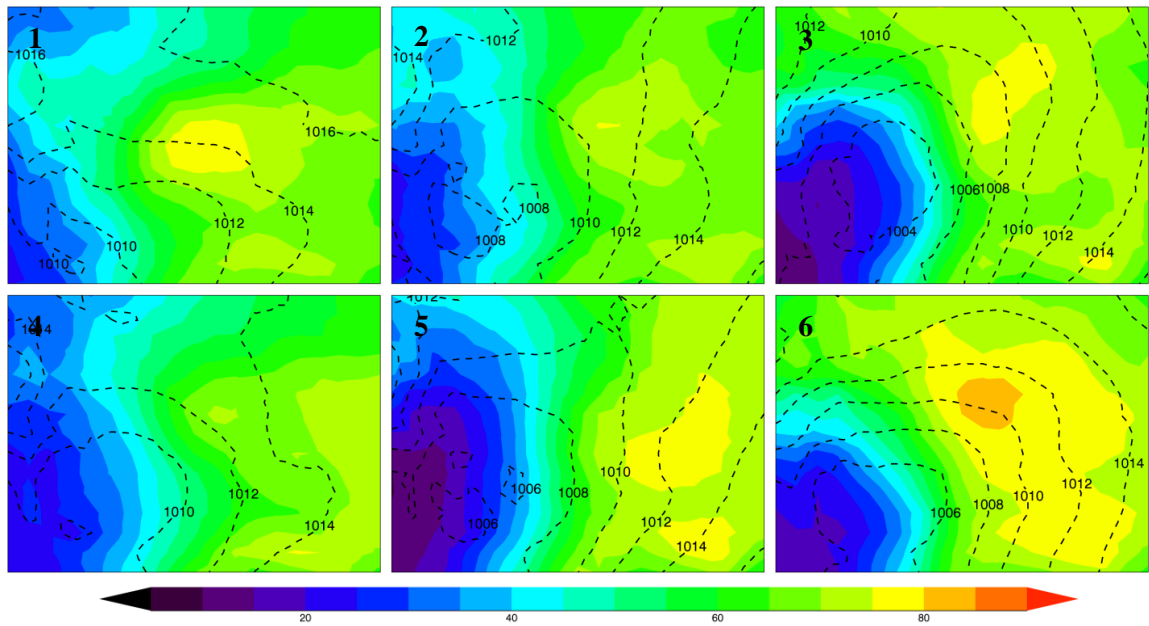


Figure 37. Near-surface analyses for the SGP CDF50 averaged SOM.

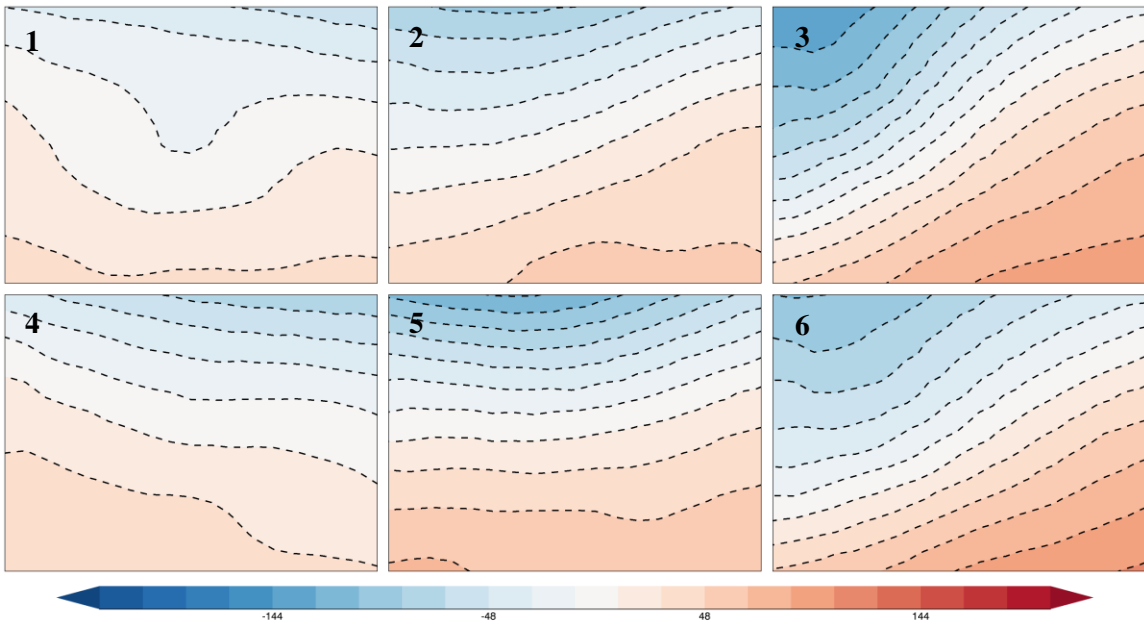


Figure 38. The 500 hPa geopotential height anomalies for the SGP CDF50 averaged SOM.

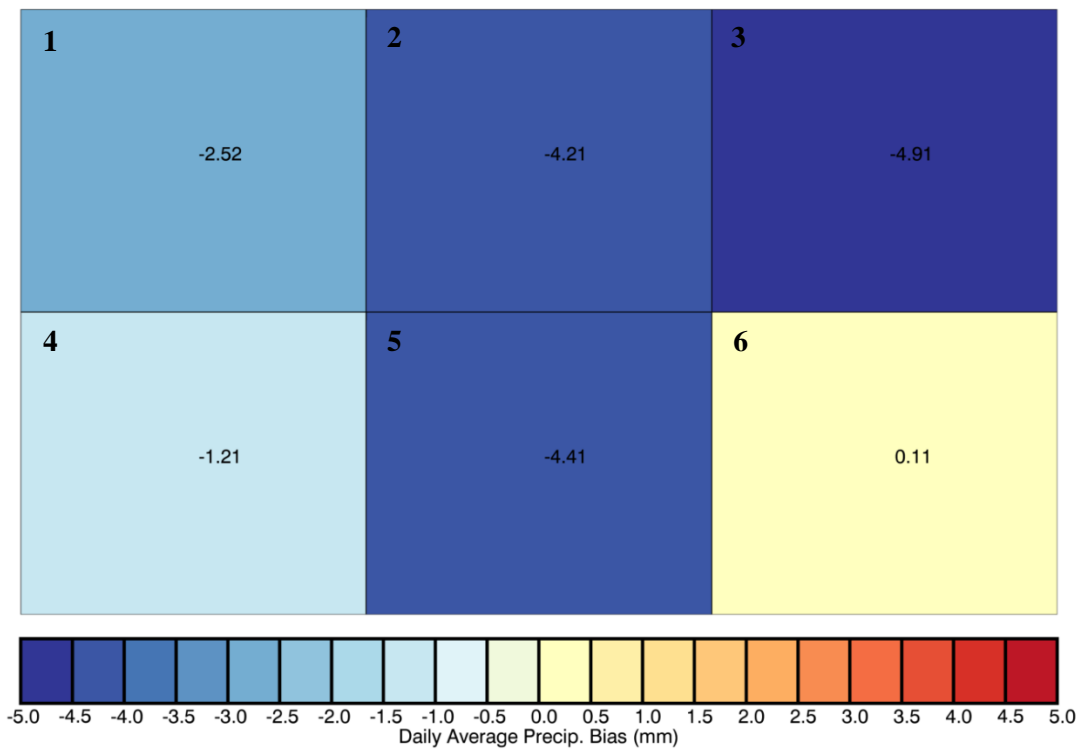


Figure 39. Daily average precipitation bias for the SGP CDF50 averaged SOM.

Table 5. Z-Test statistic results for the SGP CDF50 SOM. Statistically significant results in **bold**.

Comparison Class	1	2	3	4	5	6
1	0	-.598	-.820	.513	-.706	.789
2	.598	0	-.307	1.697	-.102	1.555
3	.820	.307	0	1.946	.240	1.752
4	-.513	-1.1697	-1.946	0	-2.119	.534
5	.706	.102	-.240	2.119	0	1.723
6	-.789	-1.555	-1.752	-.534	-1.723	0

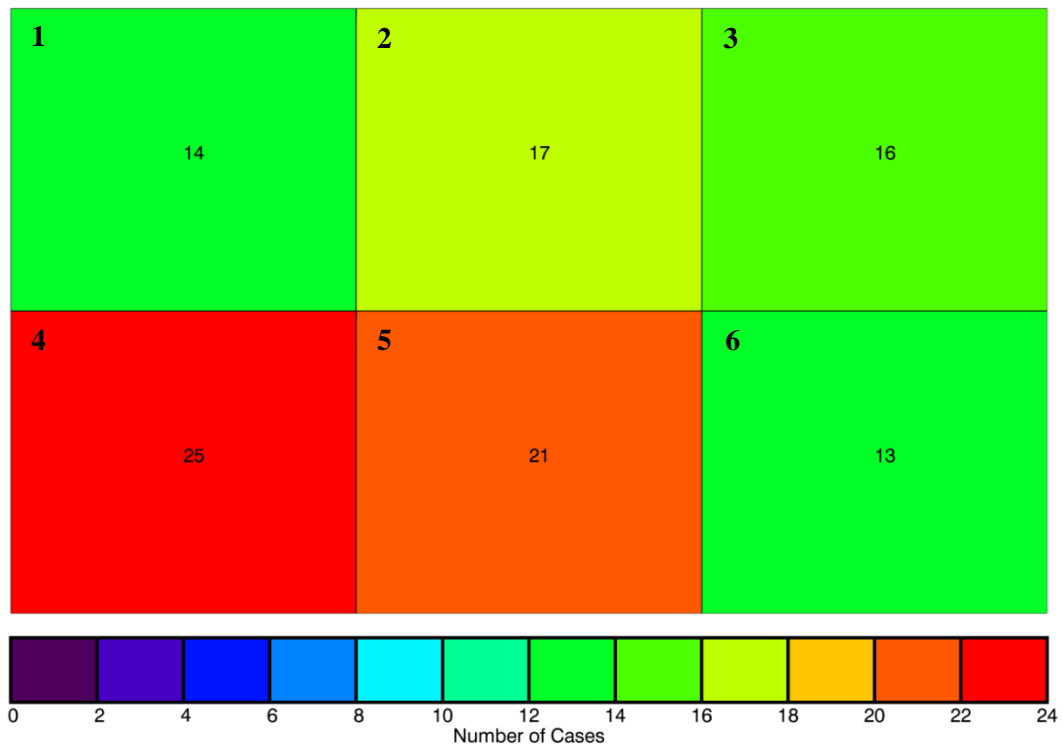


Figure 40. Number of cases in each class of the SGP CDF50 averaged SOM.

Discussion

Overall, the model displays diurnal tendencies with an over prediction of daytime precipitation and an under prediction of nocturnal precipitation. Daytime precipitation occurs in very small amounts for several classes but the classes with the higher precipitation events are largely due to patterns where high RH and surface warm fronts coincide with upper-level troughs. Nocturnal precipitation is driven by patterns with

strong forcing at the surface and aloft (tighter pressure or height gradients) or by patterns with weaker surface forcing and northwest flow aloft. Meteorologically, these are believed to be MCSs that originated well away from the domain, then propagated over the region during these late hours, a feature seen in Goines and Kennedy (2017). It is well documented that MCSs account for a significant amount of warm season precipitation across the plains, particularly during June – August (Fritsch et al. 1986) and that these MCSs have a large nocturnal component (Maddox 1980; Carbone and Tuttle 2008). This supports the hypothesis of MCSs being responsible for the nighttime negative model bias in that the cases contributing to these patterns occur later in the year when MCS activity is more common (Figure 41) and the 6-hour time period with highest precipitation totals for these patterns is 0600-1100 UTC (Figure 24). In Figure 24 the patterns in the upper right with strong forcing at the surface and aloft also have a high percentage of cases with maximum precipitation occurring overnight. For these cases, it is hypothesized that the convection began as isolated convection and grew upscale as the day progressed. These patterns also have a higher negative precipitation bias overnight (Figure 23).

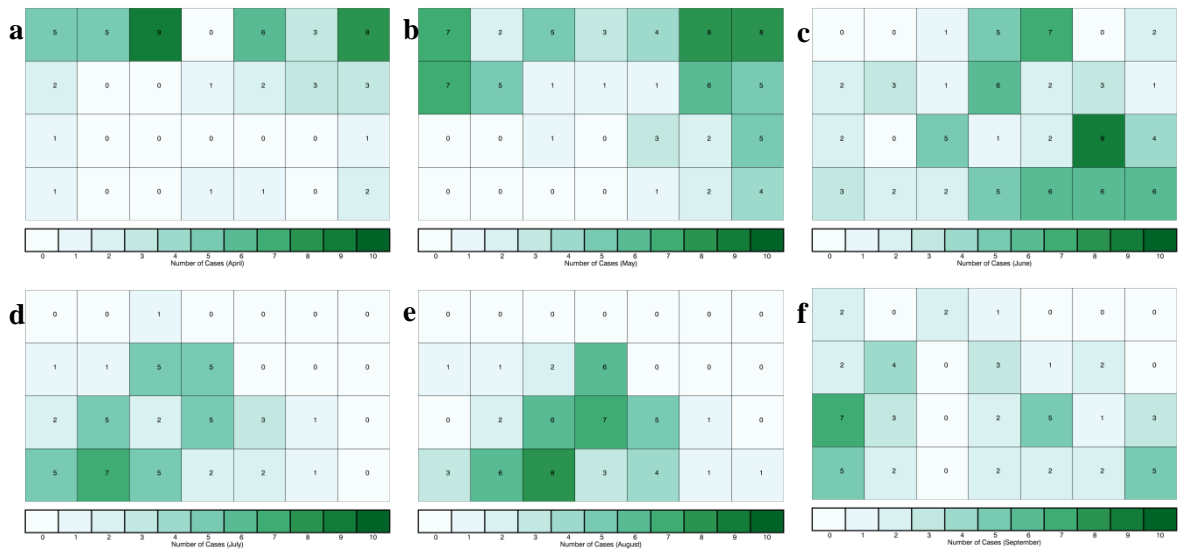


Figure 41. Distribution of SGP CDF90 cases that occur for each pattern by month. April, (a), May (b), June (c), July (d), August (e), September (f).

Northern Plains (NP)

The NP region experiences a range of patterns, from strong and moist mid-latitude cyclones with shortwave troughs or flow out of the southwest aloft to weaker and drier patterns with zonal flow aloft (Figures 42 and 43). Compared to SGP, the mid-latitude cyclones have higher relative humidity and stronger upper-level winds. This region is not prone to dry lines, and thus, strong gradients in humidity are not common within the SOM. Precipitating patterns are primarily confined to the lower-left area of the SOM (Figure 44). While precipitation is largely confined to cases with strong southwesterly flow aloft, precipitation increases for several classes with northwest flow aloft, similar to what was found at SGP. Unlike the former location, however, this local maximum is less pronounced. These are the same patterns that have more precipitation occurring in the late afternoon and evening (Figure 45). Patterns with heaviest precipitation amounts are also the classes with the highest precipitation bias in the model

(Figure 46). While the diurnal trend is less obvious compared to SGP, there is clearly more positive bias during the daytime and negative bias overnight (Figure 47).

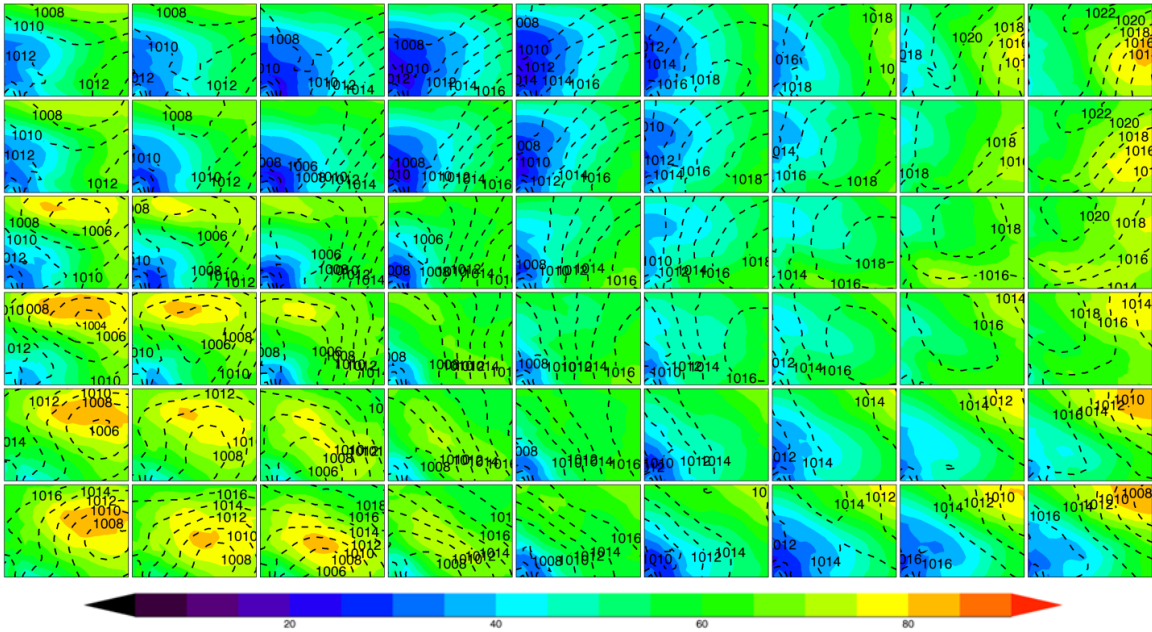


Figure 42. Near-surface analyses for the 9x6 (54-class) NP climatology SOM.

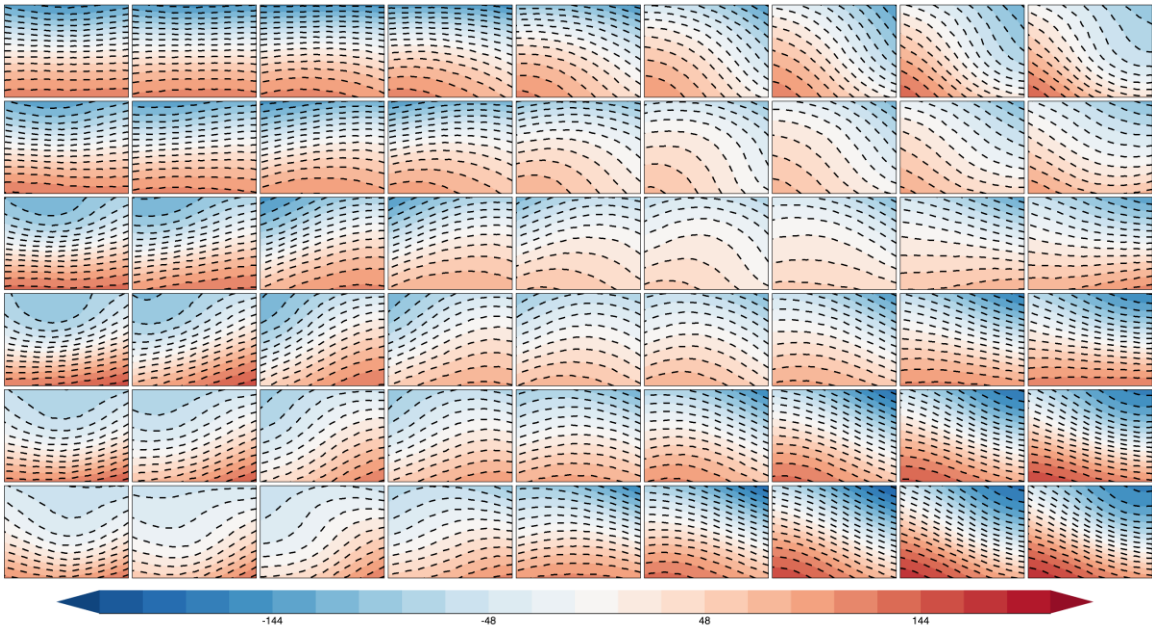


Figure 43. Climatology SOM for NP 500 hPa geopotential height anomalies.

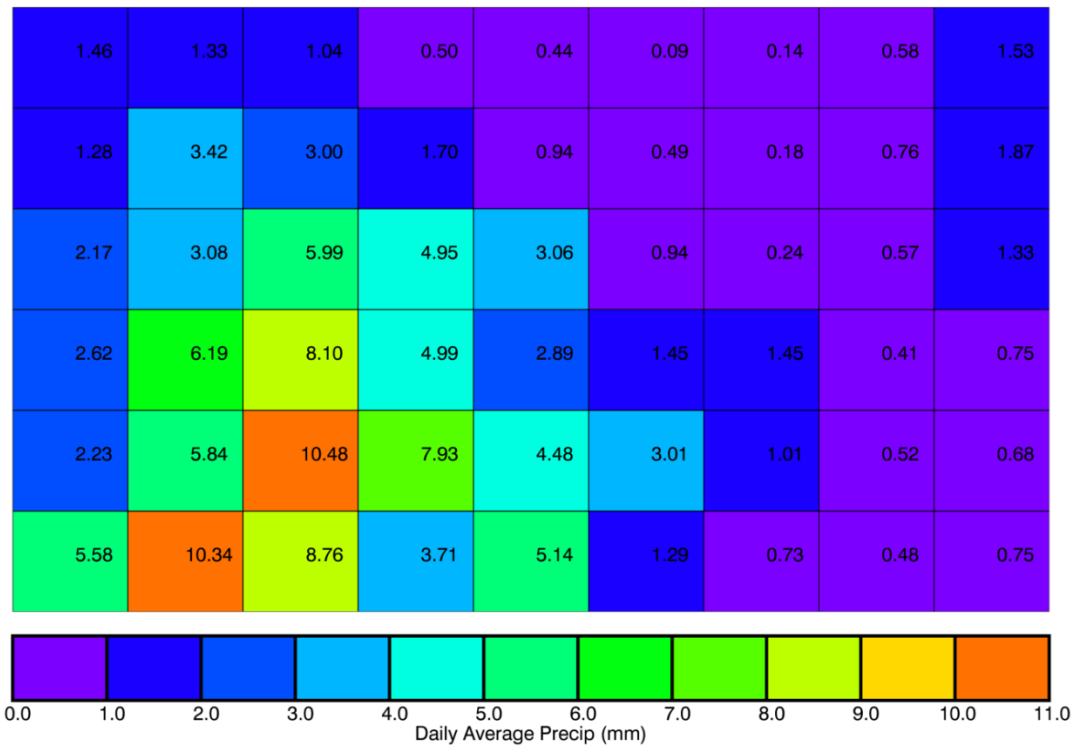


Figure 44. Daily average precipitation for the NP climatology SOM.

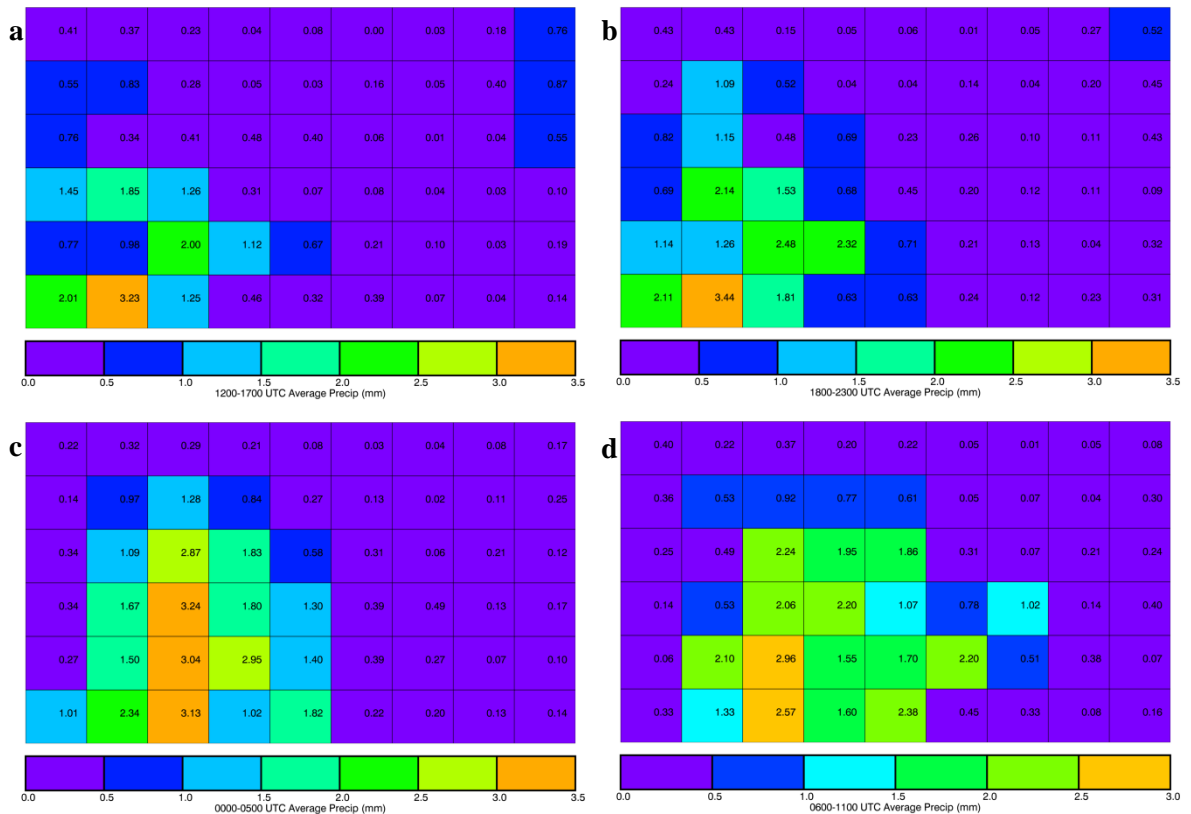


Figure 45. Average precipitation bias in six-hour increments for the NP climatology SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), and 0600-1100 UTC (d).

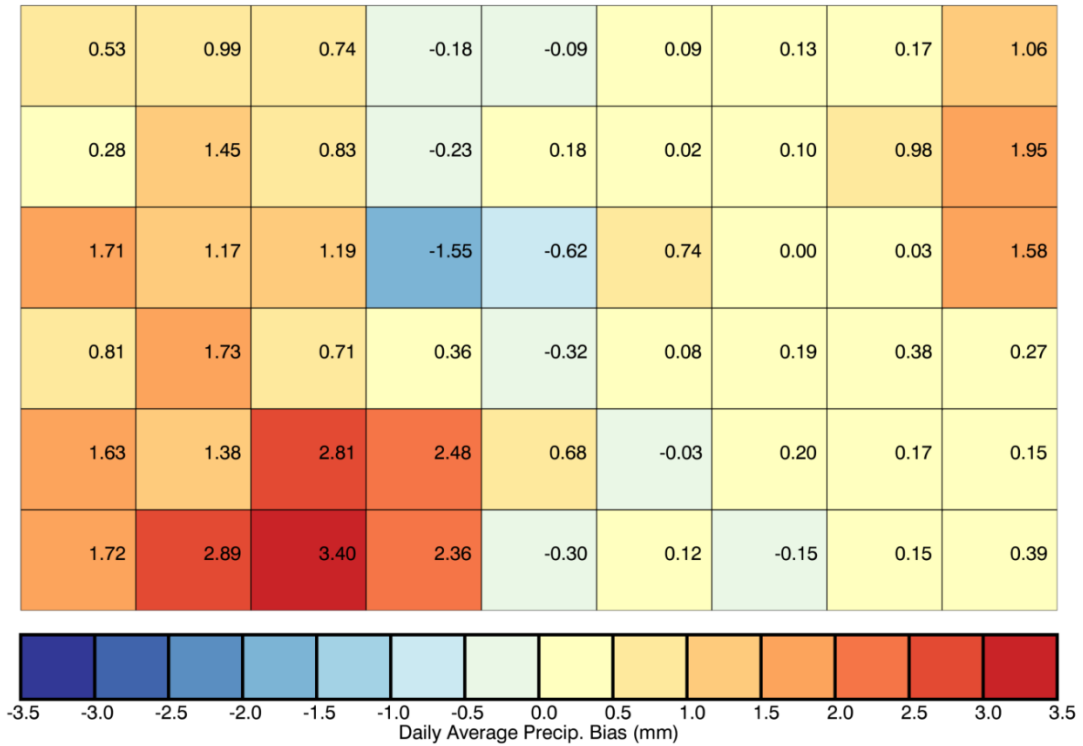


Figure 46. Daily average precipitation bias for the NP climatology SOM.

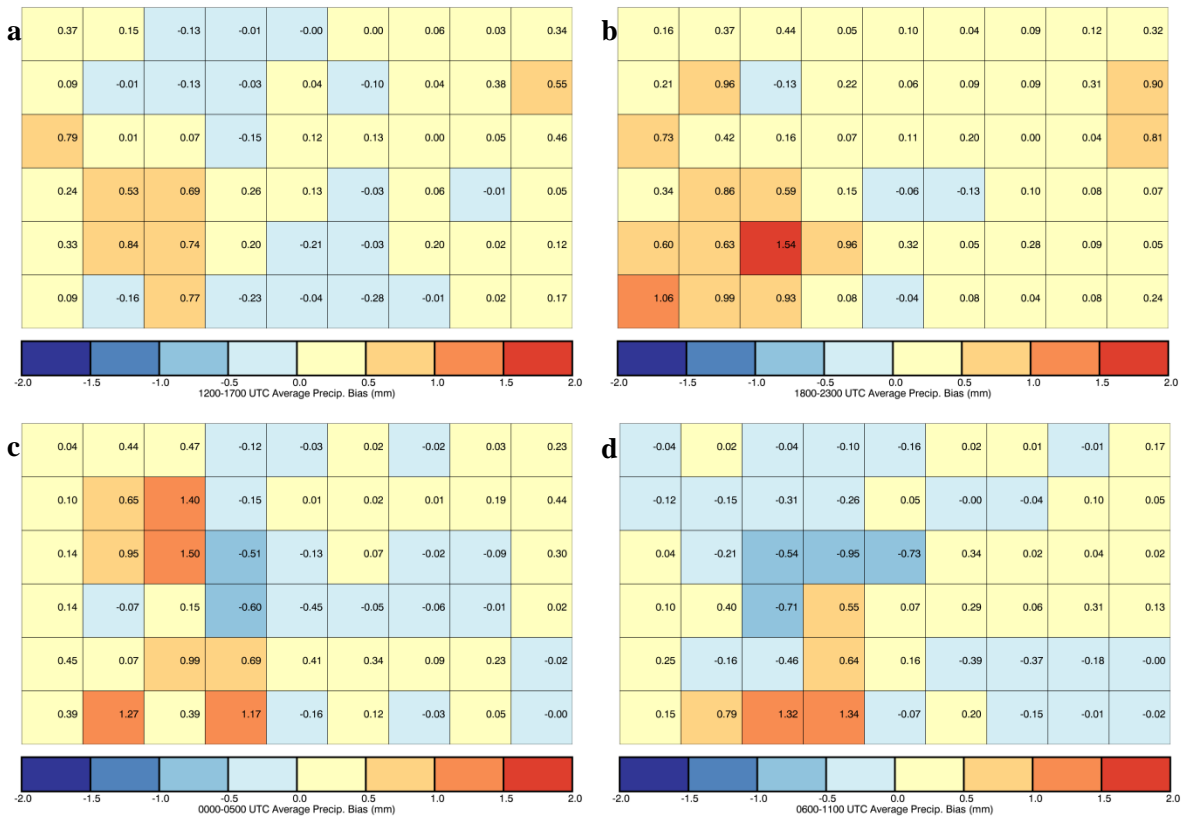


Figure 47. Average precipitation bias in six-hour increments for the NP climatology SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), and 0600-1100 UTC (d).

The classes that contribute to the CDF90 and CDF50 SOMs for the region (Figures 48 and 49) are generally in the lower-left region of the SOM. However, several cases come from the right side of the SOM, where many of the classes have a high percentage of cases with precipitation occurring (Figure 50). For the selected CDF50 cases, these come from slightly right by a column in the SOM than the CDF90 cases. This also happens to be where the few negative or lower precipitation bias classes are, suggesting that these heavier precipitation cases are the nocturnal MCS events. In SGP many of the nocturnal events occurred for cases with southwesterly flow aloft and surface cyclones. While the cyclones for this region don't have the classic dryline setup seen in SGP, the same upper level flow pattern over low level cyclones exhibits the negative precipitation bias which was seen for the higher precipitation cases in the SGP.

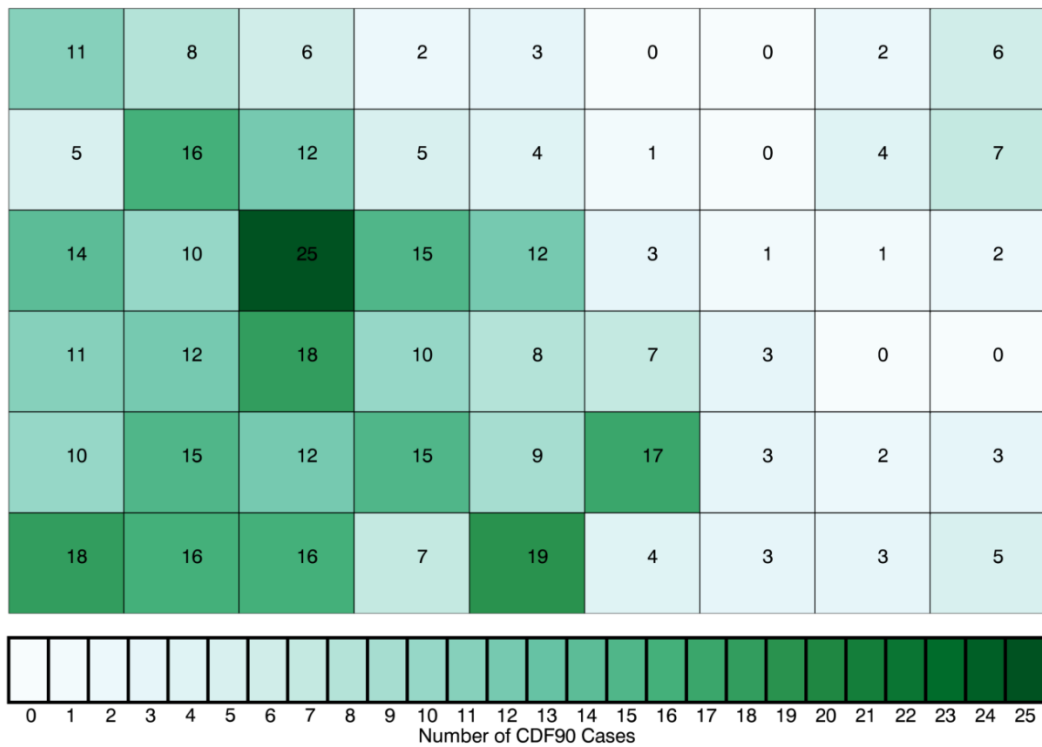


Figure 48. Distribution of CDF90 classes within the NP climatology SOM.

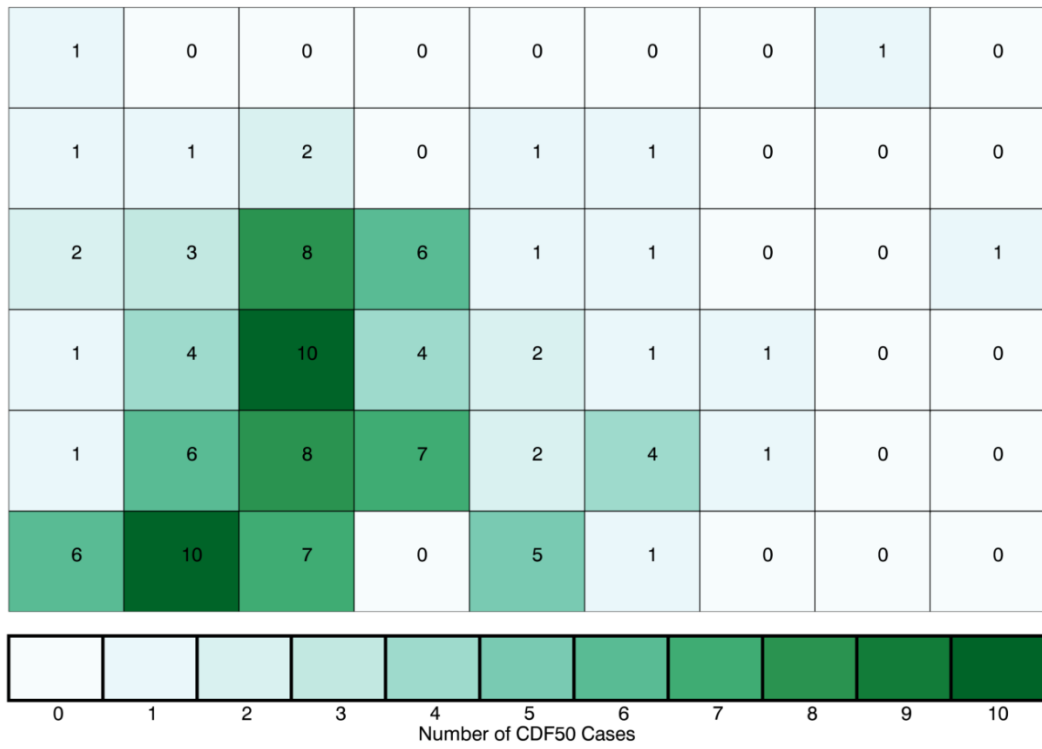


Figure 49. Distribution of CDF50 classes within the NP climatology SOM.

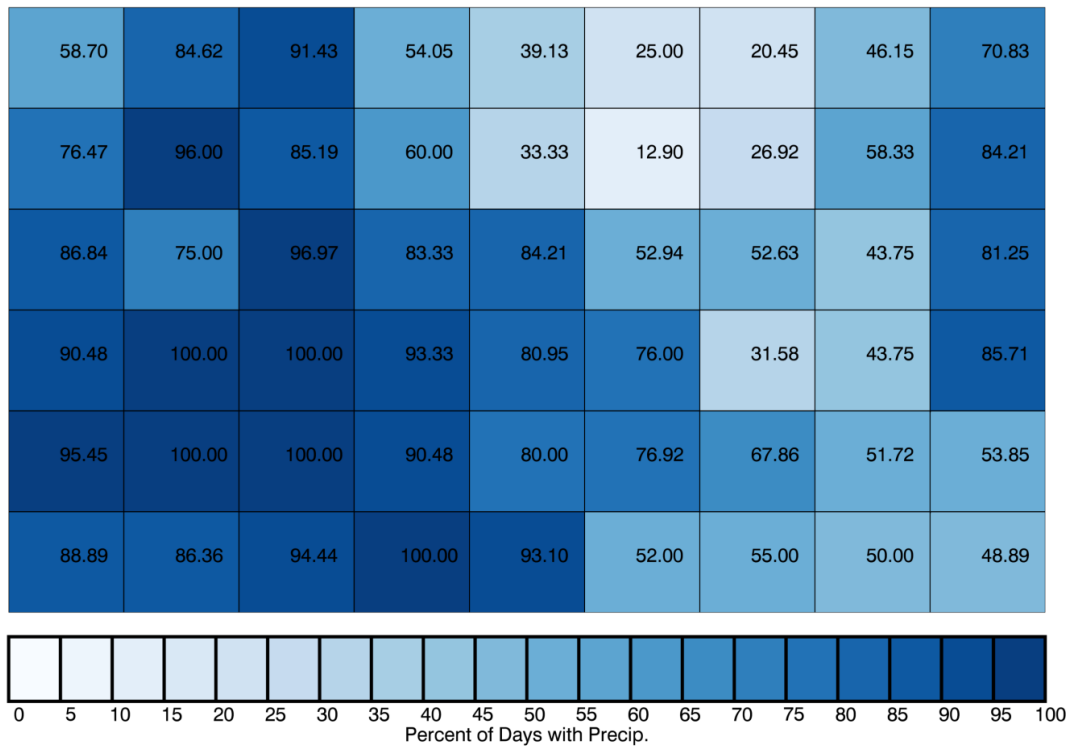


Figure 50. Percent of days in each class with precipitation.

Patterns for the NP CDF90 SOM are similar to the climatology SOM and range from mid-latitude cyclones with upper-level troughs on the right side of the SOM to weaker forced surface patterns with zonal or southwesterly flow aloft (Figures 51 and 52). Regimes with higher precipitation totals (Figure 53) occur in the upper right corner of the SOM, where lower level patterns indicate high moisture levels on the north-to-northwest side of a low pressure center with flow aloft primarily out of the southwest.

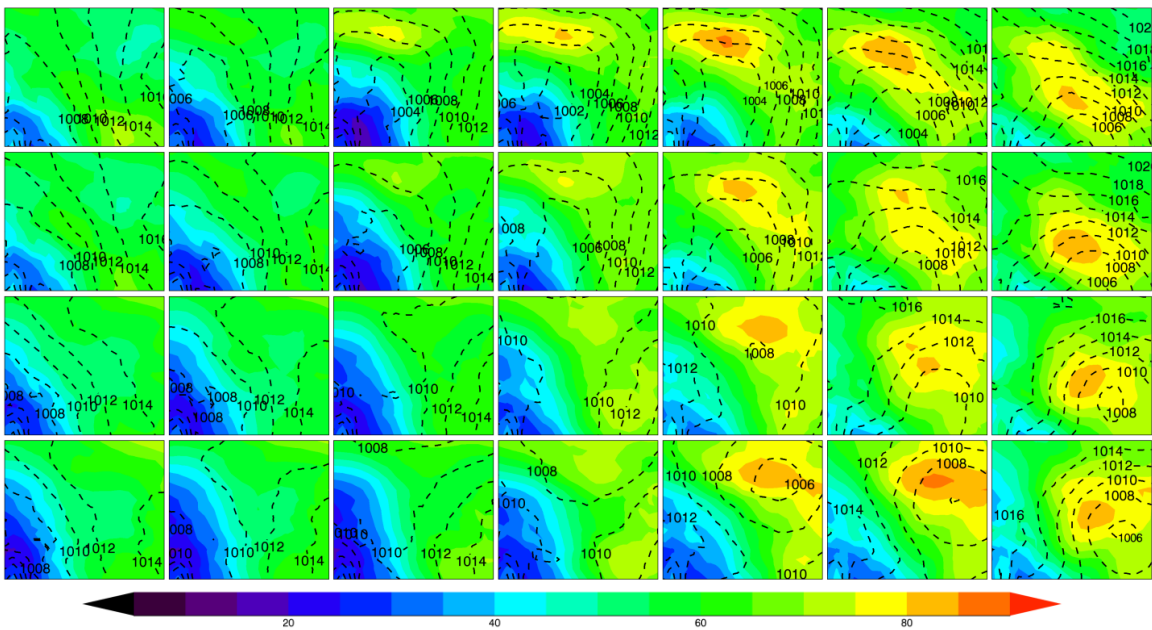


Figure 51. Near-surface analyses for the NP CDF90 SOM.

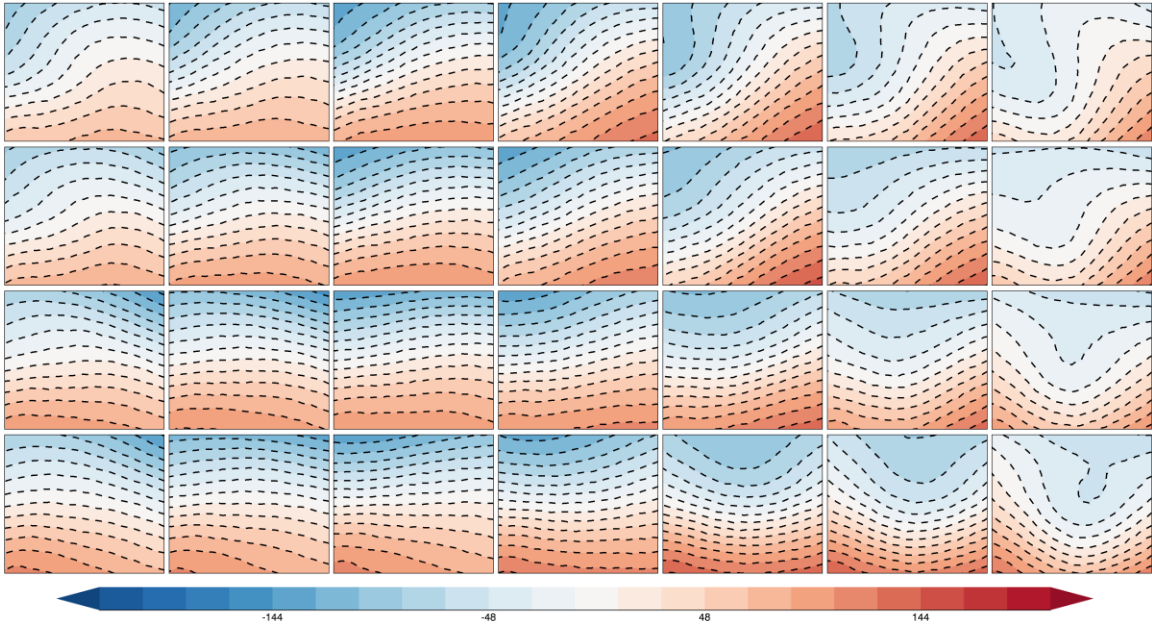


Figure 52. The 500 hPa geopotential height anomalies for the NP CDF90 SOM.

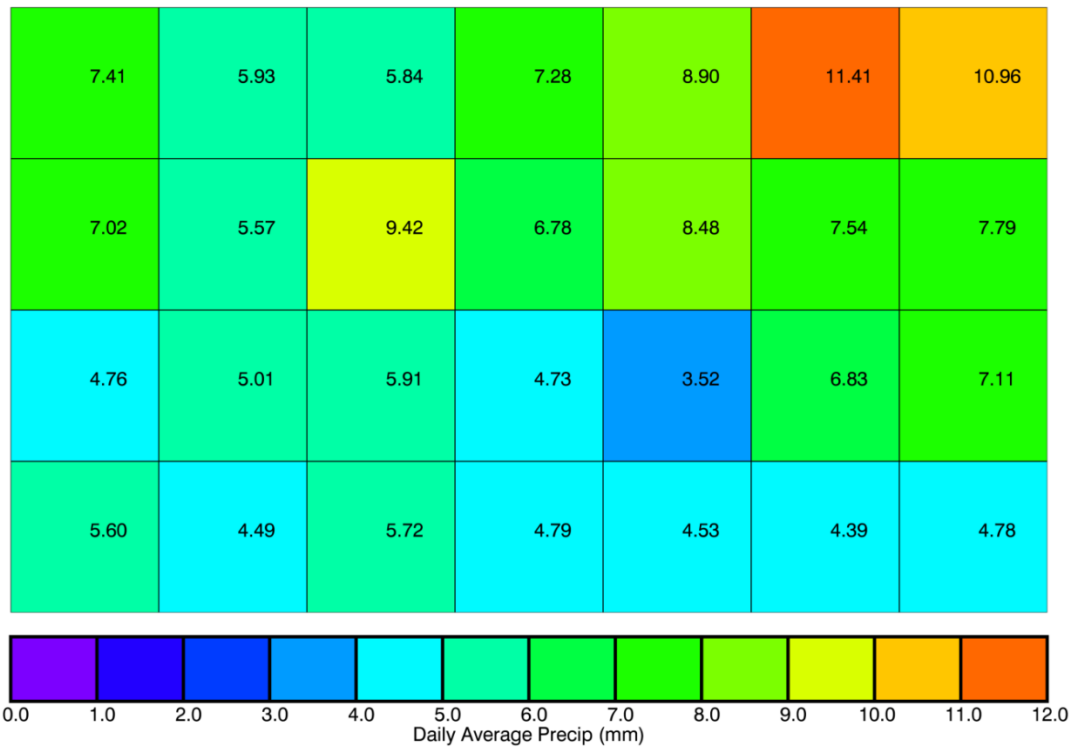


Figure 53. Daily average precipitation for the NP CDF90 SOM.

Precipitation biases are overwhelmingly positive for this region (Figure 54), as expected based on Table 3. Patterns that include cases which primarily occur overnight

contribute to the few classes that have negative biases, on the left side of the SOM. Understanding the negative bias in this case requires a closer look at the patterns. While near-surface patterns are similar across most of the left side of the SOM, upper level flow has some variability. The upper left corner features southwesterly flow and the lower left corner features northwesterly flow, while the middle patterns which still had the positive bias tend to have more zonal flow aloft. These changes in upper level flow have a noticeable impact on model bias while lower level patterns remain the same.

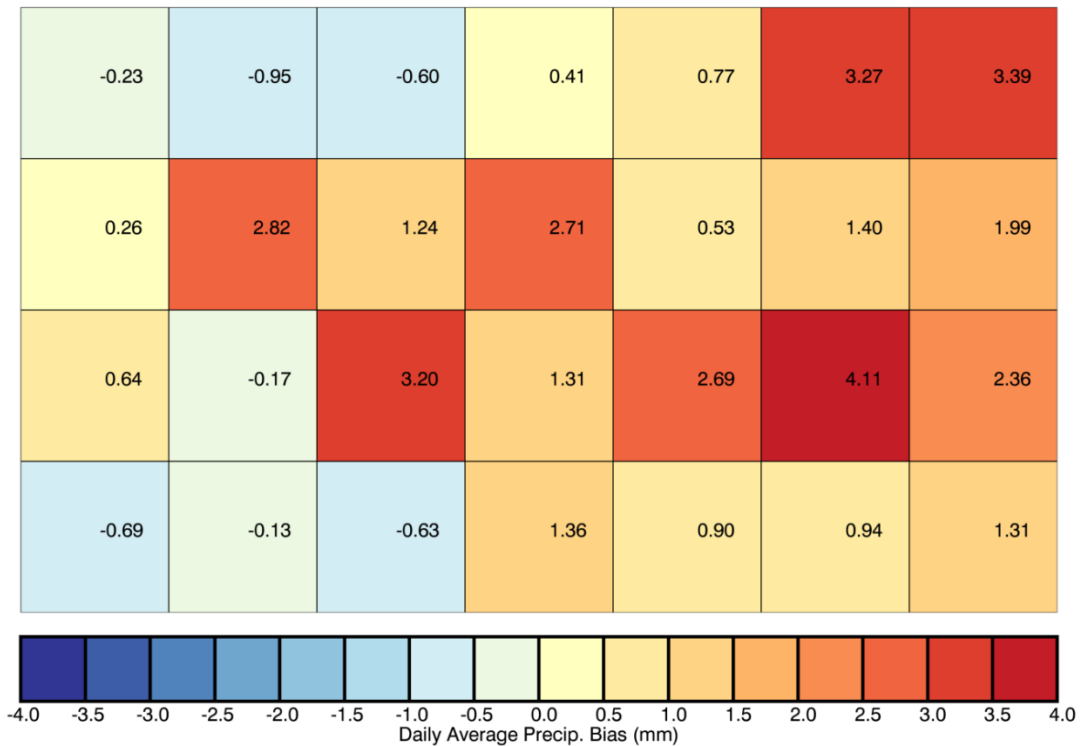


Figure 54. Daily average precipitation bias for the NP CDF90 SOM.

While the signal for negative precipitation bias associated with MCS activity is harder to discern, the NP does in fact see this property as well. Precipitation plots for the SOM show that precipitation occurs during the daytime for the warm front classes on the right side of the SOM and becomes nocturnal for classes on the left side, which are the

weaker forced classes (Figure 55). Precipitation bias throughout the day evolves from nearly completely positive during the daytime to a portion on the left side of the SOM with a strong negative bias from 0600-1100 UTC (Figure 56). Cases for SOM tend to be evenly distributed, except for the lower-left corner of the SOM (Figure 57) where the weaker surface patterns with zonal to northwest flow aloft are located. Here a class such as the very lower-left corner has over 10 more cases classified to it than the neighboring classes. This indicates that the SOM is most likely not capturing the full variability in patterns.

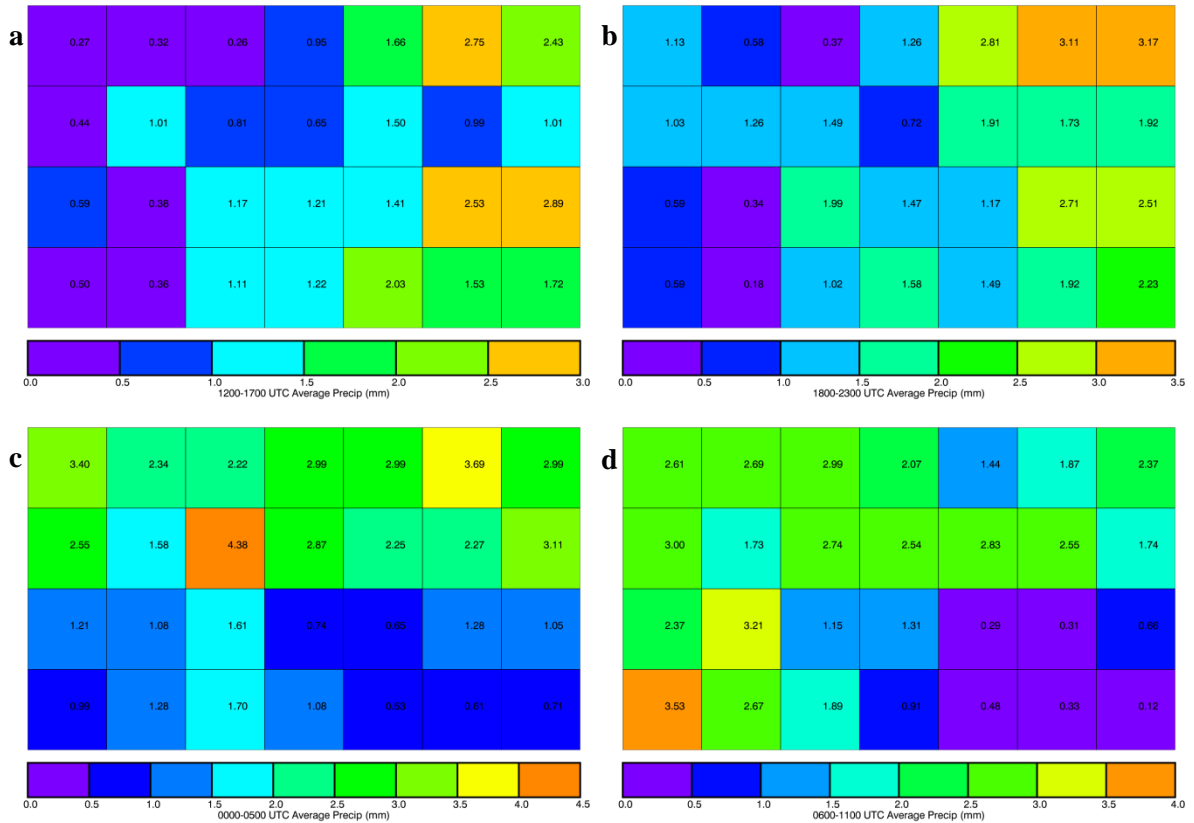


Figure 55. Average precipitation in six-hour increments for the NP CDF90 SOM

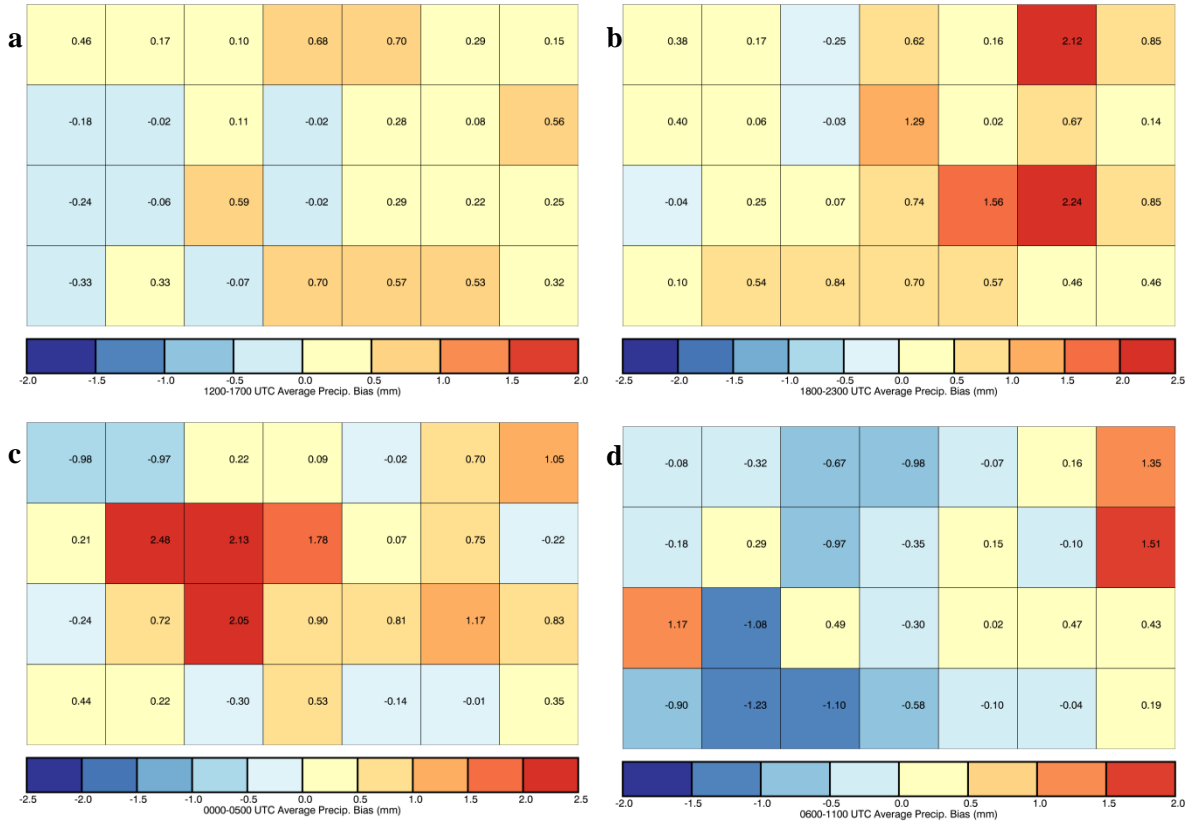


Figure 56. Average precipitation bias in six-hour increments for the NP CDF90 SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), and 0600-1100 UTC (d).

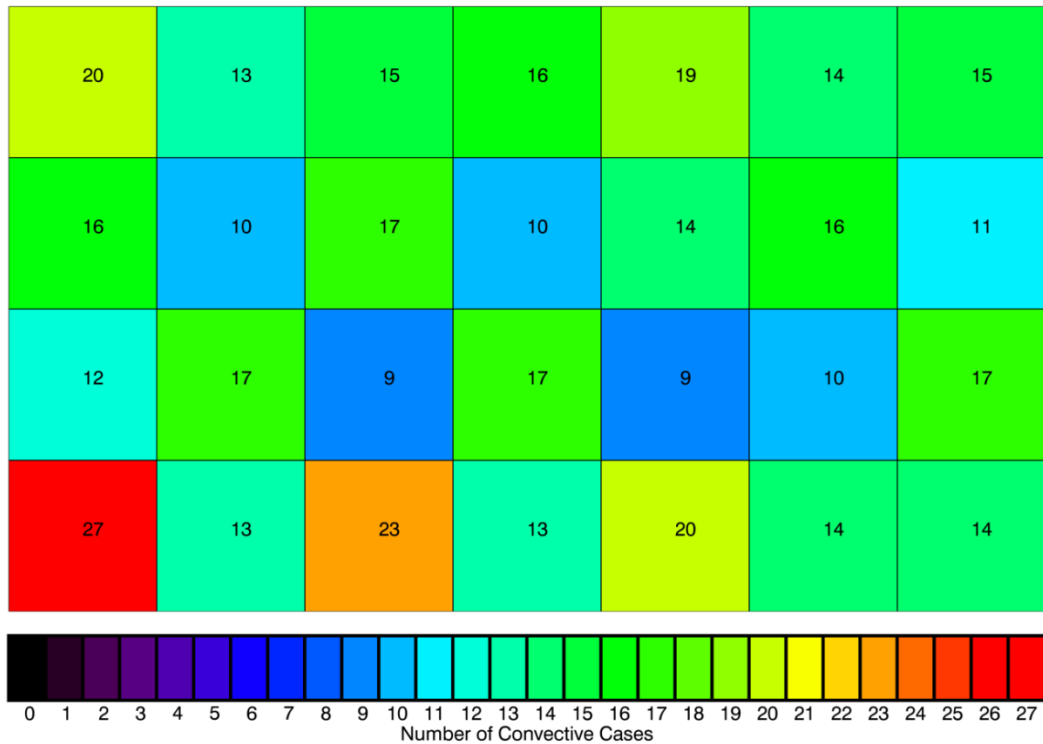


Figure 57. Number of cases in each class of the NP CDF90 SOM.

Statistical Analysis

Grouping the patterns retains the same general pattern properties that are seen in the larger SOM, as was also the case for the SGP region (Figures 58 and 59). The stronger cyclone classes with higher moisture and warm fronts are on the right side of the SOM while the weaker surface patterns are to the left (Figure 58). Pronounced troughs remain in the averaged SOM (Figure 59) and bias tendencies are as expected from previous results with a gradient in negative to slightly positive biases on the LHS and strong positive biases on the RHS (Figure 60).

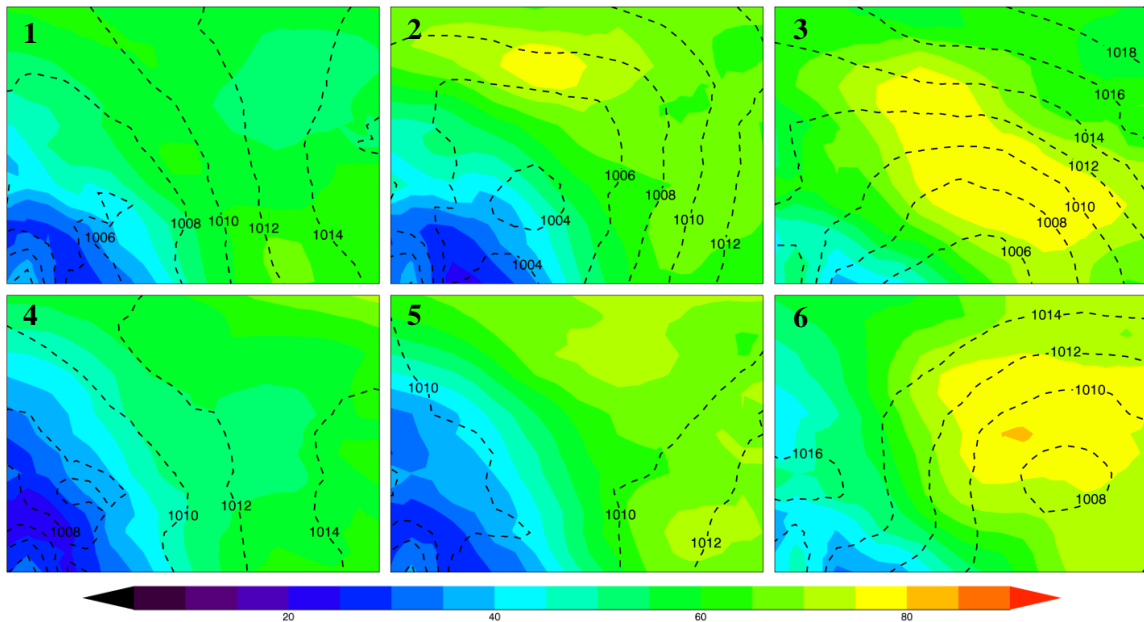


Figure 58. Near-surface analyses for the NP CDF90 averaged SOM.

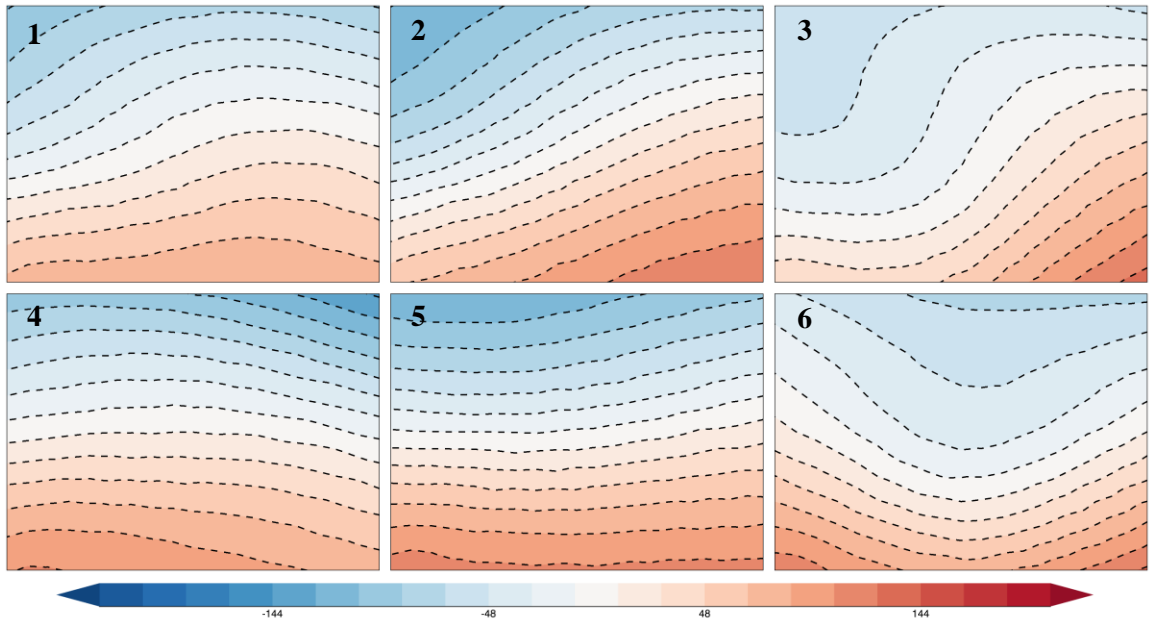


Figure 59. The 500 hPa geopotential height anomalies for the NP CDF90 averaged SOM.

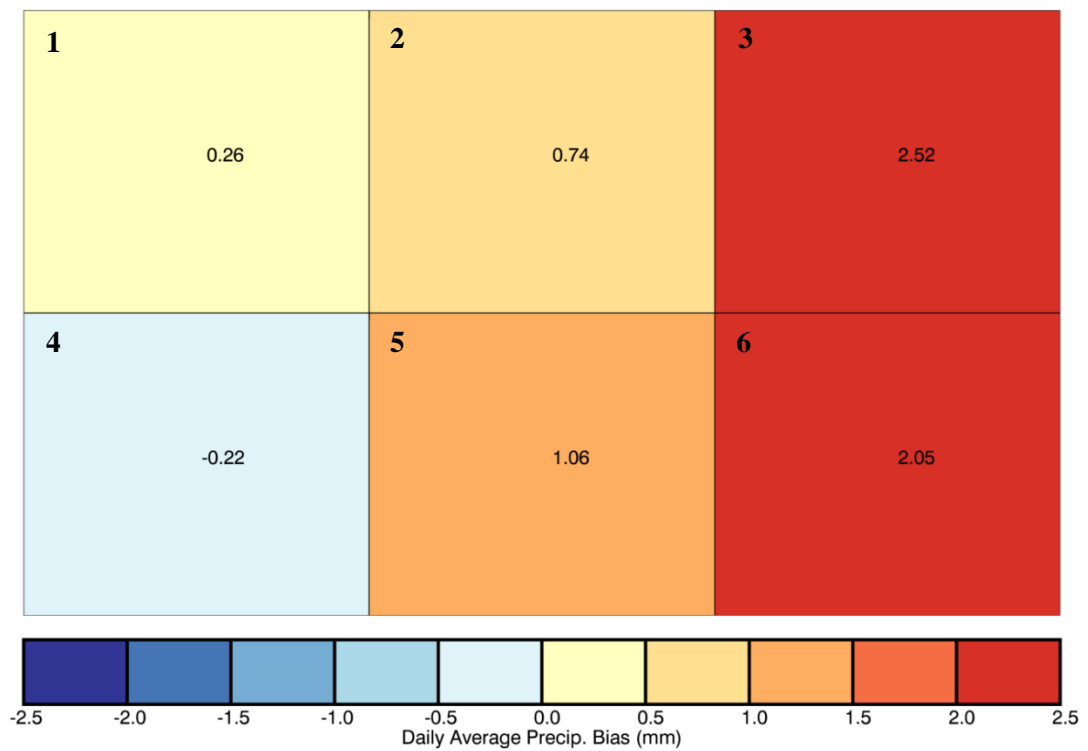


Figure 60. Daily average precipitation bias for the NP CDF90 averaged SOM.

Statistically, classes 3 and 6, the classes with the highest positive bias, are significantly different for the other classes in the averaged SOM, excluding each other (Table 6). Class 4, the negative precipitation bias class, is significantly different from the three highest positive bias classes. It is likely that the small group of negative bias that was seen for the NP CDF90 SOM in the upper left (Figure 54) is smoothed out because of some grouping with positive bias classes. Because there are fewer classes with negative precipitation biases in the NP region, the signal averages out when the classes are grouped together for analysis. The number of cases contributing to each class is shown in Figure 61.

Table 6. Z-Test statistic results for the NP CDF90 SOM. Statistically significant results in **bold**.

Comparison Class	1	2	3	4	5	6
1	0	.667	2.803	-.685	1.202	2.450
2	-.667	0	2.461	-1.574	.564	2.053
3	-2.803	-2.461	0	-3.859	-2.209	-.643
4	.685	1.574	3.859	0	2.358	3.642
5	-1.202	-.564	2.209	-2.358	0	1.748
6	-2.450	-2.053	.643	-3.642	-1.748	0

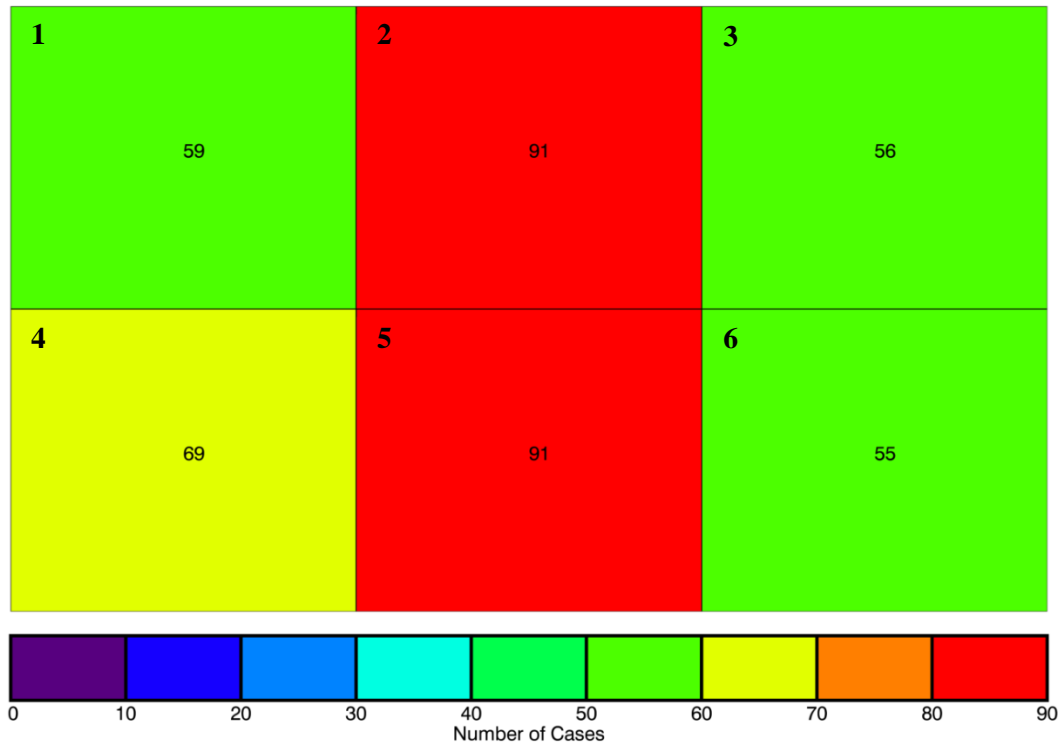


Figure 61. Number of cases in each class of the NP CDF90 averaged SOM.

CDF50

The CDF50 cases for the NP occur within the CDF90 SOM in the upper-right of the SOM (Figure 62). This also shows that NP tends to see less nocturnal precipitation when compared to the SGP region. The patterns represented in the SOM are strong mid-latitude cyclones with high moisture accompanied by flow aloft out of the south or southwest, or with shortwave troughs (Figures 63 and 64). Many of the cyclone patterns on the left side of the SOM are warm front classes. The drier surface patterns represented are accompanied by more zonal flow aloft.

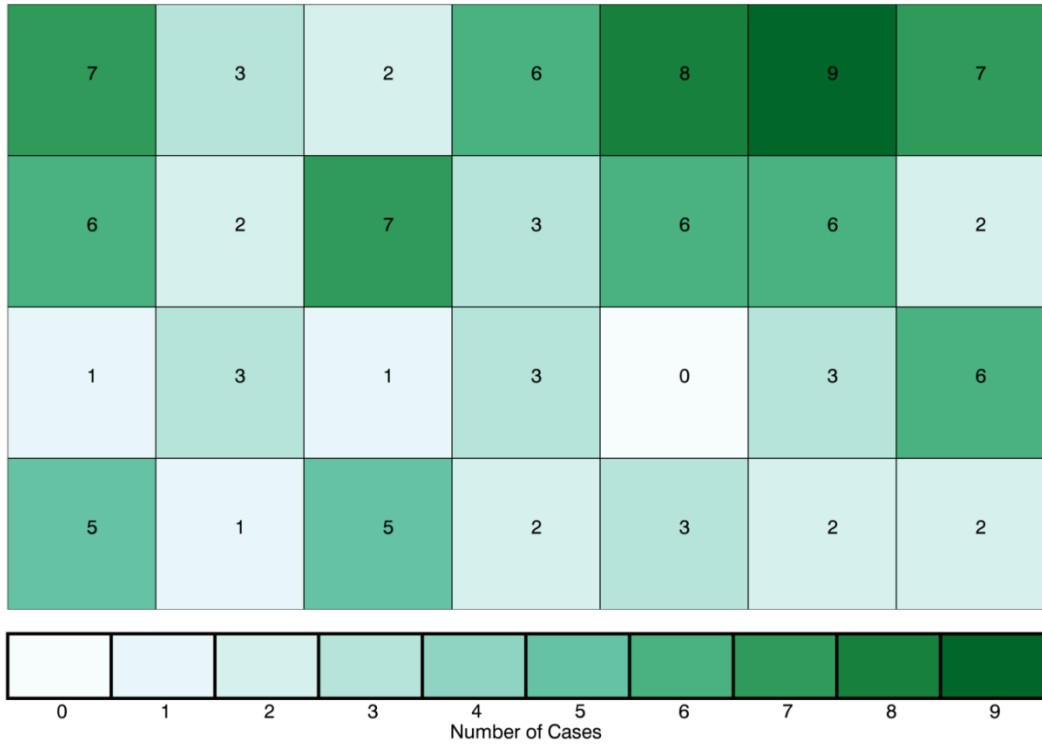


Figure 62. Plot depicting where the NP CDF50 cases are located on the NP CDF90 SOM.

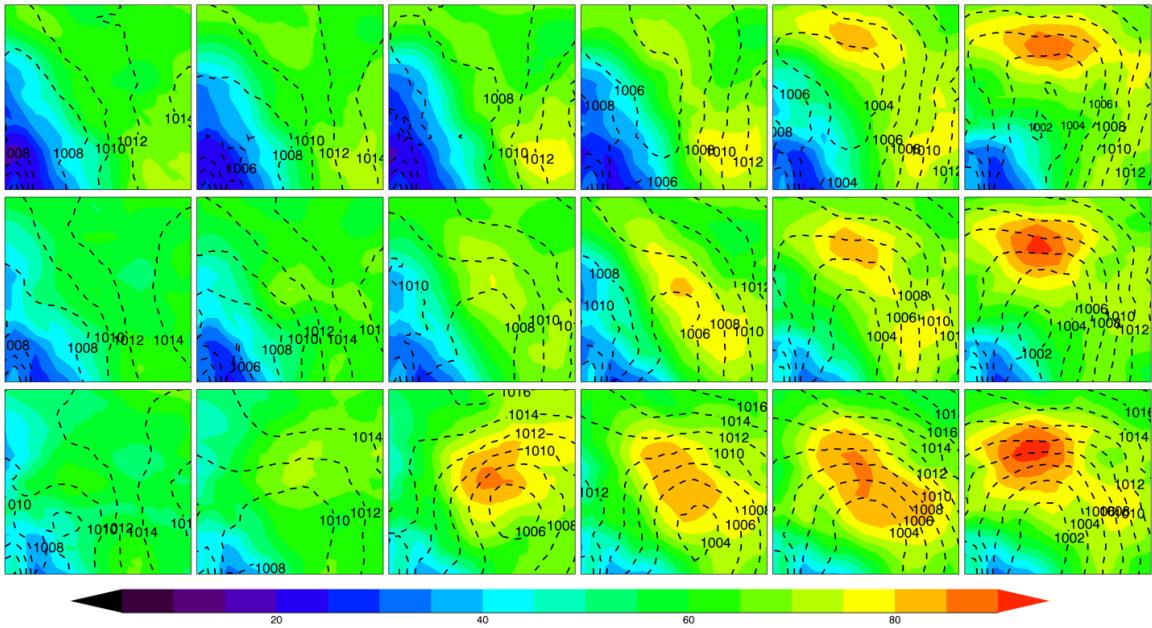


Figure 63. Near-surface analyses for the NP CDF50 SOM.

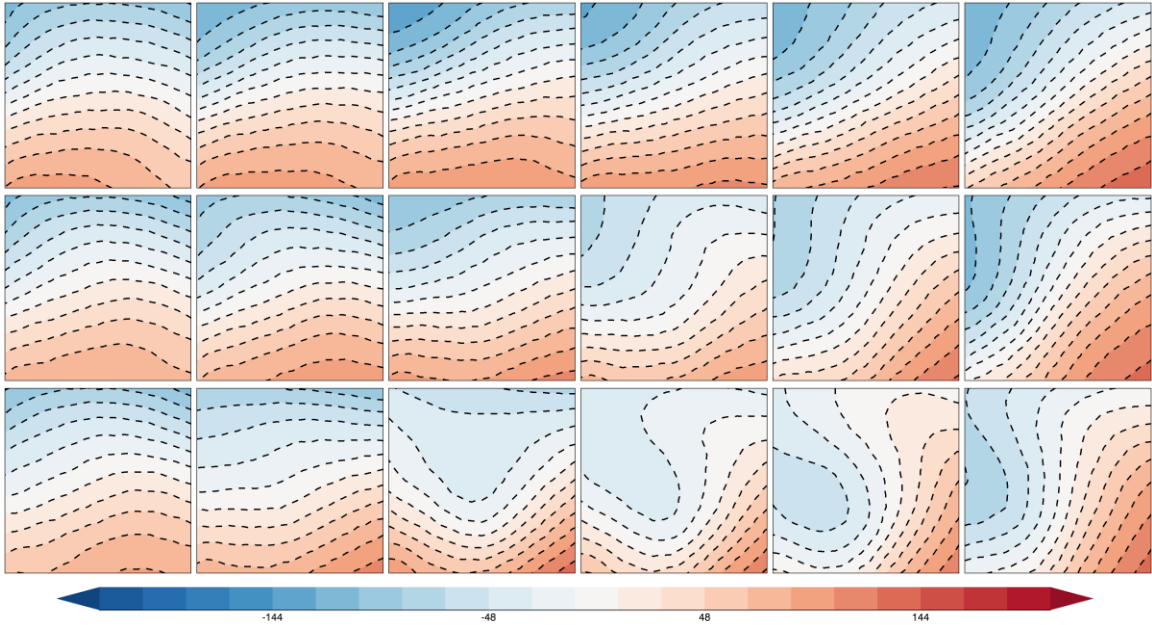


Figure 64. The 500 hPa geopotential height anomalies for the NP CDF50 SOM.

Precipitation amounts tend to be similar across the SOM, with the cyclones in the lower right producing a little more precipitation than the other classes (Figure 65). Two classes in the lower right of the SOM exhibit precipitation values much higher than other classes in the SOM. The highest class in the lower center likely sees the highest precipitation values because of optimal placement of the warm front (Figure 63) for this class, whereas other classes see the warm front shifted slightly farther north. The other class in the far lower right likely sees higher precipitation amounts because of the higher RH values (Figure 63) seen here compared to the other classes. The bias values are of higher magnitude for these CDF50 cases, with larger values for both the positive and negative precipitation bias patterns (Figure 66). While the total bias in Table 3 does not become negative here as it did for SGP, it does decrease. This is not because there is less bias for this SOM, but rather that the negative overnight bias is not enough to overcome the positive bias seen across most of the patterns for the rest of the day. This SOM also sees more wide spread precipitation during the afternoon (0000-0500 UTC) across most

of the patterns. However, the cyclones on the right do see increased precipitation amounts during the daytime (1800-2300 UTC) and the classes on the left with weaker surface forcing see increased precipitation amounts during the overnight hours (0600-1100 UTC; Figure 67). Even with increased precipitation amounts for 0000-0500 UTC, the negative precipitation bias continues to dominate with overnight precipitation, indicating that the model does not handle nocturnal precipitation well (Figure 68). Many of the strongest negative bias classes are on the LHS of the SOM, where the synoptic patterns and heavy 0600-1100 UTC precipitation amounts indicate that these are likely propagating nocturnal MCSs. This SOM also shows a lot of variability in the distribution of cases amongst classes (Figure 69). Again, this indicates more variability in the patterns represented in this SOM.

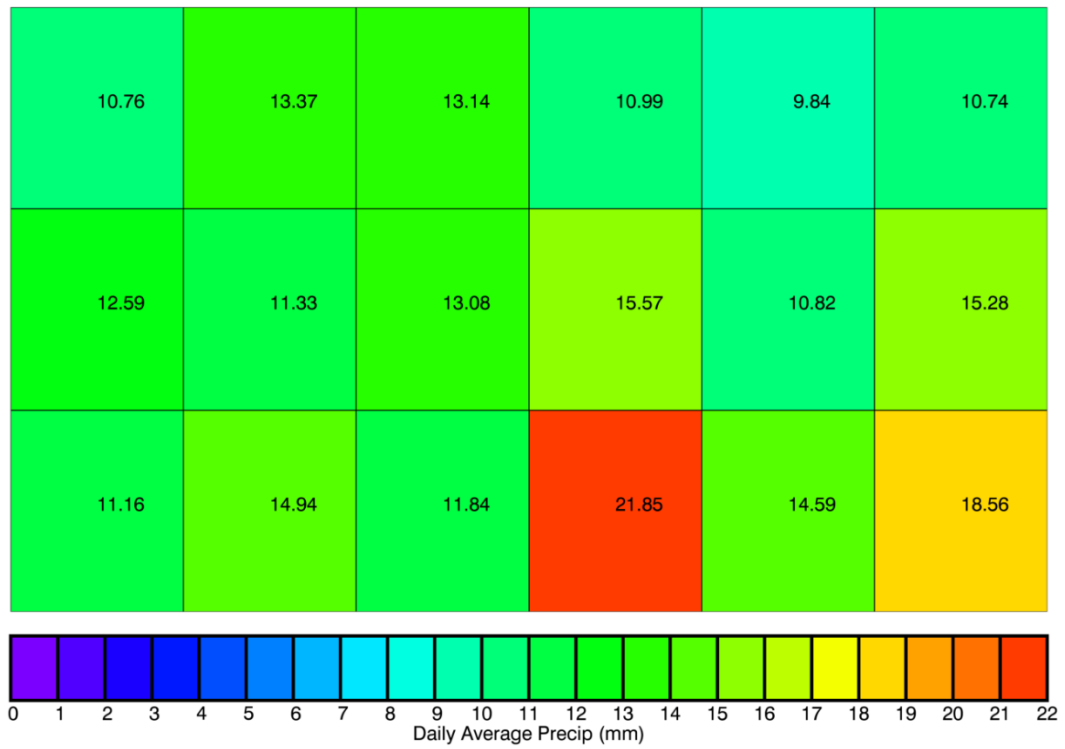


Figure 65. Daily average precipitation for the NP CDF50 SOM.

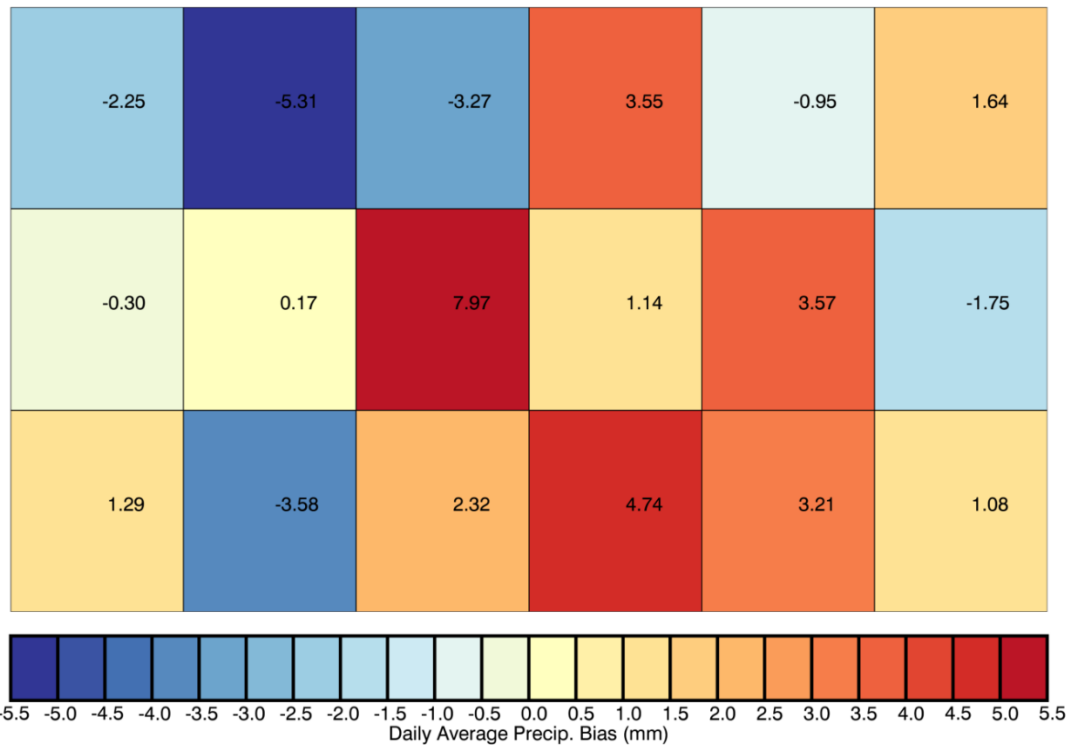


Figure 66. Daily average precipitation bias for the NP CDF50 SOM.

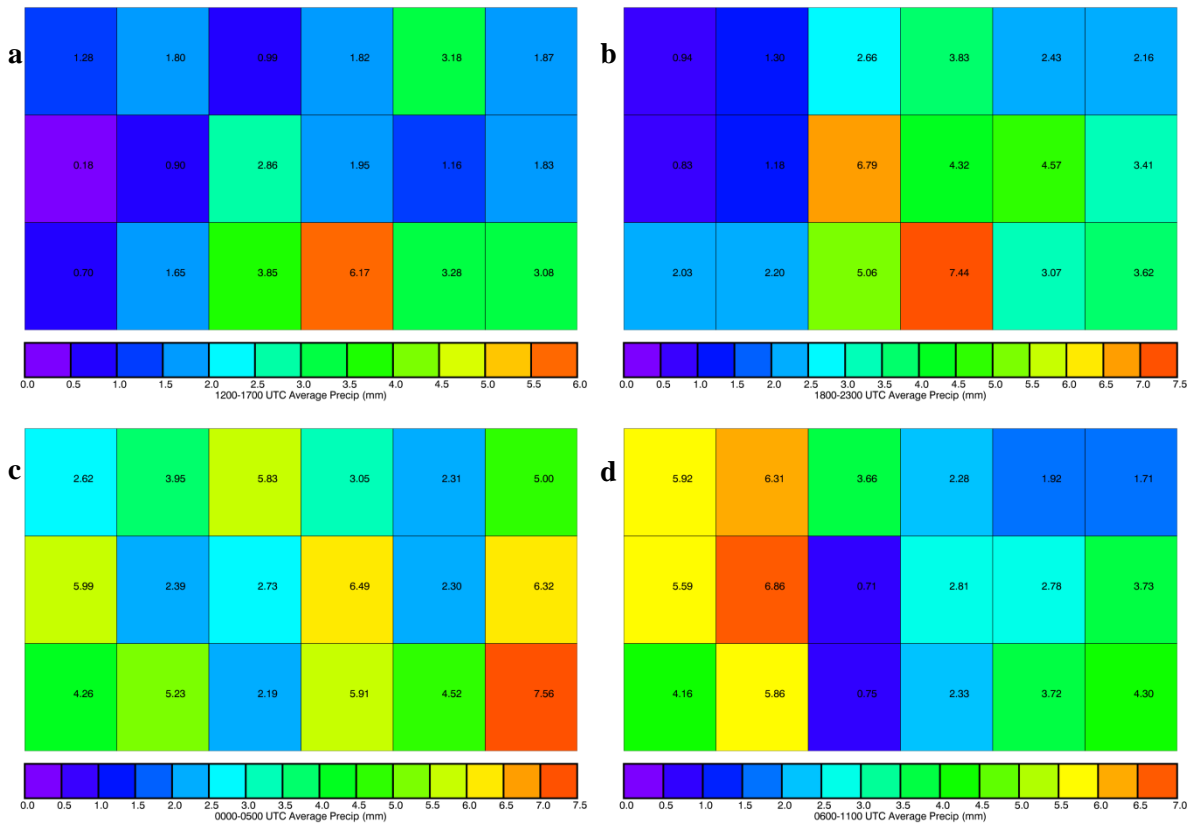


Figure 67. Average precipitation in six-hour increments for the NP CDF50 SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), and 0600-1100 UTC (d).

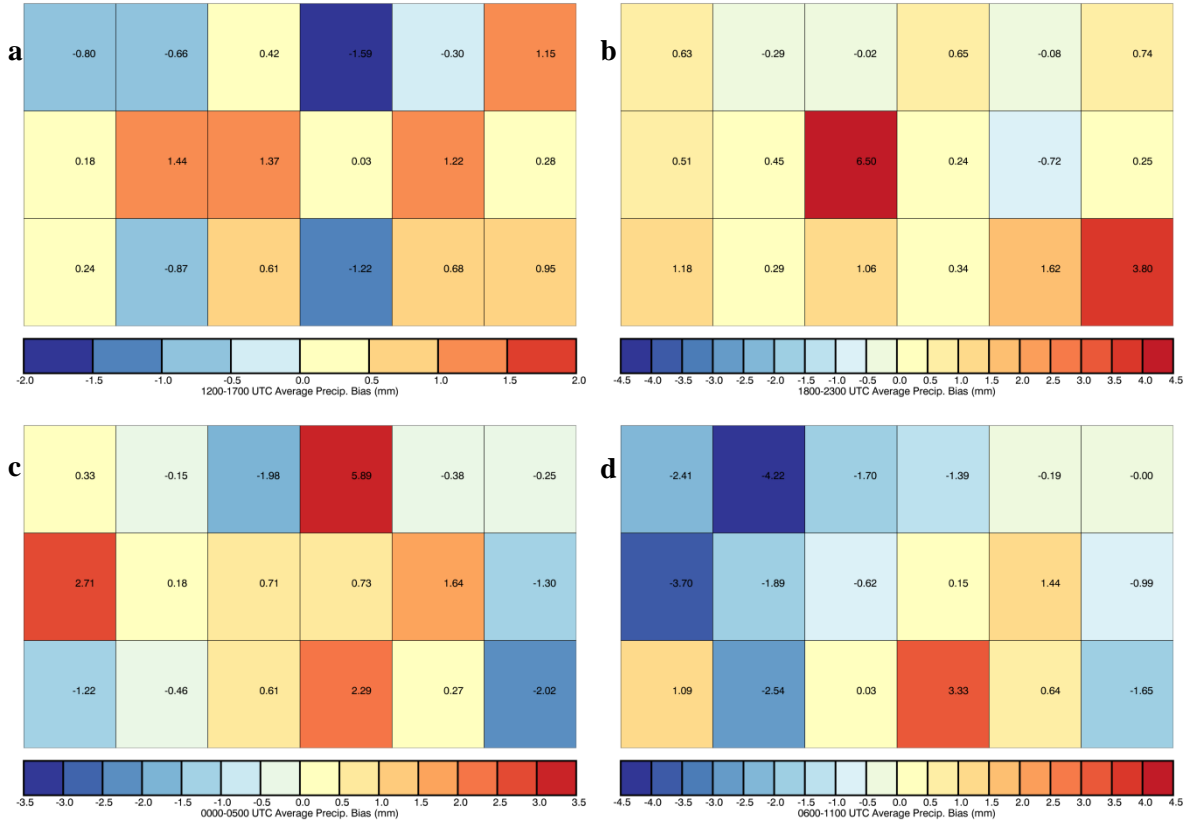


Figure 68. Average precipitation bias in six-hour increments for the NP CDF50 SOM. 1200-1700 UTC (a), 1800-2300 UTC (b), 0000-0500 UTC (c), and 0600-1100 UTC (d).

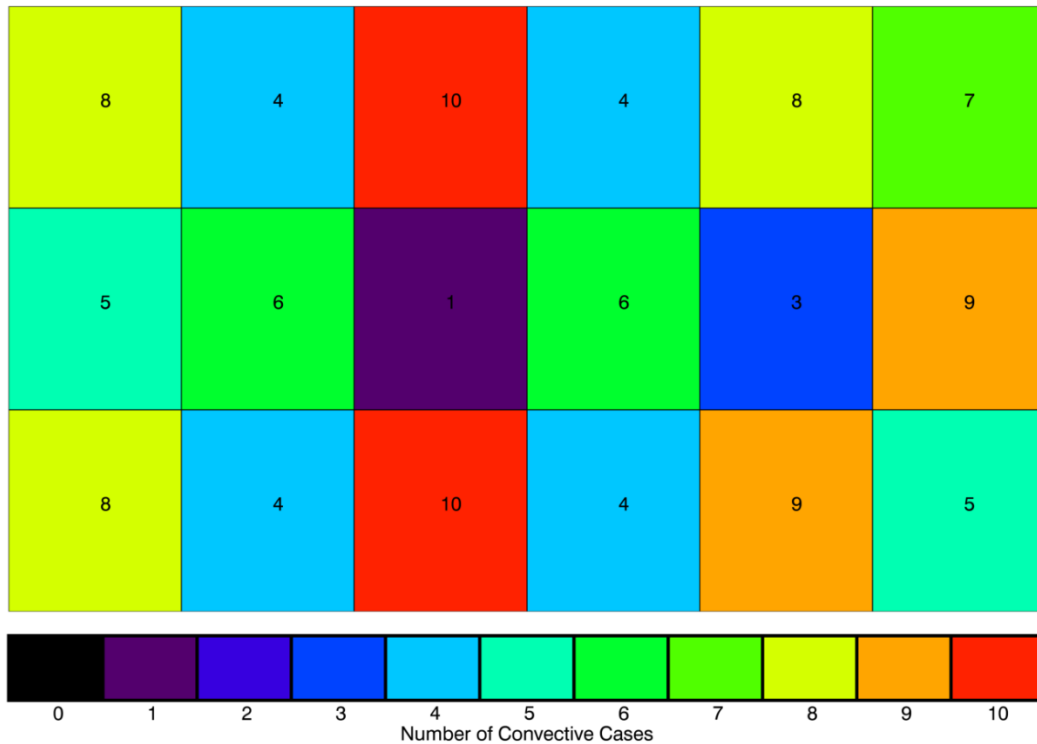


Figure 69. Number of cases in each class of the NP CDF50 SOM.

Statistical Analysis

The warm front cases on the right and the zonal/northwest flow cases on the left (Figures 70 and 71) keep the strong precipitation bias signal that was associated with the patterns in the larger SOM (Figure 72). Many of these classes are significantly different from one another, something not seen in the SGP CDF50 grouped SOM. While there are approximately the same number of cases (Figure 73) in the SGP and NP SOMs, the classes here have a higher range in bias values, whereas the SGP CDF50 was largely negative. This helps increase the significance of the NP CDF50 averaged classes. Statistically, the strong negative class is significantly different from the three positive bias classes meaning that the WRF model performs differently under these patterns (Table 7).

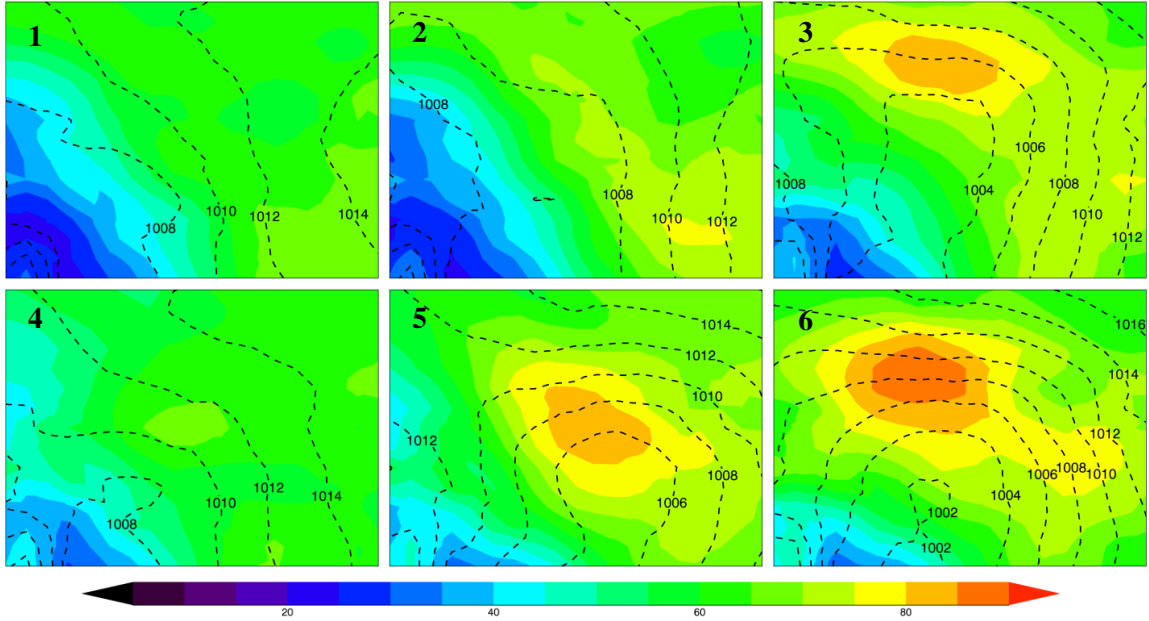


Figure 70. Near-surface analyses for the NP CDF50 averaged SOM.

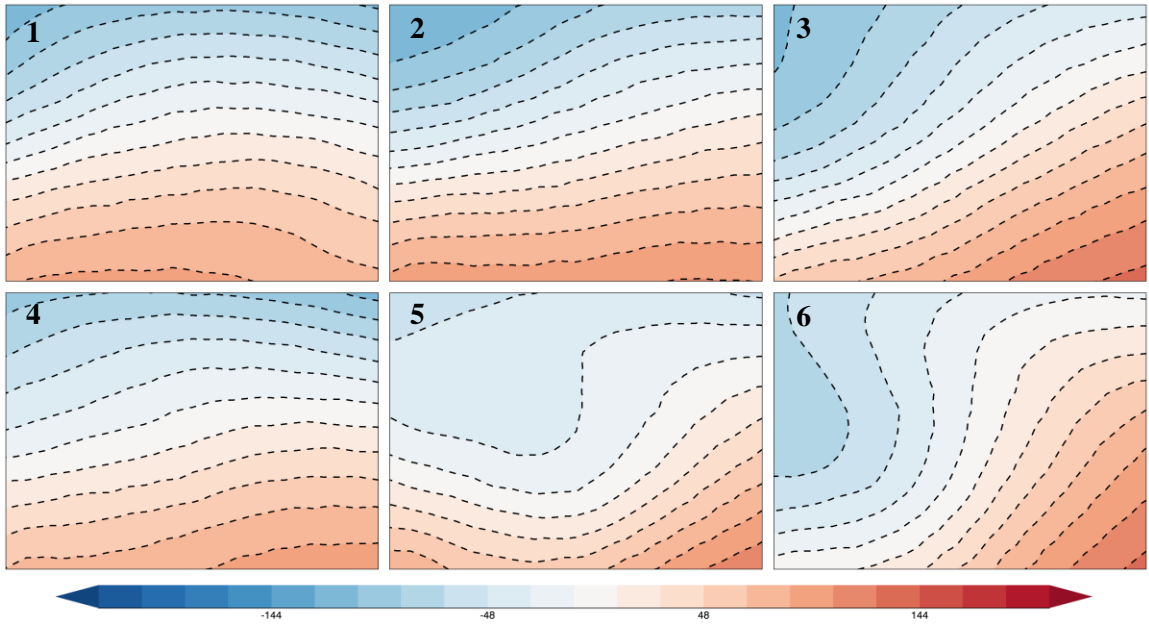


Figure 71. The 500 hPa geopotential height anomalies for the NP CDF50 averaged SOM.

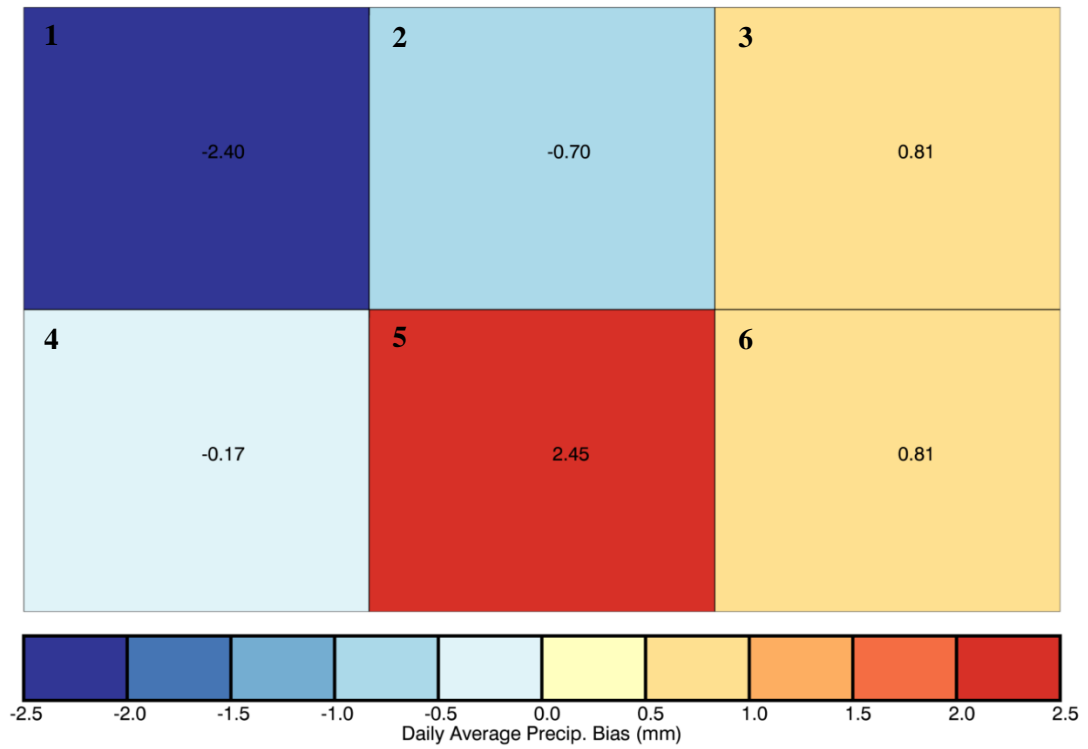


Figure 72. Daily average precipitation bias for the NP CDF90 averaged SOM.

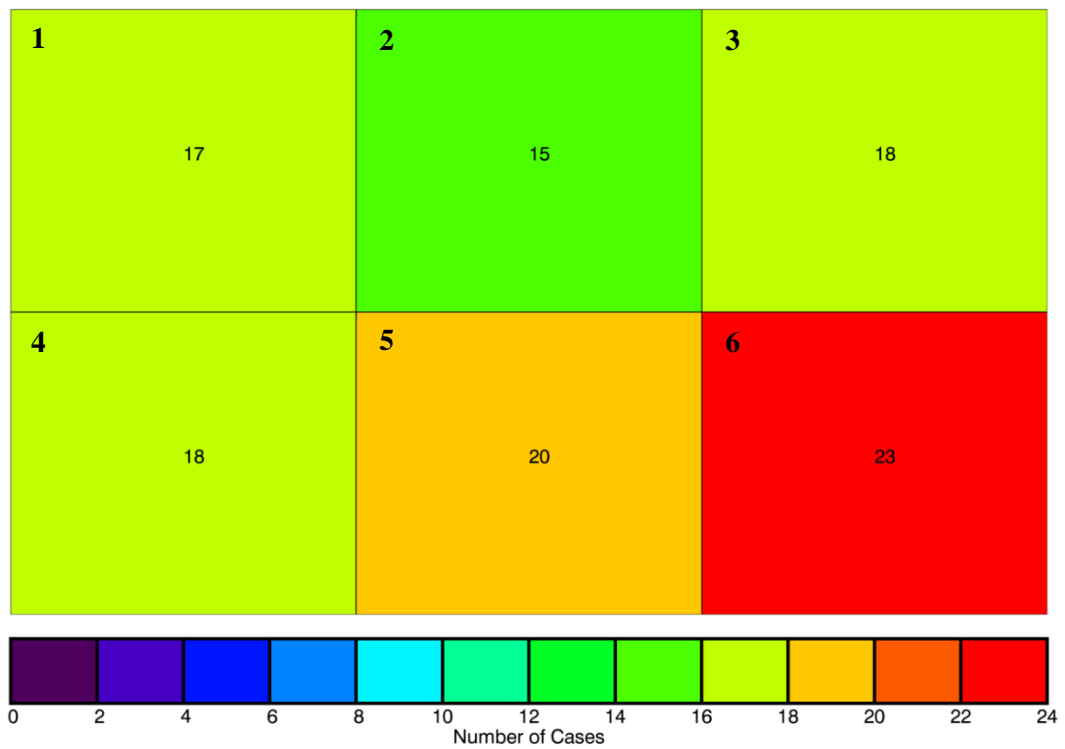


Figure 73. Number of cases in each class of the NP CDF50 averaged SOM.

Table 7. Z-Test statistic results for the NP CDF50 SOM. Statistically significant results in **bold**.

Comparison Class	1	2	3	4	5	6
1	0	.852	2.124	1.190	2.855	2.108
2	-.852	0	.843	.255	1.610	.836
3	-2.124	-.843	0	-.585	1.113	-.005
4	-1.190	-.255	.585	0	1.419	.578
5	-2.855	-1.610	-1.113	-1.419	0	-1.112
6	2.108	-.836	.005	-.578	1.112	0

Discussion

For the NP, precipitation occurs for classes with surface cyclones with flow aloft out of the southwest or weaker surface patterns with flow out of the northwest aloft. Overall the forcing aloft is stronger than was previously shown for SGP, due to the fact that this stronger forcing is necessary to carry moisture far enough into the region to have convection occur. Variability in the upper level flow impacts whether the model will over predict or under predict precipitation when surface patterns are similar. The model over predicts precipitation amounts for warm front patterns and under predicts precipitation for the weaker surface patterns with northwest flow aloft. These patterns with flow out of the northwest and weaker surface features are common patterns for MCSs. These cases also occur most frequently in July and August (Figure 74), when MCS activity is more frequent (Fritsch et al. 1986). The timing of 6-hr precipitation totals supports this hypothesis as these patterns experience their maximum 6-hour precipitation from 0600-1100 UTC (Figure 75).

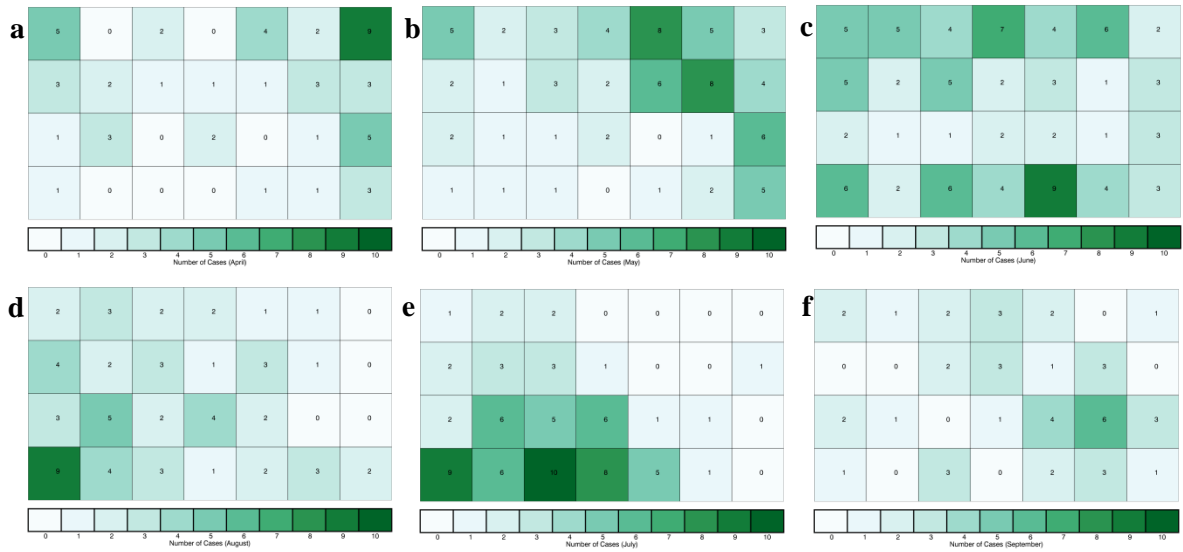


Figure 74. Distribution of NP CDF90 cases that occur for each pattern by month. April, (a), May (b), June (c), July (d), August (e), September (f).

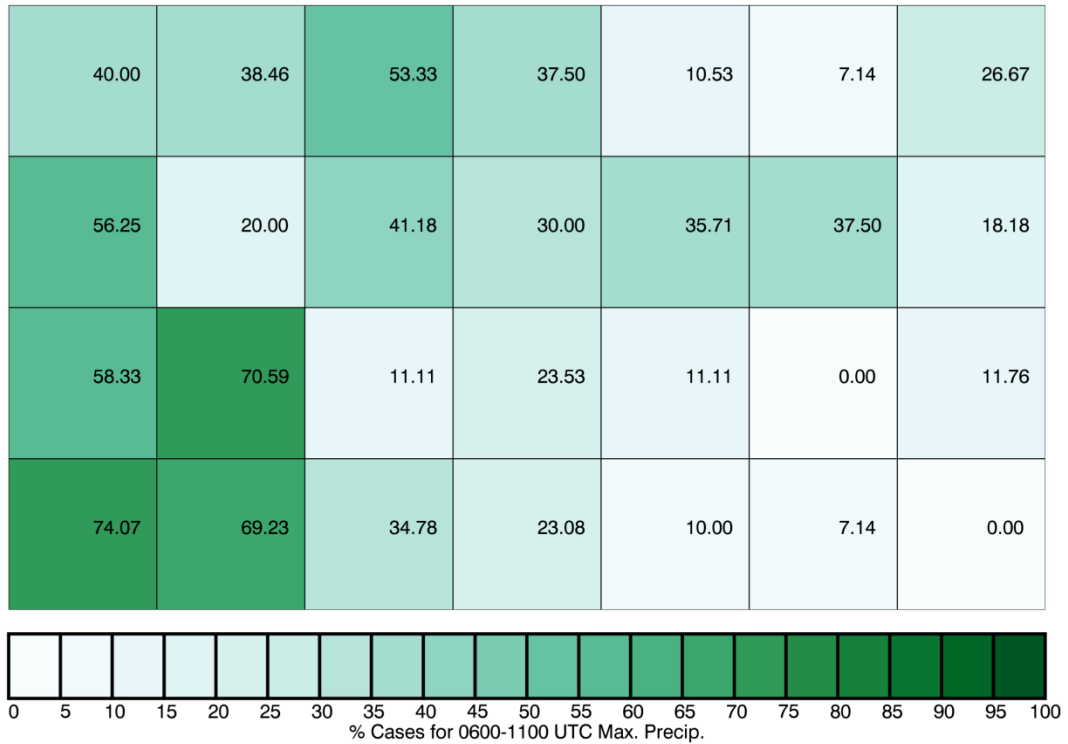


Figure 75. NP CDF90 percent of cases in each class with maximum six-hour accumulated precipitation occurring for 0600-1100 UTC.

Case Study

To test the usefulness of these results in an applied setting, a case study was conducted. Observed data at the surface and upper levels were taken at 0000 UTC on 13 June 2017 and 24-hour accumulated precipitation from Stage IV and NSSL WRF were obtained from NCEP's Environmental Modeling Center (EMC) precipitation verification page. Synoptic patterns from SPC (Figure 76) indicate that the patterns most closely resemble class 1 from the NP CDF90 averaged SOM (Figures 58 and 59). Figure 60 indicated that the model should slightly over predict precipitation amounts for this regime; however, the precipitation verification (Figure 77) indicates that precipitation amounts were under estimated. It is important to note that these patterns did exhibit a negative bias in the full NP CDF90 SOM (Figure 54) but these features were smoothed out in the averaging process, which was discussed in the NP CDF90 statistical analysis section.

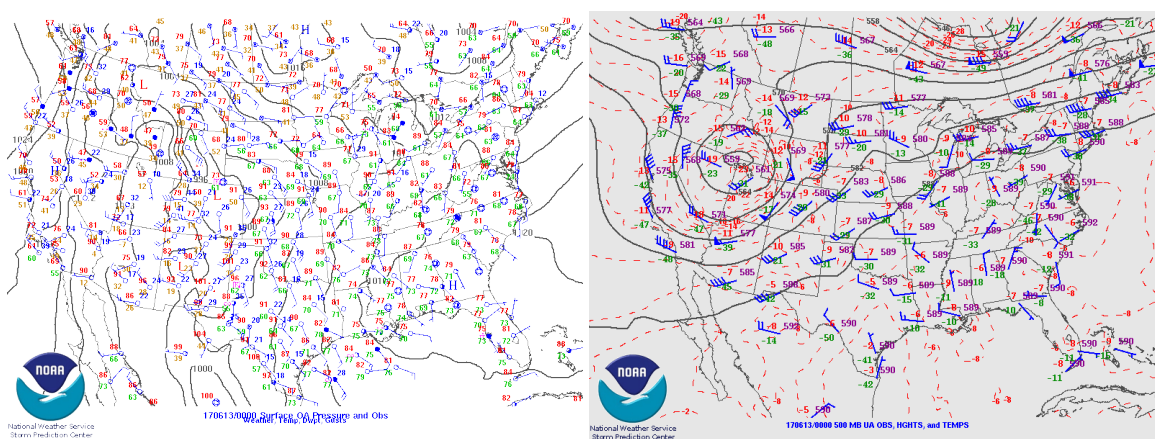


Figure 76. Surface (left) and 500 hPa (right) analysis for the 13 June 2017 case study.

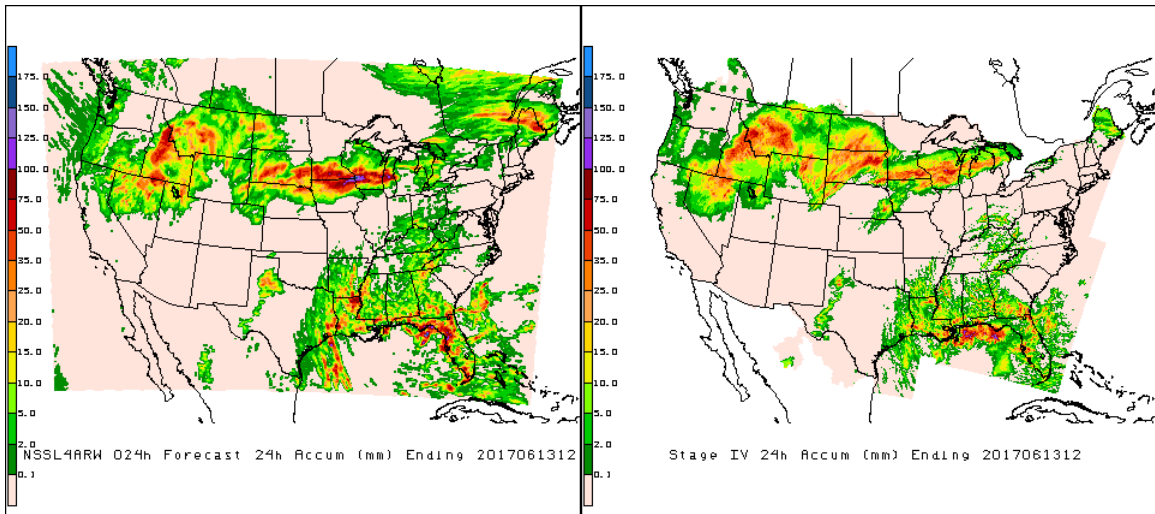


Figure 77. Accumulated precipitation amounts for NSSL WRF (left) and Stage IV (right) for the 13 June 2017 case study.

CHAPTER IV

CONCLUSIONS

This study evaluated the skill of the NSSL-WRF model at simulating precipitation amounts for varying synoptic regimes as an alternative to traditional verification methods. Analysis utilized NARR data for constructing synoptic patterns and Stage-IV data for observational precipitation. The SOM technique allowed for objective grouping of selected cases for analysis by regime. SOMs were trained using NARR data over the region of interest and iterated over to converge to a final solution where a continuum of synoptic regimes was represented. Classifications within the SOM were then grouped together for statistical analysis to determine the significance of the results using a z-test.

In the SGP region, precipitation biases are largely dependent on forcing and timing. Warm front cases produce a positive bias and precipitation occurs throughout the day. Cases with the most negative bias are those with strong cyclones and upper-level southwest flow. These are the cases where convection begins 0000-0500 UTC, likely as isolated convection, and continues through 0600-1100 UTC, likely growing upscale into and MCS. Negative bias is also prevalent for the cases of propagating MCSs which exhibit northwesterly flow aloft and are typically nocturnal for the region. Higher precipitation events yielded overwhelmingly negative bias across patterns with greater bias amount for these cases. It was shown that not only does the model struggle more

with the higher precipitation events, but these cases also occur more often overnight when the model struggles to produce enough precipitation.

In the NP region, the model predominately has a high bias of precipitation. For this region, precipitation is driven by more strongly forced events as this is essential in transporting enough moisture into the region to support significant precipitation. Small variations in upper level patterns can yield significant differences in model bias over the region. While several classes exhibit weaker surface patterns that are generally the same, variations in the upper level flow lead to a negative precipitation bias for cases with upper level flow out of the southwest or northwest, but positive bias for cases with more zonal flow. The negative trend in nocturnal precipitation is still apparent, though less common in the region. The higher precipitation cases show higher magnitude in bias, both positive and negative, indicating that while the model struggle to simulate precipitation amounts for high impact precipitation events. When those events happen to be nocturnal, precipitation is largely underestimated and when events occur during the daytime, which is more often the case in the NP, the model over predicts precipitation amounts.

Overall, the model displays diurnal tendencies in both regions with an over prediction of daytime precipitation and an under prediction of nocturnal precipitation. Daytime precipitation is largely due to patterns with high RH and surface warm fronts with upper-level troughs. Nocturnal precipitation is driven either by patterns with strong forcing at the surface and aloft in SGP or by patterns with weaker surface forcing and northwest flow aloft in both regions. Meteorologically, the latter are believed to be MCSs that originated well away from the domain, then propagated over the region during these

late hours, a trend seen to be under predicted in Goines and Kennedy (2017), while the former are likely cases of isolated convection which grew upscale as the event progressed. This hypothesis is supported by research that has identified that not only are MCSs largely nocturnal (Maddox 1980), but that they are also responsible for the nocturnal precipitation maxima seen across the plains during the warm season (Carbone and Tuttle 2008). The problems with the model under predicting for the MCS patterns is of particular concern because this type of convection is responsible for a large portion of warm season convection, particularly later in the season in June – August (Fritsch et al. 1986). In both regions the model struggles more with higher precipitation events. While the model struggles to produce enough precipitation overnight in both regions, these nocturnal events are more common in the SGP region. Results for the study determined that the patterns with the greatest bias amounts were significantly different from other classes after statistical testing.

Future Work

While model biases have been identified, questions still remain regarding the reasons for these biases. Is larger scale precipitation displacement occurring in the model? Is the model improperly simulating convective mode? To address this, this work can be expanded to other areas to further examine model performance by synoptic regime, such as an object based analysis using Method for Object-Based Diagnostic Evaluation-Time Domain (MODE-TD). Such work could identify whether the model correctly predicts rainfall area, regions of convective cores and stratiform precipitation, or number of precipitation objects. This analysis would be useful in diagnosing how the model performs in simulating the appropriate convective mode for synoptic patterns. An

additional area of interest is applying this method for different microphysical schemes to determine whether these results are consistent or vary. Both will likely utilize simulations run for another facet of the Research to Operations grant, which funded this study.

REFERENCES

- Baldwin, M. E., S. Lakshmiarahan, and J. S. Kain, 2001: Verification of mesoscale features in NWP models. Preprints, *Ninth Conf. on Mesoscale Processes*, Fort Lauderdale, FL, Amer. Meteor. Soc., 255-258.
- Becker, E. J., E. H. Berbery, and R. W. Higgins, 2009: Understanding the characteristics of daily precipitation over the United States using the North American Regional Reanalysis. *J. Climate*, **22**, 6268–6286, doi:10.1175/2009JCLI2838.1.
- Breidenbach, J. P., D.-J. Seo, and R. A. Fulton, 1998: Stage II and III post processing of NEXRAD precipitation estimates in the modernized Weather Service. Preprints, *14th Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, Phoenix, AZ, Amer. Meteor. Soc., 263–266. [Available online at: <http://www.nws.noaa.gov/ohd/hrl/papers/ams/ams98-1.htm>]
- Bukovsky, M. S., and D. J. Karoly, 2007: A brief evaluation of precipitation from the North American Regional Reanalysis. *J. Hydrometeor.*, **8**, 837–846, doi:10.1175/JHM595.1.
- Carbone, R. E., and J. D. Tuttle, 2008: Rainfall occurrence in the U.S. warm season: The diurnal cycle. *J. Climate*, **21**, 4132–4146, doi:10.1175/2008JCLI2275.1.

- Carbone, R. E., J. D. Tuttle, D. A. Ahijevych, and S. B. Trier, 2002: Inferences of predictability associated with warm season precipitation episodes. *J. Atmos. Sci.*, **59**(13), 2033–2056, doi:10.1175/1520-0469(2002)059,2033:IOPAWW.2.0.CO;2.
- Cassano, J. J., P. Uotila, and A. Lynch, 2006: Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, part 1: Arctic. *Int. J. Climatol.*, **26**: 1027–1049, doi:10.1002/joc.1306.
- Cavazos, T., 1999. Large-scale circulation anomalies conducive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas. *J. Climate*, **12**: 1506–1523, doi:10.1175/15200442(1999)012<1506:LSCACT>2.0.CO;2.
- Clark, A. J., W. A. Gallus Jr., and T. C. Chen, 2007: Comparison of the diurnal precipitation cycle in convection-resolving and non-convection-resolving mesoscale models. *Mon. Wea. Rev.*, **135**, 3456–3473, doi:10.1175/MWR3467.1.
- Done, J., C. A. Davis, and M. Weisman, 2004: The next generation of NWP: explicit forecasts of convection using the weather research and forecasting (WRF) model. *Atmosph. Sci. Lett.*, **5**, 110–117, doi:10.1002/asl.72.
- Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.
- Fritsch, J. M., R. J. Kane, and C. R. Chelius, 1986: The contribution of mesoscale convective weather systems to the warm season precipitation in the United States.

J. Climate Appl. Meteor., **25**, 1333–1345, doi:10.1175/1520-0450(1986)025,1333:TCOMCW.2.0.CO;2.

Goines, D., and A. Kennedy, 2017: Precipitation in a multi-year database of convection-allowing WRF simulations. *Submitted to J. Geophys. Res.*

Gourley, J. J., H. Yang, Z. L. Flamig, L. Li, and W. Jiahu, 2010: Intercomparison of Rainfall Estimates from Radar, Satellite, Gauge, and Combinations for a Season of Record Rainfall. *J. Appl. Meteor. Climatol.*, **49**(3), 437-452, doi:10.1175/2009JAMC2302.1.

Hahn, R. S., and C. F. Mass, 2009: The impact of positive-definite moisture advection and low-level moisture flux bias over orography. *Mon. Wea. Rev.*, **137**, 3055–3071, doi:10.1175/2009MWR2873.1.

Hewitson, B. C., and R. G. Crane, 2002: Self-organizing maps: Applications to synoptic climatology. *Climate Res.*, **22**, 13–26, doi:10.3354/cr022013.

Hong, S-Y., and J-O. J. Lim, 2006: The WRF single-moment microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.

Janjic, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945, doi:10.1175/1520-0493(1994)122,0927:TSMECM.2.0.CO;2.

Kain, J. S., S. J. Weiss, J. J. Levit, M. E. Baldwin, and D. R. Bright, 2006: Examination of convection-allowing configurations of the WRF model for the prediction of

- severe convective weather: The SPC/NSSL spring program 2004. *Wea. Forecasting*, **21**, 167–181, doi:10.1175/WAF906.1.
- Kennedy, A. D., X. Dong, and B. Xi, 2016: Cloud fraction at the ARM SGP site: reducing uncertainty with self-organizing maps. *Theor. Appl. Climatol.*, **124**, 43–54., doi:10.1007/s00704-015-1384-3.
- Kennedy, A., X. Dong, B. Xi, S. Xie, Y. Zhang, and J. Chen, 2011: A comparison of MERRA and NARR Reanalyses with the DOE ARM SGP data. *J. Climate*, **24**, 4541–4557, doi:10.1175/2011JCLI3978.1.
- Kohonen T., J. Hynninen, J. Kangas, and J. Laaksonen, 1996: SOM_PAK: the self-organizing map program package. Espoo: Helsinki University of Technology, Laboratory of Computer and Information Science.
- Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications. Preprints, *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., CD-ROM, 1.2.
- Maddox, R. A., 1980: Mesoscale convective complexes. *Bull. Amer. Meteor. Soc.*, **61**, 1374–1387, doi:10.1175/1520-0477(1980)061<1374:MCC.2.0.CO;2.
- Marchand, R., N. Beagley, and T. P. Ackerman, 2009: Evaluation of hydrometeor occurrence profiles in the multiscale modeling framework climate model using atmospheric classification. *J. Climate*, **22**, 4557–4573, doi:10.1175/2009JCLI2638.1.

- Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430, doi:10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2.
- Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, doi:10.1175/BAMS-87-3-343.
- Mitchell, K., 2005: The community Noah land-surface model. User's Guide Public Release Version 2.7.1, Environmental Modeling Center, NOAA/NCEP, 26 pp.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16663-16682, doi:10.1029/97JD0023.
- Parker, M. D., and R. H. Johnson, 2000: Organizational modes of midlatitude mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 3413–3436, doi:10.1175/1520-0493(2001)129<3413:OMOMMC>2.0.CO;2.
- Prat, O. P. and B. R. Nelson, 2015: Evaluation of precipitation estimates over CONUS derived from satellite, radar, and rain gauge data sets at daily to annual scales (2002–2012). *Hydrol. Earth Syst. Sci.*, **19**, 2037-2056, doi:10.5194/hess-19-2037-2015.
- Schwartz, B. and S. Benjamin, 2000: Verification of RUC2 Precipitation Forecasts using the NCEP Multisensor Analysis, *Fourth Symp. on Integrated Observing Systems*,

- Long Beach, CA, Amer. Meteor. Soc., 5.2. [Available online at: https://ams.confex.com/ams/annual2000/techprogram/paper_6732.htm]
- Sheridan, S. C. and C. C. Lee, 2011: The self-organizing map in synoptic climatological research. *Prog. Phys. Geogr.*, **35**(1), 109–119, doi:10.1177/0309133310397582.
- Smith, J.A., and W.F. Krajewski, 1991: Estimation of the Mean Field Bias of Radar Rainfall Estimates. *J. Appl. Meteor.*, **30**, 397-412, doi:10.1175/1520-0450(1991)030<0397:EOTMFB>2.0.CO;2.
- Theobald, A., H. McGowan, J. Speirs, and N. Callow, 2015: A synoptic classification of inflow-generating precipitation in the Snowy Mountains, Australia. *J. Appl. Meteorol. Climatol.*, doi:10.1175/JAMC-D-14-0278.1.
- Tselioudis, G., and C. Jakob, 2002: Evaluation of midlatitude cloud properties in a weather and a climate model: Dependence on dynamic regime and spatial resolution. *J. Geophys. Res.*, **107**(D24), 4781, doi:10.1029/2002JD002259.
- Weisman, M. L., W. C. Skamarock, J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.*, **125**, 527–548, doi:10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2
- Westcott, N. E., H. V. Knapp, and S. D. Hilberg, 2008: Comparison of gauge and multi-sensor precipitation estimates over a range of spatial and temporal scales in the Midwestern United States. *J. Hydrol.*, **351**, 1–12, doi:10.1016/j.jhydrol.2007.10.057.

Wu, W., D. Kitzmiller, and S. Wu, 2012: Evaluation of radar precipitation estimates from the National Mosaic and Multisensor Quantitative Precipitation Estimation System and the WSR-88D Precipitation Processing System over the conterminous United States. *J. Hydrometeor.*, **13**, 1080–1093, doi:10.1175/JHM-D-11-064.1.