

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A PERFECT PROG APPROACH TO FORECASTING DRY
THUNDERSTORMS OVER THE CONUS AND ALASKA

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN METEOROLOGY

By

LINDSEY M. RICHARDSON

Norman, Oklahoma

2013

A PERFECT PROG APPROACH TO FORECASTING DRY
THUNDERSTORMS OVER THE CONUS AND ALASKA

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY

Dr. Phillip D. Bothwell, Co-chair

Dr. Michael B. Richman, Co-chair

Dr. William H. Beasley

Acknowledgments

This work was graciously funded by the Joint Fire Sciences Program under JFSP Project 07-2-1-42. I sincerely thank my advisor, Dr. Phillip Bothwell, for convincing me to work on this project and for his dedication and support throughout. Dr. Michael Richman has my gratitude for his help in keeping the statistics straight and for his support during this project. I thank Dr. William Beasley for his help on making this project and paper official to a higher degree.

This project would not have been possible without the technical resources and IT support from the Storm Prediction Center and the National Severe Storms Lab. Their assistance and allowance of computer space helped ensure adequate data collection for processing, and this project would be very different without such benefits. I would also not have been able to work on this project without the administrative support from the Cooperative Institute for Mesoscale Meteorological Studies.

I owe many thanks to my Family, Dwayne Richardson, Teresa Richardson, and Todd Richardson for letting me dream big and for all of their continued support throughout my life. Mariah Gumfory and Autumn Losey have my thanks for healthy insanity, and I would not have achieved this opportunity without the continued support from those at the Radar Operations Center Application Branch. Thank you, to all of my mentors and friends, for helping prove that we can persevere and reach our dreams.

Table of Contents

List of Tables	vii
List of Figures	viii
Chapter 1: A Legacy of Lightning Prediction	1
1.1: The Lightning Wildfire Problem	1
1.2: Early Lightning Prediction Systems	3
1.3: Alaska Development	5
1.4: Dry Thunderstorm Focus	8
Chapter 2: Development Data	12
2.1: Definition of a Development Set	12
2.2: NARR Data	13
2.1.1: Predictor Variables	13
2.1.2: Precipitation Data	16
2.3: Lightning Data	18
2.2.1: CONUS Lightning	18
2.2.2: Alaska Lightning	20
2.4: Predictand Categories	21
2.5: Climatology	24
2.5.1: CONUS Climatology	25
2.5.2: Alaska Climatology	38
Chapter 3: Perfect Prognosis & Principal Component Analysis	48
3.1: Perfect Prog Technique	48
3.2: Introduction to PCA	49
3.3: Describing or Naming PCs	55
3.4: Building Equations and Forecasts	63
Chapter 4: Verification & Case Studies	66
4.1: Contingency Statistics	66
4.2: Performance Diagrams	78
4.3: Case Studies	88
4.3.1: CONUS Bonita Complex Fires	88
4.3.2: CONUS Mustang Complex Fire	92
4.3.3: AK Uvgoon Creek # 1	95
4.3.4: AK 120708/0000	100

Chapter 5: Summary and Future Work	103
5.1: Summary and Discussion	103
5.2: Future Work	106
References	110
Appendix A: Acronyms and Predictors	120
Appendix B: Further Climatology Examples	125
Appendix C: Further PCA Explanation and Examples	158

List of Tables

Table 2.1: List of Predictands	22
Table 4.1: Category Bins	67
Table 4.2: Contingency Table Format	67
Table 4.3: Contingency Metrics	68
Table 4.4: Forecast Hour/Day Notation	69
Table A1: Acronyms	120
Table A2: List of Predictors	121
Table A3: List of Predictors Continued	122
Table A4: List of Predictors Continued	123
Table A5: List of Predictors Continued	124
Table C1: CONUS PC Groups May-Jul	162
Table C2: CONUS PC Groups Aug-Sep	163
Table C3: AK PC Groups May-Jul	164
Table C4: AK PC Groups Aug-Sep	165

List of Figures

Figure 1.1: CONUS Vegetation Regimes	3
Figure 1.2: AK vs. CONUS size	5
Figure 1.3: AK Vegetation Regimes	6
Figure 1.4: Historic AK Fires	6
Figure 1.5: 40-km CONUS grid	11
Figure 1.6: 10-km AK grid	11
Figure 2.1: NLDN Sensors	19
Figure 2.2: NLDN Detection Efficiency	19
Figure 2.3: ALDN Sensors	20
Figure 2.4: AVETOTF vs. AVEFONE	26
Figure 2.5: AVEPTOT vs. AVEPHDTH	27
Figure 2.6: AVEFONE vs. AVEDRYTH1	28
Figure 2.7: AVEFONE vs. AVEDRYTH2	28
Figure 2.8: CONUS AVEDRYTH1 vs. AVEDRYTH2	29
Figure 2.9: CONUS AVEDRYTH1 early May-early Jun	30
Figure 2.10: CONUS AVEDRYTH1 mid Jun-mid Jul	31
Figure 2.11: CONUS AVEDRYTH1 late Jul-late Aug	32
Figure 2.12: CONUS AVEDRYTH1 Sep	33
Figure 2.13: CONUS AVEDRYTH2 early May-early Jun	34
Figure 2.14: CONUS AVEDRYTH2 mid Jun-mid Jul	35
Figure 2.15: CONUS AVEDRYTH2 late Jul-late Aug	36
Figure 2.16: CONUS AVEDRYTH2 Sep	37
Figure 2.17: AK AVEDRYTH1 vs. AVEDRYTH2	39
Figure 2.18: AK AVEDRYTH1 early May-early Jun	40
Figure 2.19: AK AVEDRYTH1 mid Jun-mid Jul	41
Figure 2.20: AK AVEDRYTH1 late Jul-late Aug	42
Figure 2.21: AK AVEDRYTH1 Sep	43
Figure 2.22: AK AVEDRYTH2 early May-early Jun	44
Figure 2.23: AK AVEDRYTH2 mid Jun-mid Jul	45
Figure 2.24: AK AVEDRYTH2 late Jul-late Aug	46
Figure 2.25: AK AVEDRYTH2 Sep	47
Figure 3.1: Scree plot example	52
Figure 3.2: CONUS Jun CC Chart	54
Figure 3.3: AK Jun CC Chart	55
Figure 3.4: Sample of Rotated Loadings	56

Figure 3.5: Commonly seen Processes extracted by PCA	57
Figure 3.6: CONUS Jun RPC2 Scores	58
Figure 3.7: CONUS Jun RPC11 Scores	59
Figure 3.8: AK Jun RPC1 Scores	60
Figure 3.9: CONUS Jun RPC3 Scores	61
Figure 3.10: AK Jun RPC9 Scores	62
Figure 4.1: East and West domains	70
Figure 4.2: CONUS POD and POFD	71
Figure 4.3: CONUS FAR and FOH	72
Figure 4.4: AK POD, POFD, FAR, and FOH	73
Figure 4.5: CONUS HSS and HKSS	74
Figure 4.6: AK HSS and HKSS	75
Figure 4.7: CONUS EDI	76
Figure 4.8: AK EDI	76
Figure 4.9: CONUS BS and BSS	77
Figure 4.10: AK BS and BSS	78
Figure 4.11: CONUS Attributes Diagrams	79
Figure 4.12: AK Attributes Diagrams	81
Figure 4.13: CONUS and AK Attributes Bin Highlight	82
Figure 4.14: CONUS Performance Diagrams	83
Figure 4.15: East and West Performance Diagrams	84
Figure 4.16: AK Attributes Diagrams	85
Figure 4.17: CONUS ROC Diagrams	86
Figure 4.18: AK ROC Diagrams	87
Figure 4.19: Bonita Complex Location	88
Figure 4.20: Bonita Day 7-Day 4 DRYTH1 Forecasts	90
Figure 4.21: Bonita Day 3-Day 1 DRYTH1 Forecasts	90
Figure 4.22: Bonita Forecasts vs. Climatology	91
Figure 4.23: Mustang Complex Location	92
Figure 4.24: Mustang Day 7-Day 4 Forecasts	93
Figure 4.25: Mustang Day 3-Day 1 Forecasts	93
Figure 4.26: Mustang Forecast vs. Climatology	94
Figure 4.27: Uvgoon 00 and 06 UTC Forecasts	96
Figure 4.28: Uvgoon 12 and 18 UTC Forecasts	97
Figure 4.29: Uvgoon Event Forecasts and Climatology	99
Figure 4.30: AK 120708 Prior Forecasts	101
Figure 4.31: AK 120708 Day 7 and Day 1 Comparison	102
Figure B1: CONUS AVETOTF early May-early Jun	126

Figure B2: CONUS AVETOTF mid Jun-mid Jul	127
Figure B3: CONUS AVETOTF late Jul-late Aug	128
Figure B4: CONUS AVETOTF Sep	129
Figure B5: CONUS AVEFONE early May-early Jun	130
Figure B6: CONUS AVEFONE mid Jun-mid Jul	131
Figure B7: CONUS AVEFONE late Jul-late Aug	132
Figure B8: CONUS AVEFONE Sep	133
Figure B9: CONUS AVEPTOT early May-early Jun	134
Figure B10: CONUS AVEPTOT mid Jun-mid Jul	135
Figure B11: CONUS AVEPTOT late Jul-late Aug	136
Figure B12: CONUS AVEPTOT Sep	137
Figure B13: CONUS AVEPHDTH early May-early Jun	138
Figure B14: CONUS AVEPHDTH mid Jun-mid Jul	139
Figure B15: CONUS AVEPHDTH late Jul-late Aug	140
Figure B16: CONUS AVEPHDTH Sep	141
Figure B17: AK AVETOTF early May-early Jun	142
Figure B18: AK AVETOTF mid Jun-mid Jul	143
Figure B19: AK AVETOTF late Jul-late Aug	144
Figure B20: AK AVETOTF Sep	145
Figure B21: AK AVEFONE early May-early Jun	146
Figure B22: AK AVEFONE mid Jun-mid Jul	147
Figure B23: AK AVEFONE late Jul-late Aug	148
Figure B24: AK AVEFONE Sep	149
Figure B25: AK AVEPTOT early May-early Jun	150
Figure B26: AK AVEPTOT mid Jun-mid Jul	151
Figure B27: AK AVEPTOT late Jun-late Jul	152
Figure B28: AK AVEPTOT Sep	153
Figure B29: AK AVEPHDTH early May-early Jun	154
Figure B30: AK AVEPHDTH mid Jun-mid Jul	155
Figure B31: AK AVEPHDTH late Jul-late Aug	156
Figure B32: AK AVEPHDTH Sep	157
Figure C1: Flattening the data	158
Figure C2: Example of orthogonal rotation	160
Figure C3: CONUS CC Charts	166
Figure C4: AK CC Charts	167
Figure C5: CONUS Scores RPC 1-4	168
Figure C6: CONUS Scores RPC 5-8	169
Figure C7: CONUS Scores RPC 9-12	170

Figure C8: AK Scores RPC 1-4	171
Figure C9: AK RPC Scores 5-8	172
Figure C10: AK RPC Scores 9-12	173
Figure C11: CONUS Pair plots	174
Figure C12: CONUS Pair plots continued	175
Figure C13: CONUS Pair plots continued	176
Figure C14: CONUS Pair plots continued	177
Figure C15: CONUS Pair plots continued	178
Figure C16: CONUS Pair plots continued	179
Figure C17: AK Pair plots	180
Figure C18: AK Pair plots continued	181
Figure C19: AK Pair plots continued	182
Figure C20: AK Pair plots continued	183
Figure C21: AK Pair plots continued	184
Figure C22: AK Pair plots continued	185

Abstract

Cloud-to-Ground (CG) lightning, in particular when it occurs in dry thunderstorms, is a major cause of wildfire initiation each year. A thunderstorm is considered dry when little to no precipitation reaches the ground. There have been several studies that have developed methods to predict CG, but little has been done to take precipitation into account. The perfect prognosis technique is applied to a 40-km Contiguous United States (CONUS) domain and 10-km Alaska (AK) domain to try to predict dry thunderstorms. Twelve years (2000-2011) of warm season (May-Sep) North American Regional Reanalysis (NARR) data are used as a development set. Principal Component Analysis (PCA) is used to identify the most meaningful atmospheric parameters associated with dry thunderstorms for the predictive equations. Predictive equations are created in S-PLUS using a generalized linear model structure. Global Forecast System (GFS) data are fed into these equations to create forecasts over the two domains during the summer of 2012. Verification scores and reliability assessments reveal that the dry thunder forecasts contain skill appropriate for their rare-event nature, but some results have noticeable impacts from GFS model biases. Case studies show how forecasts, even with the uncertainty, can benefit users up to a week in advance, especially if dry thunder forecasts are coupled with information about other atmospheric and fuel conditions.

CHAPTER 1

A Legacy of Lightning Prediction

1.1 The Lightning Wildfire Problem

From choking smoke to blazing fields and structures, wildfires impact many parts of the world, including the United States, every year. A significant fraction, roughly 13%, of wildfires in the United States are started by cloud-to-ground (CG) lightning, but wildfires started by lightning account for at least 60% of the total acreage burned (based on 2001-2011 data; National Interagency Fire Center 2013). Within the Contiguous United States (CONUS) and Alaska (AK) there are some regions in which lightning-started fires are the main wildfire concern. In particular, dry thunderstorms, thunderstorms with little to no precipitation reaching the surface (less than 0.10" or less than 2.54-mm), pose considerable threat of fire ignition.

In these regions, forecasts of dry thunderstorms are in essence forecasts of where CG lightning could start wildfires. It follows that advanced knowledge of where CG lightning is likely allows emergency managers and fire crews to make preparations for possible fires sooner than if they did not have the forecast. This means resources and crews can be sent to risk areas to contain fires at early stages – the most important stage for fire control. Fire control during initial attack also helps reduce overall costs of firefighting and possibly structure damage by decreasing the suppression time. Furthermore, lightning and/or dry thunderstorm guidance benefits fire weather forecasters at places such as the Storm Prediction Center (SPC).

Observations and probabilistic fields from model output data make up a majority of guidance products that highlight threat areas. These guidance products play critical roles in developing and sending accurate forecasts to those in the field before and during a fire event. Thus, high interest exists for lightning and dry thunderstorm guidance.

It usually takes more than lightning to start a wildfire even with high lightning flash rates. Antecedent drought status, fuel type, and moisture content, along with surrounding atmospheric variables such as wind speed, storm motion, and relative humidity play important roles in determination of whether there will be wildfires and how severe they might become. Generally, moderate winds coupled with low relative humidity and dry fuels set up prime wildfire conditions.

As weather observation networks became more consistent and widespread, meteorologists and fire crew managers looked to forecasting lightning as a possible tool to help predict lightning-started fires. Forecasters in regions with generally drier fuel types (Fig. 1.1) considered this a priority even at the early stages of observational/model forecasting. Note how most of the western region is covered with shrub/scrub and some forests - prime fire fuels. Other cropland and forests, including wetland forests, can become fire hazards as well depending on their moisture content.

With all of the dry fuels and North American Monsoon-generated lightning, forecasters in the western CONUS initiated lightning-fire prediction. For example, Beals (1923) explains that a High in the North Pacific region

and a Low around Arizona often led to thunderstorms that generated wildfires. The author introduces the suggestion that studies of thunderstorm formation in different wind regimes could be beneficial – decades before lightning detection equipment development.

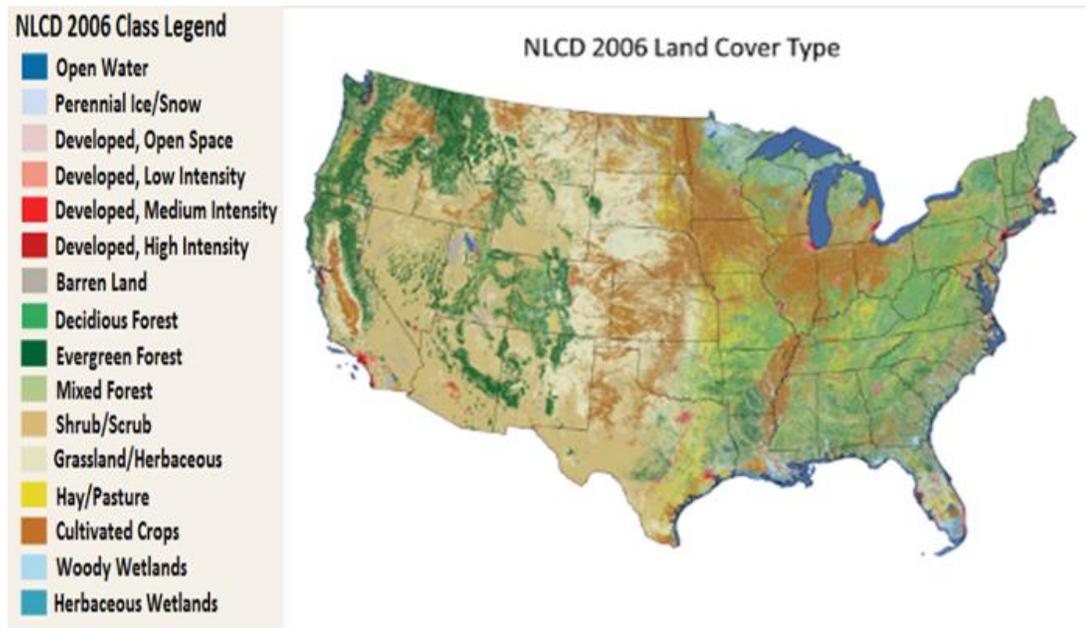


Figure 1.1: Example of vegetation type (fuels) across the CONUS from the National Land Cover Database (NLCD) – see Fry et al. (2011) for more information about the NLCD.

1.2 Early Lightning Prediction Systems

Initially, prediction of thunderstorms often focused on severe weather with lightning considered only for the definition of a thunderstorm. Reap and Foster (1979) required that “an observer must hear thunder” for a storm to count as a thunderstorm in their attempt to associate thunderstorms with certain Video Integrator and Processor (VIP) radar data to predict storms and severe weather. Regardless of the validity of their method (see Weiss et al.

1980), this example is one of many studies focused on thunderstorms without specific interest in lightning or fire weather. Lightning-specific research and prediction did not start until lightning location sensors, like Direction Finders (DFs), and computer models entered the scene.

Numerical weather prediction has advanced thunderstorm prediction considerably in the last several decades. Watson et al. (1991) studied the low-level convergence concept suggested by Byers and Rhoades (1948) and Sullivan (1963) by using Lightning Location and Protection (LLP) Company's sensors in the lightning hotbed of Florida. In particular, Watson et al. focused on CG strikes near the Kennedy Space Center. Fuelberg and Biggar (1994) took these environmental concepts further to the Florida Panhandle. Back to a more basic thunderstorm focus, they discovered that low-level humidity and wind direction can play a big role in thunderstorm development in the region. Reap (1994) shows similar results by dividing environmental patterns over Florida into different map types. Other studies came with a Model Output Statistics (MOS) predictive system (Hughes 2001; 2004) which associates certain model data to lightning observations by linear regression. This technique has built-in bias correction depending on the model used for development.

Most of the previously mentioned studies did not focus on fire weather concerns explicitly. Bothwell (2002) addressed the problem with a lightning prediction system applicable to western regions of the CONUS. His technique of perfect prognosis (or perfect prog), principal component

analysis, and logistic regression (covered more in Chapter 3) set the stage for lightning forecasts as a possible fire weather prediction tool. Coverage was expanded to the entire CONUS within a few years.

1.3 Alaska Development

As mentioned previously, the CONUS does not bear the lightning-fire burden alone. The state of Alaska experiences many weather patterns similar to the CONUS because of its large size, gulf coast, and mountainous regions. Figure 1.2 shows a comparison between AK and CONUS areas.

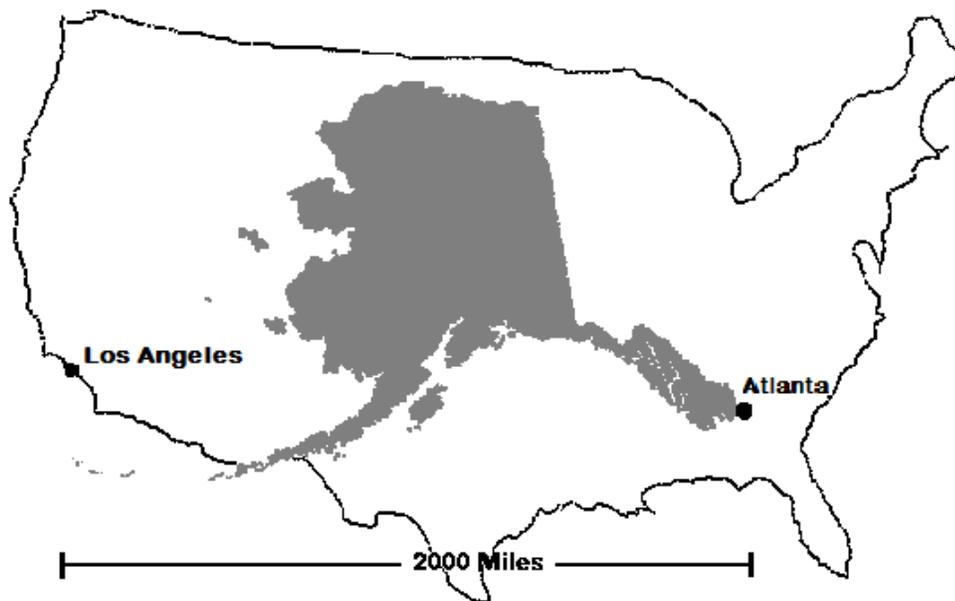


Figure 1.2: The size of AK compared to the CONUS.

Parts of Alaska, especially mainland in the boreal forests, deal with fire threats each year (Todd and Jewkes 2006). Figures 1.3 and 1.4 show this boreal forest coverage across the interior region of the state, especially along the Yukon River valley, where many large fires have occurred.

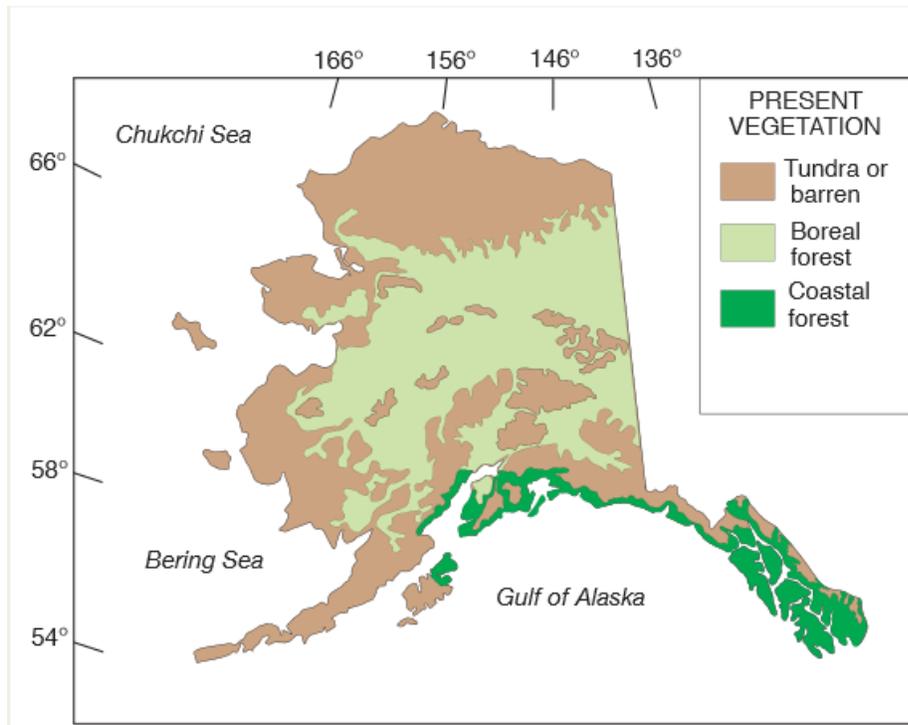


Figure 1.3: Basic overview of vegetation coverage in Alaska - Fig. 2 in Todd and Jewkes (2006).

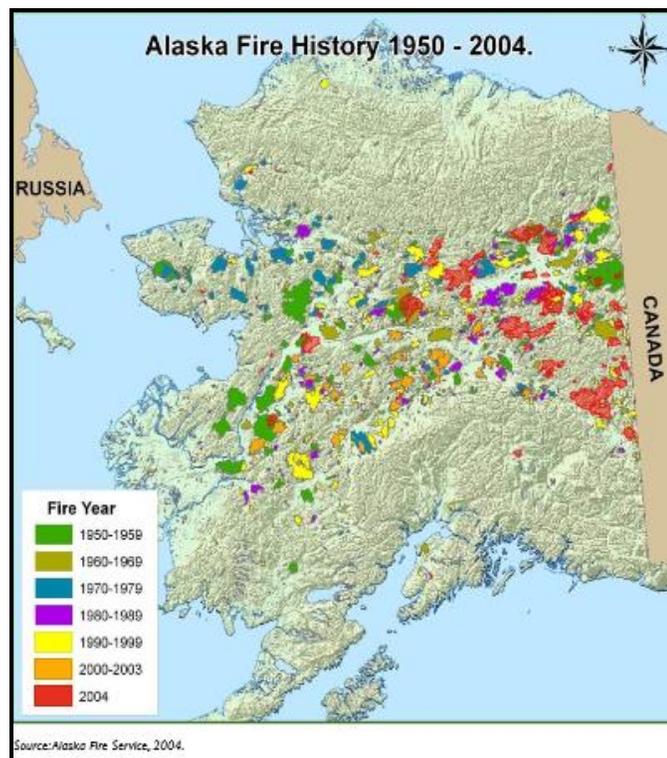


Figure 1.4: Historically large fires and the areas burned from 1950-2004 - Fig. 11 in Todd and Jewkes (2006).

Prior to Alaska building their own lightning detection network, techniques like using large fire starts as an indicator for thunderstorms (Sullivan 1963), helped start research in Alaska. The U.S. Bureau of Land Management (BLM) installed an LLP lightning location system in Alaska in 1978. Using data from that network, Reap (1991) performed a climatological study of lightning across the state. This study reaffirms the affinity for lightning in regions of elevated terrain and along the major interior rivers.

Reap's study included early work for thunderstorm prediction over Alaska using a linear regression and MOS technique. Such predictive systems persisted as the BLM slowly updated their detection network. Eventually, people began using weather models like the Global Forecast System (GFS) with a MOS approach (Shafer and Gilbert 2008). This updated method improved areal and temporal resolution and accuracy of forecasts, but more work was needed for verification of the overall benefits. Using a reasonably large grid spacing (48-km) helped with the scoring of lightning in the limited sectors, but still lacked in detail around certain mountains and river valleys. Buckey (2009) brought the scale down to 45-km and then 10-km while implementing the perfect prognosis technique in the Alaska domain. The efficacy of this technique was noticeably better, showing improvement in precision of areas with high lightning probabilities.

All methods mentioned thus far have an overforecasting bias regardless of grid size. This bias is generally accepted because lightning resembles rare events (similar to severe weather) where a miss seems

worse to a user than a false alarm. To ensure events are not missed, overforecasting is deemed acceptable by most users.

1.4 Dry Thunderstorm Focus

Though some studies address the question of precipitation amount to distinguish between wet and dry thunderstorms (e.g., Rorig and Ferguson 1999; Hughes 2004; Hostetler et al. 2005), none of the aforementioned research focuses directly on a dry thunderstorm predictor. A paper by Rorig et al. (2007) reports initial work towards that end, defining a dry thunderstorm as one or more lightning flashes and less than 0.10" (2.54-mm) of accumulated precipitation. Only those storms associated with large fire starts were considered. Their method used the Pennsylvania State University/National Center for Atmospheric Research (NCAR) mesoscale model (MM5) on a western CONUS domain. Overprediction still remained an issue along with their fairly small sample size, but results showed relatively successful forecasts for most cases presented.

Wallmann (2004) developed a Dry Thunderstorm Procedure that used atmospheric variables, such as jet streaks and a dynamic tropopause, to distinguish dry thunderstorm areas. This technique was advanced further to attempt to predict the large 21 June 2008 fire outbreak in California with a successful hind-cast prediction of lightning 30-h prior to the first flash (Wallmann et al. 2010). Instead of using an explicit dry thunderstorm predictand, Bothwell (2008) expanded his perfect prog technique to include a

Dry Thunder Potential Index (DTPI) that takes cloud base height (Above Ground Level, AGL) and sub-cloud relative humidity into account for fire weather potential. Coupling DTPI with the probabilistic lightning forecasts assisted in distinguishing drier air masses from wetter ones. Distinction between wet and dry thunderstorms generally comes from a 24-hr rainfall perspective and the common 0.10" (2.54-mm) threshold.

In contrast to the previous studies, observationally based studies of dry thunderstorms have not been neglected. Dowdy and Mills (2012) explore the combination of atmospheric environments and fuel setups that are generally associated with lightning-started fires. Their research reinforces the concept that low relative humidity has significant impact on fire potential, perhaps more so than does warm temperature.

Despite the consideration of the larger environment and total precipitation, studies with an explicit dry thunderstorm predictor are rare, and there have been no climatological studies for dry thunderstorms (with a formal definition) to date. This study proposes to explore prediction of an explicit dry thunderstorm parameter developed using the perfect prog technique established by Bothwell (2002) and used by Buckley (2009). An overview of system development and performance of their work in operations can be found in Bothwell (2010). It is hypothesized that a dry thunderstorm probabilistic field will help simplify the amount of data ingest for forecasters and fire crews preparing for fire weather. Specifically, rather than assessing all of the atmospheric variables separately, the product developed herein can

assist in pinpointing regions of main fire threat because the storm-dryness factor is inherently included.

Owing to high interest in long-term forecasts, a system that uses the GFS data is constructed to generate forecasts out to 1 week (0-180 hours). Testing occurs on two different domains: a 40-km CONUS grid (Fig. 1.5) and 10-km AK grid (Fig 1.6). Separate forecast equations are developed for each warm season month (May-Sep) for multiple predictands (see Section 3.4). These months make up the majority (entirety) of the convective season in the CONUS (AK) (Reap 1991; Hostetler et al. 2005; Buckley 2009). While it is noted that the overall environment and fuel moisture play major roles in sustaining a fire (e.g., Rorig and Ferguson 2002; Calef et al. 2008), the focus herein is mostly ignition due to combinations of atmospheric parameters for lightning and precipitation potential.

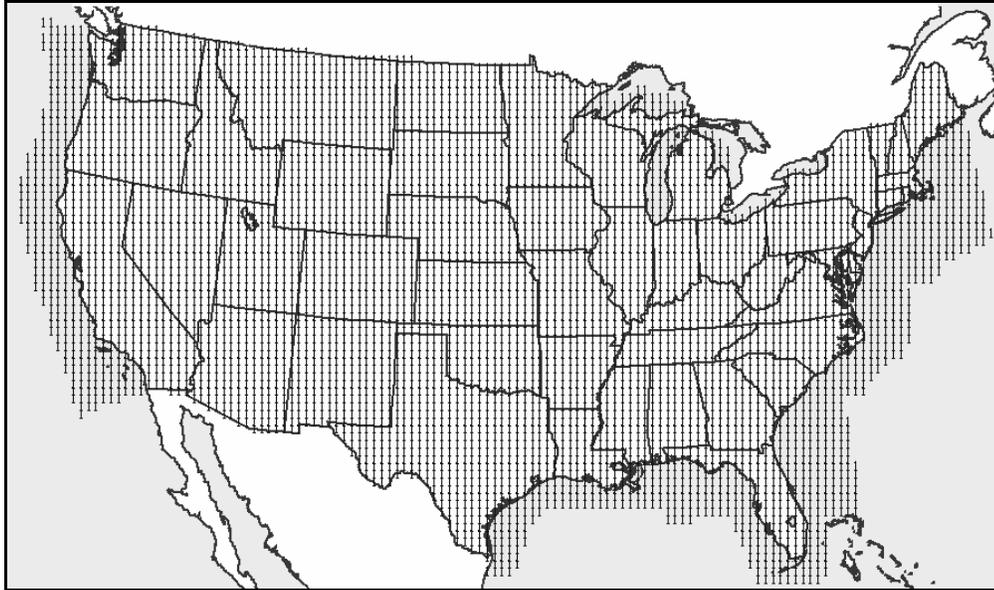


Figure 1.5: Coverage map of 40-km CONUS grid - demarcated by the symbol "1". Some coastal waters are included, but areas of Canada and Mexico are excluded. Lambert Conic Conformal (LCC) projection used mostly throughout unless specified otherwise.

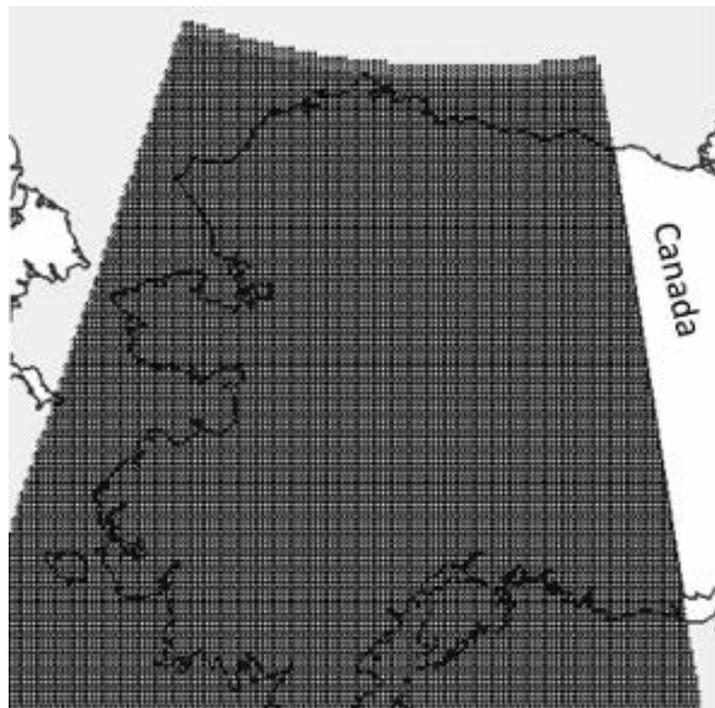


Figure 1.6: Coverage map of 10-km Alaska grid. The arc shape is the result of the North Polar Stereographic (NPS) projection – used throughout unless specified otherwise. The right edge of the image demarks the Alaska/Canada border (141° longitude).

CHAPTER 2

Development Data

2.1 Definition of a Development Set

One of the most common methods of creating a forecast comes from developing a model/set of equations that determines the probability of an event. Instead of elementary probability, complex systems such as the atmosphere work in the realm of conditional probability. Equation 1 describes elementary probability where $p(x)$ is the probability of an event x , and Σx is the set of all possible events. Equation 2 describes conditional probability in relation to Bayes' Theorem $p(x|f)$ represents the probability of an event happening given a forecast f , $p(f|x)$ is the probability of a forecast given an event, and $p(f)$ is the probability of a forecast. That is, an event may

$$p(x) = \frac{x}{\sum x} \quad (1)$$

$$p(x|f) = \frac{p(f|x)p(x)}{p(f)} \quad (2)$$

be more likely given a certain forecast. For example, an area with a 30% probability of dry thunder is more likely to see a dry thunderstorm than an area with 1% probability.

Additionally, a model needs information about events and conditional parameters in a large enough sample that describes the climatology and conditional probability domain of a particular event. This adequately-sized sample of conditional parameters, the development data set, can be used to train the model to a specific event such as dry thunderstorms. Models can

then be used to predict said event on a different assortment of data – the testing data set (see Chapter 3). Twelve years of data (2000-2011) from the warm season months (May1 – Sep 30) were used for the development set, and forecasts are generated during the 2012 warm season (testing set).

2.2 NARR Data

The National Centers for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR; Mesinger et al. 2006) data set is suitably reliable and accurate for the development set of interest. Because dry thunderstorms are conditional to many parameters in the atmosphere, a sufficient set of standard and derived parameters is desired for the development data set. This study uses information from the NARR for its spatial and temporal resolution and the fact that NARR data covers both the CONUS and AK domains. NARR data contains many variables on a 32-km grid in 45 vertical Eta levels - levels described by horizontal planes of the atmosphere.

2.2.1 Predictor Variables

Much of the development information for NARR comes from the Global Reanalysis (GR) data set which uses data from radiosondes, dropsondes, pibals, surface stations, and satellite data (Shafran et al. 2004). Temperature, moisture, wind components, radiative fluxes, and other variables are archived in NARR on a 45-layer system. To improve resolution around mid-levels, the program GDVINT from the General Meteorological

Package (GEMPAK) version 6.2.0 on a Linux server is used to vertically interpolate data to every 25-hPa between 650 to 300-hPa. Thus, all data variables can be assessed at the surface (2-m and 10-m readings) and pressure levels from 1000 to 100-hPa with a granularity of 25-hPa – a usable resolution for the spatial scale in this study. These variables are grouped into 3-hourly (3-h) time bins, which is a common forecast time for many operational models. NARR 3-h bins for each of our warm season months, for the twelve year span mentioned above, were downloaded from Research Data Archive of the Computational and Informational Systems Library at the National Center for Atmospheric Research (NCAR). GDBIINT, a bilinear interpolation program included with GEMPAK, was used to convert the 32-km NARR data to each of the domains in Figures 1.5 and 1.6.

Variables such as temperature, dewpoint, wind components, and lapse rate come directly from the NARR. It has been noted that the 2-m temperature bias is smaller compared to the Global Reanalysis (Mitchel et al. 2005), but there is a slight negative 10-m wind bias (Messinger et al. 2006). Other predictors of interest such as Most Unstable Convective Available Potential Energy (MUCAPE), divergence, and Lifted Index (LIFT) are calculated using GDDIAG in GEMPAK – see Tables A2-5 for a full list of predictors in this study. Many of these calculated variables such as advection, moisture convergence, and frontogenesis have been shown to be associated with lightning and dry thunderstorms in previous studies (e.g.,

Sullivan 1963; Reap 1991; Watson et al. 1991; Klein et al. 1996; Rorig and Ferguson 1999; Bothwell 2002; Rorig et al. 2007).

Instead of using variables or indices as stand-alone terms, it has been shown (e.g., Reap 1979; Hughes 2004) that combining them into unique predictors can strengthen predictability. These Interactive Predictors benefit forecasting systems because terms related to a similar phenomenon, say Convective Available Potential Energy (CAPE) and lightning flashes to denote thunderstorm potential, are grouped into a single, stronger term (Reap 1979; Bothwell 2002; Hughes 2004; Hughes and Trimarco 2004; Buckey 2009). That is, statistical techniques can extract these interactive predictors as some of the highest contributors to meteorological signal compared to individual parameters.

MUCAPE multiplied by Lightning Climatology (sometimes called lightning frequency) has been shown to be a skillful interactive predictor for predicting lightning flashes (Bothwell 2002; Buckey 2009). These and other terms developed by Bothwell (2002) are used in this study, and are notable for capturing much of the variance of lightning prediction, as discussed in the next chapter. The list of interactive predictors can be found in Table A5 near the end of the predictor list.

For this study, MUCAPE is calculated by finding the maximum CAPE value calculated at each 25-hPa level from 1000 to 525-hPa and the surface. One benefit to using MUCAPE includes the fact that the term is less prone to errors noted in CAPE and surface-based CAPE (Doswell and Rasmussen

1994; Rochette et al. 1999; Craven et al. 2002). MUCAPE is also often the main choice to describe environments for deep convection and/or elevated storms, which contributes highly to dry thunderstorm formation. That is, MUCAPE can distinguish information about the most unstable parcel even when said parcel is elevated.

2.1.2 Precipitation Data

Another benefit to using NARR data comes from the continuous precipitation data set. The NARR total precipitation (in units of mm converted to inches) comprises the rainfall totals in this study. While the precipitation does not come from direct assimilation, latent heating profiles are derived from multiple sources to coincide with observed precipitation closely. Sources include rain gauges, radar measurements, and analysis with the Parameter-Elevation Regressions on Independent Slopes (PRISM; see Daly et al. 2004) technique over terrain in the CONUS domain. These assimilation techniques have been shown to reduce the precipitation overforecasting bias of the Eta model (Ruane 2010a; Ruane 2010b).

A study by Bukovsky and Karoly (2007) suggest that the NARR precipitation values match the Climate Prediction Center's (CPC) observed values closer than the other reanalysis examined (i.e. the NCEP-DOE and ERA-40, refer to their paper for more information on these systems), yet it has been shown that the NARR tends to underforecast heavy precipitation events and slightly overforecast light ones (Shafran et al. 2004; Becker et al. 2009). Also, Spurious Grid-Scale Precipitation (SGSP) can create problems

in isolated extreme events in the Eta model. SGSP is most commonly identified by a single grid point containing a high amount of precipitation (West et al. 2007; Becker et al. 2009). A single point with a very high precipitation value could affect the dry thunderstorm forecasts. That is, a grid point may have an erroneously high value for precipitation that would exceed the thresholds used to distinguish dry thunderstorms in this study (see Sec. 2.3). This study does not attempt to correct problems related to SGSP in the training data due to their overall rarity.

Bi-linear interpolation was applied to the NARR precipitation grid along with the other NARR variables mentioned above. While it is noted that other budgeting techniques may conserve precipitation better, bi-linear interpolation conserves precipitation across areas well enough that it suffices for the 32 to 10-km conversion (Brill and Ling 2011, personal communication). Error values from this 32 to 10-km conversion fall $\sim\pm 0.004$ " (± 0.1 -mm) – about an order of magnitude smaller than the scale of interest (0.01"). Moving information to the larger 40-km grid is acceptable for this study as error values fall $\sim\pm 0.040$ " (± 1.0 -mm) – about the same order of magnitude as the scale of interest. Thus, the NARR atmospheric variables can be associated to the "observed" precipitation values at the same grid size. The NARR data set includes three precipitation fields: Total Precipitation, Large Scale Precipitation, and Convective Precipitation. Only the Total Precipitation field is considered in this study.

2.3 Lightning Data

Precipitation data alone will not explain dry thunderstorms, thus observed lightning data becomes a necessity. Total flash data (CG only), as opposed to strike data, is used to match the previous Bothwell (2002) and Buckey (2009) studies. CG flashes contain the most interest for fire weather because the act of striking a fuel is required for ignition. Consideration for flash data over strikes comes from the fact that there can be multiple strikes in a single flash (Valine and Krider 2002). Furthermore, most detection systems recorded and offered flash data up to 2013.

2.3.1 CONUS Lightning Data

As mentioned in the previous chapter, the two domains utilize different lightning detection networks. Vaisala, Inc.'s National Lightning Detection Network (NLDN) sensors are spatially distributed across the CONUS region (see Fig 2.1). This network of 187 lightning sensors, which consists of Improved Accuracy from Combined Technology (IMPACT) and Time of Arrival (TOA) sensors, detects CG lightning with $\geq 95\%$ efficiency over the CONUS with decreasing efficiency over adjacent coastal waters as seen in Fig. 2.2 (Cummins et al. 1998; Orville 2008).

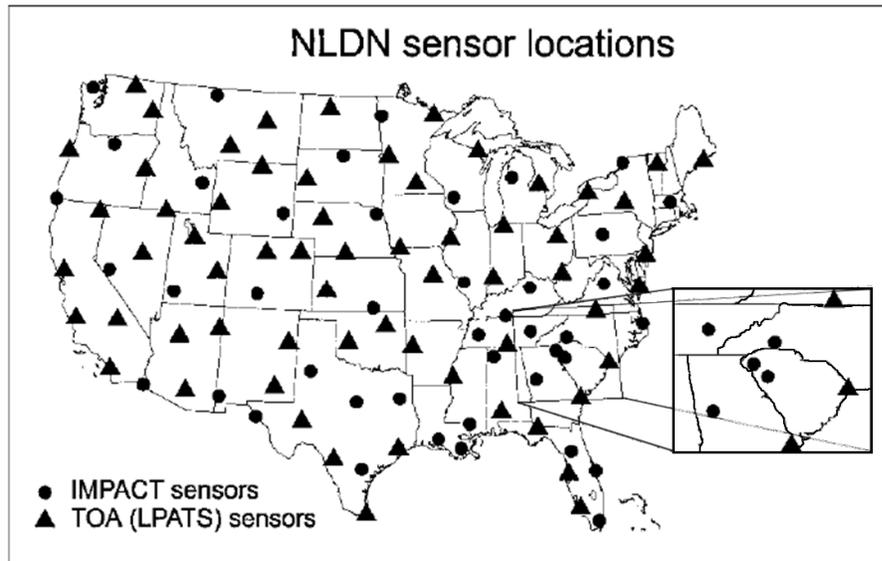


Figure 2.1: Layout of National Lightning Detection Network sensors across the CONUS - Fig. 1 in Orville and Huffines (2001).

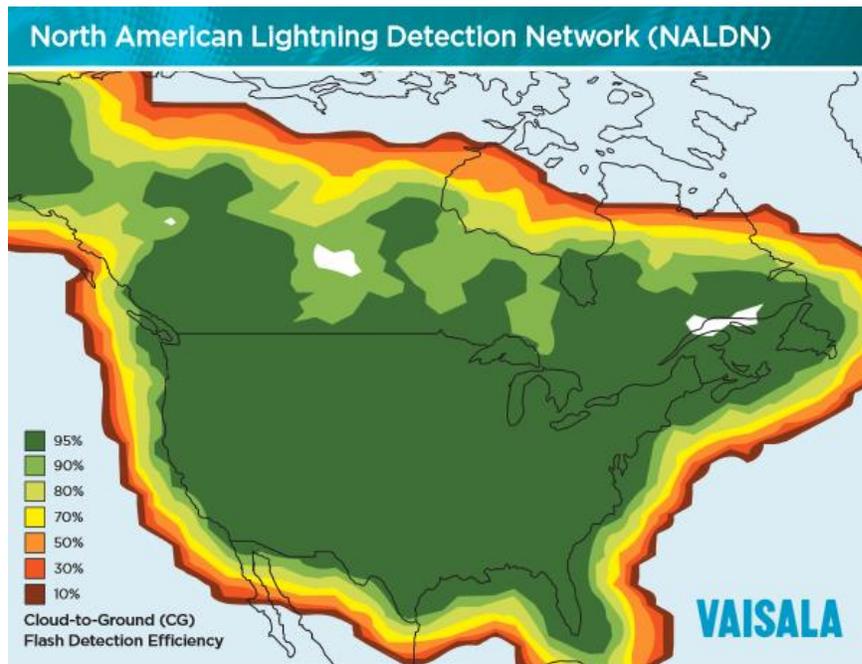


Figure 2.2: NLDN CG detection efficiency from 2011. Map available on the Vaisala website at: http://www.vaisala.com/Vaisala%20Documents/Brochures%20and%20Datasheets/0537_WCO-WEN-G-DE-Lightning%20Map%20for%20Detection%20Efficiency%20%28zoom%29.pdf

CG flash data is supplied in 3-h intervals which match with the NARR data. Consistent NLDN system operations since 1989 and previous studies ensure reliable accuracy of the network's lightning data for use as official observations. Data for this study come in the form of ASCII files gridded to a GEMPAK file by Lat/Lon coordinates with time stamps in Coordinated Universal Time (UTC). Part of the gridding process filters out flashes < 10 kA due to the likelihood of them being In-Cloud (IC) flashes (Cummins et al. 1998; Orville and Huffines 2001).

2.3.2 Alaska Lightning Data

Alaska slowly developed their separate network years later than the CONUS. The first Alaska lightning detection network consisted of only 9 DFs

from Lightning Location and Detection, Inc. (LLP) across the state (Reap 1991). These LLP sensors were used prior to the NDLN in the CONUS, so Alaska made use of this system. The AKBLM Lightning Detection Network (sometimes referred to as ALDN) evolved into 12 sensors across the mainland region (see Fig. 2.3) –

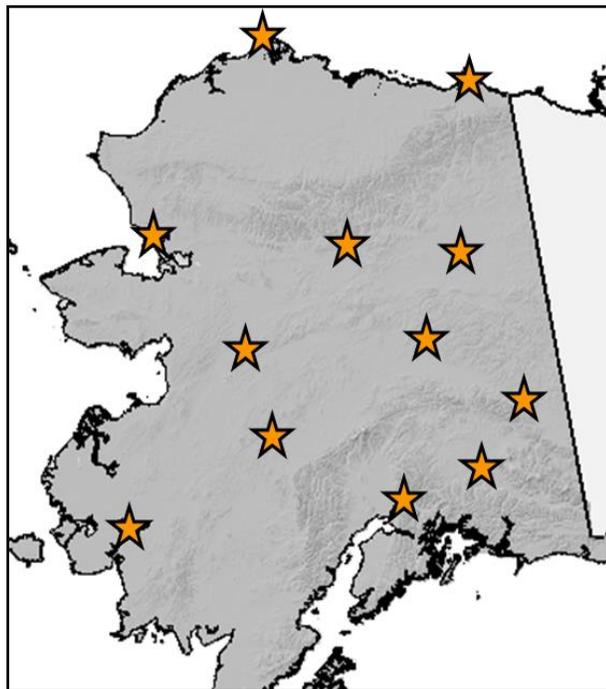


Figure 2.3: AKBLM Lightning Detection Network with stars denoting sensors. Stations in this basic configuration were in operation from 2000-2012. Map is a color-modified version of the one on their website listed in the text.

reducing the coastal bias mentioned by Reap (1991). The network records CG flashes of positive and negative strikes as part of the detection, which matches the NLDN method, and makes it suitable for observations over Alaska. A lower network detection efficiency and range over Alaska could cause some location bias at the 10-km scale though the 45-km grid used in Buckey's (2009) work showed no such issue.

These data were freely available from the AKBLM website for historical lightning which could be downloaded as shapefiles/textfiles (http://afsmaps.blm.gov/imf_customlight/imf.jsp?site=customlight at the time of writing). Data for the entirety of May-Sep from 2000-2011 were downloaded and modified for gridding purposes to replicate the NDLN style. The time stamp is in local time (Alaska Daylight Time, in the warm season). All AK lightning times were converted to UTC by adding 8 hours. The <10 kA filter used for the CONUS is not applied to the AK domain because the overall magnitude of flash intensity is lower in Alaska, and the network is slightly less reliable. Any flashes recorded east of 141° longitude (in Canada) were excluded from this study to match the domain shown in Fig. 1.6.

2.4 Predictand Categories

Instead of focusing on total amounts of lightning and precipitation, total lightning flashes and total precipitation are placed into non-exclusive categories (Table 2.1). Bothwell (2002) and Buckey's (2009) method for

lightning categories is repeated here as a continuation study, and new categories are created for precipitation. All resulting fields are binary (1 or 0; yes or no). For example, FONE denotes areas where one or more flash has occurred ($TOTF \geq 1$). Lightning categories range from low to high amounts of flash density where the higher categories represent “significant” lightning.

Table 2.1: List of Predictands based on lightning flashes and precipitation. All fields are binary and unitless except for TOTF and PTOT which are measured in total flash count and hundredths of inches respectively.

	Name	Description	Threshold
Lightning	TOTF	Total CG lightning flashes	All CG lightning
	FONE	One or more flashes	$TOTF \geq 1$
	FTHR	Three or more flashes	$TOTF \geq 3$
	FTEN	Ten or more flashes	$TOTF \geq 10$
	FTHT	Thirty or more flashes	$TOTF \geq 30$
	FHUN	One hundred or more flashes	$TOTF \geq 100$
Precipitation	PTOT	Total precipitation in inches	All rainfall
	PHDTH	Hundredth of an inch or more of precipitation	$PTOT \geq 0.01''$
	PTNTH	Tenth of an inch or more of precipitation	$PTOT \geq 0.10''$
	PQTR	Quarter of an inch or more of precipitation	$PTOT \geq 0.25''$
	PHLF	Half of an inch or more of precipitation	$PTOT \geq 0.50''$
	P3QTR	Three Quarters of an inch or more of precipitation	$PTOT \geq 0.75''$
	PONE	One inch or more of precipitation	$PTOT \geq 1.00''$
	PLTN	Less than a tenth of an inch of precipitation	$PTOT < 0.10''$
	PLQT	Less than a quarter of an inch of precipitation	$PTOT < 0.25''$
Dry Thunder	DRYTH1	One or more flashes and less than a tenth of an inch of precipitation	$FONE * PLTN$
	DRYTH2	One or more flashes and less than a quarter of an inch precipitation	$FONE * PLQT$

The precipitation data are processed similar to the lightning with a couple of special fields included for the dry thunder consideration. The choice to start at 0.01" (0.25-mm) is consistent the NWS definition of measurable precipitation. That is, all values at the Trace level are counted as not measurable and assigned a value of 0.00". Most of the categories fit accepted NWS measurement thresholds up to the 1.00" (25.40-mm) mark. Events with 3-h precipitation rates more than 1.00" are considered too wet for the threat of wildfire ignition and will not be a major focus here.

Two other rainfall-related fields test for the *lack* of heavy precipitation. One category, Precipitation Less than a Tenth of an inch (PLTN), matches the common threshold for a dry thunderstorm. Rainfall less than 0.10" (2.54-mm) is often associated with rains that either evaporate before reaching the ground or may not wet the area enough to prevent ignition. Some lightning flashes occur outside of the rain region, which can spark fires, or very dry fuels mean that rain greater than 0.10" in a storm will not wet the vegetation enough. The Precipitation Less Than a Quarter of an inch (PLQT) category is used to try and account for very dry fuels or strikes outside of the rain region. Both of these categories can include a value of 0.00".

With these fields that account for low amounts of precipitation, specific dry thunder fields can be created. The first field, Dry Thunder 1 (DRYTH1), follows the NWS definition that a dry thunderstorm consists of one or more flash and rainfall measuring less than 0.10". Dry Thunder 2 (DRYTH2) adds

in the slight extra precipitation allowance by using the PLQT field for reasons mentioned above. Because the lightning and precipitation categories are binary, simple multiplication works for creating the dry thunder fields.

2.5 Climatology

As several comprehensive lightning and precipitation climatology studies exist (e.g. Reap 1991; Orville and Huffines 2001; Hostetler et al 2005; Richman and Lamb 1985; Walsh et al. 1982; Trenberth et al 2003; Becker et al. 2009) only a small sample will be reviewed here. Instead of a seasonal or monthly climatology, a pentad (5-day) climatology is developed for each of the main predictands listed in Table 2.1. A pentad consists of information from five days, with the date being set as the middle (or 3rd) day. Pentads are created for each 3-h interval (e.g., 00-03 UTC) every five day period. This process is repeated for each of the twelve training years (2000-2011). Averaging over these 5-day groups results in the 12-year “climatology” value. Utilizing pentads has been shown to resolve variability better than daily or seasonal trends, and can be especially effective descriptors temporally for deep convection (Wang and Xu 1997).

Peak heating (insolation) occurs in the afternoon to evening hours in both domains (around 21 UTC for the CONUS and 00 UTC for Alaska), generally corresponding with maximum convection and lightning activity. For consistency, 00 UTC data (covering 00-03 UTC) will be the focus of the development and subsequently used for the climatology examples (see Sec

3.4); only examples of 00 UTC climatology data from both domains are covered here. The common time of 00 UTC across both domains is selected because it contains the most current information about the atmosphere from input such as soundings. A Gaussian weighted filter of 10 grid points has been applied to all climatology fields as this amount of filtering has been found to reduce noise (Bothwell 2002; Buckey 2009). Any values below a given magnitude (generally within 0.001 units) are considered noise in this case.

2.5.1 CONUS Climatology

Lightning climatologies have been generated in multiple studies across the CONUS domain (e.g. Reap 1994; Rorig and Ferguson 1999; Orville and Huffines 2001; Bothwell 2002; Hostetler et al. 2005); many refer to NLDN as the data source. Lightning flashes start off concentrated in the Plains in early May and move toward the Gulf Coast and Rocky Mountains by early June (Fig. B1). The North American Monsoon is responsible for this intra-seasonal increase in lightning across the Inter-Mountain west during the summer months (Rorig and Ferguson 1999; Bothwell 2002; Hostetler et al 2005). From June through July, the monsoon starts to move further west, and lightning chances decrease in a line from Oklahoma to Indiana (Fig. B2). Monsoon activity in the western states dominates much of the lightning activity in August with some enhanced activity in the southeast in southern Georgia and Alabama into Florida in mid-August (Fig. B3). Most lightning activity begins to decrease in the month of September (Fig. B4).

A climatology built off of binary data, such as one-or-more flashes (FONE), varies slightly in magnitude compared to total flashes (TOTF). This is particularly based on the fact that the climatology of TOTF (AVETOTF) is in units of flashes while the climatology of FONE (AVEFONE) is in units of percent. Certain regions are highlighted with the percentage unit more than the flashes unit. Figure 2.4 shows one example of the differences between AVETOTF and AVEFONE from early July. Notice how the Inter-Mountain west is more prominent in AVEFONE compared to AVETOTF. Also, values of AVEFONE are higher in parts of Vermont and Maine while the AVETOTF values are small or close to 0. The figure indicates the area between Vermont and Maine has a 7% chance of seeing at least 1 lightning flash, but it is likely not much more than 1 because the total flash climatology is close to or less than this value. More AVEFONE images can be found in Appendix B following the AVETOTF images.

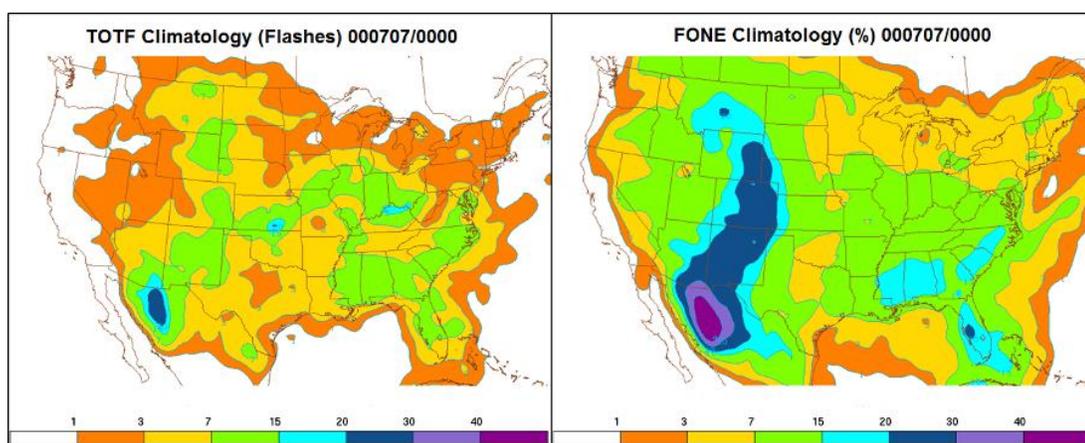


Figure 2.4: AVETOTF (left) representing average number of flashes and AVEFONE (right) representing the chance of getting one or more flashes for the pentad centered on 7 July (5-9 Jul). The timestamp format (e.g. 000707/0000) represents the year, month, day, and hour (yyymmdd/hhhh); all climatology grids are placed at a year 2000 date.

Intra-seasonal precipitation trends highlight the monsoon track during the summer, and a couple of other regions such as the Ohio Valley and Carolina Coast appear as local maxima according to the total precipitation climatology (AVEPTOT). In comparison the climatological average of (PHDTH) seems higher because it is measured in percent instead of hundredths of an inch (see Fig. 2.5 – note the difference in scales). Illustrations for each can be found in Appendix B.

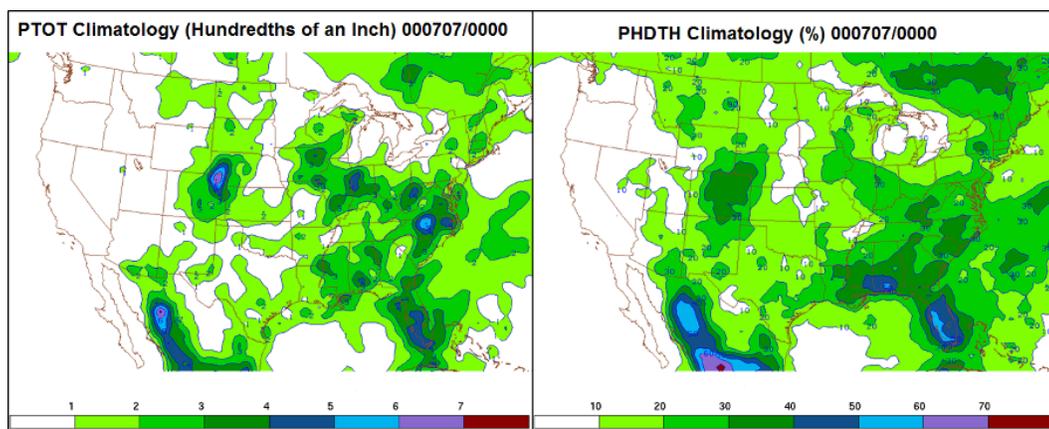


Figure 2.5: Example of different viewpoints of climatology between AVEPTOT (left) and AVEPHDTH (right).

Instead of focusing on lightning and precipitation separately, a dry thunder climatology will be more beneficial for the predictive system as part of the development data set. Figures 2.6 and 2.7 show how the dry thunder climatology fields follow closely with the lightning climatology – lightning is the ultimate identifier for a thunderstorm (wet or dry). Areas that receive

more rain from phenomena like Mesoscale Convective Systems or deep sea-breeze convection generally have lower dry thunder climatology values.

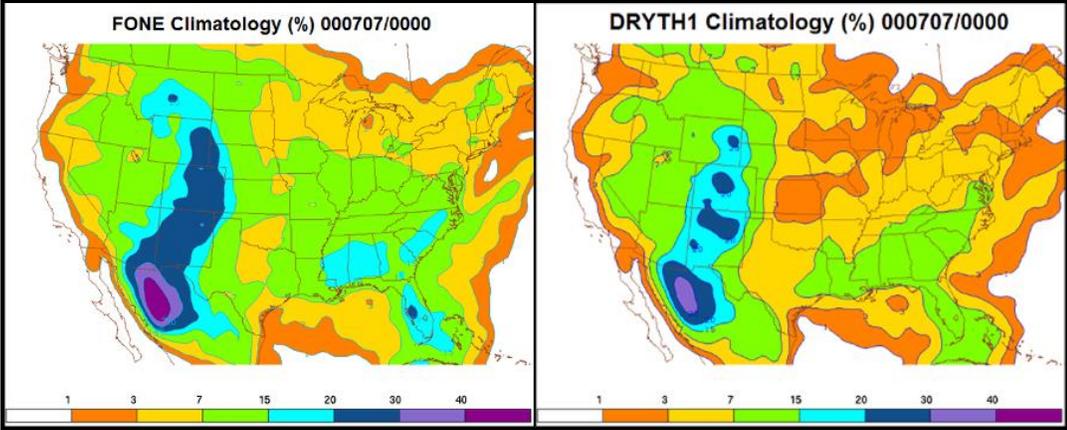


Figure 2.6: Example of FONE climatology (left) vs. DRYTH1 climatology (right) for the CONUS.

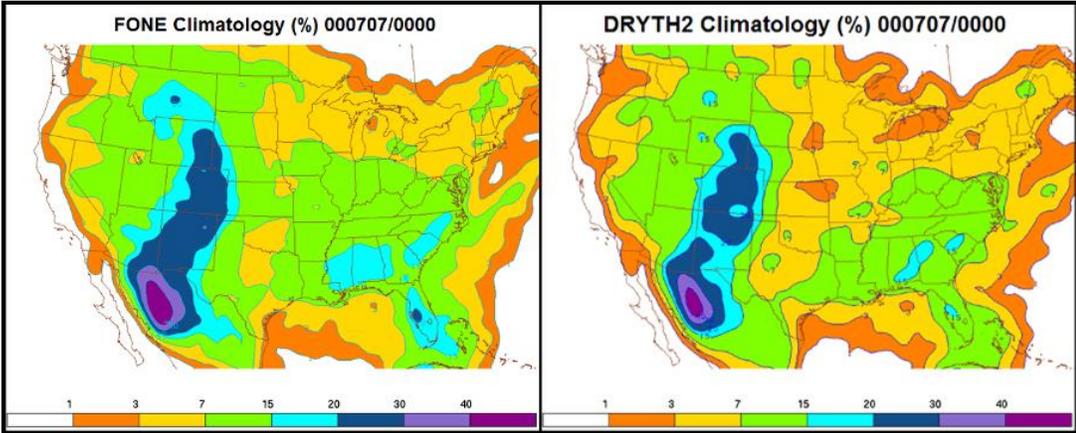


Figure 2.7: Example of FONE climatology (left) vs. DRYTH2 climatology (right) for the CONUS.

Climatology values of DRYTH2 (AVEDRYTH2) are generally higher than the climatology values of DRYTH1 (AVEDRYTH1), as more grid points receive at least a tenth but less than a quarter of an inch of rain. This could be important in places such as Florida where the threat for lightning-started fires outside of the rain shaft can be a bigger problem than not getting

enough wetting rain. Figure 2.8 displays how DRYTH2 has climatologically higher values than DRYTH1 in southeastern states (e.g., Georgia and Florida) and in western states (e.g., New Mexico and Colorado). Figures 2.9-2.16 show more examples of AVEDRYTH1 and AVEDRYTH2 from May-Sep.

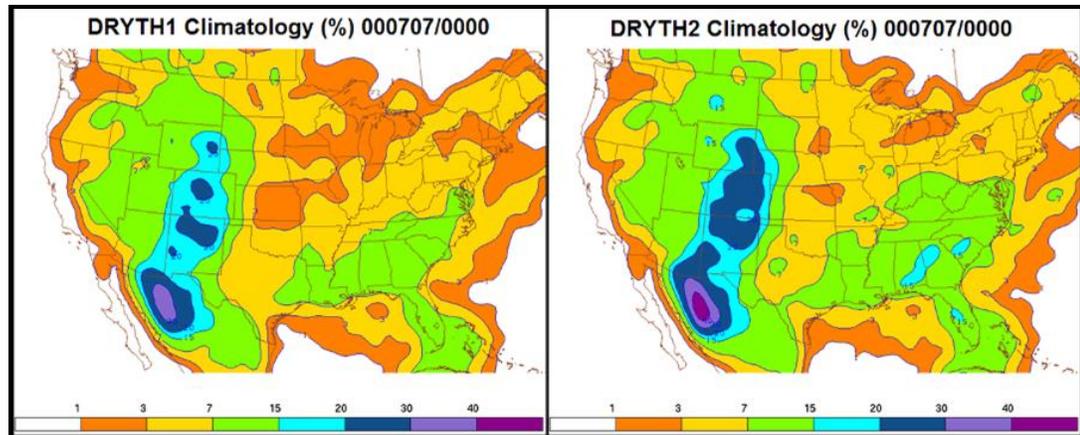
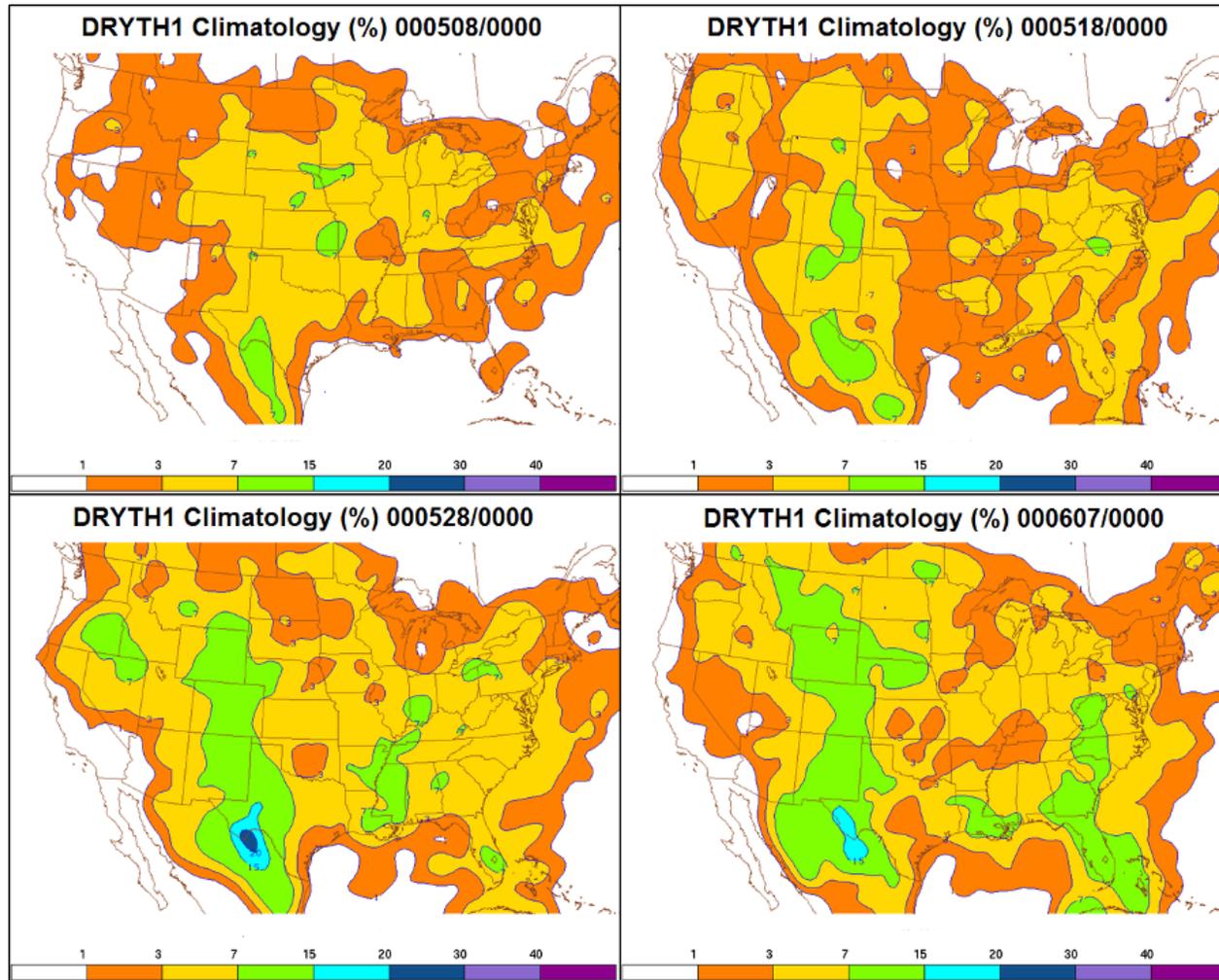


Figure 2.8: Example of DRYTH1 climatology (left) vs. DRYTH2 climatology (right) for the CONUS.



30 Figure 2.9: Average Dry Thunder 1 (AVEDRYTH1) pentad examples from early May through early June – 00 UTC data.

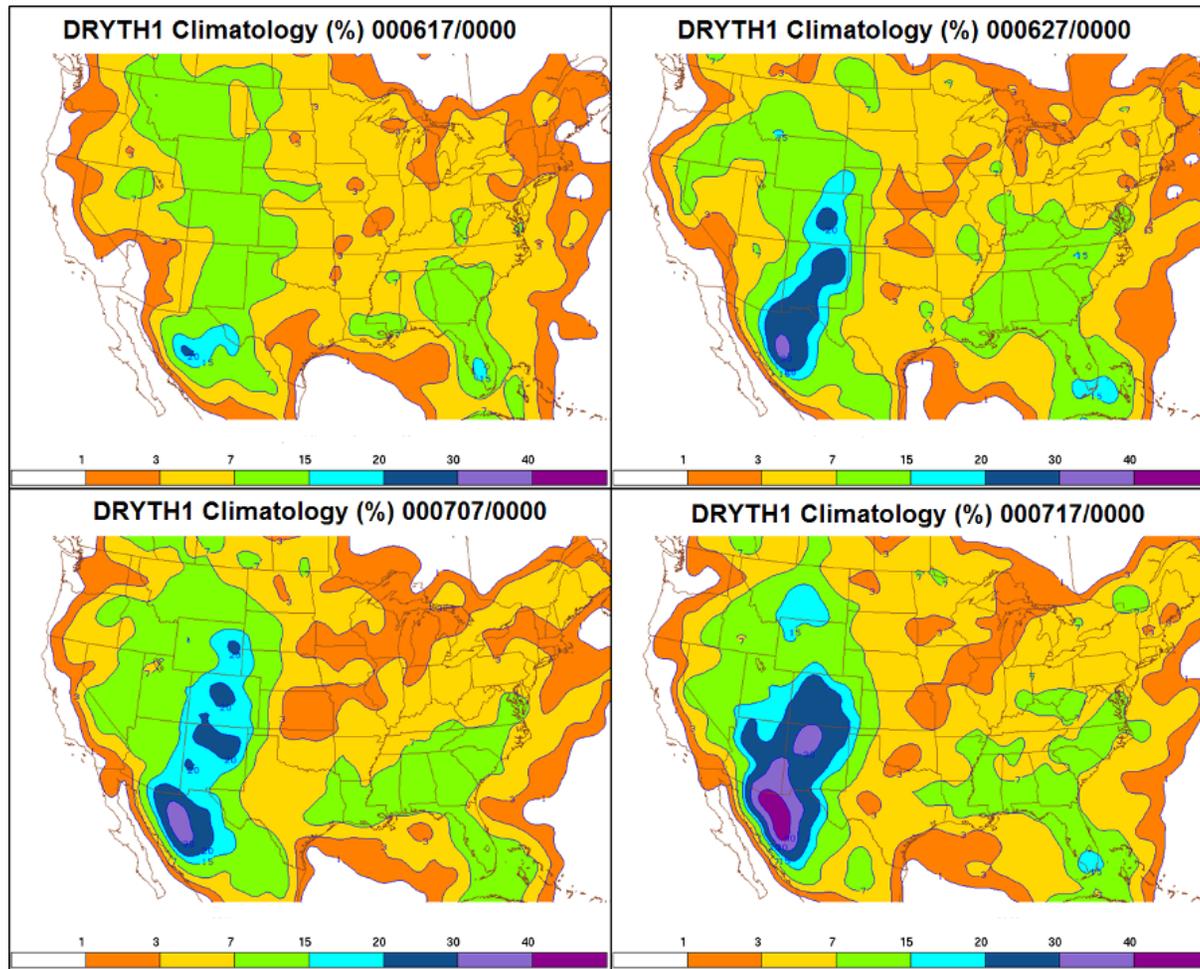


Figure 2.10: AVEDRYTH1 pentad examples from mid June through mid July – 00 UTC data.

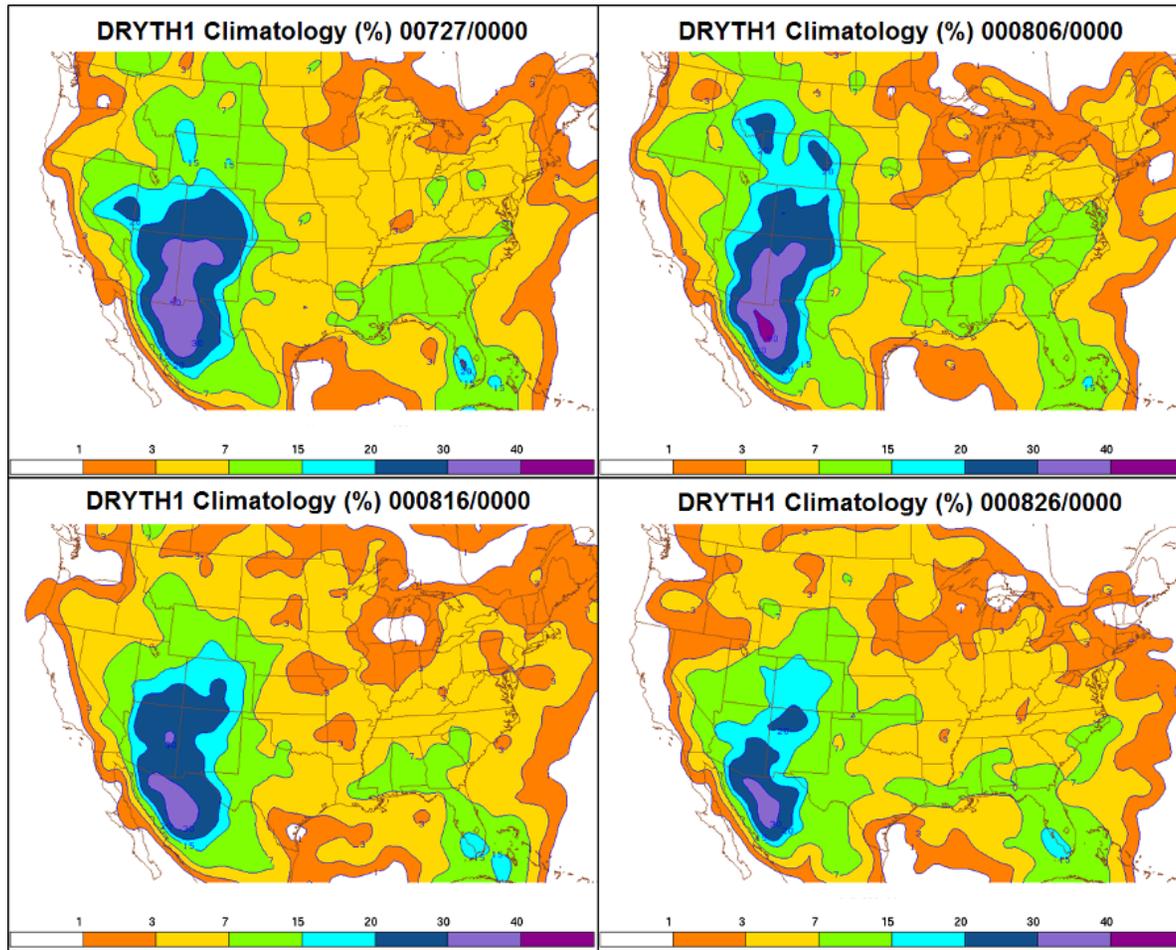


Figure 2.11: AVEDRYTH1 pentad examples from late June through late August – 00 UTC data.

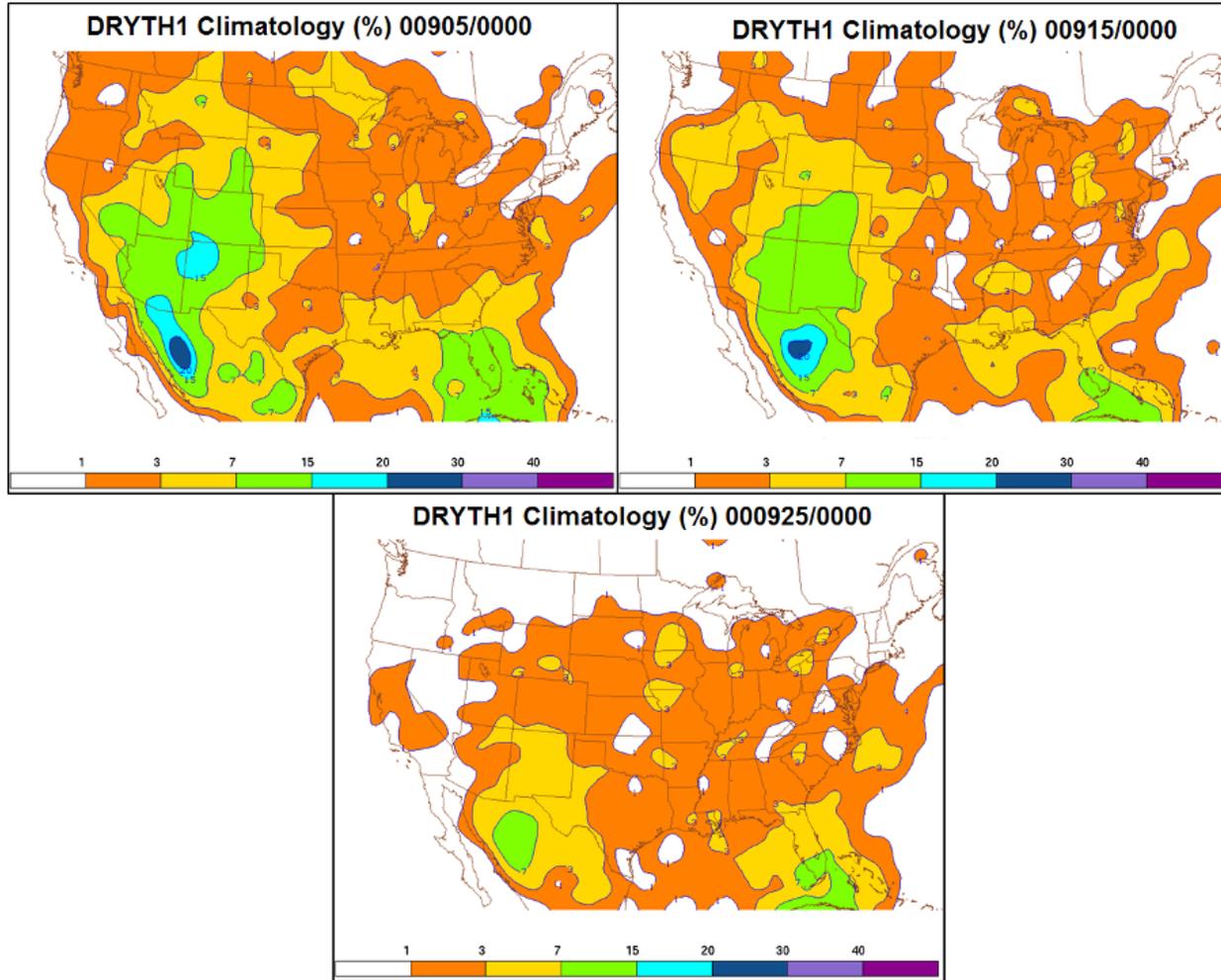
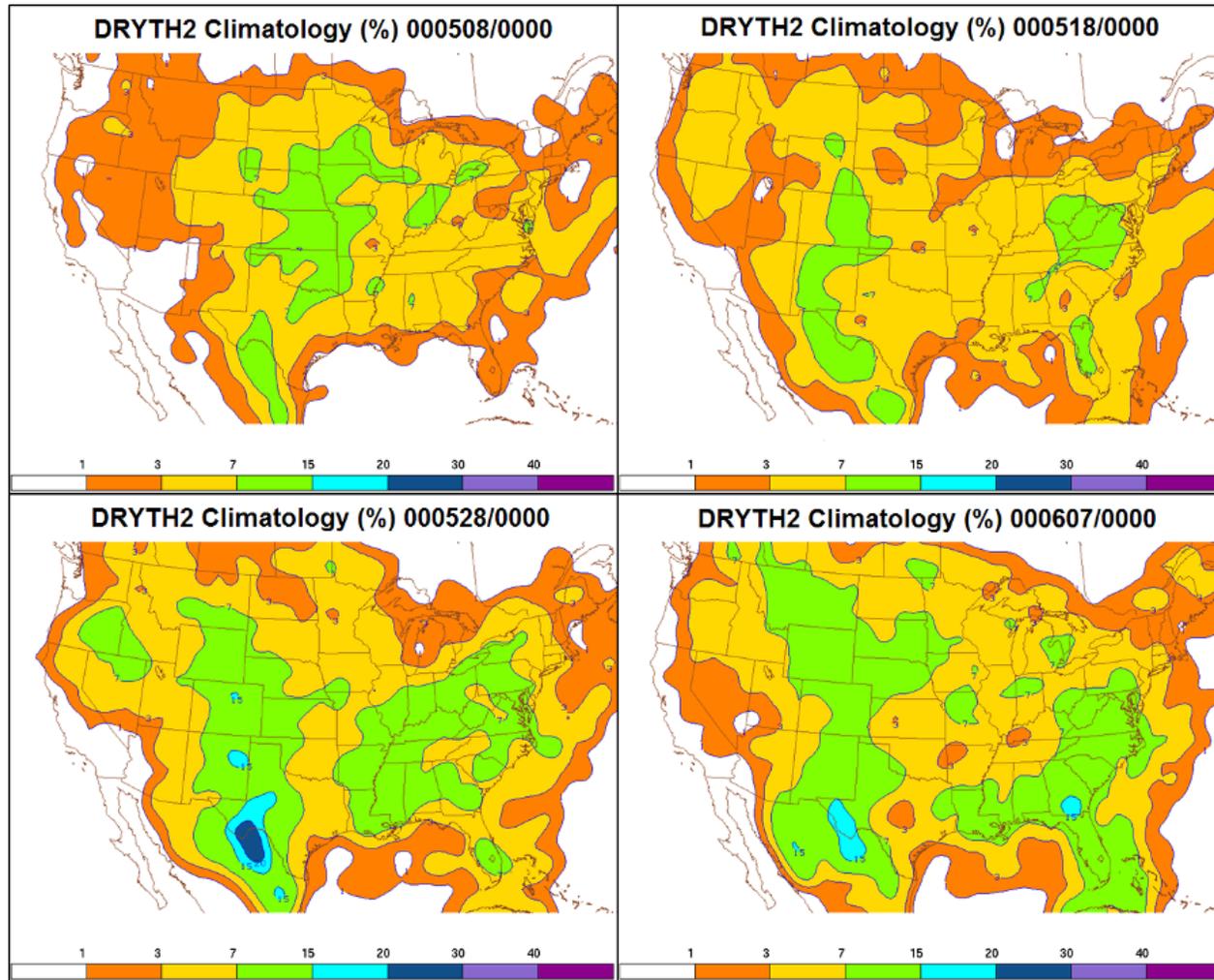


Figure 2.12: AVEDRYTH1 pentad examples from Sep – 00 UTC data. Convection diminishes in general during this time frame.



34 Figure 2.13: Average Dry Thunder 2 (AVEDRYTH2) pentad examples from early May through early June for the CONUS – 00 UTC data.

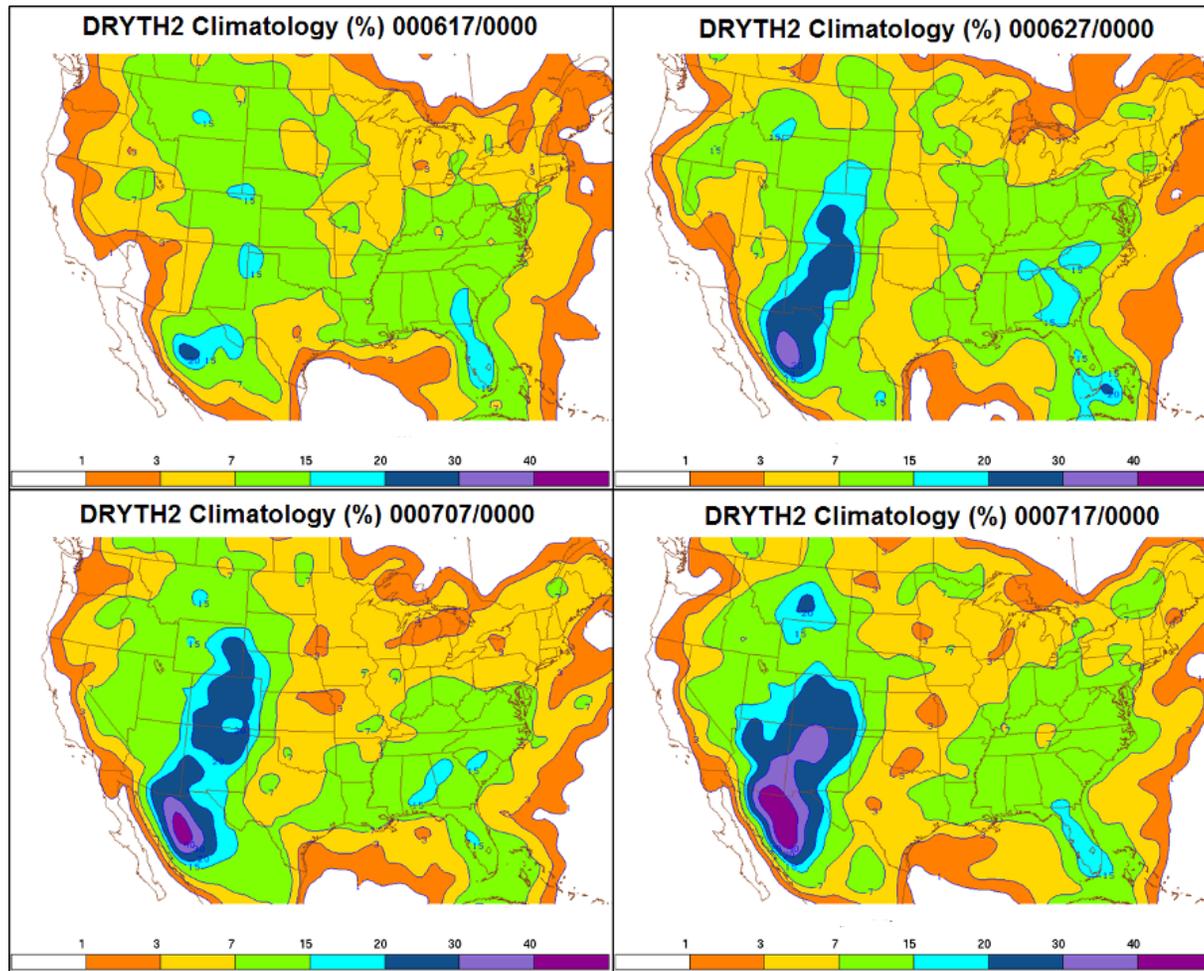


Figure 2.14: AVEDRYTH2 pentad examples from mid June through mid July for the CONUS – 00 UTC data.

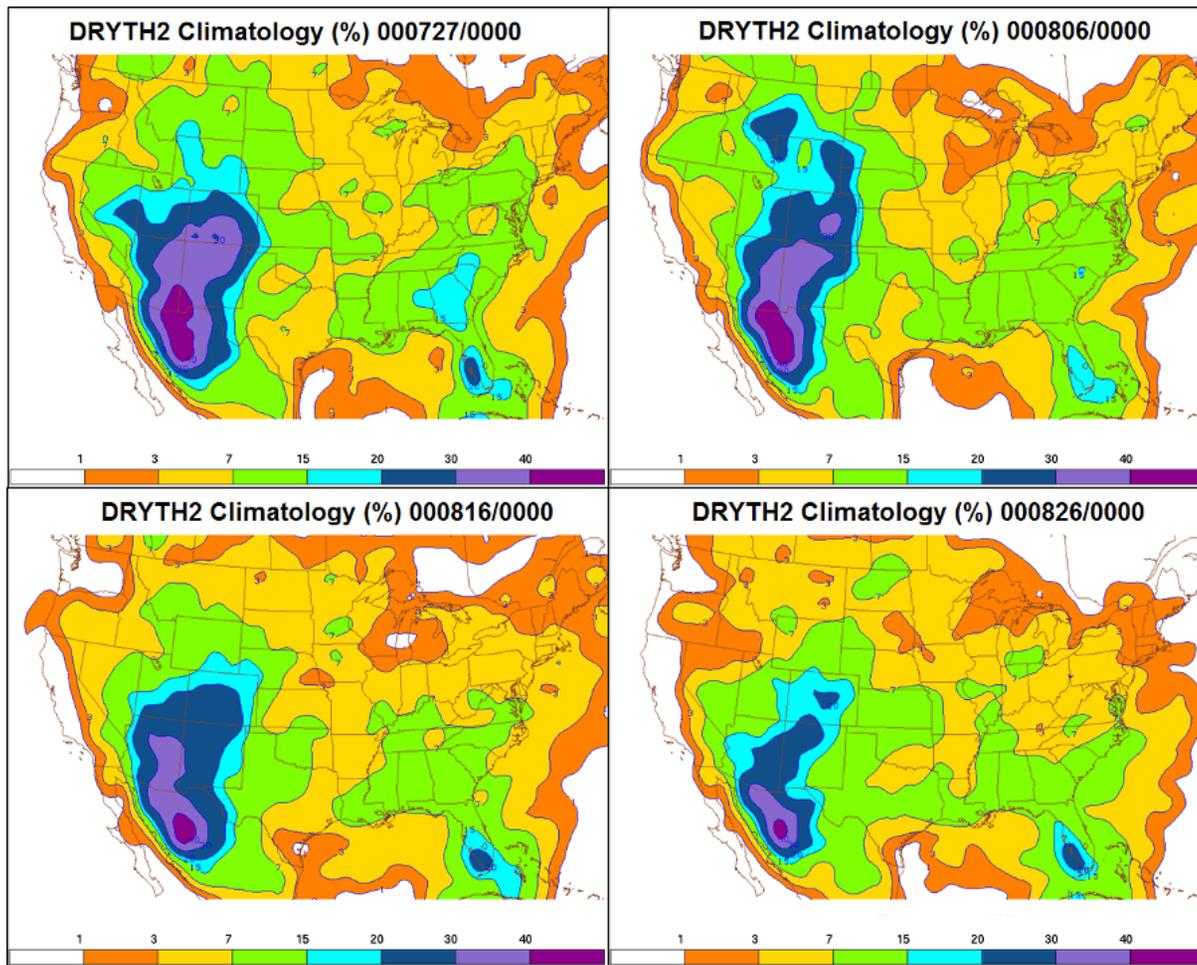
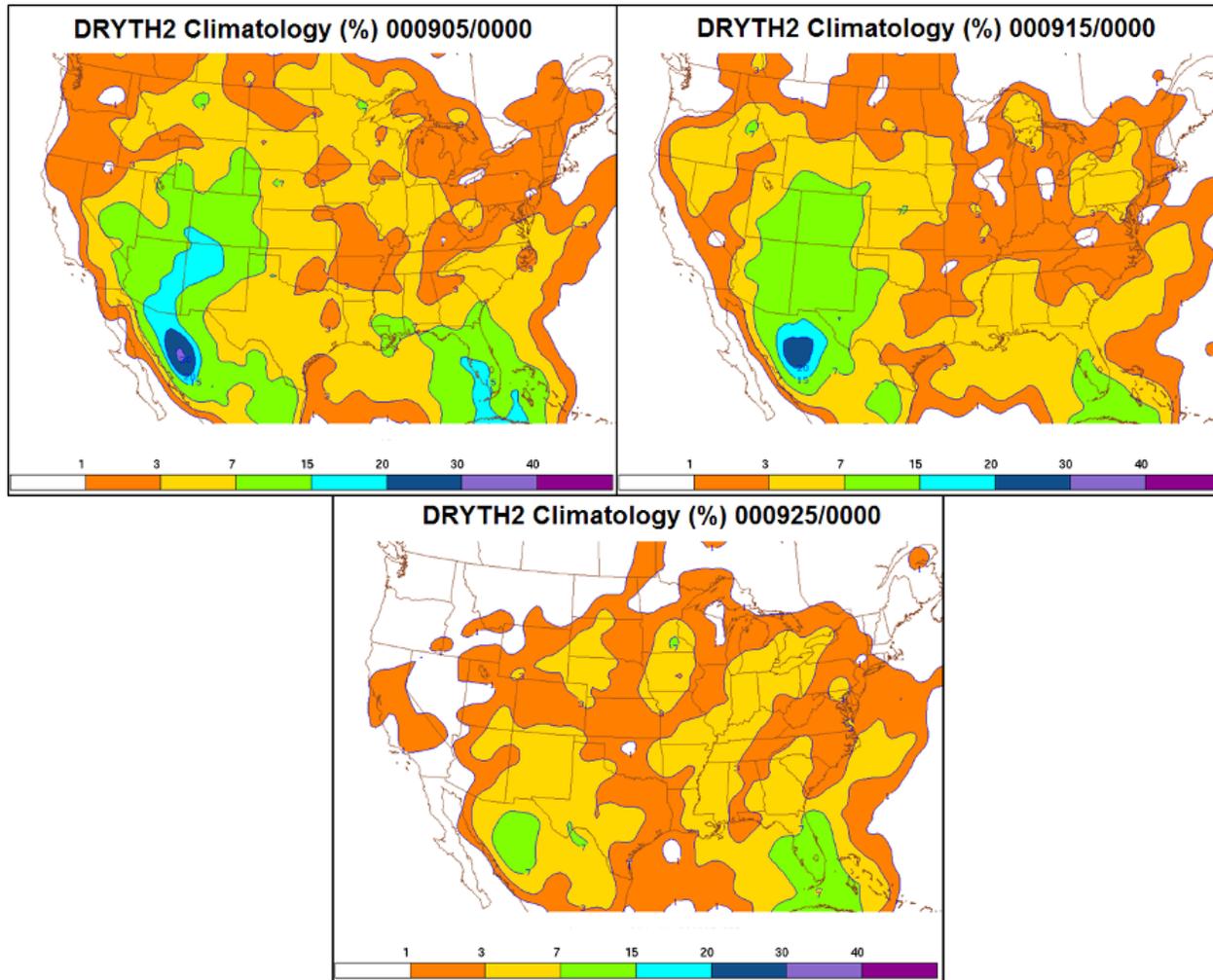


Figure 2.15: AVEDRYTH2 pentad examples from late July through late August for the CONUS – 00 UTC data.



37 Figure 2.16: AVEDRYTH2 pentad examples from Sep for the CONUS – 00 UTC data.

2.5.2 Alaska Climatology

An intra-seasonal pentad climatology for AK lightning has been investigated thoroughly by Buckey and Bothwell (2009). Their work reveals an increase in lightning into June, a slight decrease mid-to-late June, and another increase into July. The lightning season tapers off quickly by August and September. Climatology values of TOTF and/or FONE are lower in Alaska compared to the CONUS because of the lower amount of total lightning received. Additionally, the climatology values shown here are smaller than those covered in Buckey and Bothwell (2009) because of the differences of resolution (45-km in Buckey and Bothwell compared to 10-km in this study). Fewer flashes fall into a bin on the 10-km grid due to the splitting of the total amount seen at a coarser resolution. Fields of thirty-or-more flashes (FTHT) and one-hundred-or-more flashes (FHUN) are excluded from consideration in Alaska as these amounts are rare at this smaller grid size (and in Alaska, in general). Refer to Figs. B17-B25 in Appendix B for examples of Alaska lightning climatology.

This finer resolution also affects precipitation (specifically rainfall) values in AK. Intra-seasonal trends show precipitation regions shift from the Gulf of Alaska, northward along the Alaska/Canada border, and into the mainland forest regions. Southwestern regions see an increase in precipitation moving into the warmer months of July and August. Precipitation fields excluded from consideration in Alaska include rainfall three-quarters-of-an-inch-or-more (P3QTR) and one-inch-or-more (PONE)

due to the smaller grid size. Examples of AVEPTOT and AVEPHDTH can be found in Appendix B (Figs. B25-B32).

Examination of the dry thunder climatology reveals pertinent information. Again, the chances for dry thunder (AVEDRYTH1 and AVEDRYTH2) generally resemble the chances for lightning (AVEFONE). Certain coastal regions and areas of lower terrain have higher climatology values of DRYTH2 compared to climatology values of DRYTH1 (see Fig. 2.17). The red ovals highlight one of the more notable areas of differences between AVEDRYTH1 and AVEDRYTH2. Lower elevation and coastal sites generally receive more than a tenth of an inch of rain compared to mountainous or sloped areas, so the chances for DRYTH1 would be lower. Some of these lower regions contain higher chances of a dry thunderstorm or become apparent when considering AVEDRYTH2. Figures 2.18-2.25 show more example of AVEDRYTH1 and AVEDRYTH2 climatology over Alaska.

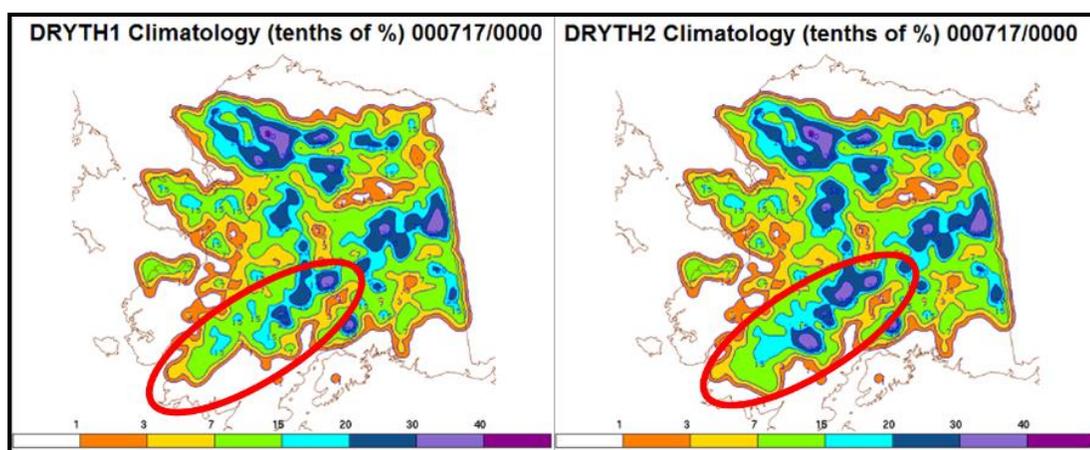


Figure 2.17: Example of DRYTH1 climatology (left) vs. DRYTH2 climatology (right) in AK.

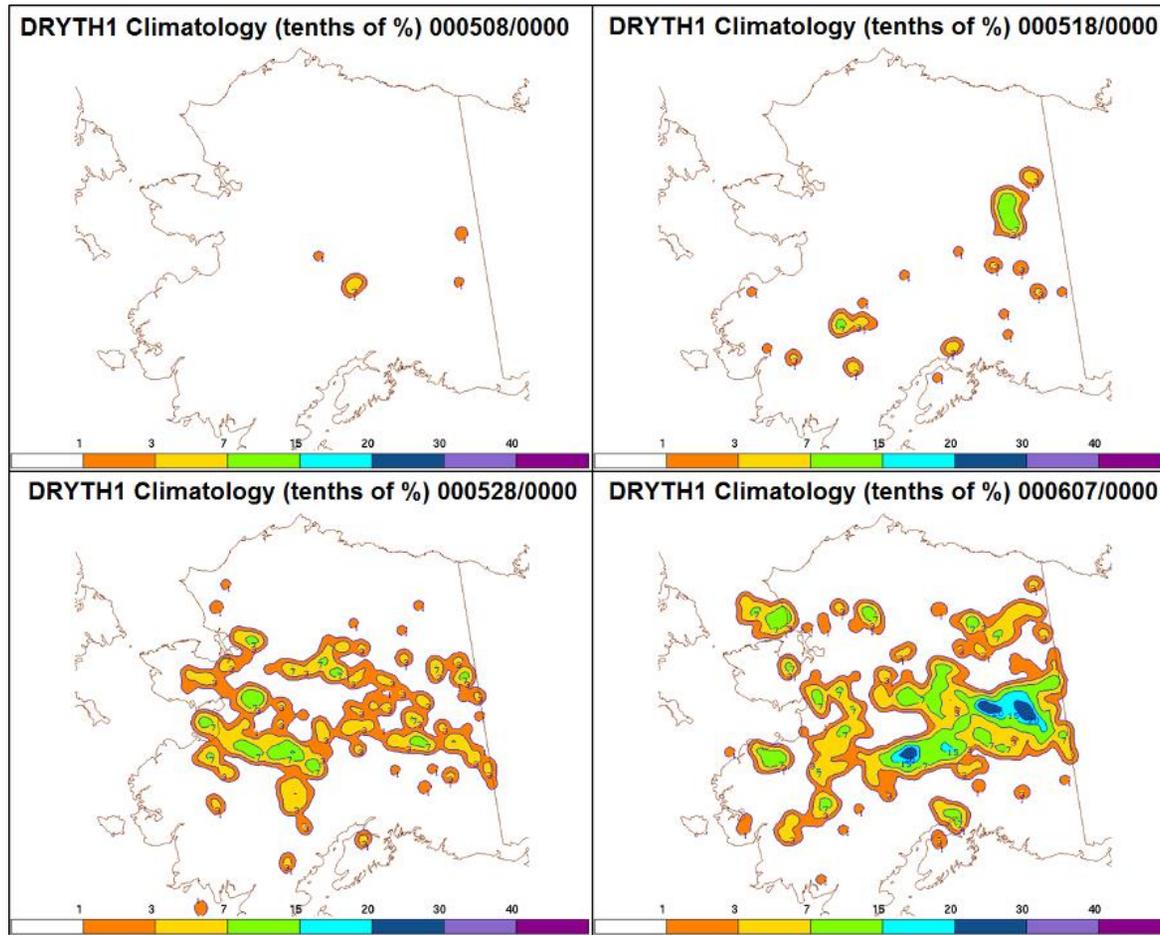


Figure 2.18: Average Dry Thunder 1 (AVEDRYTH1) pentad examples from early May through early June for Alaska – 00 UTC data. Values are in tenths of an inch, so, e.g., values of 30 = 3%.

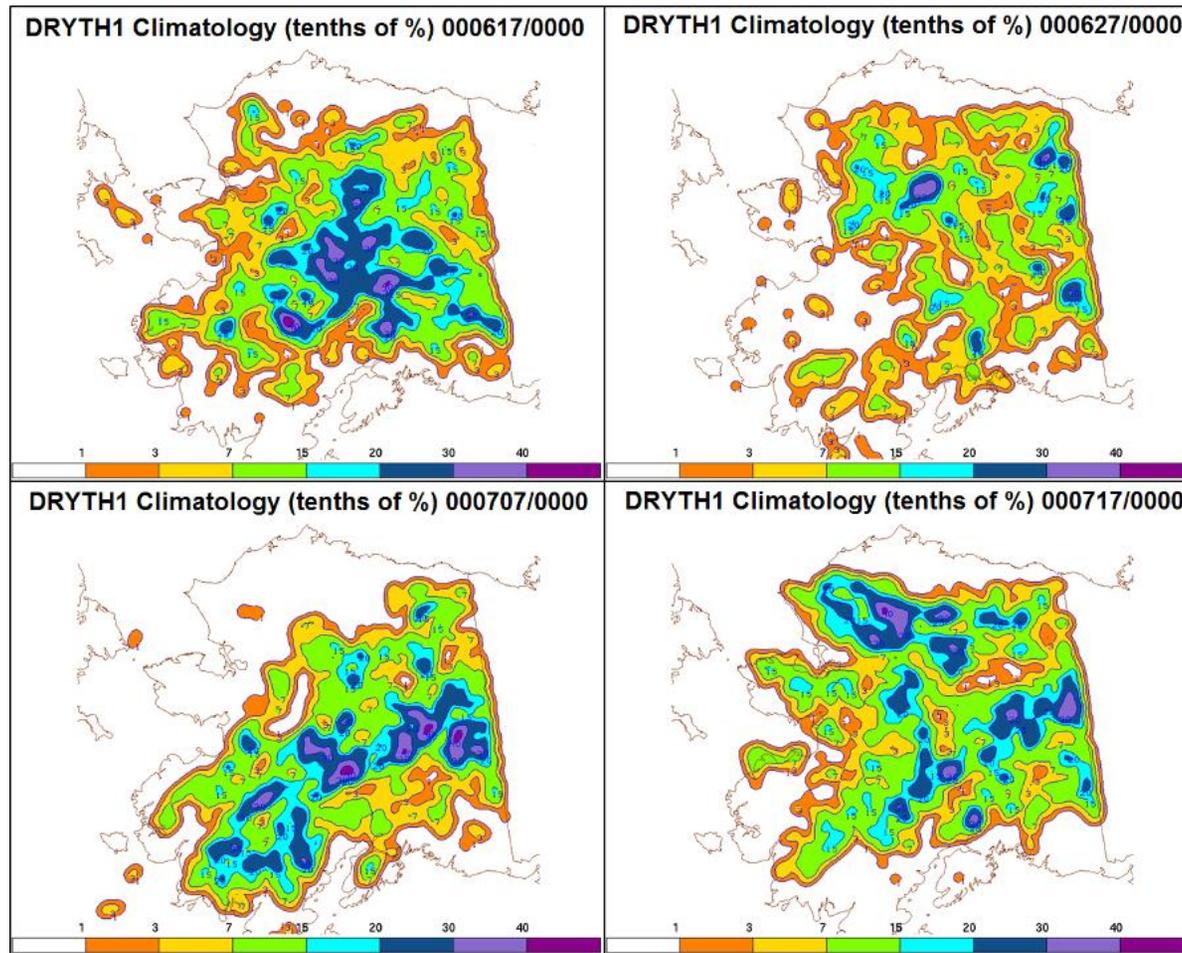


Figure 2.19: AVEDRYTH1 pentad examples from mid June through mid July for Alaska – 00 UTC data.

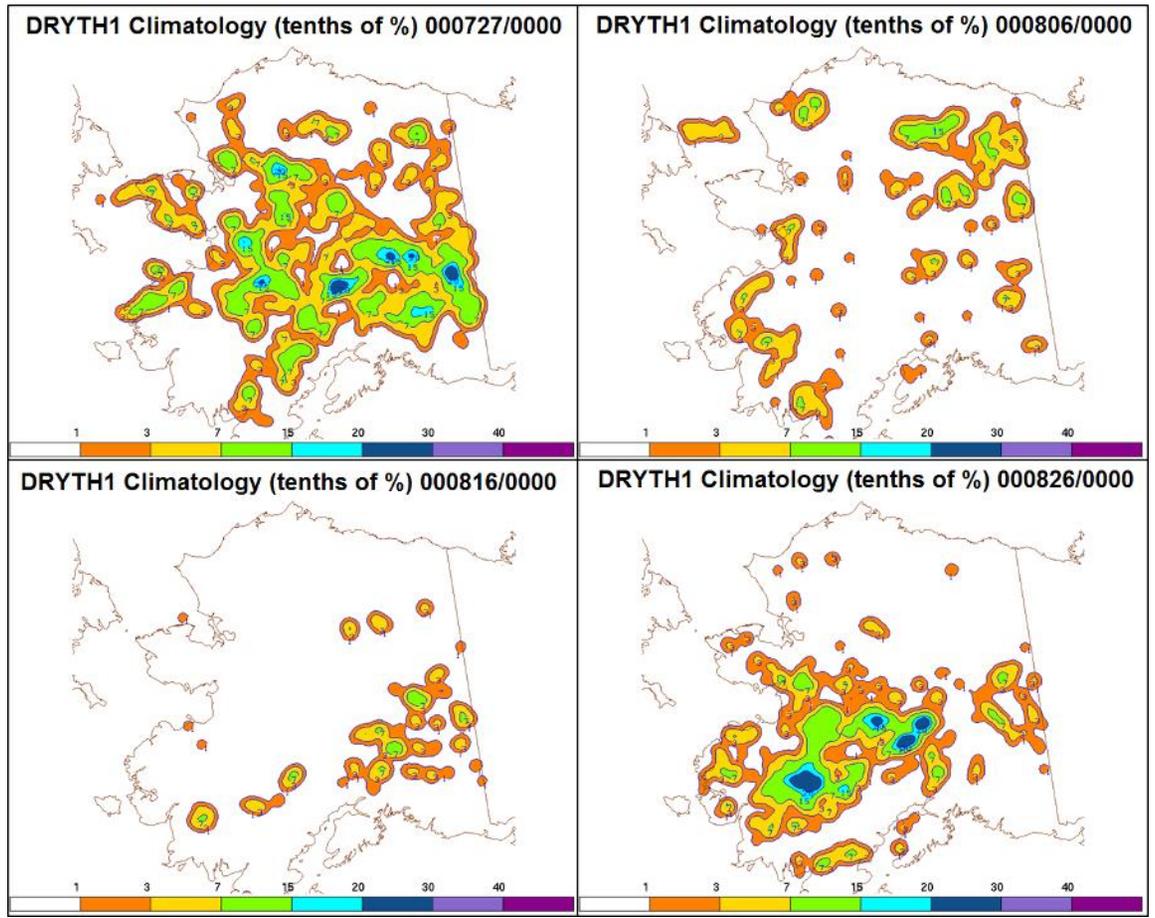


Figure 2.20: AVEDRYTH1 pentad examples from late July through late August for Alaska – 00 UTC data.

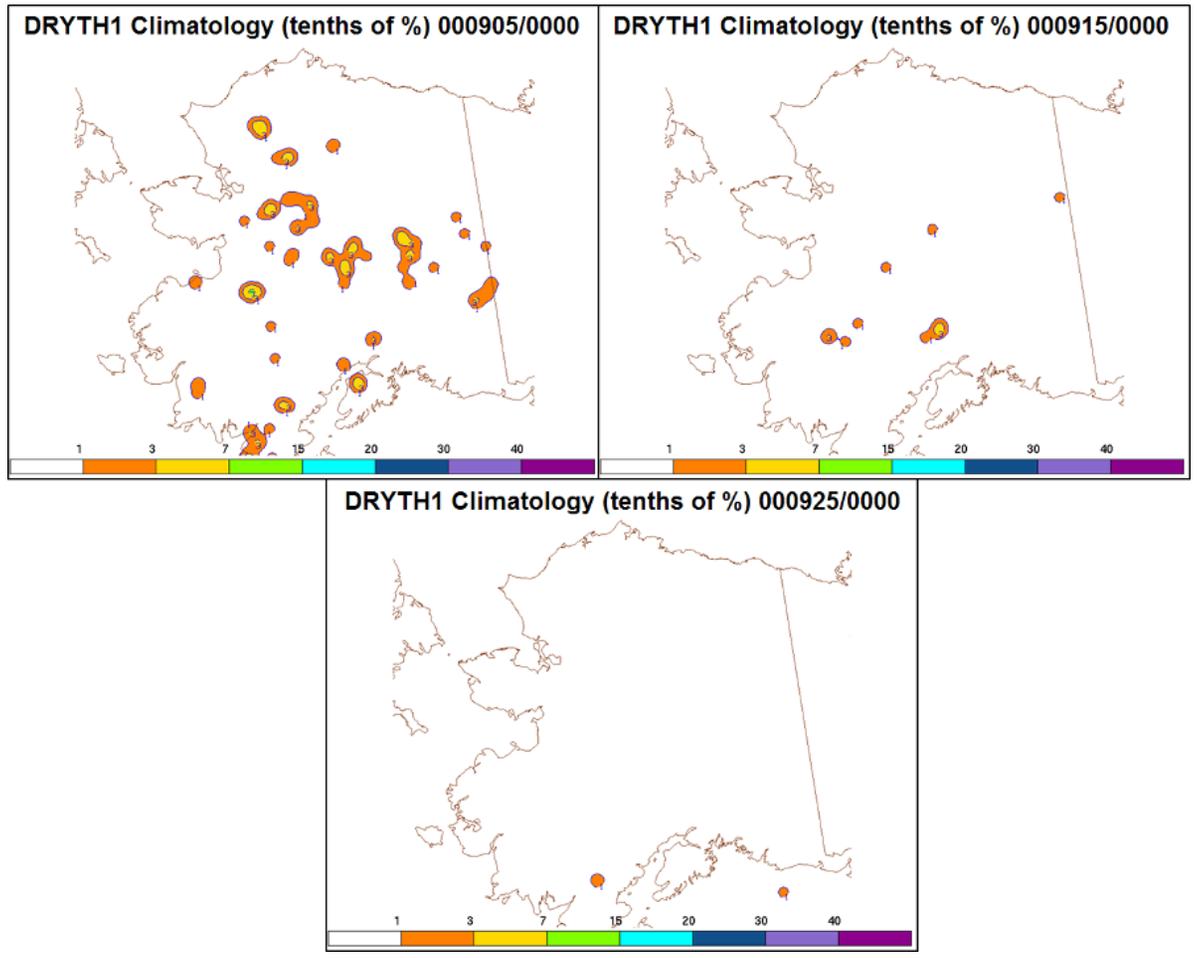
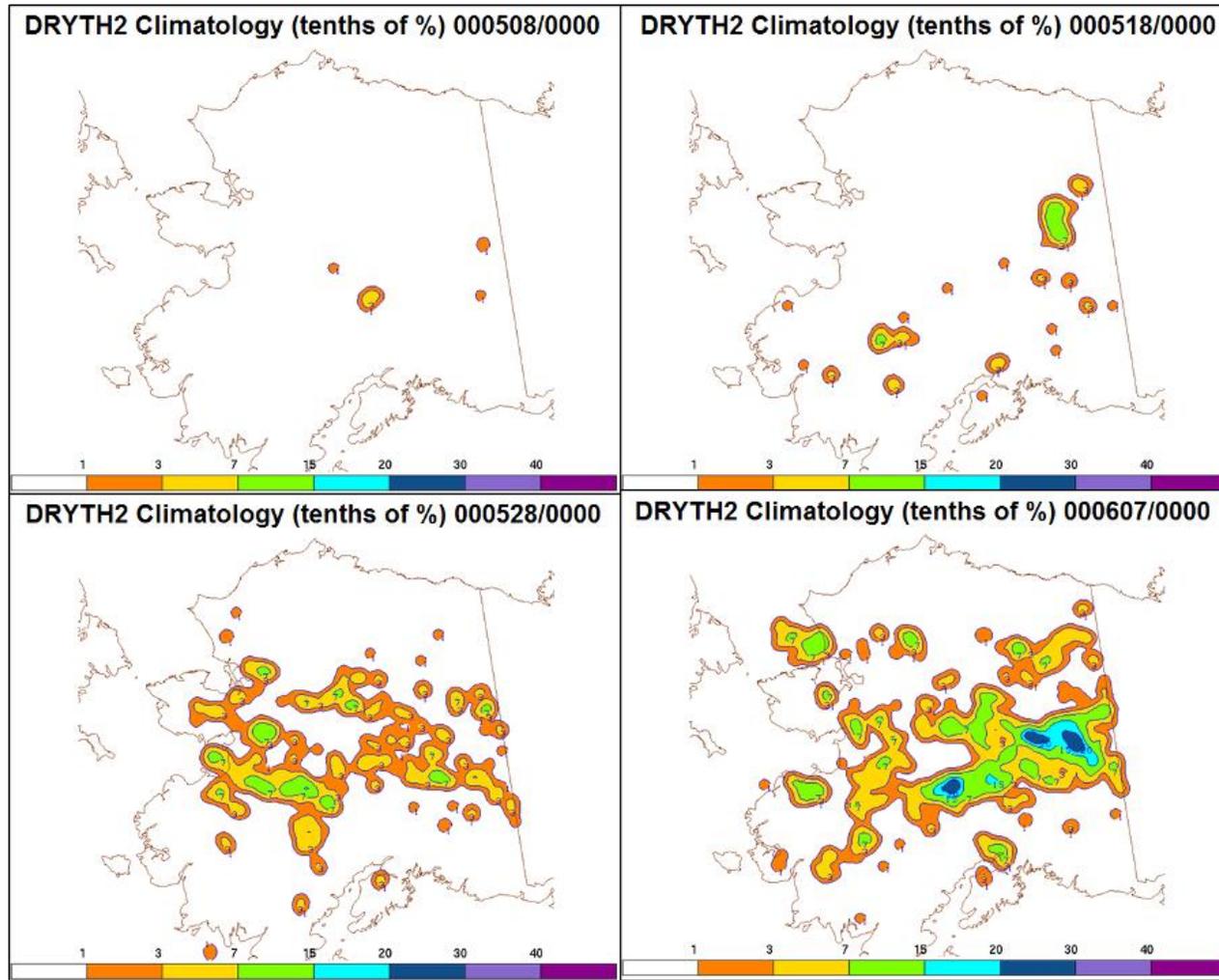


Figure 2.21: AVEDRYTH1 pentad examples from Sep for Alaska – 00 UTC data.



44 Figure 2.22: Average Dry Thunder 2 (AVEDRYTH2) pentad examples from early May through early June for Alaska – 00 UTC data.

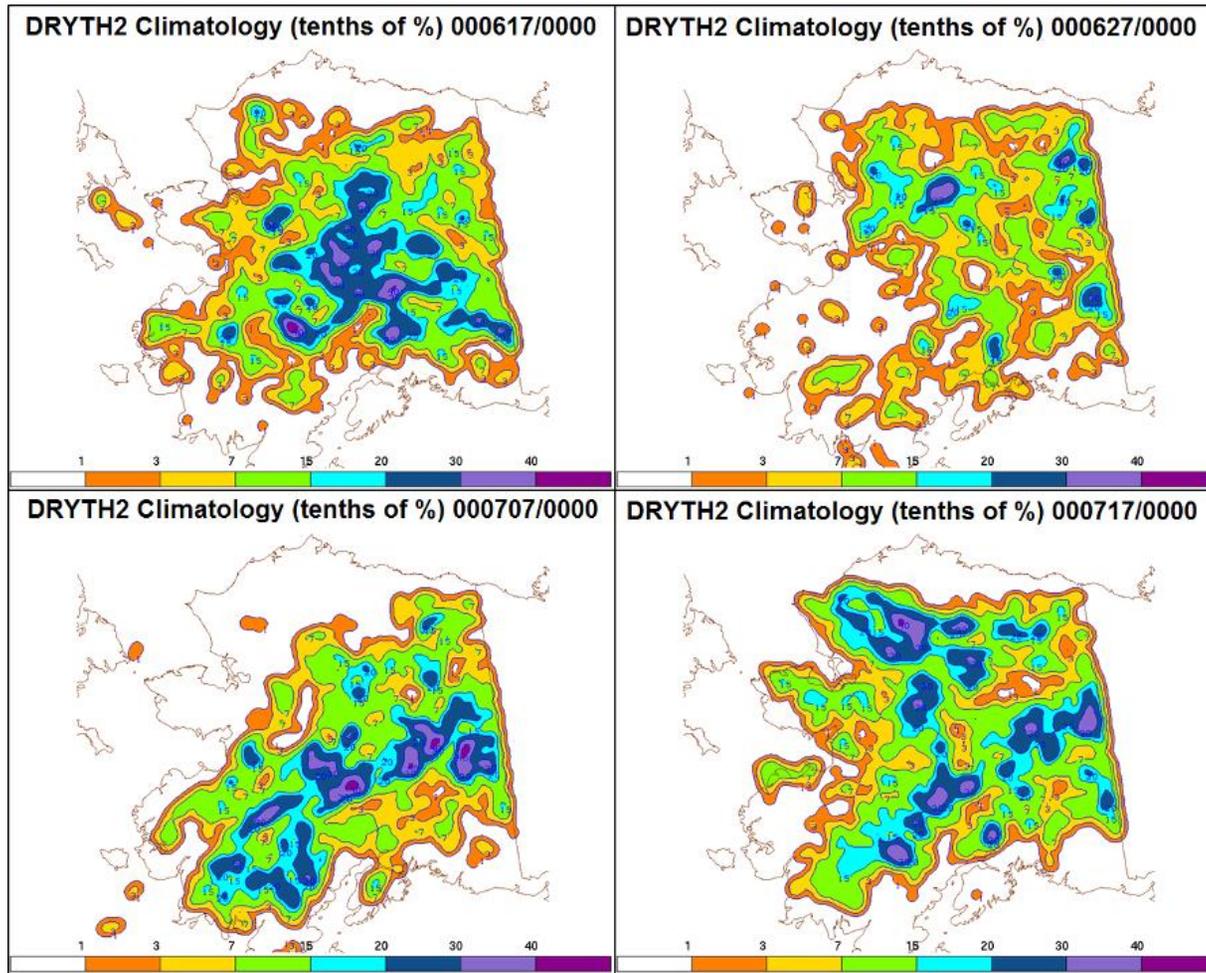


Figure 2.23: AVEDRYTH2 pentad examples from mid June through mid July for Alaska – 00 UTC data.

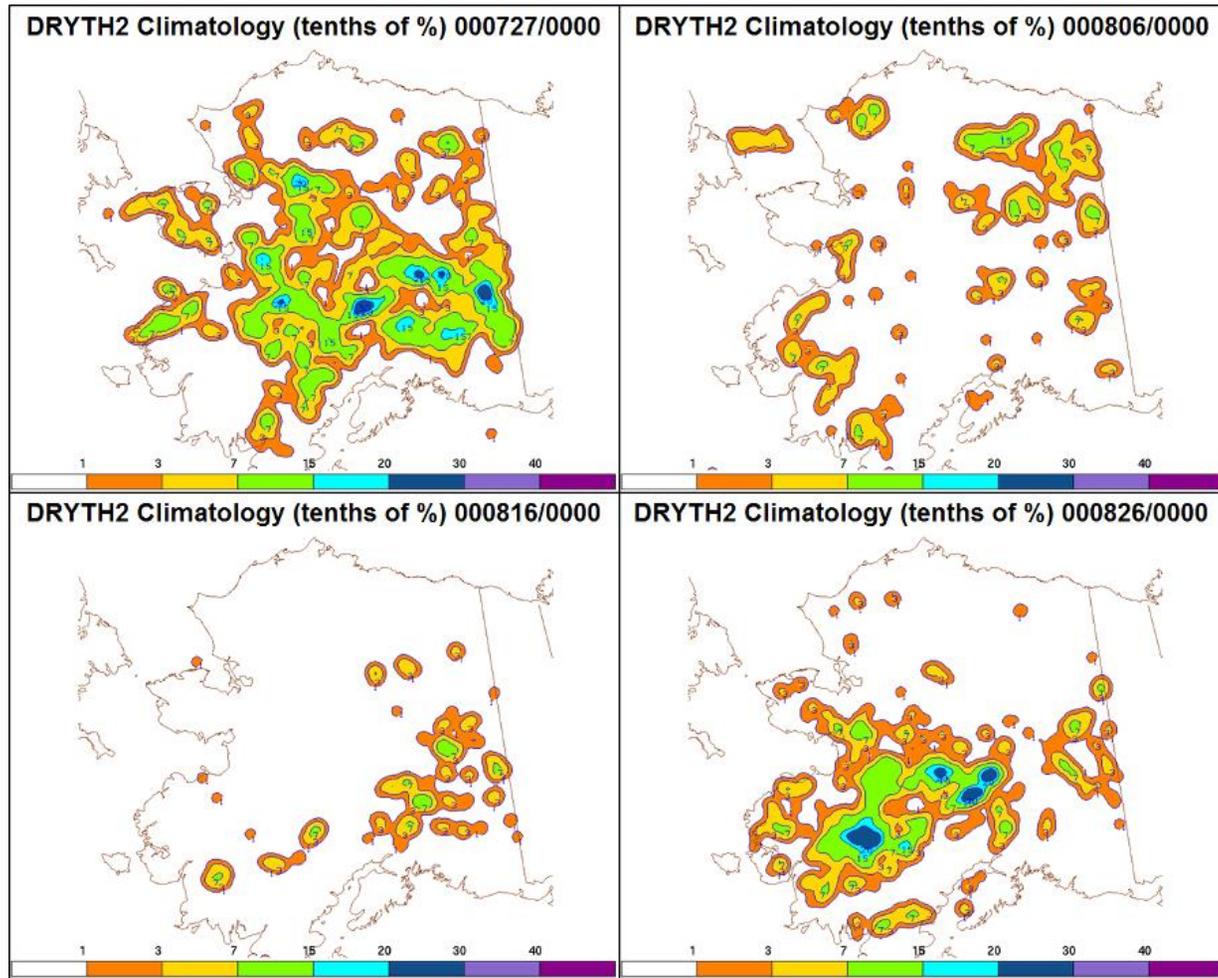


Figure 2.24: AVEDRYTH2 pentad examples from late July through late August for Alaska – 00 UTC data.

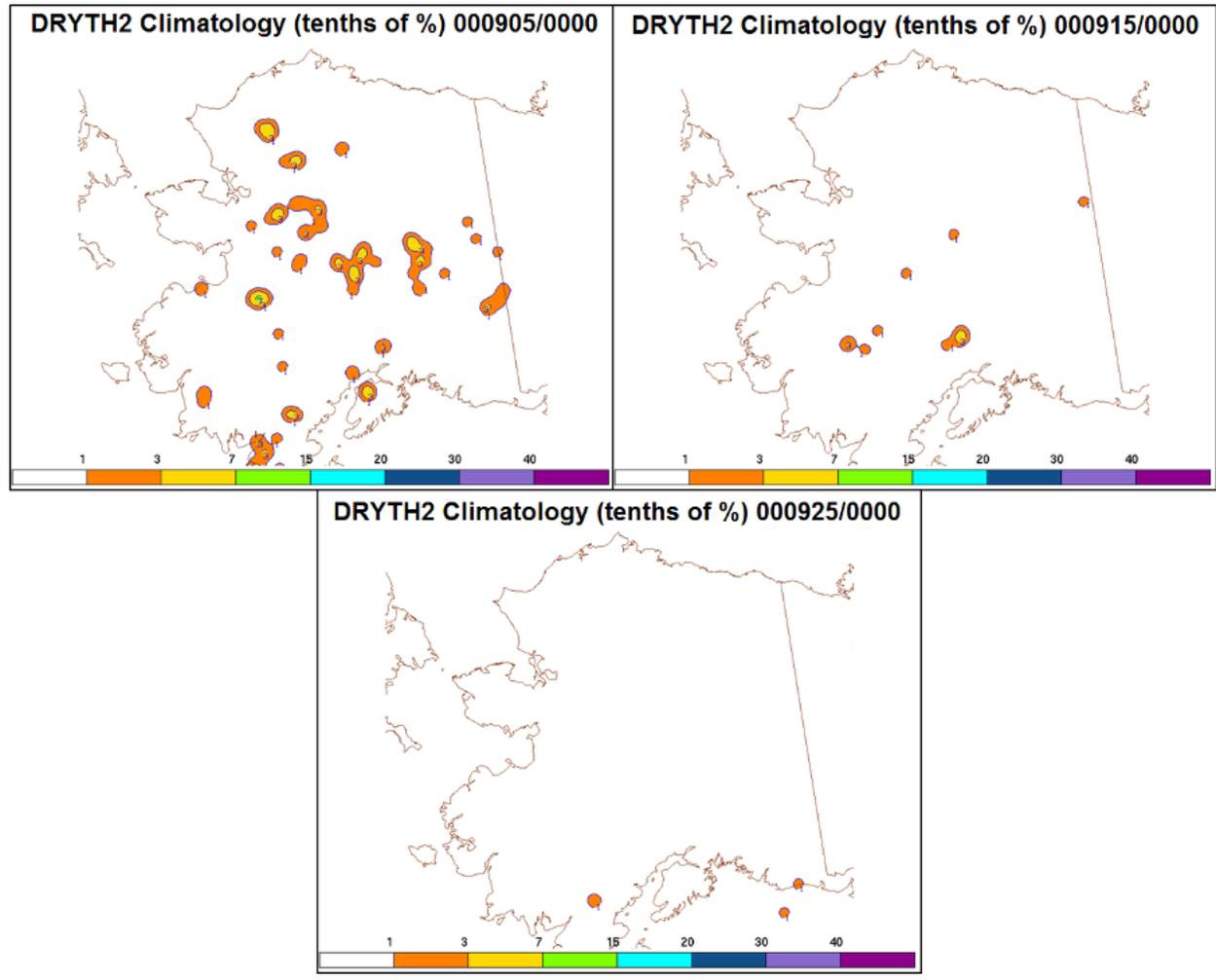


Figure 2.25: AVEDRYTH2 pentad examples from Sep for Alaska – 00 UTC data.

CHAPTER 3

Perfect Prognosis & Principal Component Analysis

3.1 Perfect Prog Technique

There are several methods to create a relationship between model variables (predictors) and the observed data (predictands or response values). Perfect prognosis (perfect prog or PP) assumes the resulting model forecast fields are perfect once a relationship is established (Klein et al. 1959; Wilks 2006). As such, it comes with all of the parent model biases. No bias correction occurs in this technique, but it can be used with output from any model because it is not dependent on parent model physics. Therefore, perfect prog equations do not have to be rederived if the parent model changes, and the forecasts should improve as the driving input improves (Wilks 2006).

Though some studies claim that MOS is more reliable than perfect prog due to its built in bias correction (e.g., Vislocky and Young 1989; Brunet et al. 1998, Wilson and Vallée 2003), its dependence on a large, stable sample of model variables and observations can make it difficult to maintain and update. Brunet et al. (1998) mention that “[t]he PP forecasts are sharper, and the MOS forecasts are more reliable” out to 72 hours for precipitation and temperature forecasts. A model with better sharpness is more likely to correctly forecast a rare event instead of converging to a forecast of climatology at longer time ranges (Brunet et al. 1998; Marzban et al. 2006; Shafer and Fuelberg 2008). This ability to correctly forecast rare

events compared to climatology makes PP the method of choice to predict dry thunderstorms instead of MOS in this study.

Both analysis variables and observed values go into a logistic regression technique. A total of 139 predictors created from the NARR data can be related to the lightning and precipitation data (Tables A2-A5). Several of these variables explain more about the inherent signal than others at the noise level. Another statistical technique can be used prior to logistic regression to draw out the groupings of meteorological patterns/variables that are interrelated and explain the most variance of a data set.

3.2 Introduction to PCA

Principal Component Analysis (PCA) can effectively extract signal vs. noise in complicated and mixed patterns, such as variables that measure the atmosphere. In particular, the variables highlighted in this procedure explain the *most* variance about a data set. PCs are similar to (and sometimes called) Empirical Orthogonal Functions (EOFs), as EOFs are unit length PCs. Refer to Lorenz (1956) for more information on EOFs. PCA decomposes a similarity matrix (e.g., correlation matrix), relating the aforementioned atmospheric variables, into two displays: PC loadings and PC scores. Equation 3 describes the basic formula on relating, in this case, the PC Scores (**F**) and the transpose of the Loadings (**A**) matrix to a set of effective standard scores (Z-Scores), or scores standardized to the mean.

$$\mathbf{Z} = \mathbf{FA}^T \quad (3)$$

See Appendix C for more details and plots not covered in this overview. All analysis for PCA was performed in the TIBCO Spotfire S-Plus 8.2.0 statistical software package on a Linux workstation with RedHat 5.

Using (3) as a template, the construction of \mathbf{Z} will be discussed. First, a data matrix, \mathbf{X} , is constructed for each month. Each column contains information about the predictors at each grid point. Each row represents the value of the variable for a given day, for a given month. As such, this type of analysis is similar to mixed P/R-Mode, but ultimately is most similar to R-mode because variables are columns that are assessed spatially. Data in the data matrix, \mathbf{X} , is scaled because of the mixed units across the variables. The orders of magnitude difference between variables, such as pressure and omega, require that all variables are scaled to a relative threshold. That is, each variable is adjusted to a mean of 0 and a standard deviation of 1.

A correlation matrix, \mathbf{R} , is created to measure the linear similarity among predictors. The process of correlating the variables removes the mean from each variable and makes each variable unit length (standardization). This standardization allows variables measured on different metrics to be compared. Because each variable is perfectly linearly related to itself, \mathbf{R} has 1's along the diagonal. Any points (rows) with missing data are omitted from the correlation calculation. Eigenvectors (\mathbf{V}) and Eigenvalues (λ) are then calculated from \mathbf{R} and used to form the PC loading matrix (\mathbf{A}) by postmultiplying \mathbf{V} by a diagonal matrix (\mathbf{D}) of the square root of the eigenvalues. This loading matrix displays “weights” of variables along a

PC, and because \mathbf{R} is decomposed, those weights in \mathbf{A} are in correlation units.

Rotating the PCs makes the results more stable and less susceptible to sampling errors, as shown in North et al. (1982), and more consistent with the similarities measured in the correlation matrix (Richman 1986). Assuming the correlation matrix is a valid representation of the similarities among the variables and the variables express meaningful relationships between the atmosphere and dry thunderstorms, the rotated PC loadings (\mathbf{B}) are more interpretable physically or meteorologically than the initial (unrotated) loadings (\mathbf{A}) (Richman 1986). Such an analysis is often referred to as Rotated Principal Component Analysis (RPCA). See Appendix C for more mathematical background behind the rotation process. O'Lenic and Livezy (1988) tested the RPCA method on upper-air heights to find discrepancies between data sets. RPCA was applied to lightning research by Bothwell (2002) over the western United States and was utilized for Alaska by Buckey (2009). Given the predictive skill in these analyses and the statistical and interpretational advantages of RPCA, it will be applied as part of this study, using the Varimax (orthogonal) rotation method.

Because one of the key advantages of any PCA (unrotated or rotated) is efficient data reduction, PCA seeks to compress as much variability of the variables into as few dimensions (k) as possible. Additionally, the rotation of PCs requires that a number of PCs (k) be identified prior to the rotation process. Several tests are applied to help determine how many PCs to keep

for rotation. While keeping all of the PCs explains all of the variance, a few can be found to explain most of the variance. North et al. (1982) remark that “...for a given n no other basis set can explain more of the average variance...” where n refers to a certain number of PCs kept.

One test comes from the Scree diagram. Scree refers to rock rubble that falls down cliff faces and collects near the bottom. Thus, the diagram displays how PC variance explained (a proxy for relevance) starts off high and then falls off the cliff as the number of PCs retained (k) goes from 1 to 2 and onward. Unfortunately this test alone, even when coupled with other Scree-related tests such as North’s test (not shown), makes no clear distinction on the number of PCs to retain because it has never been demonstrated that the amount of variance explained by a PC relates to its meaningfulness. Moreover, as shown in Fig. 3.1, in many analyses, multiple breaks occur. The first PC explains a lot of variance as seen by its high value. Moving down the slope means that each additional PC contributes less and less to additional variance explained until it starts to level off around 9-15 PCs. The actual contribution to explaining the variance is difficult to distinguish from this chart.

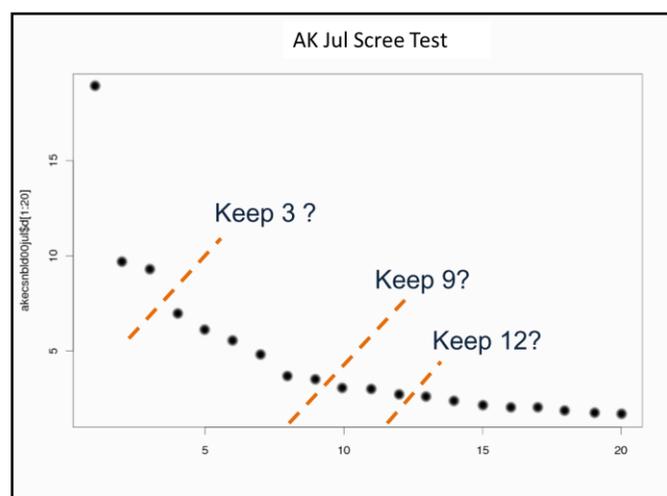


Figure 3.1: Scree Test showing how PC contributions slope down from PC1 (top left) to PC20 (bottom right).

Another method, the Congruence Coefficient (CC), constructed to relate two sets of PC loadings, is shown in Equation 4 (Harman 1976). CC

$$g_{AB} = \frac{\sum_{j=1}^n b_{jA} b_{jB}}{\left[\left(\sum_{j=1}^n b_{jA}^2 \right) \left(\sum_{j=1}^n b_{jB}^2 \right) \right]^{1/2}} \quad (4)$$

was introduced in this form by Harman and later adapted by Richman (1986) who related the loadings (matrix A) to the correlation matrix from which the loadings have been extracted (matrix B). A congruence coefficient value is calculated for each PC, and varies depending on the number of PCs kept. Values fall between -1 and 1. Because the signs of the loadings are arbitrary (i.e. any loading vector can be multiplied by -1 with no loss in interpretation) the absolute value of the CC is evaluated. For any PC, the largest absolute PC loading is identified and the variable number (row number) associated with that largest value is noted. That variable number indexes the appropriate column of the correlation matrix for the CC calculation. CC values between 0.82 and 1.00 indicate a good to excellent match between the two vectors.

Examples from the CONUS and AK domain are shown in Figs. 3.2 and 3.3. Keeping 12 PCs shows good matches down each PC value for both domains as the blue line falls between the perfect and good match thresholds. The variance explained (dashed-red) line starts to flatten as

more PCs are kept and less additional variance is explained. Keeping 7 PCs explains 50% (0.5) of the variance and keeping 12 explains a bit over 60% (0.6) over the CONUS domain. About 55% of the variance is explained when 12 PCs are kept in Alaska. Twelve PCs are kept across each domain, for each month, of this study for consistency. Plots for the other months can be found in Appendix C.

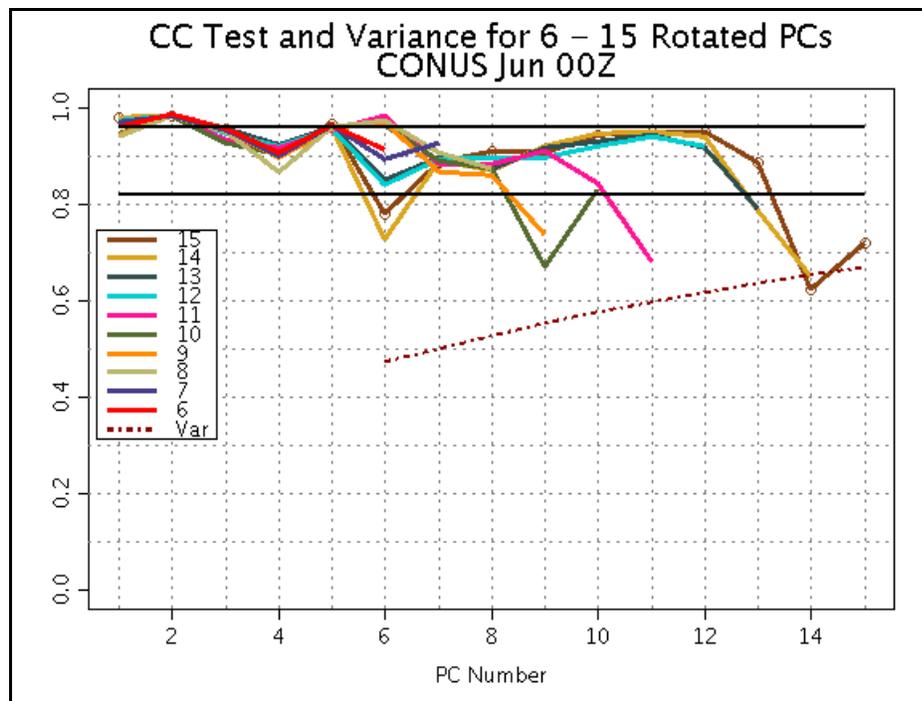


Figure 3.2: Congruence coefficient lines when keeping different numbers of PCs (6-15). The dark black lines near the top represent a perfect (≥ 0.9998) and poor (< 0.82) match. It is considered a good match when moist points fall between these lines for a given number of PCs. The light-blue line representing 12 PCs is a good match here. The dashed red line represents the variance explained by a kept number of PCs.

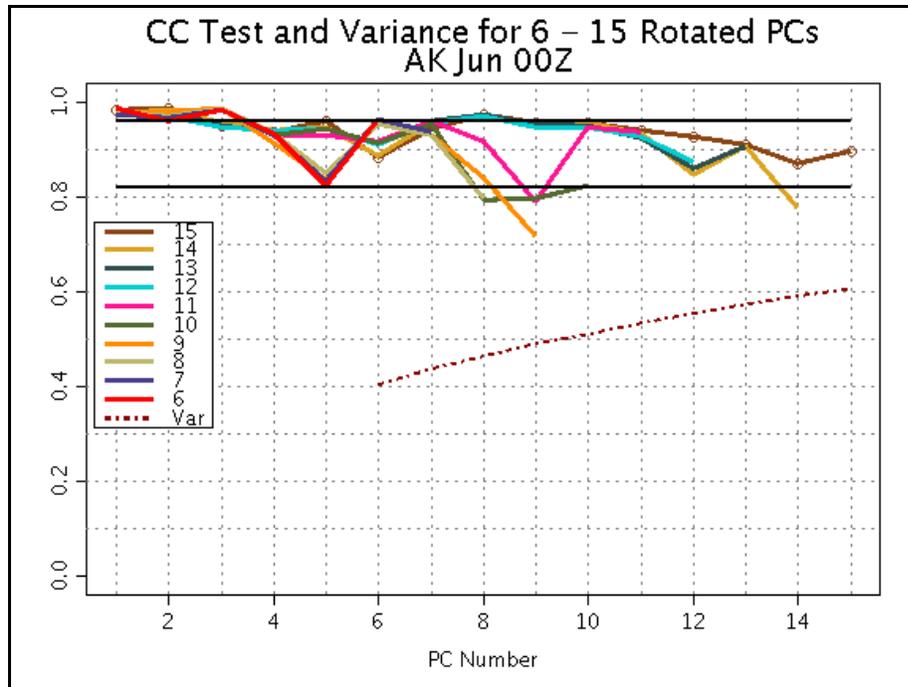


Figure 3.3: Same as Fig. 3.2 for Alaska. Again, 12 PCs (light-blue) appears to be a good match of the data overall and explains ~55% of the variance.

3.3 Describing or Naming PCs

PC loading values (b_{ij}), formed from a correlation matrix, have a range of -1 and 1 with a higher contribution or importance the further away from zero. After rotating the first 12 column vectors of PC loadings, variables with large (not near zero or not in the hyperplane) values are noted and the commonality among variables in the groupings noted and used to name the physical process underlying that PC. Richman and Gong (1999) describe that values below a small PC loading cutoff are essentially very close to 0 (zero-projection on a PC). This concept is often referred to as simple structure (Richman 1986; Richman and Gong 1999). Maximum amount of simple structure is obtained when most points fall along the axes of a pair plot (See Appendix C). Previous work by Bothwell (2002) found that 0.4 is

an appropriate hyperplane cutoff for the loadings in lightning research. That is, anything with a magnitude less than 0.4 is ignored during the grouping and naming process. The cutoff of 0.4 relates to a variance squared weight of 0.16, so anything below this value is considered to have an insufficient amount of weight to contribute to the PC group.

An example of rotated loadings shown in Fig. 3.4 highlights how the variable weights fall on certain PC columns. The yellow highlights groupings of lightning climatology, and the purple highlights strong contributions from vorticity and the Laplacian of the Geopotential Height field. Refer to Tables C1-C4 to see a detailed layout of variables above this cutoff for each month/region. Figure 3.5 shows some of the most commonly seen groups noted across both domains. As expected, predictors such as the lightning climatology (LTGCLIMO) and the associated interactive predictors of MUCAPE*LTGCLIMO consistently appear. Other terms related to storm production, such as low level forcing and moisture, are present often.

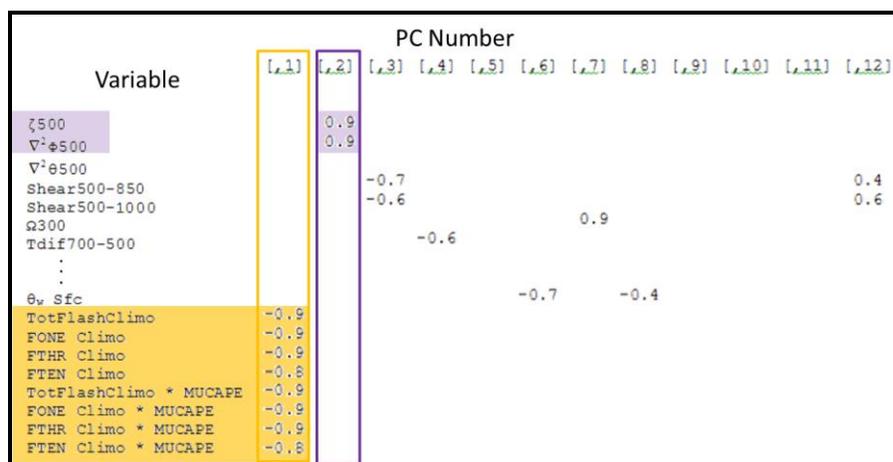


Figure 3.4: Example of loadings from Alaska during the month of July. A small sample of the 139 loadings are shown here. Parameters with a magnitude ≥ 0.4 can be grouped down each column (PC).

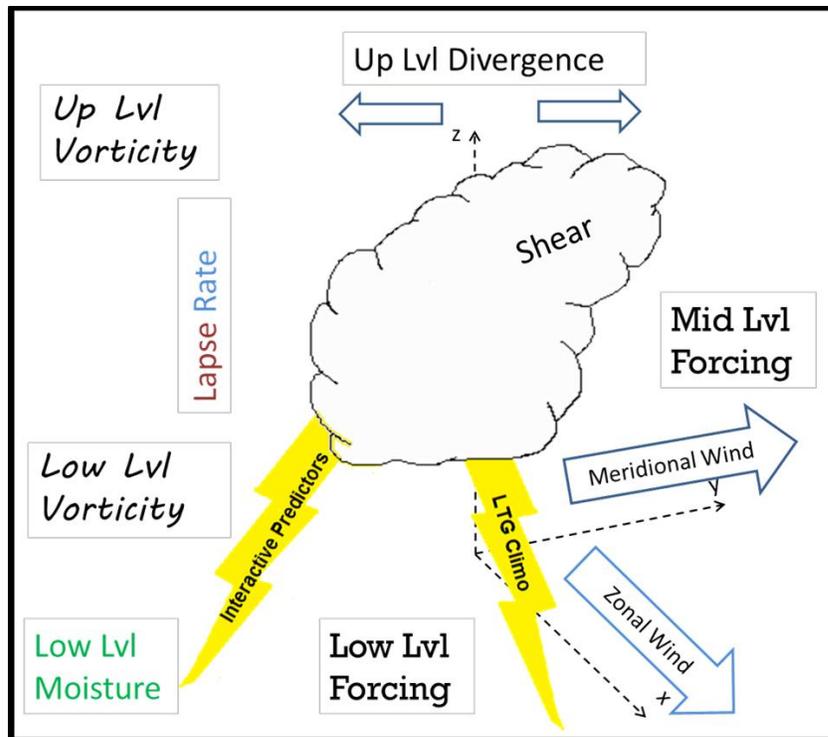


Figure 3.5: Processes seen often across all warm season months for both regions.

Another step in the PCA process involves generating PC scores by algebraically combining the standardized data with the rotated loadings. These scores display the standard deviations of data away from a mean of 0. Owing to the linearity of the model, the actual sign of the PC loadings (and hence the PC scores) is irrelevant, so any PC loading vector can be multiplied by -1 and the corresponding PC score vector also multiplied by -1 to aid in the meteorological interpretation.

For this hybrid-R-mode analysis, scores reveal regions of excess moisture, overall wind flow, and other important terms. For example, PC2 scores over the CONUS highlight the drier western regions (Fig. 3.6). Negative numbers (blue) represent areas that are generally drier than the

mean, so the New Mexico/Arizona border is generally drier than other parts of the country in June.

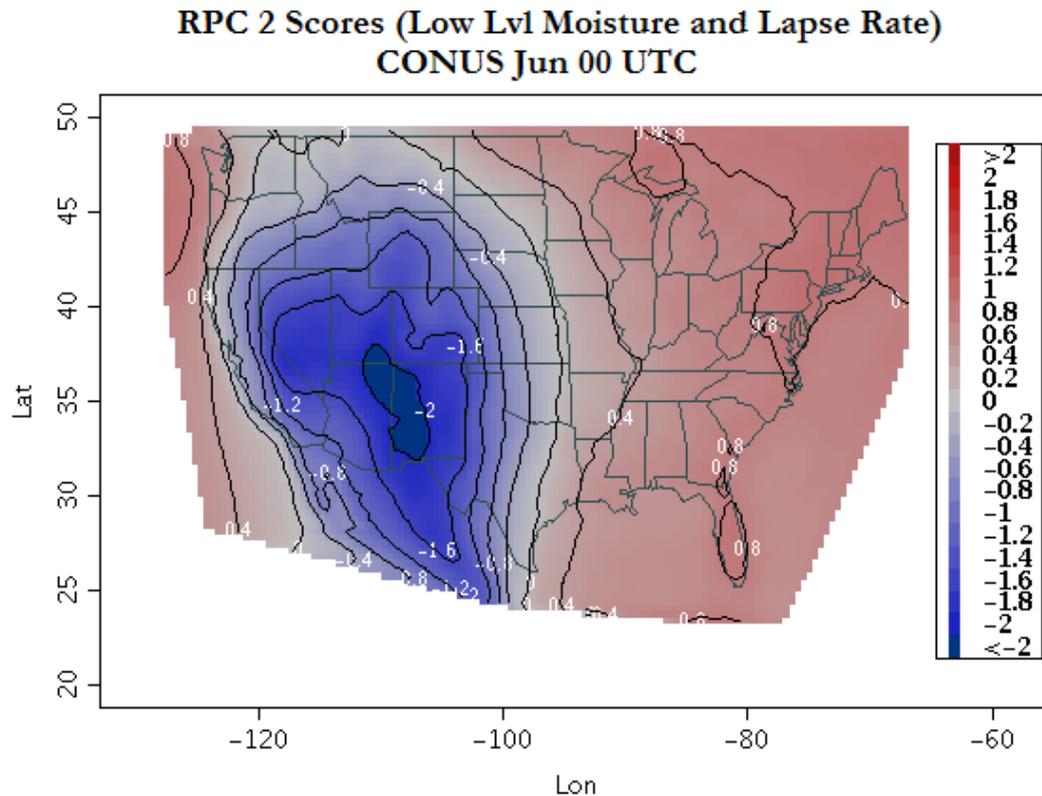


Figure 3.6: Scores from the CONUS domain that represent low level moisture and lapse rate.

Scores on PC11 look similar to lightning climatology around this month as shown in Fig. 3.7. Areas of higher lightning climatology amounts generally correspond spatially with higher score deviances (red). Similar results fall on PC1 in Alaska as both images display higher values in the central part of the state (Fig. 3.8). Meanwhile, the scores on AK PC3 and CONUS PC9 highlight areas of terrain based on relative vorticity differences and the Laplacian of the Geopotential Height field (Figures 3.9 and 3.10). See Appendix C for further examples of scores from June over both domains; other months show similar results.

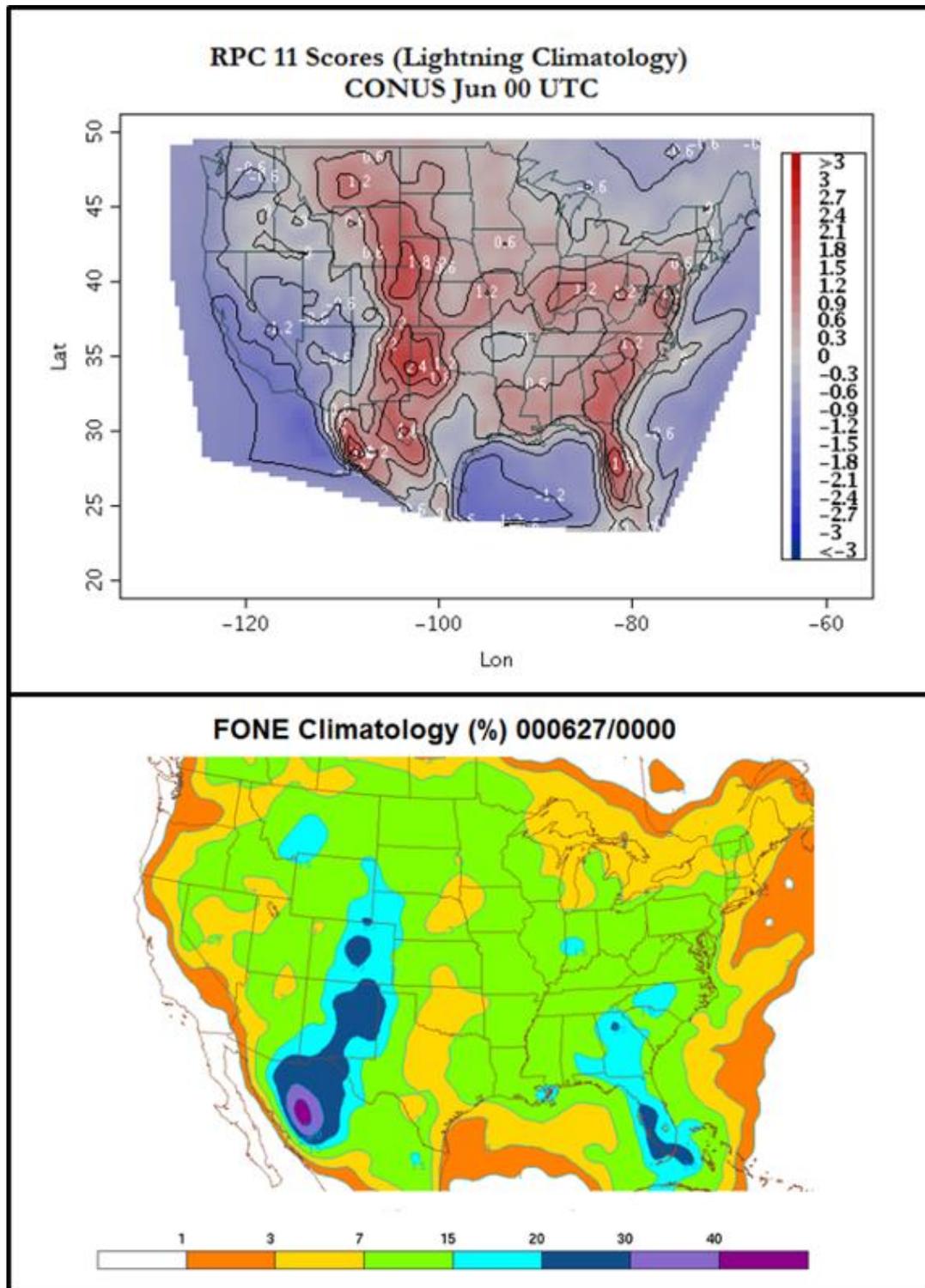


Figure 3.7: An example of scores (top) that correspond with high contributions from lightning climatology as seen in the example of FONE Climatology (bottom). Note that the Scores are plotted using an Equidistant Cylindrical projection and the lightning climatology is plotted using a Lambert Conic Conformal projection.

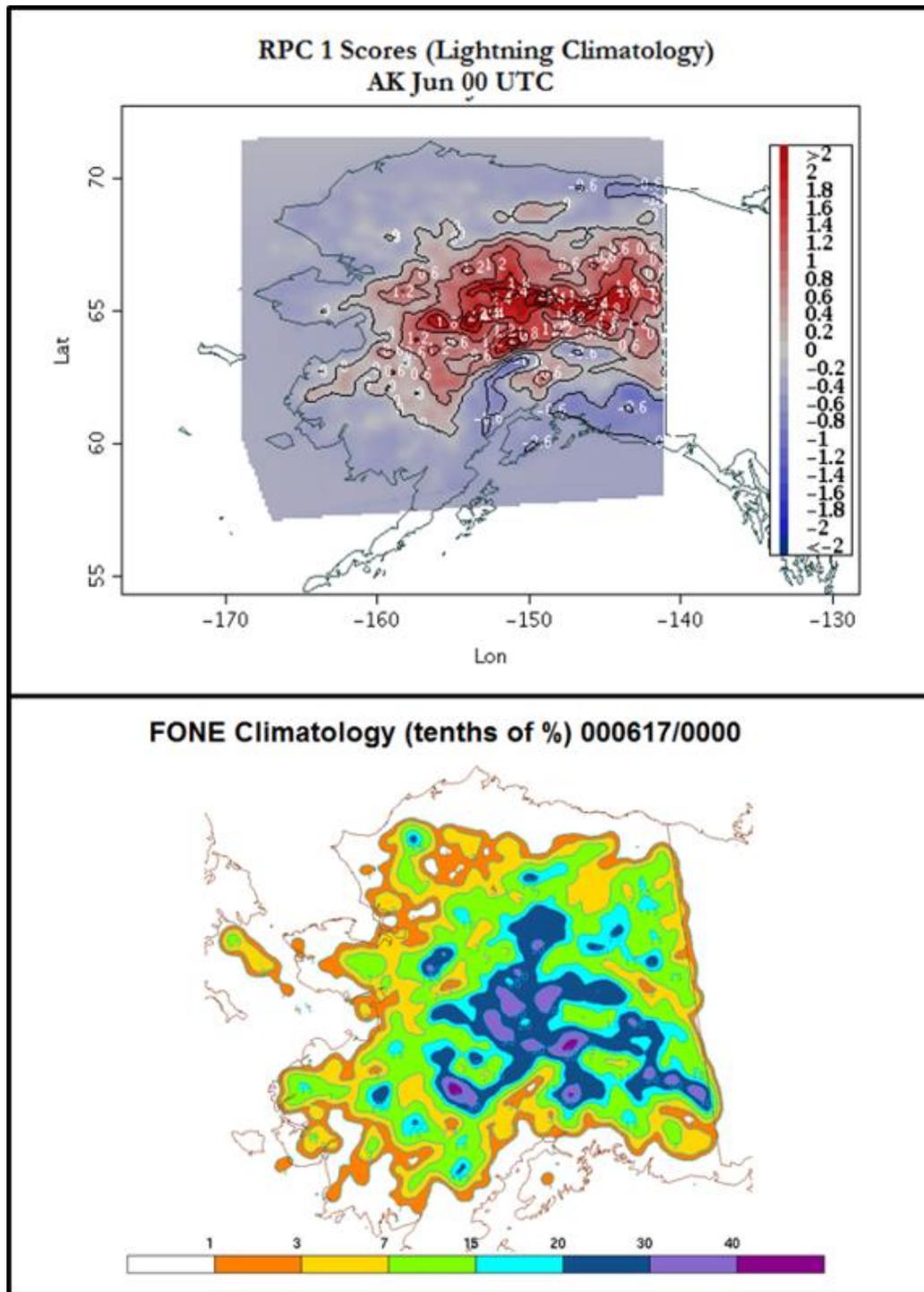


Figure 3.8: Lightning climatology and interactive predictors ($\text{MUCAPE} \times \text{lightning climatology}$) contribute the most to these scores. Note that the Scores (top) are plotted using an Equidistant Cylindrical projection and the lightning climatology (bottom) are plotted using a North Polar Stereographic projection.

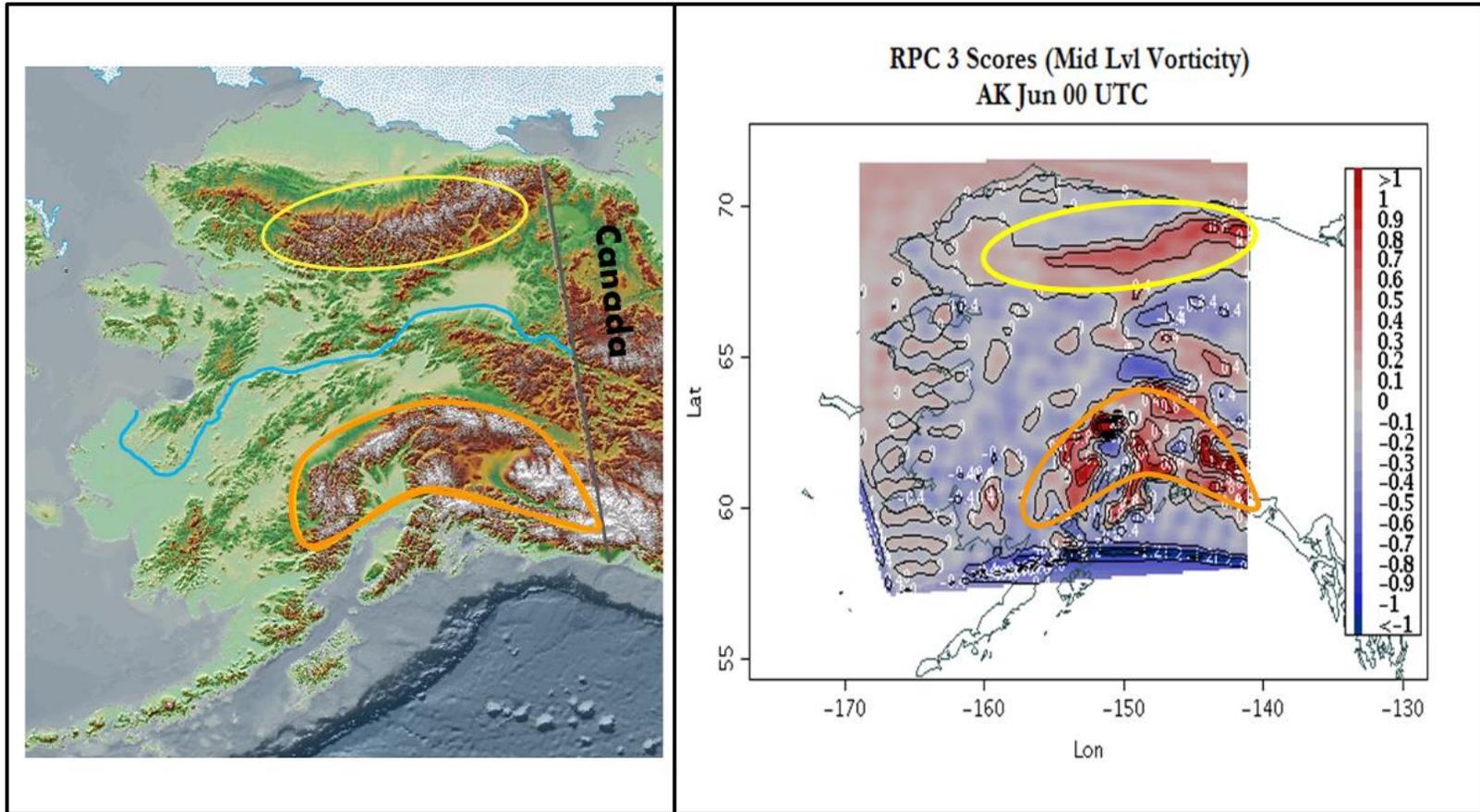


Figure 3.9: Score example that highlights terrain (right). The Lambert Conic Conformal relief map on the left from Armap.org contains shape overlays denoting the Brooks Range in the north (yellow), Yukon River (blue), and the Alaska Range in the south (orange). Mountainous regions, or regions with sharp terrain differences, appear as the high red standard deviations in the score plot on the right. Note that the score plot uses an Equidistant Cylindrical projection.

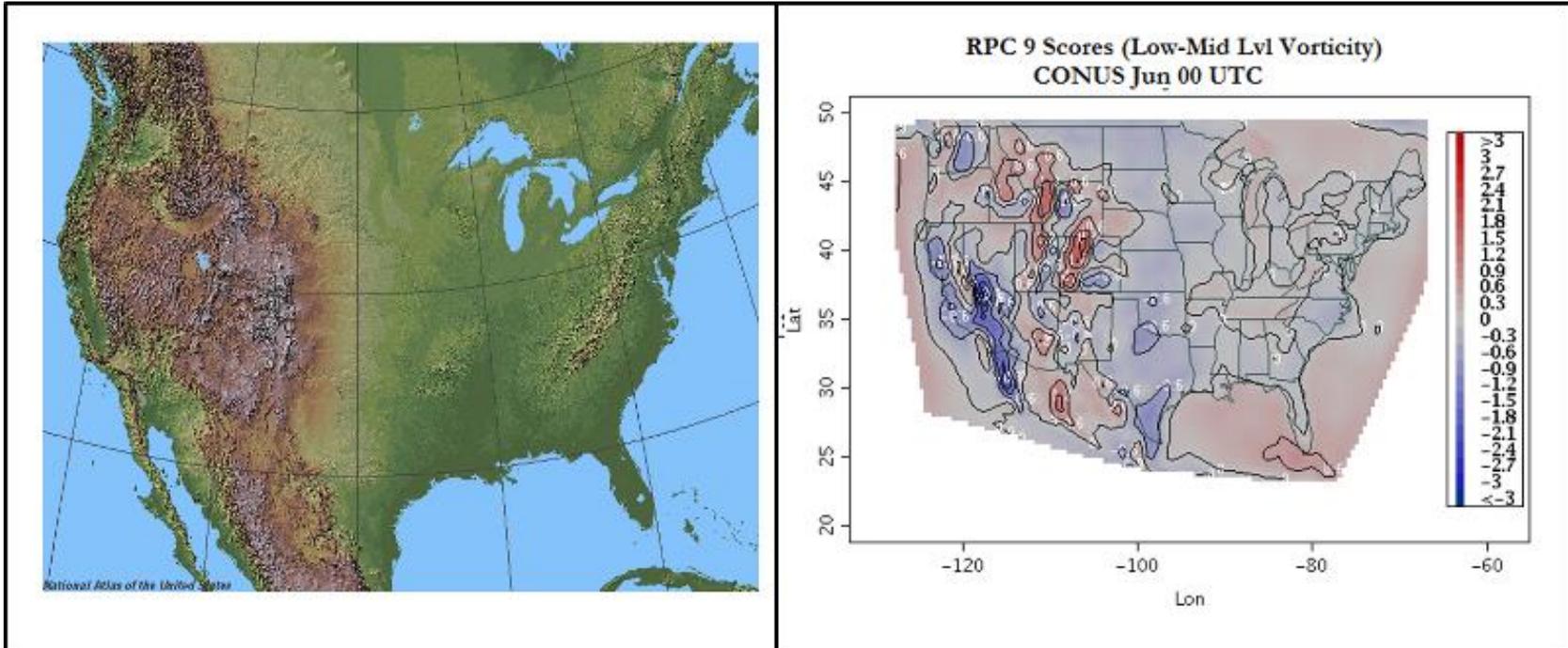


Figure 3.10: Scores showing high contributions from elevation changes (right). This CONUS example shows that elevation changes play more of a role than actual elevation as seen with the narrow blue strip of scores between higher mountain peaks denoted by reds. A Lambert Conic Conformal projection is used for the relief map (left), and an Equidistant Cylindrical projection is used for the scores (right).

3.4 Building Equations and Forecasts

Predictive equations were built for each binary predictand - all of the terms in Table 2.1 excluding TOTF and PTOT. Scores (in units of standard deviations) are related to observed values with a Generalized Linear Model (GLM) which relates such “dependent” and “independent” data by a certain family relationship (Nelder and Wedderburn 1972). The binomial family is chosen for logistic regression due to the binary nature of the observations. Nelder and Wedderburn suggest that independent binomial and independent Poisson models are similar, so the binomial method should suffice for this complex atmospheric data. Complexity in this case refers to the many variables in the atmosphere and the number of predictors used for this study in particular. Using a GLM ensures an accurate fit where the fewest number of parameters explain the most variability in the response variable.

Equations were created per response variable (or predictand) for each month because the predictand categories are not mutually exclusive. That is, equations are independent of each other (e.g., Equations 5 and 6) Predictand observations (e.g., FONE and DRYTH1) and PC Scores (S1-S12) are passed into a logistic regression model to create an associative equation.

$$\text{Eqtn}_{\text{FONE}} = \text{FONE} + \text{S1} + \text{S2} + \text{S3} + \dots + \text{S12} \quad (5)$$

$$\text{Eqtn}_{\text{DRYTH1}} = \text{DRYTH1} + \text{S1} + \text{S2} + \text{S3} + \dots + \text{S12} \quad (6)$$

For example, a predicted value for dry thunder (DRYTH1) does not know the predicted value for rainfall greater than a tenth of an inch. Dry thunder fields do not depend on the predicted values of FONE or PLTN/PLQT because the

dry thunder fields were passed in as distinctive independent data with uniquely developed equations. It is hoped that the fields will show similar trends, say low precipitation chances in regions of high dry thunder chances.

Only 00 UTC data was used to create equations due to peak heating and convective activity interest as mentioned earlier. Another reason 00 UTC has the focus is that model and reanalysis data contain the most up-to-date information of actual observations (e.g., from soundings) than other times of peak convective activity. This may introduce a slight low bias around 12 UTC (convective minimum time) and result in more misses because of the higher 00 UTC threshold. For example, CAPE and temperatures may be higher around 00 UTC, so the equations expect these higher values to predict dry thunder. Lower temperature values more likely around 12 UTC may not be enough to generate dry thunder probabilities due to the 00 UTC structure even if the environmental setup is reasonable for dry thunder at that time. However, it is hoped that the climatology amounts hold enough weight in development (as shown with PCA) that this potential bias will be reduced further. That is, the equation pulls information about the associated climatology value for the given forecast hour. A forecast for 12 UTC contains information about the associated pentad 12-15 UTC climatology.

GFS model data were chosen as input to the equations to create forecasts. This model forecasts over both domains used in this research, which means each region is subject to inherent GFS biases (as per the PP

definition). A standard operational 40-km grid covers the CONUS and needs no modification. The Pacific 35-km grid over Alaska is converted to a 10-km grid with the same interpolation method used on the NARR development data. Some differences from the development data to note: 1) Only the standard atmospheric variables are available instead of every 25-hPa, 2) CAPE is calculated using the virtual temperature correction with MUCAPE on a pressure difference layer from the surface to 300-hPa above the surface. Details about convergence, forcing, and wind may be reduced with the limited amount of pressure level information. Meanwhile, the CAPE adjustment may seem larger than the development data and contribute slightly to overprediction. This may be more of a problem at the 10-km Alaska scale. Generating forecasts from GFS data takes about 1-h to run and grid (2-h for both domains).

A CAPE and CIN filter is applied to the resulting FONE probabilities. Grid points with $CAPE \leq 50$ J/kg and/or $CIN \leq -50$ J/kg in the CONUS are set to 0 in the lightning probabilities for one or more flash in the CONUS. The filter of FONE in AK uses $CAPE \leq 10$ J/kg and/or $CIN \leq -50$ J/kg. Additionally, this predicted FONE field works as a mask for higher lightning values. That is, if the chance of FONE is predicted to be 0%, FTNR, FTEN, etc. are set to 0% as well. Locations are unlikely to receive FTNR if they do not receive FONE. This limit is separately applied as a post-filter – the equations are still independent due to the non-exclusive category definitions. None of these filters are applied to precipitation or dry thunder fields.

CHAPTER 4

Verification & Case Studies

4.1 Contingency Statistics

Lightning flash/dry thunderstorm forecasts follow similar guidelines and traits of other rare-event forecasts, such as the ones for severe weather. A general goal of such forecasts lies in predicting an area of interest several days in advance. Most of these predictions come in the form of probabilistic fields. One benefit to probabilistic forecasts comes from convenient comparison to climatological data. Forecast probabilities higher than climatology can suggest an event is more likely to occur as proposed by Brier (1950). Probabilistic fields also convey the uncertainty of a forecast. Antolik (2000) states, “[p]robabilistic formulation not only makes it possible to convey information regarding the uncertainty of forecasts, but also allows individual guidance users to decide how best to react, given this uncertainty.” All forecasts in this study use probabilistic output that range from 0 – 100% chance of an event/phenomenon occurring at a particular grid point.

Verification of continuous variables presents concerns with dimensionality as described by Murphy (1991). The amount of data needed for verification becomes so large that it either takes large amounts of temporal data (e.g., years) and/or many grid points to verify. Binning the forecast output, probabilistic in this case, into separate bins reduces this problem. Dimensionality would be too high to do every 1% category even with dichotomous observations (yes/no). Categories are binned into the

categories shown in Table 4.1. Dimensionality equals one less than the forecast categories (bins) multiplied by the resulting categories. With nine categories and two possible outcomes, the dimensionality is calculated as: $D=(9*2)-1=17$. This number represents the absolute minimum amount of observations/forecasts that can accurately verify a forecast system. Each available day over a given month counts toward this dimensionality data requirement, and doing point-to-point verification ensures adequate samples on the 10-km and 40-km grids.

One of the most common verification metrics comes from the Contingency Table (or confusion matrix) as shown in Table 4.2. Term *a* represent the number of hits; term *b* represents false alarms; term *c* describes misses, and term

Table 4.1: Binned categories and the exact ranges of consideration for each. Categories visually match intervals seen on forecast images.

Category (%)	Range (%)
0	0.00 – 0.99
1	1.00 – 2.99
3	3.00 – 6.99
7	7.00 – 14.99
15	15.00 – 29.99
30	30.00 – 49.99
50	50.00 – 69.99
70	70.00 – 89.99
90	90.00 – 100.00

Table 4.2: Format of the 2X2 Contingency Table used in this study: Y=yes, N= No.

		Observed	
		Y	N
Forecast	Y	a	b
	N	c	d

d counts the number of correct nulls. Calculations are made for each category listed in Table 2. Optimally, most values fall in terms a and d . Several metrics and scores derived from the contingency table support assessing forecast quality (Doswell et al 1990; Wilks 2006). Table 4.3 describes the formulation of several metrics of interest where N represents the total number of forecasts made in a particular bin.

Table 4.3: Name and formula for calculating common contingency metrics and scores.

Name	Formula
n	$a + b + c + d$
Frequency (f)	$\frac{a}{N}$
Base Rate	$\frac{(a + c)}{n}$
Probability of Detection (POD)	$\frac{a}{(a + c)}$
Frequency of Misses (FOM)	$\frac{b}{(a + c)}$
False Alarm Ratio (FAR)	$\frac{b}{(a + b)}$
Frequency of Hits (FOH) (Success Ratio)	$\frac{a}{(a + b)}$
Probability of False Detection (POFD)	$\frac{b}{(b + d)}$
Bias	$\frac{(a + b)}{(a + c)}$
Critical Success Index (CSI)	$\frac{a}{(a + b + c)}$
Heidke Skill Score (HSS)	$\frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)}$
Hanssen-Kuipers Skill Score (HKSS) (True Skill Score)	$\frac{(ad - bc)}{(a + c)(b + d)} = POD - POFD$
Extremal Dependency Index (EDI)	$\frac{\ln(POFD) - \ln(POD)}{\ln(POFD) + \ln(POD)}$

For dry thunderstorm prediction, high interests lie in metrics such as the Probability of Detection (POD), and Probability of False Detection (POFD). Metrics assessed in terms of time (forecast hours or FHR) reveal information similar to a time series. Forecasts from 00 UTC cycles are assessed at a given valid time of 00 UTC. Table 4.4 lists the forecast hour specifications that denote a Day which represents 00-00 UTC. Day 8 (FHR 168) was not included because this Day would not be within the 180-h limit at other cycles (such as 06 UTC).

June and July data are combined due to the small number of cases in June (5 days); significant dry thunder activity occurs in these months for both domains.

Table 4.4: List of forecast hours (FHR) that count as a Day for the contingency metrics.

Day	Cycle/FHR
Day 1 (D1)	00 UTC/F000
Day 2 (D2)	00 UTC/F024
Day 3 (D3)	00 UTC/F048
Day 4 (D4)	00 UTC/F072
Day 5 (D5)	00 UTC/F096
Day 6 (D6)	00 UTC/F120
Day 7 (D7)	00 UTC/F144

Though there are many categories of interest, examples will only be shown for the 15% (3%) bin for the CONUS (AK). This bin makes up an adequate portion of forecasts, and should be higher than the noise level to capture the coherent signal of a forecast. Additionally, the CONUS domain was split into western and eastern domains at 102° longitude (Figure 4.1). This division allows for comparison of DRYTH1 and DRYTH2 between the drier West and wetter East. Fire managers and crews in the West are more concerned with dry thunder fire ignition compared to the generally wetter East. Also, it was discovered that the GFS contained a significant 2-m

temperature and dewpoint bias ($\sim 10\text{-}20^\circ\text{ F}$) with most of the bias originating in the Central Plains during the summer of 2012. This bias was not corrected until September 2012 (Lapenta 2012, personal communication). Because the bias in the Central Plains would propagate eastward as the model forecasts move forward with time, scores are expected to be less accurate in the East than in the West. Dividing the full CONUS should help avoid misleading impacts from the bias.

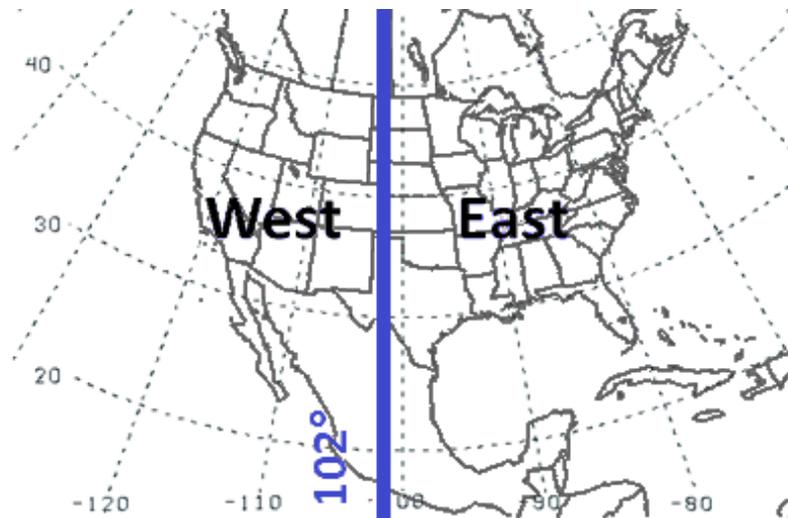


Figure 4.1: Image demarking West vs. East areas in this study.

Figure 4.2 compares POD and POFD across the full CONUS, East, and West domains for forecasts out to Day 7. FONE POD is notably higher than the dry thunder fields, and DRYTH2 is higher than DRYTH1 in CONUS and East setting. DRYTH1 POD exceeds DRYTH2 in the West. POD values decrease slightly with increasing forecast hour, as forecast accuracy decreases, in the full CONUS and East. Values of POD in the West remain relatively consistent even out to Day 7. POFD charts show very little change.

Results from the West are slightly different in that DRYTH1 has more false detections than FONE or DRYTH2. This could be a factor of overforecasting.

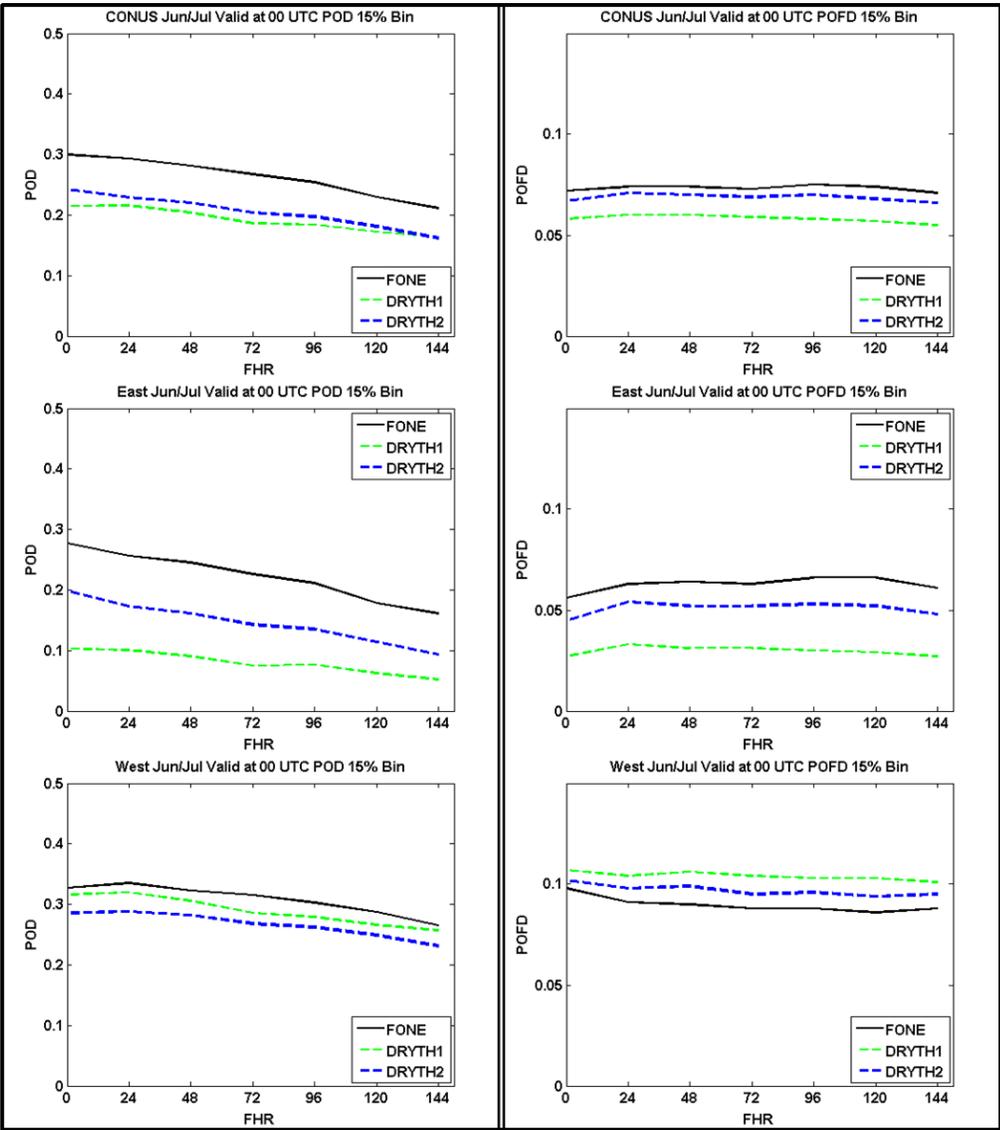


Figure 4.2: POD (left) and POFD (right) for the full CONUS (top), East (middle), and West (bottom) in the 15% bin. Note that the scales are different between POD and POFD.

FAR and FOH (Fig. 4.3) can describe a bit more about a system, but FAR results can be misleading in rare-event forecasting as high values are not uncommon (Doswell et al. 1990; Brooks 2004; Hitchens et al. 2013). As Brooks (2004) states, “[i]f missed events are considered costly, however,

much higher FAR values must be accepted.” Nevertheless, FAR values are shown here for completeness. High FAR values, a direct result from overforecasting, increase slightly as accuracy decreases with time.

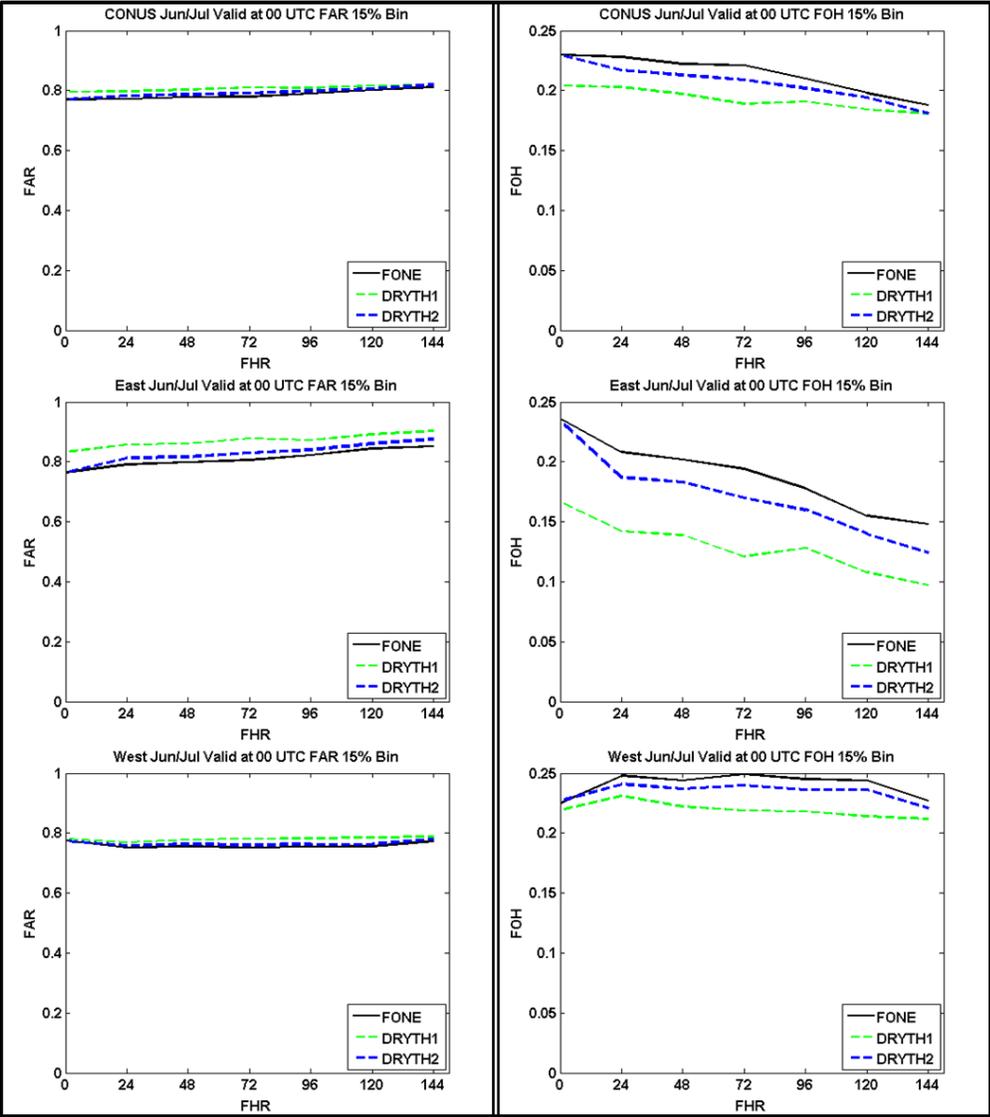


Figure 4.3: FAR (left) and FOH (right) for the full CONUS (top), East (middle), and West (bottom) in the 15% bin. Note that the vertical axis scales differ between FAR and FOH.

Results from the CONUS and East are consistently higher than the West, i.e., values for the West are better in terms of FAR. West values outperform in FOH results as well, as the hits remain fairly consistent even at

the longer lead times. The slight improvement at F024 may be associated with model spin-up. Results from the East decrease rapidly to F144 most likely due to the GFS biases.

Forecasts values at the 10-km AK grid (Fig. 4.4) follow trends similar to the CONUS with decreasing accuracy as time increases. DRYTH1 values perform better than DRYTH2 and FONE. Even though the FOH is lower in AK, the POD is higher compared to the CONUS domain. Such individual metrics reveal some information about the forecast system, but scores based on these metrics can provide further insight.

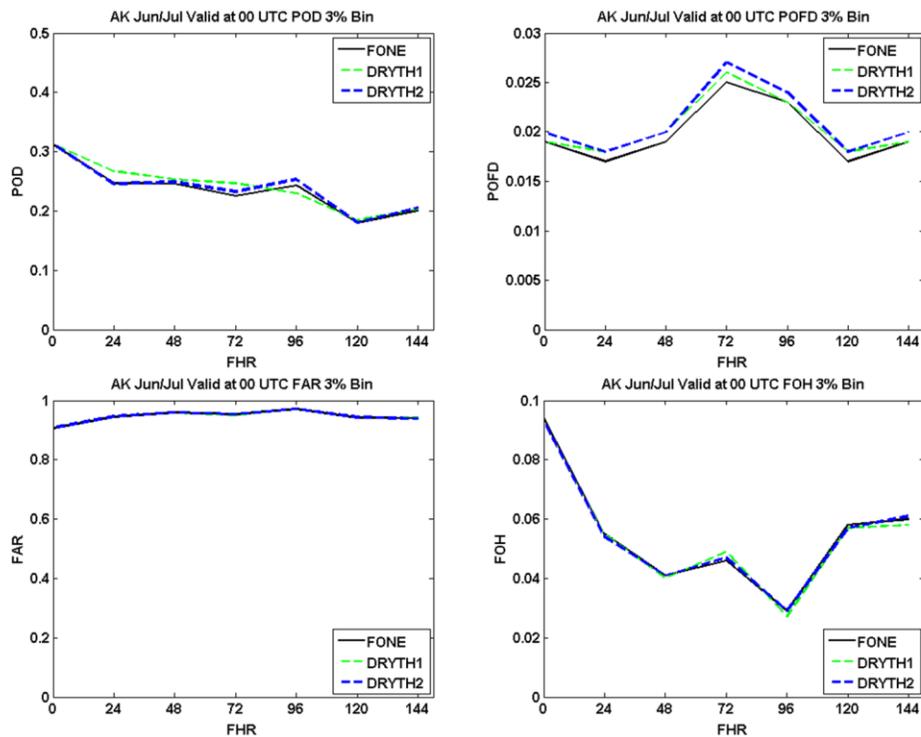


Figure 4.4: Values of POD (upper-left), POFD (upper-right), FAR (lower-left), and FOH for Jun and July data Valid at 00 UTC in the 3% bin. Note that the vertical axes are different for each term.

The Heidke Skill Score (HSS) represents a skill score from the contingency table that relates to relative accuracy. Values of 1 express perfect scores, and negative scores relate to forecasts that are worse than

the chance of random correctness (Wilks 2006). A related score, the Hanssen-Kuipers Skill Score (HKSS), equals the HSS if forecasts are unbiased. Generally, this score seems better for systems prone to overforecasting. Fig. 4.5 shows results from the CONUS, East, and West.

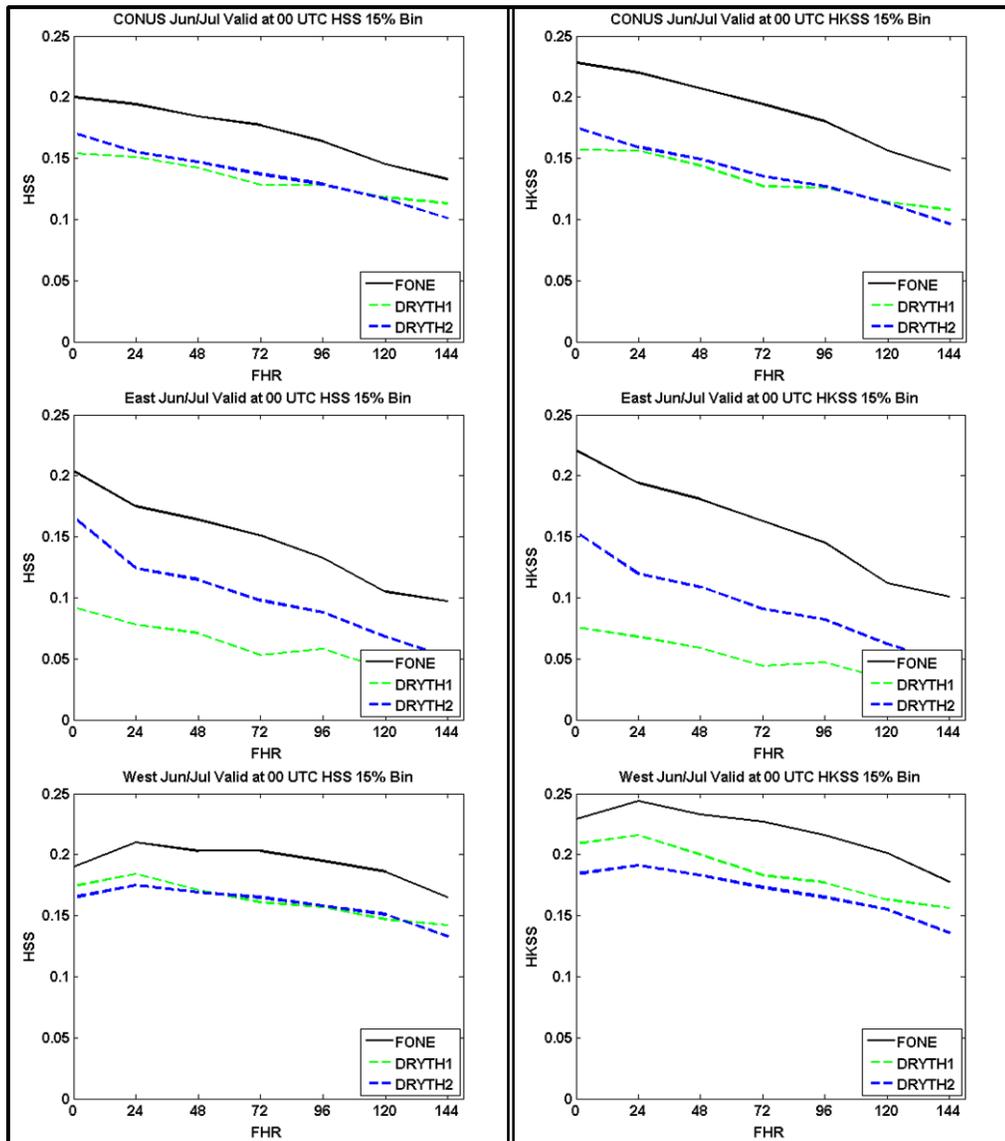


Figure 4.5: HSS (left) and HKSS (right) for the full CONUS (top), East (middle), and West (bottom) in the 15% bin.

A high bias is noted by the fact that HKSS values generally fall higher than the HSS values, especially for FONE. High values of HKSS in the West

contribute more to the higher results in the CONUS compared to the East. Values of HSS are not perfect, but they are better than random forecasts.

Results from AK (Fig. 4.6) show how the HSS follows a trend similar to FOH. Overforecasting ties directly into the higher HKSS values shown below, and the trend is similar to POD. Because the HKSS does not equal the HSS, and are larger, the forecasts contain a high bias.

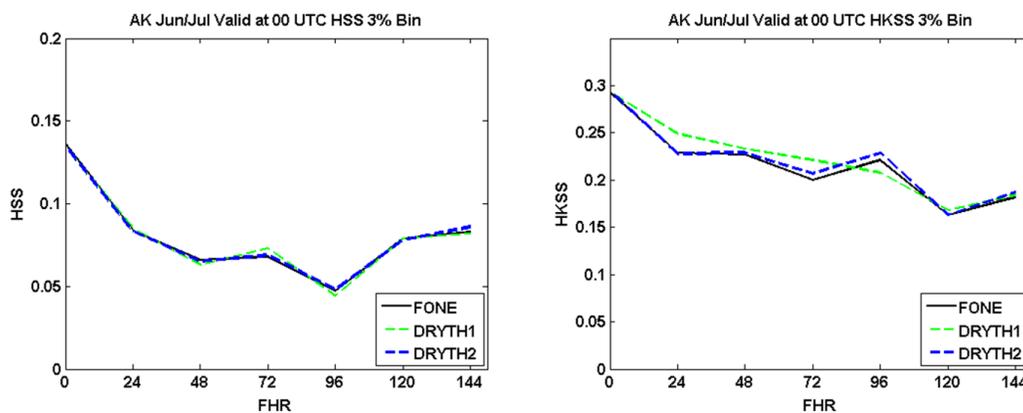


Figure 4.6: HSS and HKSS for the 3% bin for AK Jun and Jul data valid at 00 UTC. Note that the vertical axes are different.

Because lightning emulates rare-event phenomena for forecasting, the Extremal Dependency Index (EDI) is assessed (Ferro and Stephenson 2011; Wilson and Giles 2013). EDI values are related to POFD and POD as shown in Table 4.3; positive EDI relates to higher POD and negative EDI relates to higher POFD. Equal values of POD and POFD equate to 0 EDI. Results for the CONUS appear very similar to POD in the overall trend and performance between predictands (Fig. 4.7). It is encouraging that all values are positive, i.e. POD exceeds POFD even at the Day 7 time frame.

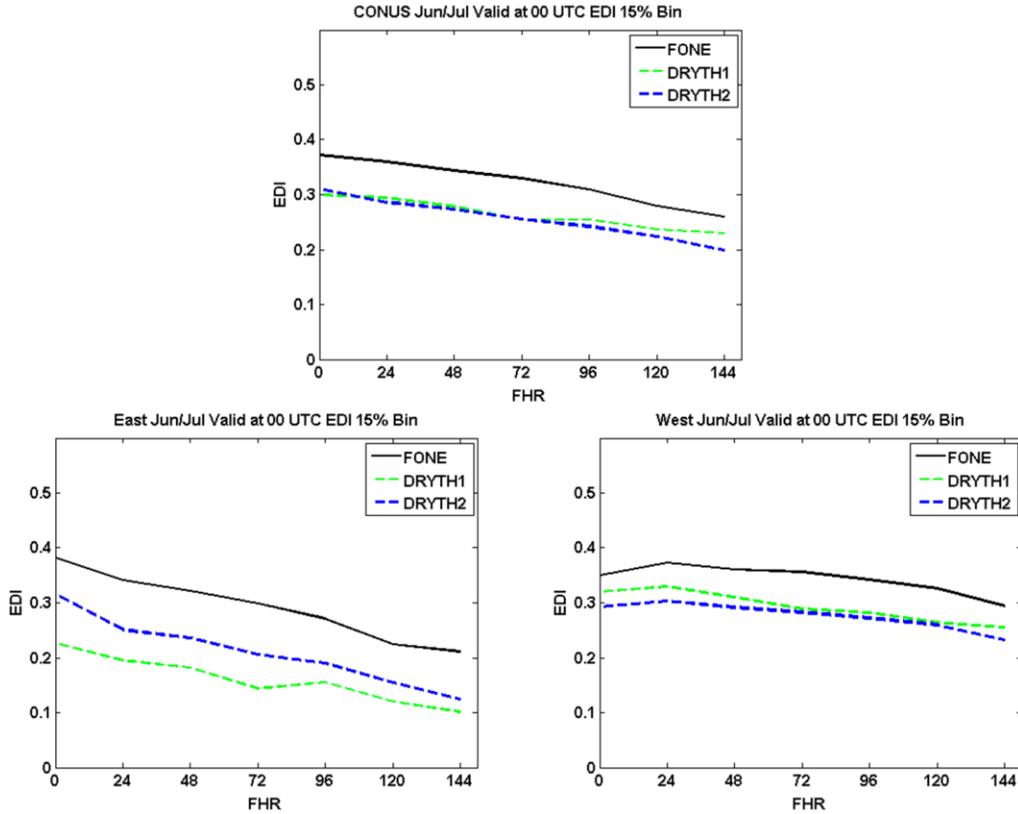


Figure 4.7: EDI values in the 15% bin for the full CONUS (top), East (bottom-left), and West (bottom-right).

Alaska values of EDI reiterate that POD is higher than POFD even at extended ranges (Fig. 4.8). This plot resembles the POD and HKSS plots, and shows that DRYTH1 typically has slightly better forecasts than DRYTH2 and/or FONE forecasts. A minor decrease in forecast accuracy with time is suggested, but consistent positive values exhibit that the forecasts contain skill even at further time ranges.

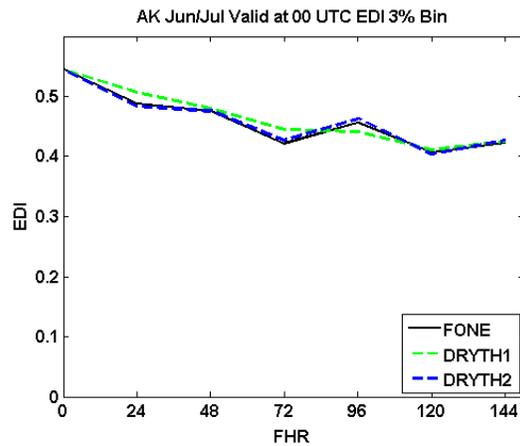


Figure 4.8: EDI plot for AK Jun and Jul data Valid at 00 UTC for the 3% bin.

The Brier Score (BS) essentially represents the mean squared error of the forecasts. Values closer to 0 denote better forecasts. Additionally, the Brier Skill Score (BSS) works as a measure of accuracy related to climatology where larger values represent superior scores (e.g., Fig 4.9).

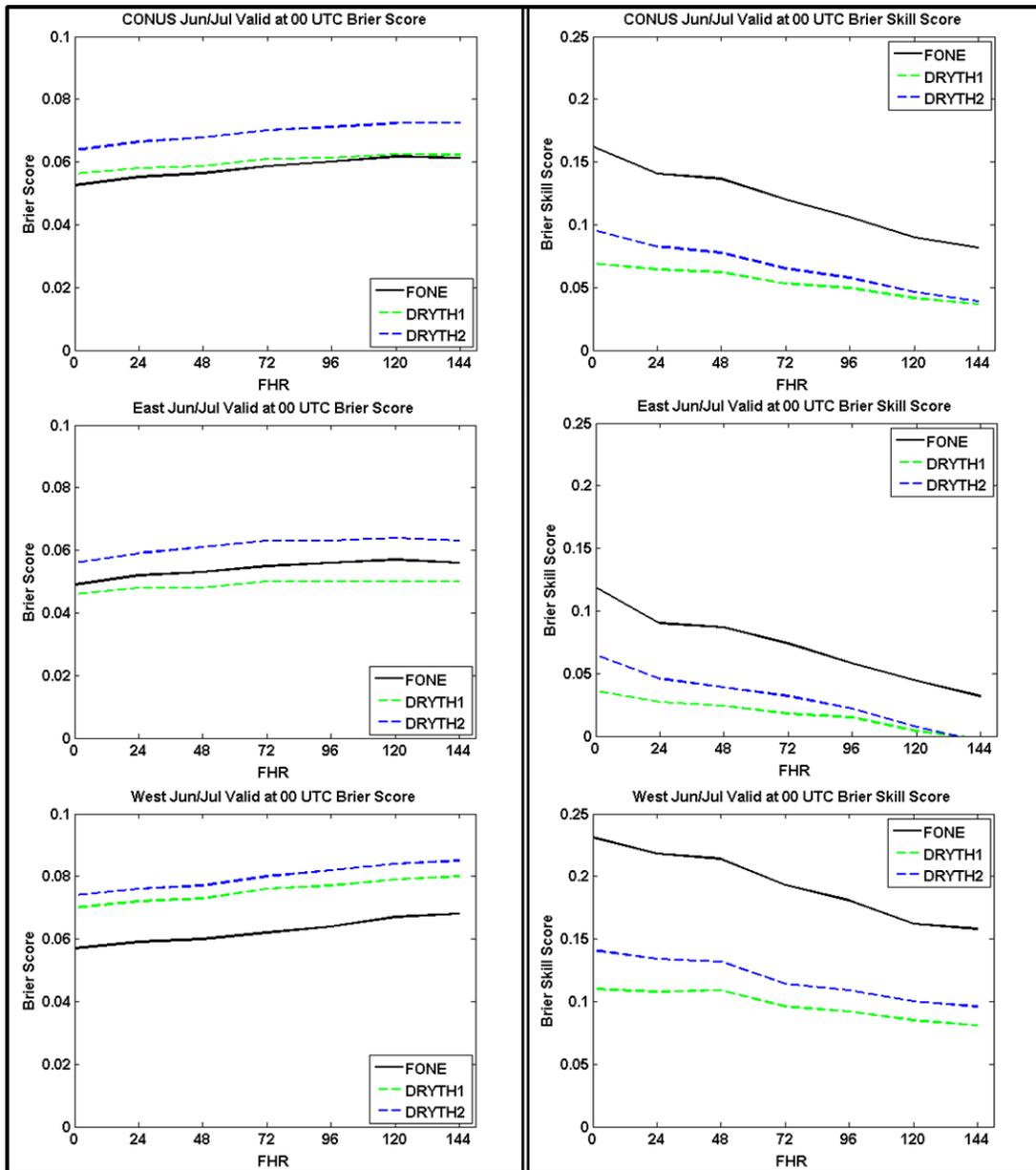


Figure 4.9: BS (left) and BSS (right) for the full CONUS (top), East (middle), and West (bottom). Note that the vertical axes are different between BS and BSS.

Values of BS in the CONUS are not too low, and the relatively low BSS values in the full CONUS and East domain suggest the forecasts could use some improvement. While the West has slightly higher BS, the BSS is also higher in this area. That is, results in the West outperform forecasts of climatology.

Plots for AK convey a decreasing reliability/resolution trend with a few interesting caveats (Fig. 4.10). BS values are surprisingly higher at F000 compared to other times until F072. One possible explanation comes from the fact that models and certain schemes need time to “spin-up” the physics of certain phenomena. For the BSS, values are highest at F000 where forecasts exhibit more reliability, and become more negative (closer to climatology) at further time scales.

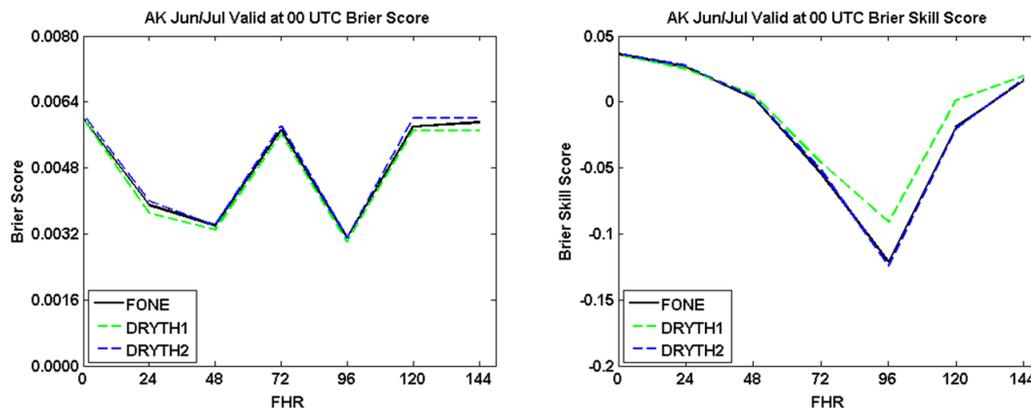


Figure 4.10: BS and BSS for AK Jun and Jul data valid at 00 UTC in the 3% bin. Again, scales are different between BS and BSS.

4.2 Performance Diagrams

Instead of the metrics and bins individually, a Reliability (or Attributes) Diagram conveys information about reliability and resolution of a given

forecast system which conveys different strengths and weaknesses of a forecast system (Wilks 2006). Because points on the reliability diagram may become jagged and misleading with a small sample size, any bins that do not contribute to 0.01% of the total forecasts (N) are not shown. A histogram of the forecasts per bin is overlaid as this information is otherwise not apparent. An example from the CONUS is displayed in Figure 4.11 below.

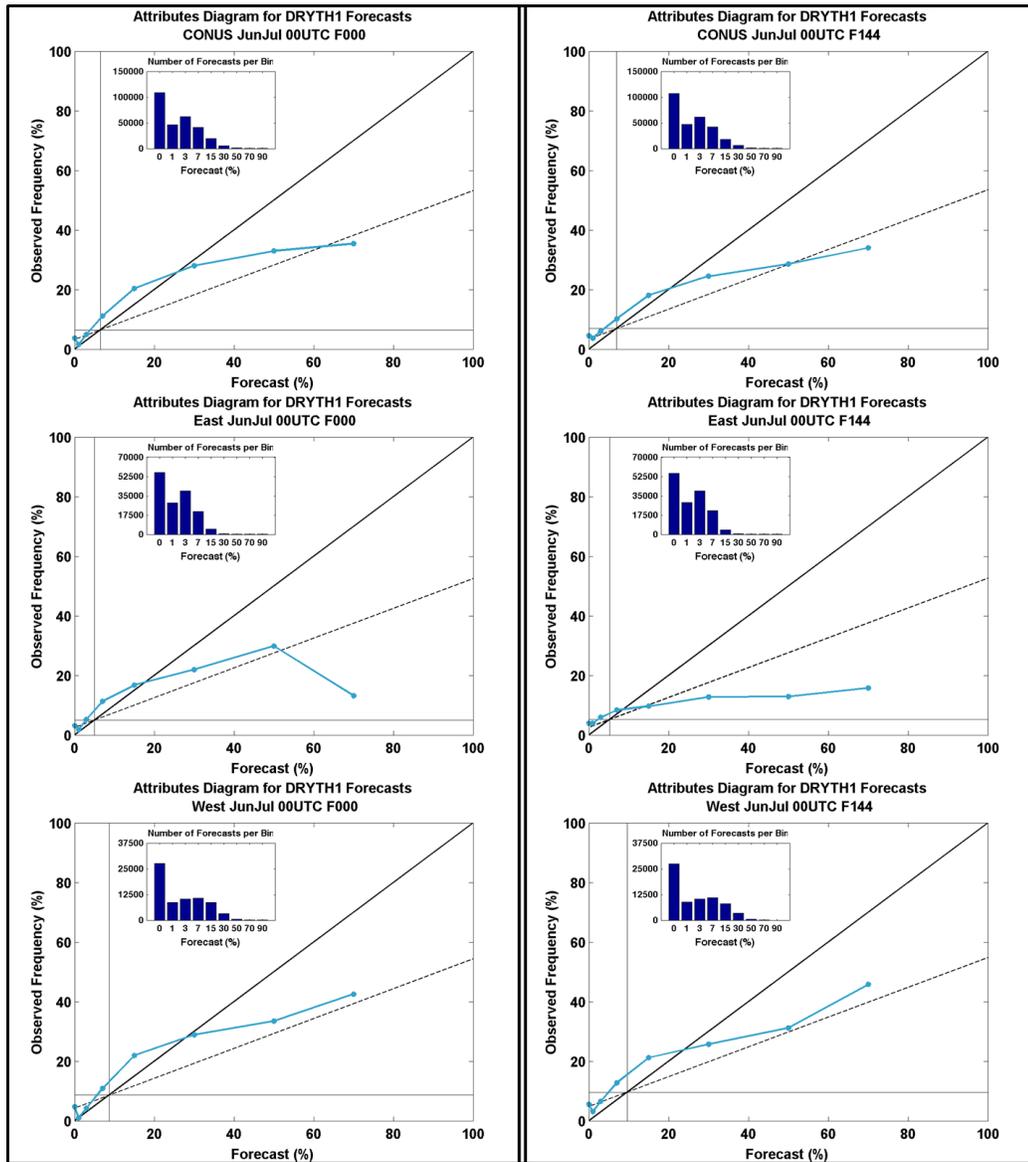


Figure 4.11: Attributes Diagrams for Day 1 (F000, left) and Day 7 (F144, right) for the full CONUS (top), East (middle), and West (bottom).

Results from the CONUS show overforecasting from 1-15%, and underforecasting for the higher bins at F000. The 90% bin is not shown because the number of forecasts falls below the 0.01% threshold utilized in this study. This trend of underforecasting at lower bins and overforecasting at higher bins is seen in the East and West, and at the later time of F144. Higher values in the CONUS are more aligned or below No Skill at F144.

The East exhibits poorer results compared to the CONUS or the West, especially at the higher categories. Total forecasts made in the 70% bin are likely very low, but just above the threshold, which causes the sharp decrease in the frequency line. By the F144 (Day 7) forecast time, most of the values fall to create a practically horizontal line. This result suggests that the forecasts cannot resolve areas any better than climatology. Meanwhile, values from the West are reasonable at the F000 and F144 time frames. Underforecasting and overforecasting occur, as observed in Bothwell (2002), but none of the bins fall below No Skill. That is, the West shows more skill than the East or full CONUS according to this diagram.

A DRYTH1 sample from AK is shown below in Figure 4.12. The forecasts are mostly reliable, but actually underpredict in the 1%, 3%, and 7% bin in the F000 time frame. By the 15% bin, overprediction drives the value down to the No Skill line at F000. Higher bins at F144 have no resolution as the line parallels the No Resolution line, so forecasts are similar to climatology. Inspection of the forecast histograms reveals why some of

the higher categories seem less reliable: the total number of forecasts in bins >1% are a small fraction of the forecasts in the 0% and 1% bins.

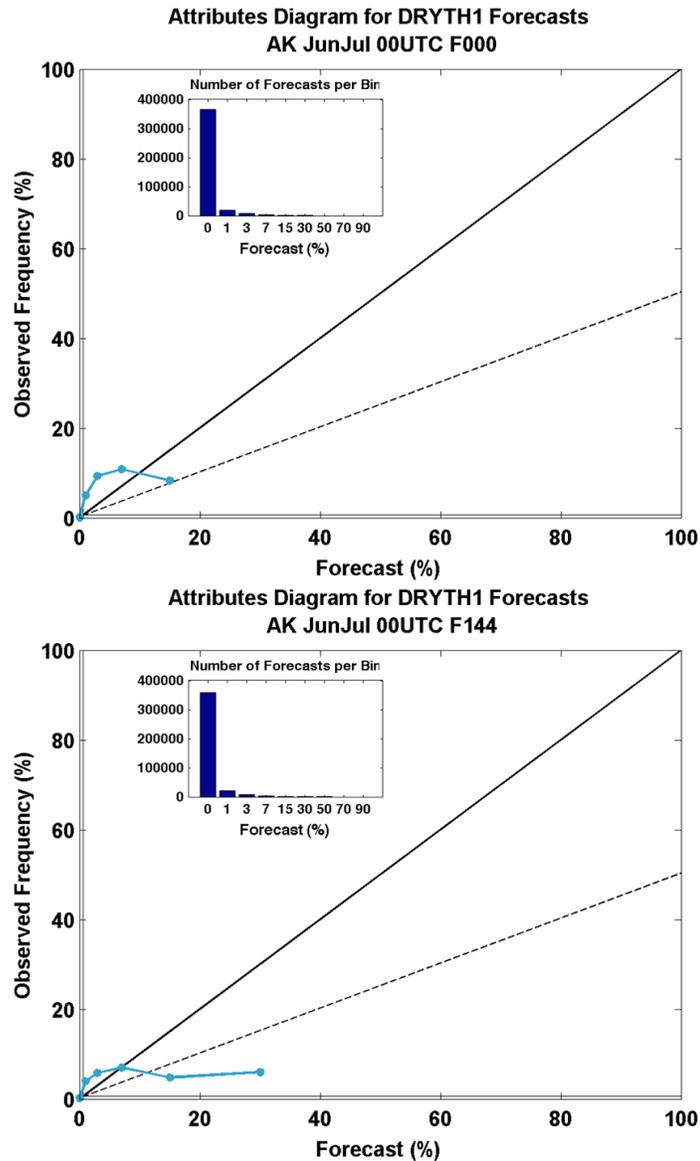


Figure 4.12: Reliability diagram for DRYTH1 in AK for Jun and Jul data Valid at 00 UTC. Points related to Day 1 (F000) are on the top, and points related to Day 7 (F144) are on the bottom.

Recall that the 15% (3%) bin was assessed for the CONUS (AK) in Section 4.1. These bins were chosen to represent coherent forecast signal,

and the Attributes Diagrams validate this statement (Fig. 4.13). More specifically, this bin is the first bin above sample climatology with adequate reliability. The bins exhibit some underforecasting, yet the histograms reveal a sufficient amount of forecasts fall within these bins. Forecasts in this bin can be considered as important because they exceed the sample climatology. Note that though these bins are greater than the sample climatology of observed events, the values may not be larger than the observed pentad climatology. Forecasts may but are not required to exceed pentad climatology to be considered significant. This will be covered more in the case studies shown in the next section.

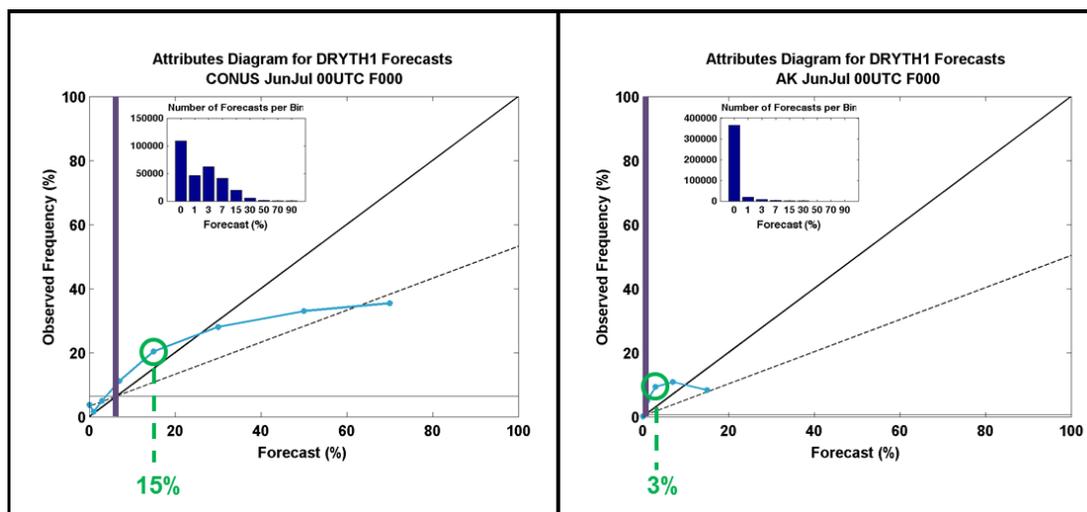


Figure 4.13: CONUS (left) and AK (right) Attributes Diagrams for 00 UTC and F000. Green circles highlight the bins that are above the sample climatology (vertical purple line).

POD, FOH, Bias, and CSI can be assessed using a Performance (Roebber) Diagram (Roebber 2009) as seen in the CONUS example below (Fig. 4.14). Values of FOH (Success Ratio) vs. POD are plotted as the main

lines on the chart. Curved reference lines represent values of constant CSI; CSI increases from (0, 0) to (1, 1). Dashed lines represent constant Bias. The 1-1 line symbolizes Bias equal to 1 (perfect); lines above represent Bias greater than 1, and vice versa for below. Performance diagrams are zoomed into a max POD/FOH of 0.5 for clarity. Values are plotted for each of the 7 Days, yet the exact order is not implicit in the chart. The individual charts of POD/FOH can assist in this case (Figs. 4.3 and 4.4). In the full CONUS domain, FONE is higher than DRYTH1 and DRYTH2 in the POD/POFD ratio. There is a slight high Bias starting out, yet the Bias and CSI decrease with increasing forecast hour – another sign that forecasts are not as skillful at longer ranges.

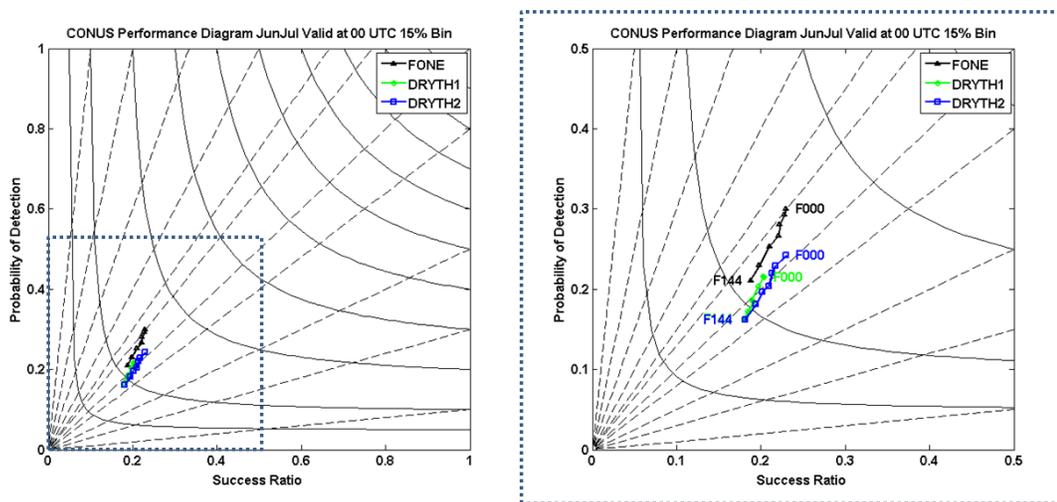


Figure 4.14: Performance Diagram for the CONUS. The full diagram is on the left, and the zoomed plot on the right is denoted by the dotted blue square. Starting and ending times are labeled.

East and West domain Performance Diagrams are shown in Figure 4.15. FONE is prominently better than DRYTH1 and DRYTH2 in the East,

and DRYTH2 outperforms DRYTH1 as well. FONE and DRYTH2 Bias values are very close to 1, while DRYTH1 has a Bias less than 1. The lower values of DRYTH1 may suggest a smaller sample size of forecasts compared to DRYTH2 or FONE, i.e., storms are generally wetter in the East, so the total number of DRYTH1 observations is expectedly lower. West results show improved FONE compared to the dry thunder fields, but DRYTH1 outperforms DRYTH2. CSI is higher, and does not decrease as much with time compared to the East.

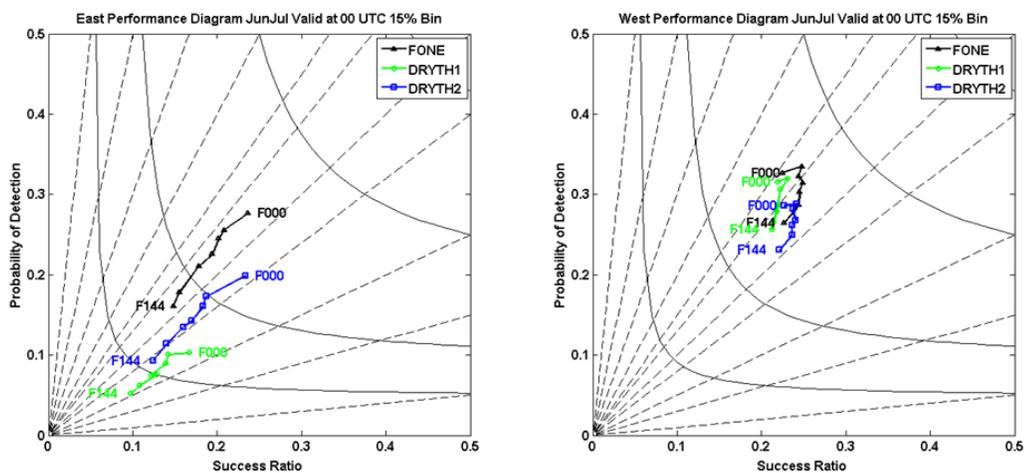


Figure 4.15: Performance diagram for the East (left) and West (right) – axes are zoomed into a subset range of the full diagram.

In Alaska, a high Bias exists in FONE and the dry thunder fields as all of the lines fall in the upper left side of the chart (Fig. 4.16). As previously noted (Fig. 4.4), POD is larger than FOH, and the decrease and increase in CSI as forecasts go out in time appears as well. Again, these higher values of POD and Bias are attributable to the overforecasting and overall lower

amounts of forecasts at the 10-km AK domain compared to the 40-km CONUS domain. Recall that most forecasts only fall within the 1-3% bins in the AK domain.

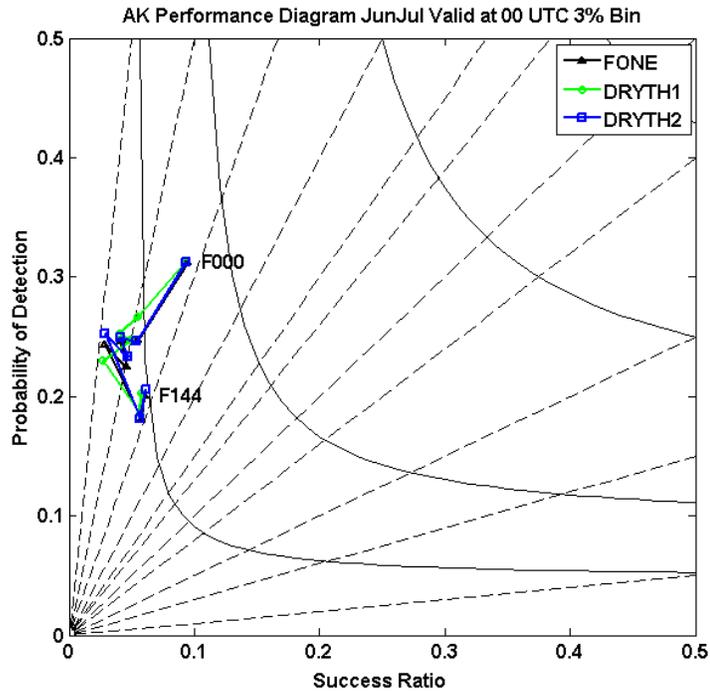


Figure 4.16: Performance Diagram for AK for F000-F144 – axes are zoomed to a subset range.

The Receiver (or Relative) Operating Characteristics (ROC) Diagram is another performance metric chart. A ROC Diagram evaluates overall performance of a system by comparing the POFD (also known as False Alarm Rate) to the POD (Mason and Graham 1999; Wilks 2006). An example from the CONUS is show in Fig. 4.17. The dashed 1-1 line represents No Skill from a forecast of always no (0, 0) to always yes (1, 1). Additionally, a performance measure, the Area Under the Curve (AUC), can be obtained from this chart. AUC values of 1 are perfect, and values of 0.5

equal the No Skill line. For this study, AUC values are found using the trapezoidal method.

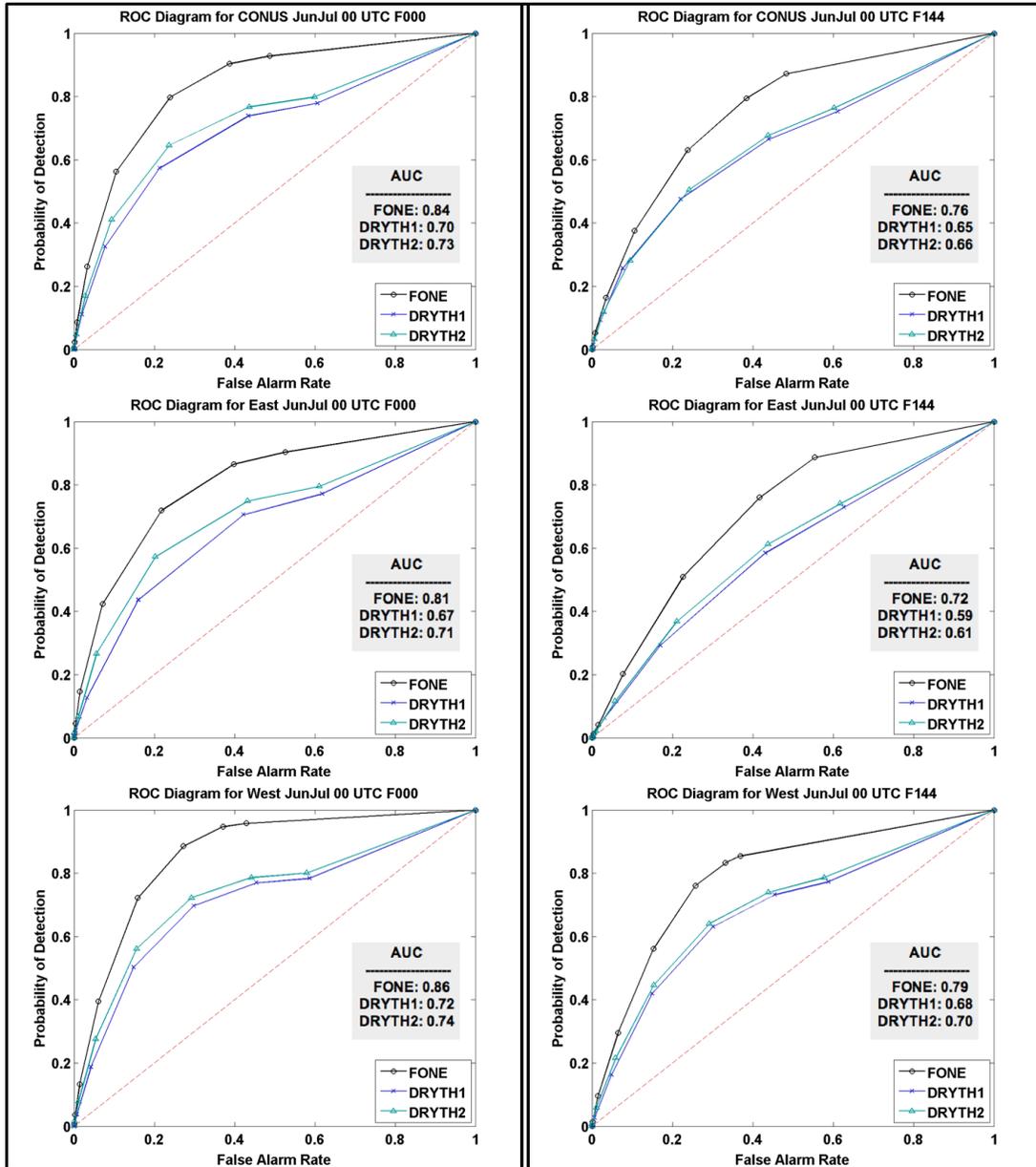


Figure 4.17: ROC diagrams for Day 1 (F000, left) and Day 7 (F144, right) for the full CONUS (top), East (middle), and West (bottom).

Values from Alaska show that the dry thunder and FONE forecasts have good performance values (Fig. 4.18), especially with AUC values around 0.88 at F000. It is seen that both dry thunder fields marginally outperform FONE as noted by the AUC value and the fact that the black line (FONE) is slightly below DRYTH1 and DRYTH2 for part of the curve. Lines and AUC values are expectedly lower at F144, but still represent good performance from lightning and dry thunder forecasts at the 10-km grid.

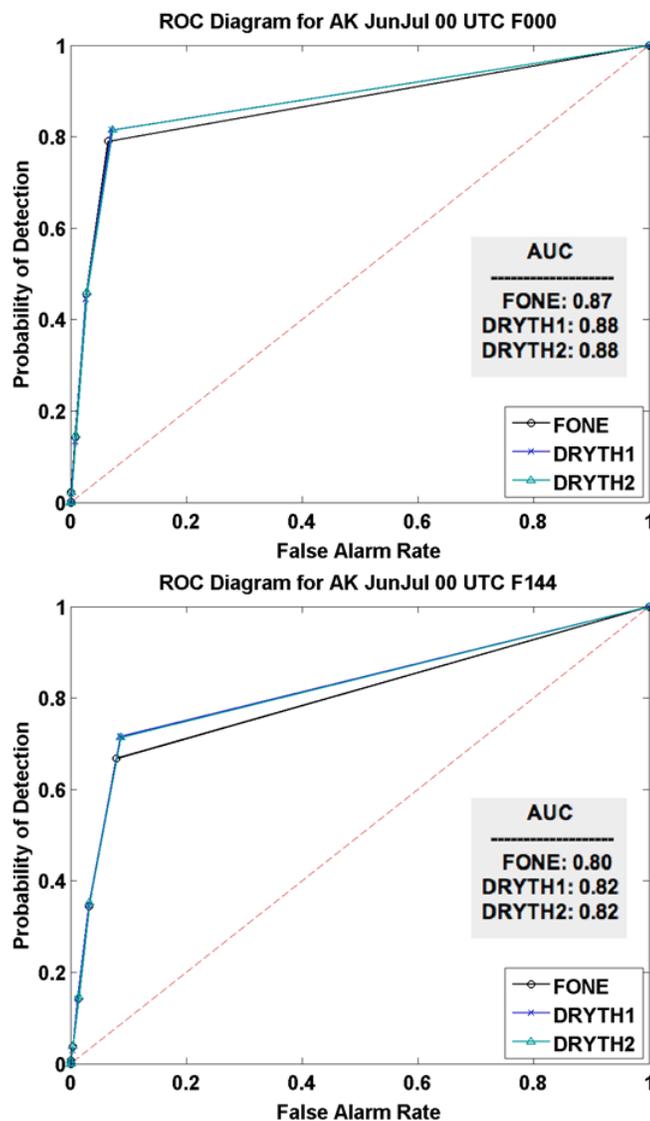


Figure 4.18: ROC diagrams for Jun and Jul data valid at 00 UTC for F000 (top) and F144 (bottom).

4.3 Case Studies

While quantitative metrics describe the true objective performance of a system, forecasters and other users often benefit from more qualitative assessments. Four case studies (2 per region) are shown as examples of the apparent visual benefit of dry thunder forecasts. Though not a requirement, most of the cases presented relate to notable fire starts, and others highlight spatial variability and accuracy.

4.3.1 CONUS Bonita Complex Fires

Several large fires plagued the CONUS during the continuing drought of 2012. Early in July, dry thunderstorms ignited two fires, the Bonita Fire and Iron Fire, in the Bonita Complex located in eastern Oregon. Figure 4.19 highlights the approximate location of these fires. Around 18,000 acres were burned in a little over a week and impacted the town of Westfall, OR (InciWeb 2012).

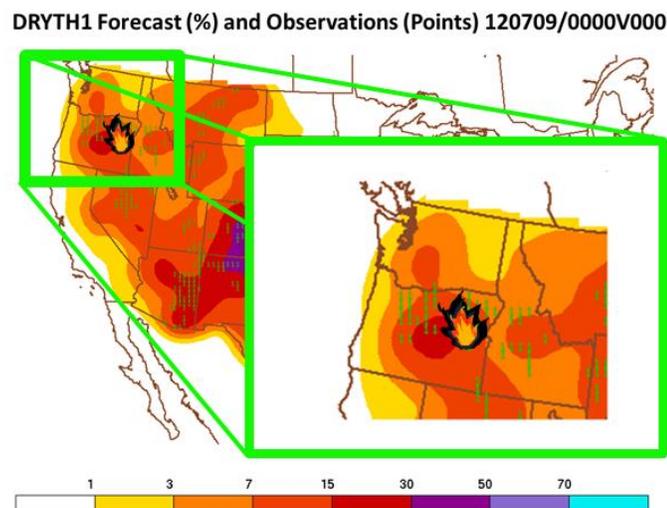


Figure 4.19: Bonita Complex DRYTH1 forecast, observations, and approximate fire location.

Figures 4.20 and 4.21 show forecasts from Day 7 (F144) up to the day of the event (Day 1, F000). Color fills represent the chance of DRYTH1 occurring at a given point. Observations of DRYTH1 are overlaid as green “1’s”, yet these do not represent total flashes within a box. Because DRYTH2 forecasts are extremely similar to DRYTH1 on the 10-km grid, only examples of DRYTH1 are presented. Forecasts are for 00 – 03 UTC.

Day 7 forecasts may not show extremely prominent values for DRYTH1, yet a noticeable area of 15% probability appears by Day 6 (F120) in central Oregon. This “bulls-eye” region remains fairly consistent for the next 3 days (through Day 3). On Day 2 (F024), the probabilities decrease slightly.

Still, the area of interest was noticeable by Day 6 and persisted up to the Day of the event. The fires started near the eastern edge of the 15% contour in the 7-15% range – refer back to Figure 4.19. Of course, the dry thunder probabilities are not directly relatable to fire start locations, so fire starts outside of the main “bulls-eye” of forecasts are acceptable. Perhaps these areas of higher probabilities are more closely related to total flashes, and that can be an area of future exploration.

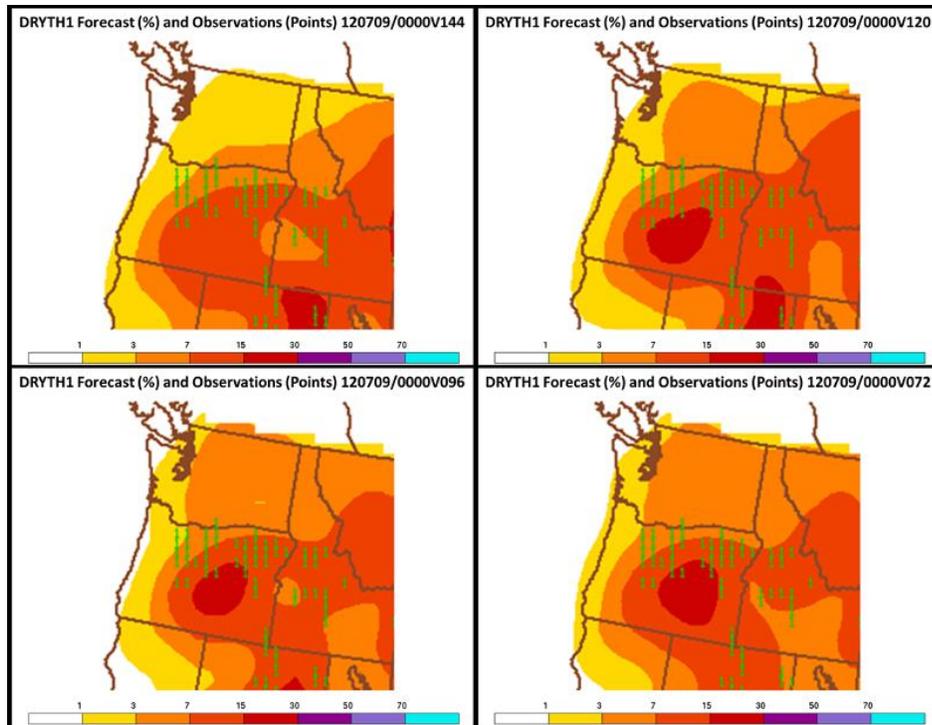


Figure 4.20: DRYTH1 forecasts and Observations for Day 7 (F144) to Day 4 (F072).

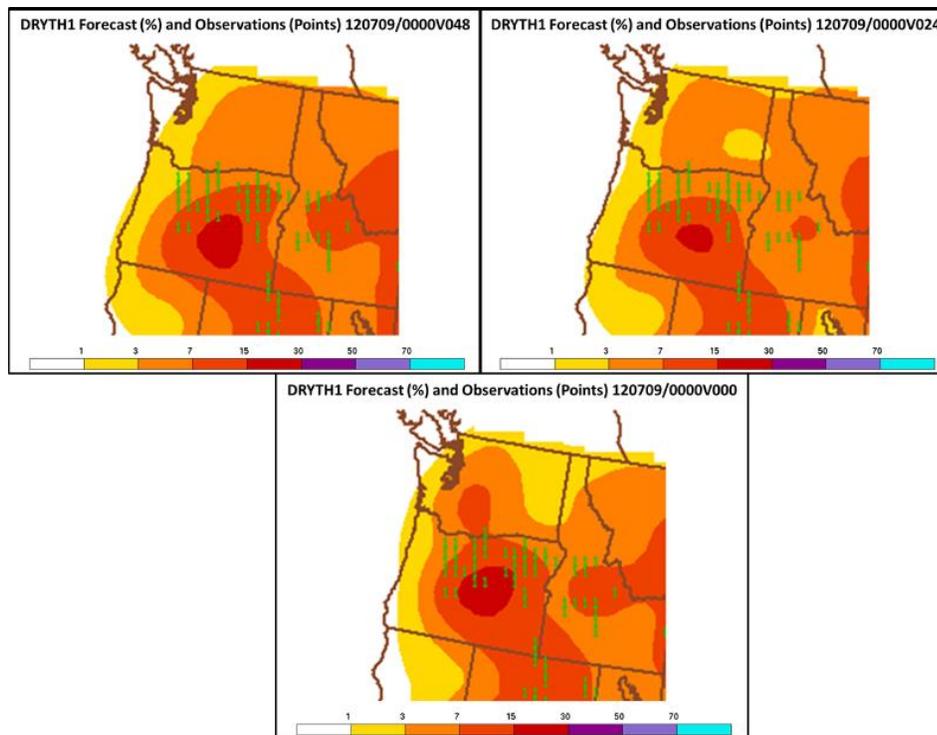


Figure 4.21: DRYTH1 Forecasts and Observations for Day 3 (F048) to Day 1 (F000) - the day of the event.

With this discussion of a DRYTH1 forecast “bulls-eye”, a comparison to climatology could reveal the actual strength of such a forecast. Figure 4.22 shows the forecast and observations from the day of the event, and also shows the pentad climatology value for the day. Fire symbols denote the approximate location of the Bonita Complex.

The Bonita Complex falls in a 7-15% forecast value and likely falls closer to the 15% range. Climatology values for this time do fall in the 7% range, but are likely closer to the low end. Values of only 3% fall in the place of the 15% forecast “bulls-eye” which may suggest that the climatology values are on the lower end of the 7-15% range. While forecast values of DRYTH1 are not significantly higher than climatology at the Bonita Complex location, forecasts are significantly higher in other portions of the state. Other DRYTH1 observations may have ignited some other fires, but focus remains on the larger starts.

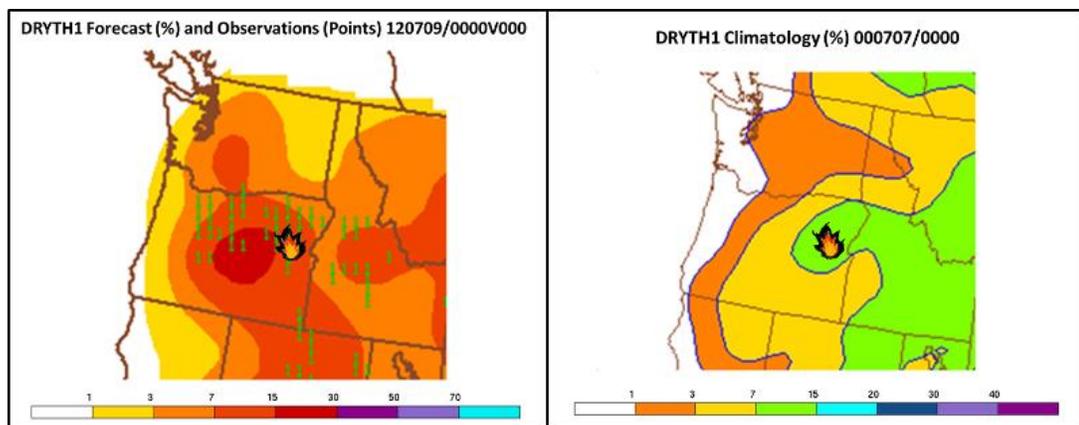


Figure 4.22: DRYTH1 Forecast for F000 on 09 July 2012 00 UTC (left) and the pentad climatology from the associated range (07 July, right). Note that the scales have similar increments and units, but the colors are different.

4.3.2 CONUS Mustang Complex Fire

Another large fire started in late July 2012 in the Pacific Northwest. Lightning ignited a fire in eastern Idaho near the Idaho/Montana border (Fig. 4.23). Dry fuels in a forested area made this fire difficult to bring under control, as it burned over 300,000 acres in a little over 3 months (InciWeb 2012). The Mustang Complex contributed more to hazardous air quality for towns east of the area compared to actual property damage from the fires.

Similar to the previous case study, Figures 4.24 and 4.25 step through the Day 7 through Day 1 forecasts leading up to 29 July 2012. The pattern is not as consistent with these forecasts as the area of interest fluctuates between 3% and 7% bins. Comparisons with the observations may suggest that the pattern is fairly consistent up to the Day 2, but is somewhat lost at Day 1 as noted by the reduction of forecast probabilities in Montana.

DRYTH1 Forecast (%) and Observations (Points) 120729/0000V000

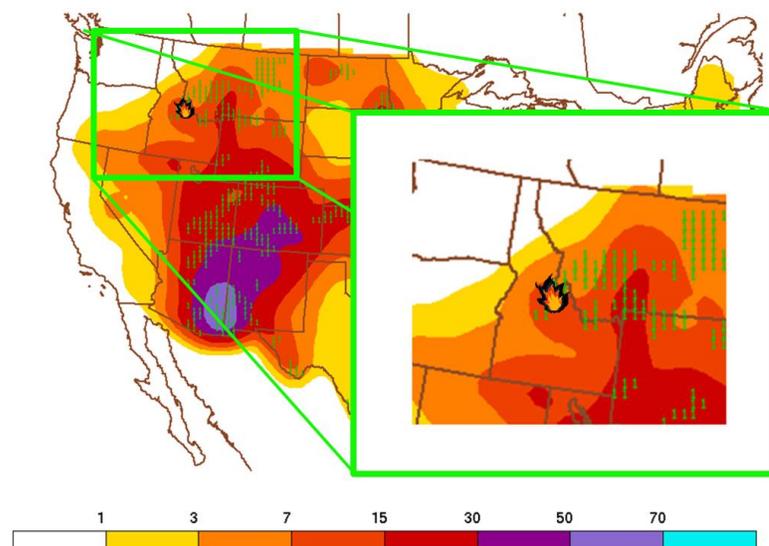


Figure 4.23: CONUS DRYTH1 forecast and observations along with a zoomed image centered on Idaho showing the approximate fire location.

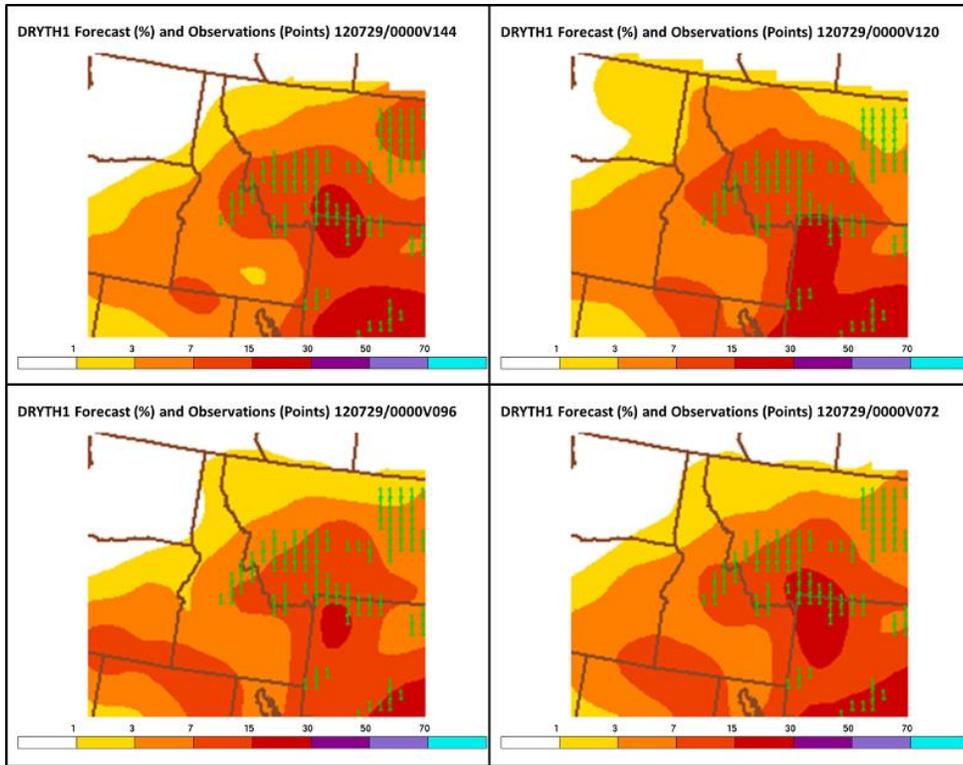


Figure 4.24: DRYTH1 00 UTC forecasts and Observations for Day 7 (F144) to Day 4 (F072).

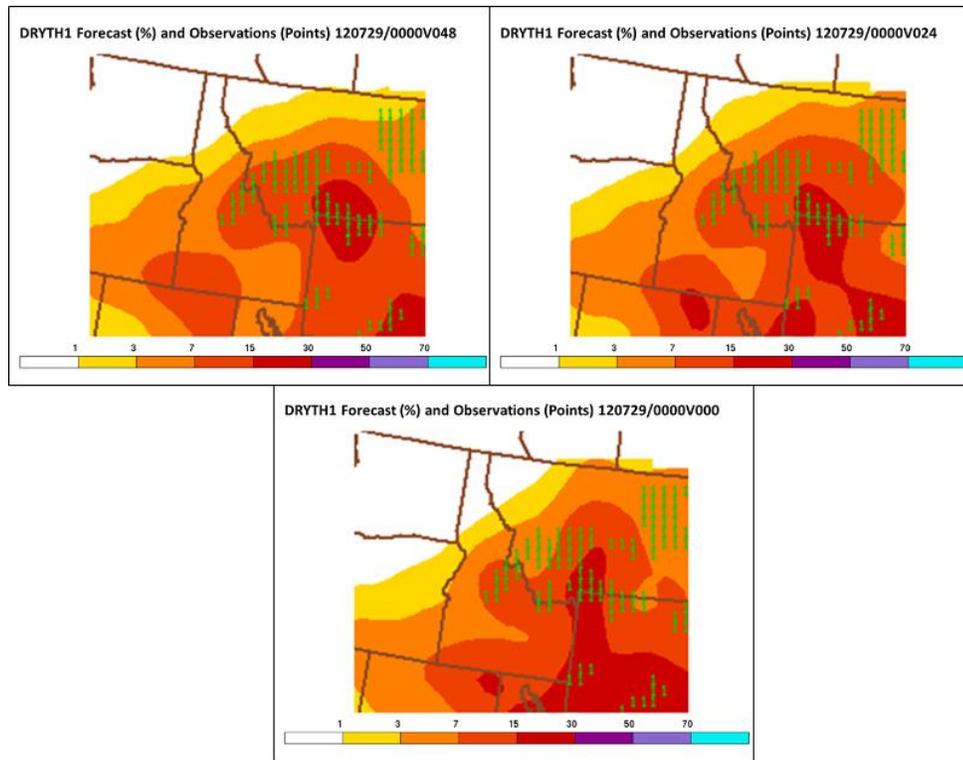


Figure 4.25: DRYTH1 00 UTC forecasts and Observations for Day 3 (F048) through Day 1 (F000).

Comparing the forecasts to the associated pentad climatology reveals that the two numbers are very close (Fig. 4.26). That is, the forecasts for DRYTH1 are not higher than climatology in this case. This example suggests the forecasts are performing as good as climatology, but not necessarily better in these isolated cases. Additionally, a case with values similar to climatology reinforces the fact that other sources of information, such as fuel dryness and drought stats, are necessary additions when making fire weather forecasts and preparations.

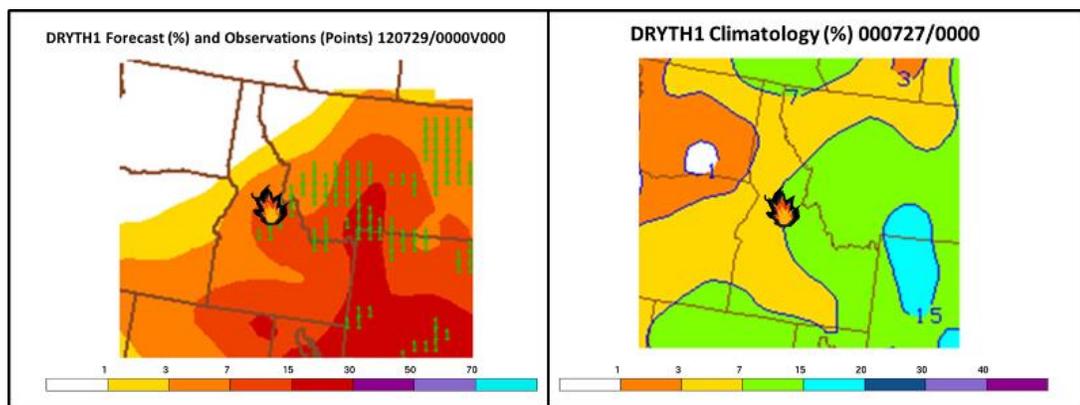


Figure 4.26: DRYTH1 00 UTC Forecast for 29 July 2012 F000 (left) and the associated pentad climatology (right).

4.3.3 AK Uvgoon Creek #1

Lightning-started fires in Alaska were lower than average in 2012, but a few cases did result in large amounts of acreage burned. One such event started around 04 July between 00 – 03 UTC. In far northwestern Alaska, a fire ignited near Uvgoon Creek ignited the tundra fuels and burned a bit more than 49,000 acres over one month (reported out on 05 August). This fire was allowed to burn because the regions remoteness presented no threat to human life or property (Alaska Interagency Coordination Center 2012).

Forecast examples from 02 July (120702/0000) and 03 July 2012 (120703/0000) in 6-h intervals are shown in Figures 4.27 and 4.28 leading up to the time of interest (Fig. 4.29). Parts of eastern Alaska contain high chances for dry thunderstorms 24-h prior at 120703/0000. Forecasts from 120702/0000 have generally higher probabilistic forecasts than the 120703/0000 forecasts because linear-regression-based forecasts have been noted to display diminished chances closer to event time (Antolik 2000). Overnight forecasts are diminished appropriately with heating (Fig. 4.27), yet the weather pattern visibly begins to shift westward across the Brooks Range in the north. Forecasts from 120703/0000 resolve the westward extension of dry thunder chances better than the forecast from 120702/0000.

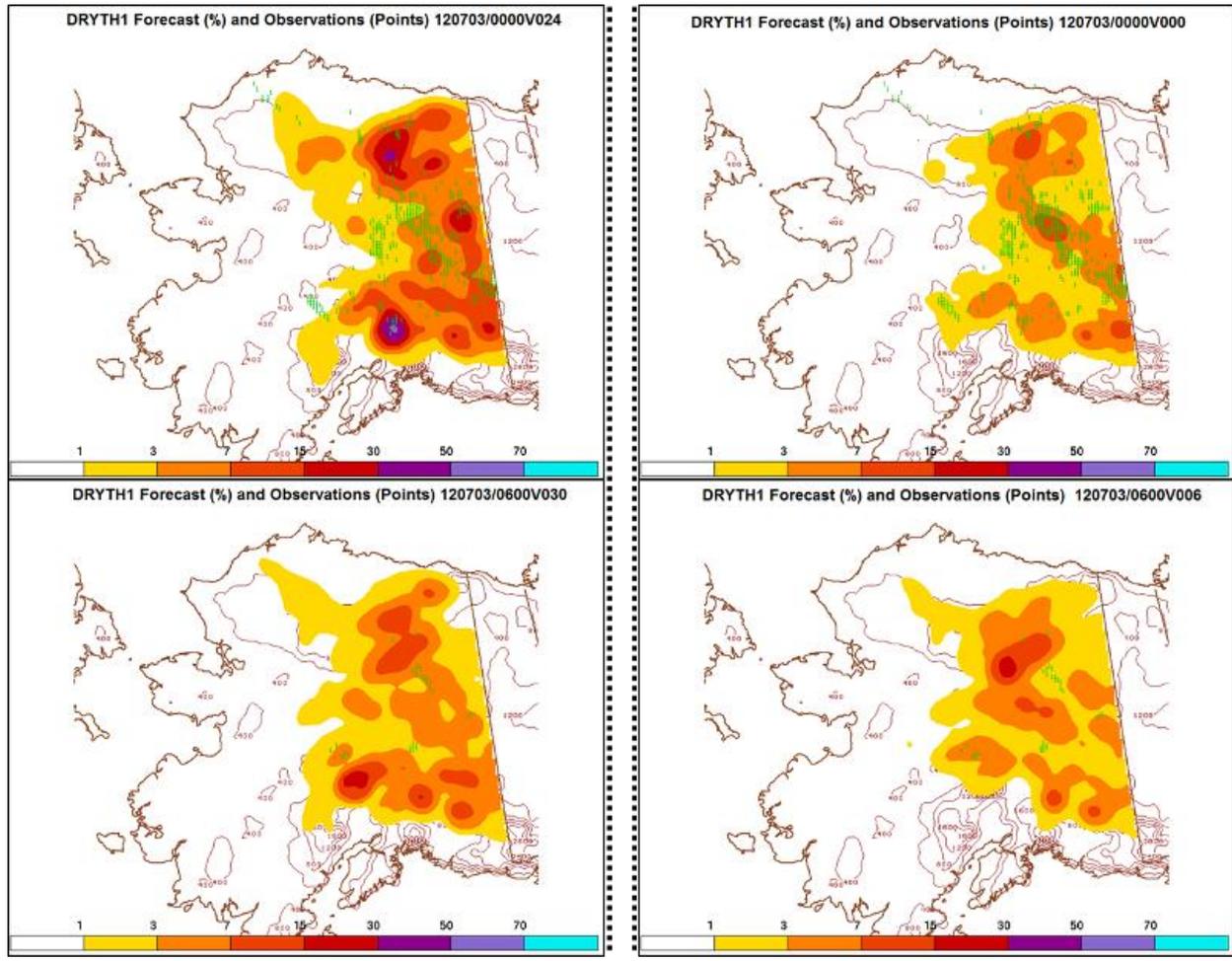


Figure 4.27: Forecasts from 120702/0000 (left) and 120703/0000 (right) runs starting on 120703/0000 (top). Forecast from 6-h later are shown on the bottom. Brown contours represent terrain. Vertical dashed lines signify the different forecast runs.

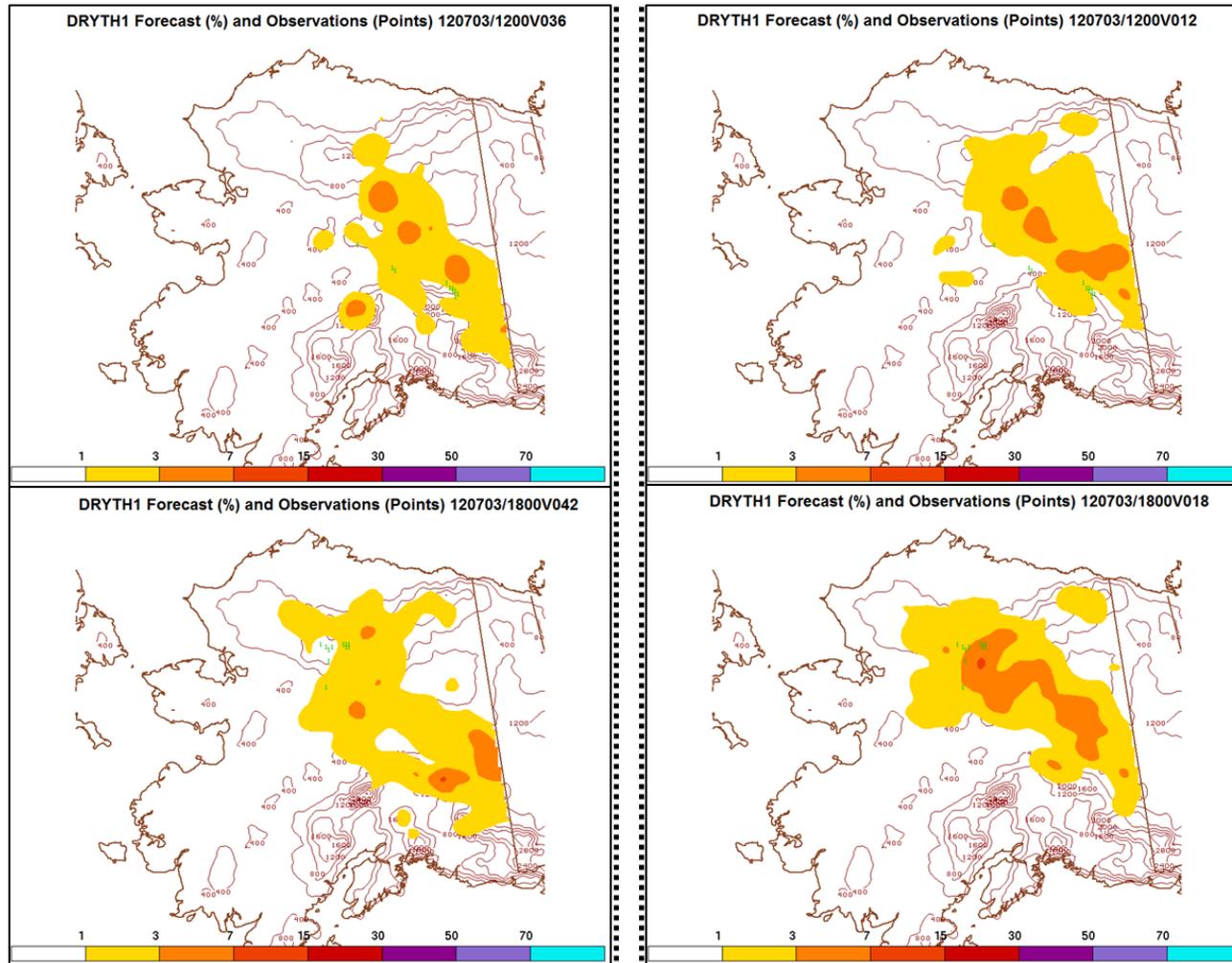


Figure 4.28: Same as Figure 4.26 valid at 12 UTC (top) – 18 UTC (bottom) hours from start time.

By the day of the event, at 120704/0000, the 48-h and 24-h forecasts show an increased chance of dry thunder occurring in northwestern Alaska and along the Brooks Range (Fig. 4.29). Visually, the overall shape of forecasts improve with the 24-h forecast from the 120703/0000 cycle run which manages to capture some of the missed forecast from the 120702/0000 cycle. Unfortunately, some of the higher probabilities on the 24-h forecast are considered misses though the areas are only slightly displaced from areas of observed DRYTH1.

Both cycles predict dry thunder chances equal to or higher than climatology for the associated pentad (the bottom left panel of Fig. 4.28) and designate areas with a noteworthy chance of dry thunder. This suggests the equations are not highly skewed toward climatology, at least at within the first two days. Additionally, both cycles manage to capture the chance of dry thunderstorms in places where fires actually ignited. The bottom-right panel in Fig. 4.28 denotes recorded fire starts as observed by the Alaska BLM (denoted with fire symbols). The southern-most fire symbol represents the approximate Uvgoon Creek ignition location.

Forecasts in AK are lower than in the CONUS due to the smaller grid size and the increased rarity for lightning in Alaska. Hughes and Trimarco (2004) mention, “[a]s we increase the resolution of the guidance, both in time and space, the magnitude of the probabilities will decrease, as the likelihood of an event at an exact time and point in space approaches zero.” Thus, it is not surprising to see fewer values above 7-15% across the AK domain.

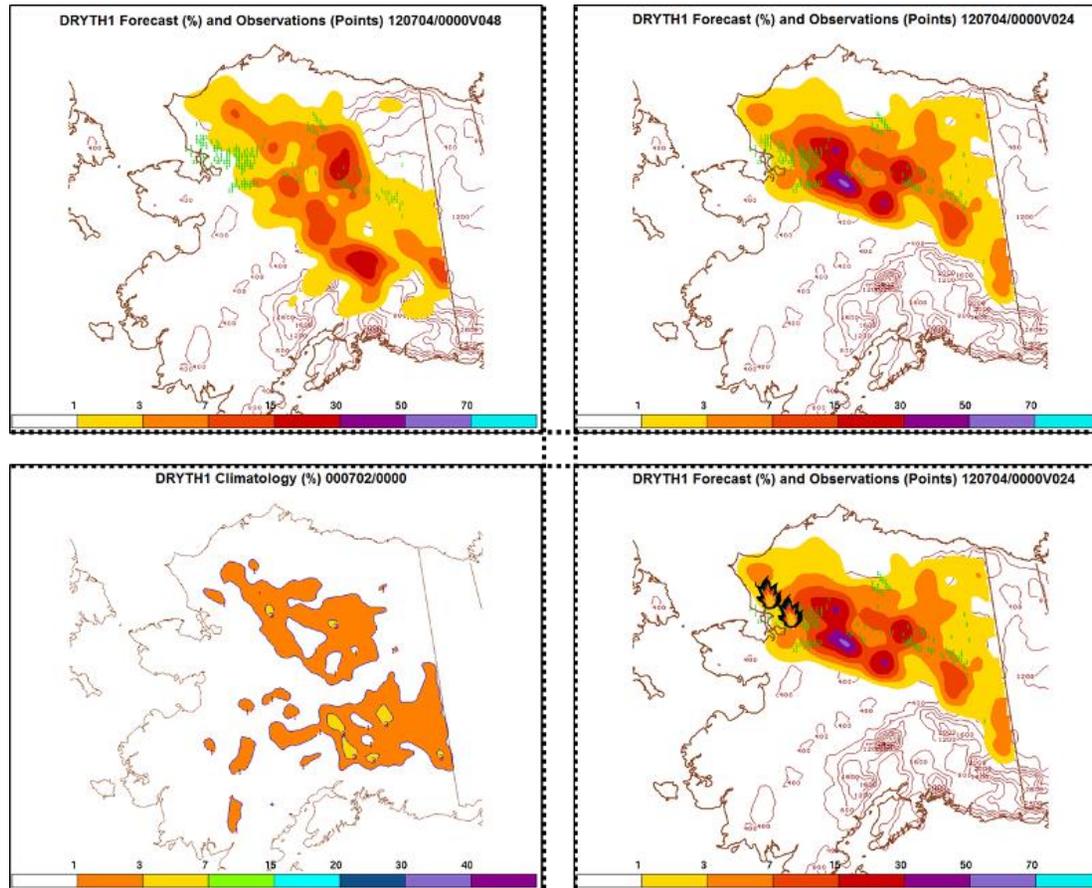


Figure 4.29: Forecasts for 120704/0000 (top), the associated climatology of DRYTH1 (bottom left), and the 24-h forecast shown again with fire locations displayed as fire symbols.

4.3.4 AK 120708 Case

Instead of focusing on a day with large fire starts, a large-scale “outbreak” of dry thunder can be assessed, i.e. a large area is expected to receive a lightning flash associated with a dry thunderstorm. Forecasts from several cycle runs valid at 08 August 2012 00 UTC are shown in Figure 4.30; again, only DRYTH1 will be shown. The 144-h (Day 7) forecast for DRYTH1 captures the overall shape and domain of dry thunder well. Some areas exhibit overprediction in places where no lightning occurred, but these regions are effectively reduced by the day of the event.

The overall pattern appears to be a strongly driven signal from the GFS data, as forecasts look solid from 144 to 120 hours prior to the event. That is, almost all of the probabilities align from northeast to southwest Alaska. A strong line of dry thunderstorms in north-central and southwestern regions is suggested as early as the 120-h forecast. Unfortunately, the 72-h (Day 3) forecast actually misses the northeast-to-southwest pattern, but the pattern reappears by the day of the event. Such a pattern change could help explain the lower POD and other scores seen in the statistics above if the trend is consistent. The unique trait exhibited here could be a further area of study for PP and/or GFS data in general.

Probability reduction from F144 to F000 appears here as well. Regardless of the low scores displayed earlier, especially at longer ranges, a user could find such an extended-range forecast to be beneficial. Having the alert for dry thunderstorms is advantageous even with overprediction.

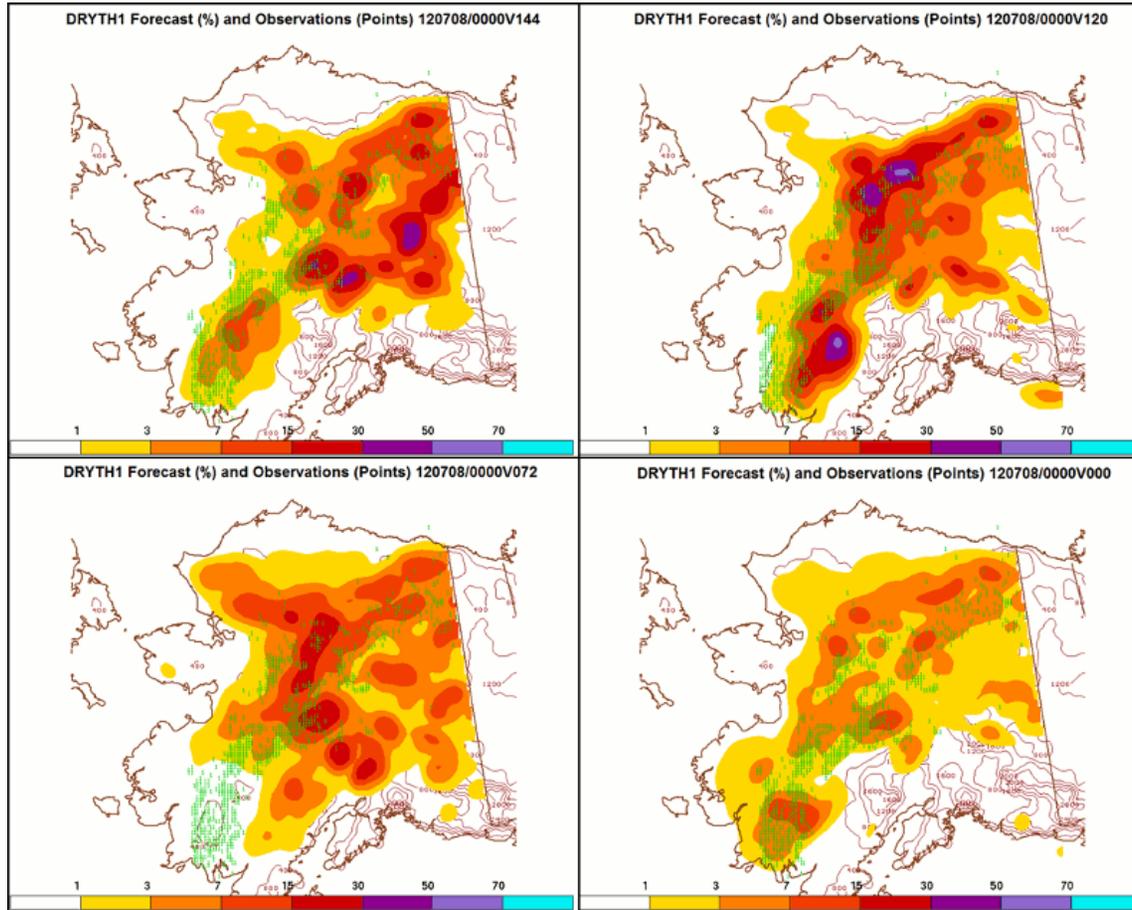


Figure 4.30: DRYTH1 forecasts valid at 120708/0000 (color-fill) from F144 prior (top-left), F120 prior (top-right), F072 prior (bottom-left), and F000 (start time; bottom-right). Observed DRYTH1 plotted with green digits.

As previously stated, the chance for dry-lightning does not necessarily correlate with fire ignition, so using other information about fuels and weather conditions could help fire crews select smaller areas to concentrate men and resources out of the larger risk area. Some of the more notable fires from 08 August 2012 (as recorded by the AK BLM) are marked on the F144 and F000 forecasts (Figure 4.31). This illustrates that high dry thunderstorm probabilities do not necessarily correspond with high chances of fire ignition, but that forecasts do emphasize certain areas of potential risk.

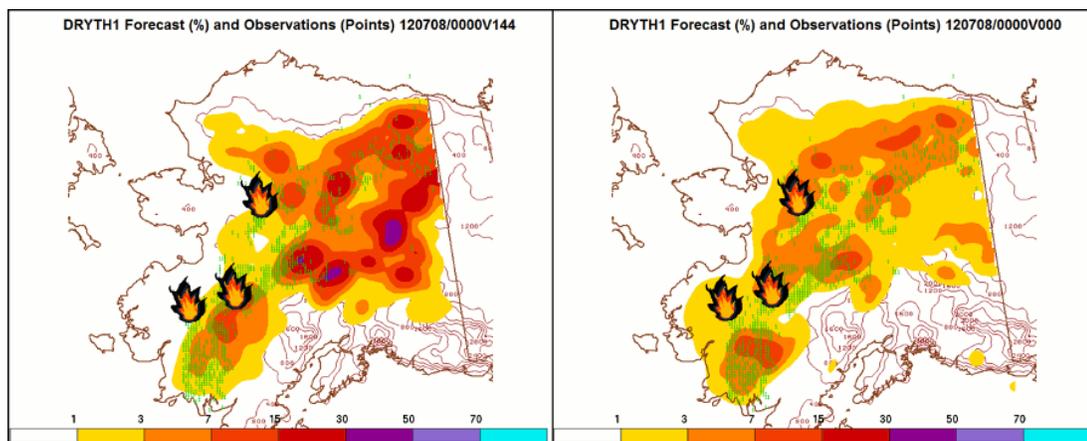


Figure 4.31: F144 and F000 forecast for 120708/0000 with recorded fire starts overlaid.

CHAPTER 5

Summary & Future Work

5.1 Summary & Discussion

A method for predicting dry thunderstorms, with regard to fire weather forecasting, by using Perfect Prognosis, Principal Component Analysis, and logistic regression is developed and tested. Dry thunder forecasts have been shown to be as, or more, successful than the lightning forecasts in the CONUS and AK domains. Such forecasts can benefit fire crews and managers, as lightning-started fires are considered a major fire weather concern across the CONUS and Alaska. Both domains are assessed separately; a 40-km CONUS grid, and a 10-km AK grid. Twelve years (2000-2011) of NARR data provided a reliable training data set. Variables of temperature, dewpoint, and wind components were used in conjunction with derived parameters, such as MUCAPE, to describe the environment.

Additionally, a dry thunderstorm climatology was developed using NARR precipitation data coupled with lightning data (NLDN for CONUS, and BLM for AK) from the twelve year data set. Climatologies were created for two different dry thunder categories: DRYTH1 and DRYTH2. DRYTH1 combines total precipitation less than a tenth of an inch (PLTN) with the product of one-or-more flashes of lightning (FONE). DRYTH2 allows a bit more precipitation by requiring precipitation less than a quarter of an inch (PLQT) and one-or-more flashes (FONE); i.e., precipitation can be a tenth or more but not a quarter or more. Both sets of climatology resemble the

climatology for FONE, yet DRYTH2 captures more variability in areas with slightly more rain (such as coastal or valley regions). The pentad method of creating a climatology set is employed in this study, matching the Bothwell (2002) and Buckey (2009) studies.

Using the perfect prognosis method, it is assumed all forecasts are perfect once a relationship between predictors and predictands has been established. Prior to making this relationship, Principal Component Analysis is used to maximize the variance explained by the predictands. After rotating the principal components (PCs) for stability, a Scree Diagram and Congruence Coefficient test found that keeping 12 PCs explained the data well for each month across both domains. Resulting PC Scores from the analysis were then used to create equations by logistic regression. Some of the strongest predictors as determined by the PCA are the lightning/dry thunder climatology and interactive predictors of lightning climatology multiplied with MUCAPE.

Because peak heating occurs in the afternoon hours, only 00 UTC equations were developed using logistic regression. An equation was created for each month across the separate domains. These equations were tested during the summer months of 2012. GFS model data are used as input, and resulting probabilistic fields of dry thunder are output. Probabilistic fields were chosen to help convey the uncertainty in the forecasts.

In general, DRYTH1 and DRYTH2 were analogous, and were both comparable to the lightning field, FONE. Sometimes, especially at the 40-km

CONUS grid, the forecasts of dry thunder accurately capture areas missed by FONE. This is most likely a result of applying a CAPE and CIN filter to the lightning fields (FONE included).

Forecasts results from June and July were combined together for verification because there were a small number of days available in June, and these two months contain the most lightning activity for both regions. Binning was used as part of the verification process, with more bins at smaller percentages because most of the forecasts fall within the first few bins. As such, the underforecasting noted for DRYTH1 and DRYTH2 on the Attributes Diagrams can be misleading. That is, if most of the hits occur within the 1% bin, then it appears that the value of 1% does not get forecasted enough to only receive 1% of the hits. ROC diagrams corresponded well with other plots by showing how skill decreases with increasing forecast hour, i.e., skill from mid to extended range (Day 3 +) forecasts contain less skill and reliability.

Two case studies are assessed per domain. Both CONUS cases relate to notable large fire events in 2012: the Bonita/Iron Complex Fire and Mustang Complex Fire. Both case studies reveal that DRYTH1 forecasts accentuate areas of dry thunderstorm probabilities up to one week in advance. Forecasts were noted to remain consistent across the 7 day time frame leading up to the event. Coupling such forecasts with other information about fuels and drought condition could have helped pinpoint the areas with the highest risk for fire ignition.

In Alaska, case studies also revealed the benefit of forecasts days in advance. One study effectively highlights regions of increased dry thunder risk in areas significantly higher than climatology that resulted in a few large fires. The second case revealed accurate predictions, at least spatially, of a large dry thunder outbreak a week in advance. Probabilities decreased as the event drew near, but this behavior has been seen in other logistic regression studies.

5.2 Future Work

Forecasts were only made for a portion of the warm season during 2012, but these forecasts are being assessed and run daily in 2013. Coupling the 2012 and 2013 data could allow for verification assessment across each month individually instead of requiring a combination of months. Additionally, it was noted that the GFS contained cool and moist biases during the summer of 2012. It should also be reiterated that the summer of 2012 was rather unique in both domains. Drought conditions across parts of the CONUS allowed for many acres to be burned from lightning-started fires. Meanwhile, AK experienced a more moist and cool summer with less dry thunderstorm activity. Continuing to run these forecasts and verification statistics should give a better estimation of overall system performance in more normal (or more abnormal) environments.

Only a small sample of products and verification metrics are assessed in this study. For example, though the perfect prog method is

applied to many predictands (Table 2.1), only the dry thunder fields are evaluated. Further work on the verification of the lightning and precipitation predictands is recommended. Lightning studies can assess the benefits of a 12-yr climatology sample, and precipitation analysis could reveal strengths and weaknesses of using perfect prog, PCA, and logistic regression on the mostly non-linear field. Work could be done to evaluate differences between using pentad climatology information versus other time averages (e.g., 3 days) and the subsequent effects on the forecasts.

Additionally, comparisons should be made between forecasts of dry thunder and the precipitation fields that exceed their thresholds. Comparisons of DRYTH1 and PTNTH could reveal if the DRYTH1 forecasts are essentially portraying a stronger bias towards lightning forecasts and need improvements in including dryness. It is expected that the DRYTH1 forecasts will be lower in areas where PTNTH forecast values are higher. Note that this does not necessarily mean that DRYTH1 forecasts will be 0 in areas with PTNTH forecasts. Comparisons for DRYTH2 could be made with PQTR and PTNTH to assess the extent of effects of the higher precipitation allowance. DRYTH2 has been shown to be more important and accurate in the East domain, so a separate East focus could be considered.

An interesting effect regarding Day 4 (F072-F090) was observed in the verification statistics and the case studies in Alaska. It is unclear at this time whether this is a common occurrence in the underlying GFS model or

associated with perfect prog/logistic regression. More research should be conducted to better understand this unique feature.

To better assess dry thunderstorms with regard to fire ignition, the inclusion of fuels as a predictor should be studied, but most operational models do not offer vegetation dryness as a field. Instead, perhaps it would be more beneficial to create additional interactive predictors that were found to be closely associated with dry thunderstorms. Some terms can be found from the groupings noted on the PC Loading analysis. For example, combining vorticity divergence and the Laplacian of the Geopotential Height into a single field could explain much of the variance in a system. Perhaps combining low level moisture terms with lapse rate could be an important predictor for the prediction of lightning/dry thunderstorms as these terms can distinguish between wet and dry thunderstorms (Rorig and Ferguson 1999). Adding DTPI as a predictor could be beneficial to distinguishing between wet and dry environments with regards to dry thunderstorm formation and potential ignition. Useful combinations can only be found through testing.

Previous studies have found that a long-continuing current of lightning (often associated with multiple strokes) is the main reason lightning sparks a fuel (Kitagawa et al. 1962, Fuquay et al 1967). Additionally, it has been found that low precipitation thunderstorms are often associated with positive CG flashes (Curran and Rust 1982; Beasley 2013, personal communication). Research that limits a dry thunderstorm to only positive CG flashes could

improve the predictions, but it should be reiterated that Rorig and Ferguson (1999) discuss that the fuel consideration is more important.

Utilizing non-linear regression techniques to develop equations could also improve accuracy of lightning and precipitation forecasts. For a different PCA technique, a T-mode analysis could reveal temporal trends of lightning/precipitation/dry thunderstorms over multiple years (not necessarily per month). As lightning, especially in Alaska, is dependent on heat and moisture, a T-mode analysis could reveal changes related to climatological warming/cooling over the region.

Regardless of the many possibilities for future research, this current method of predicting dry thunderstorms, similar to Bothwell (2002) and Buckey (2009), has shown to be beneficial to users by highlighting areas of risk up to a week prior to an event. DRYTH1 and DRYTH2 are often overpredicted, but most users agree that overprediction is more acceptable than a missed forecast in rare-event phenomena. It is hoped that such an experimental product can become a more standard guidance product for use by fire-weather forecasters (e.g., at the SPC), and by fire managers for crews and resources.

References

- Alaska Interagency Coordination Center, 2012: 2012 Alaska Fire Season. Wildland Fire Summary and Statistics Annual Report, AICC, 27 pp. [Available online at <http://fire.ak.blm.gov/content/aicc/stats/firestats.pdf>].
- Antolik, M. S., 2000: An overview of the National Weather Service's centralized statistical quantitative precipitation forecasts. *J. Hydrol.*, **239**, 306–337.
- BEALS, E. A., 1923: DISCUSSION OF THUNDERSTORMS AND FOREST FIRES IN CALIFORNIA. *Mon. Wea. Rev.*, **51**, 180–182.
doi: [http://dx.doi.org/10.1175/1520-0493\(1923\)51<180:DOTAFF>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1923)51<180:DOTAFF>2.0.CO;2)
- Becker, Emily J., Ernesto Hugo Berbery, R. Wayne Higgins, 2009: Understanding the Characteristics of Daily Precipitation over the United States Using the North American Regional Reanalysis. *J. Climate*, **22**, 6268–6286.
doi: <http://dx.doi.org/10.1175/2009JCLI2838.1>
- Bothwell, P.D., 2002: Prediction of Cloud-to-Ground Lightning in the Western United States, Ph.D., University of Oklahoma, 178pp.
- Bothwell, 2008: Predicting the location and intensity of lightning using an experimental automated statistical method. Preprints, Third Conf. on Meteorological Applications of Lightning Data, New Orleans, LA, Amer. Meteor. Soc., 6.1.
[Available online at <http://ams.confex.com/ams/pdfpapers/134374.pdf>.]
- Bothwell, P.D., 2010: Evolution of the experimental/automated perfect prog lightning forecasts at the Storm Prediction Center. *Third International Lightning Detection Conference*, April 21-22, Orlando, FL, Vaisala, Inc., Tucson, 10pp.
- BRIER, GLENN W., 1950: VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY. *Mon. Wea. Rev.*, **78**, 1–3.
doi: [http://dx.doi.org/10.1175/1520-0493\(1950\)078<0001:VOFEIT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2)
- Brooks, Harold E., 2004: TORNADO-WARNING PERFORMANCE IN THE PAST AND FUTURE: A Perspective from Signal Detection Theory. *Bull. Amer. Meteor. Soc.*, **85**, 837–843.
doi: <http://dx.doi.org/10.1175/BAMS-85-6-837>

- Brunet, N., R. Verret, N. Yacowar, 1988: An Objective Comparison of Model Output Statistics and “Perfect Prog” Systems in Producing Numerical Weather Element Forecasts. *Wea. Forecasting*, **3**, 273–283.
doi: [http://dx.doi.org/10.1175/1520-0434\(1988\)003<0273:AOCOMO>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(1988)003<0273:AOCOMO>2.0.CO;2)
- Buckey, David, 2009: Using the Perfect Prognosis Method to Forecast Cloud to Ground Lightning in Alaska. M.S. Thesis, Department of Meteorology, University of Oklahoma, 82 pp.
- Buckey, D. R. and Bothwell, P.D, 2009: A Climatology and the Intra-Seasonal Variation of Summertime Cloud-to-Ground Lightning in Mainland Alaska. Fourth Conference on Meteorological Applications of Lightning Data, Phoenix, AZ, Amer. Meteor. Soc., 7 pp.
- Bukovsky, Melissa S., David J. Karoly, 2007: A Brief Evaluation of Precipitation from the North American Regional Reanalysis. *J. Hydrometeor*, **8**, 837–846.
doi: <http://dx.doi.org/10.1175/JHM595.1>
- Byers, Horace R., Harriet R. Rodebush, 1948: CAUSES OF THUNDERSTORMS OF THE FLORIDA PENINSULA. *J. Meteor.*, **5**, 275–280.
doi: [http://dx.doi.org/10.1175/1520-0469\(1948\)005<0275:COTOTF>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1948)005<0275:COTOTF>2.0.CO;2)
- Calef, M. P., A. D. McGuire, F. S. Chapin, 2008: Human Influences on Wildfire in Alaska from 1988 through 2005: An Analysis of the Spatial Patterns of Human Impacts. *Earth Interact.*, **12**, 1–17.
doi: <http://dx.doi.org/10.1175/2007EI220.1>
- Craven, Jeffrey P., Ryan E. Jewell, Harold E. Brooks, 2002: Comparison between Observed Convective Cloud-Base Heights and Lifting Condensation Level for Two Different Lifted Parcels. *Wea. Forecasting*, **17**, 885–890.
doi: [http://dx.doi.org/10.1175/1520-0434\(2002\)017<0885:CBOCCB>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(2002)017<0885:CBOCCB>2.0.CO;2)
- Cummins, K. L., M. J. Murphy, E. A. Bardo, W. L. Hiscox, R. B. Pyle, and A. E. Pifer (1998), A Combined TOA/MDF Technology Upgrade of the U.S. National Lightning Detection Network, *J. Geophys. Res.*, 103(D8), 9035–9044, doi:[10.1029/98JD00153](https://doi.org/10.1029/98JD00153).

- Curran, E. Brian, W. David Rust, 1992: Positive Ground Flashes Produced by Low-Precipitation Thunderstorms in Oklahoma on 26 April 1984. *Mon. Wea. Rev.*, **120**, 544–553.
doi: [http://dx.doi.org/10.1175/1520-0493\(1992\)120<0544:PGFPBL>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1992)120<0544:PGFPBL>2.0.CO;2)
- Daly, Christopher, Ronald P. Neilson, Donald L. Phillips, 1994: A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *J. Appl. Meteor.*, **33**, 140–158.
doi: [http://dx.doi.org/10.1175/1520-0450\(1994\)033<0140:ASTMFM>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2)
- Doswell, Charles A., Robert Davies-Jones, David L. Keller, 1990: On Summary Measures of Skill in Rare Event Forecasting Based on Contingency Tables. *Wea. Forecasting*, **5**, 576–585.
doi: [http://dx.doi.org/10.1175/1520-0434\(1990\)005<0576:OSMOSI>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(1990)005<0576:OSMOSI>2.0.CO;2)
- Doswell, Charles A., Erik N. Rasmussen, 1994: The Effect of Neglecting the Virtual Temperature Correction on CAPE Calculations. *Wea. Forecasting*, **9**, 625–629.
doi: [http://dx.doi.org/10.1175/1520-0434\(1994\)009<0625:TEONTV>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(1994)009<0625:TEONTV>2.0.CO;2)
- Dowdy, Andrew J., Graham A. Mills, 2012: Atmospheric and Fuel Moisture Characteristics Associated with Lightning-Attributed Fires. *J. Appl. Meteor. Climatol.*, **51**, 2025–2037.
doi: <http://dx.doi.org/10.1175/JAMC-D-11-0219.1>
- Fuelberg, Henry E., David G. Biggar, 1994: The Preconvective Environment of Summer Thunderstorms over the Florida Panhandle. *Wea. Forecasting*, **9**, 316–326.
doi: [http://dx.doi.org/10.1175/1520-0434\(1994\)009<0316:TPEOST>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(1994)009<0316:TPEOST>2.0.CO;2)
- Ferro, Christopher A. T., David B. Stephenson, 2011: Extremal Dependence Indices: Improved Verification Measures for Deterministic Forecasts of Rare Binary Events. *Wea. Forecasting*, **26**, 699–713.
doi: <http://dx.doi.org/10.1175/WAF-D-10-05030.1>
- Fuquay, D. M., R. G. Baughman, A. R. Taylor, and R. G. Hawe (1967), Characteristics of seven lightning discharges that caused forest fires, *J. Geophys. Res.*, 72(24), 6371–6373, doi:[10.1029/JZ072i024p06371](https://doi.org/10.1029/JZ072i024p06371).

- Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., and Wickham, J., 2011. [Completion of the 2006 National Land Cover Database for the Conterminous United States](#), *PE&RS*, Vol. 77(9):858-864.
- Harman, H. H. 1976. *Modern Factor Analysis*. The University of Chicago Press, Chicago, IL.
- Hitchens, Nathan M., Harold E. Brooks, Michael P. Kay, 2013: Objective Limits on Forecasting Skill of Rare Events. *Wea. Forecasting*, **28**, 525–534.
doi: <http://dx.doi.org/10.1175/WAF-D-12-00113.1>
- Hostetler, S. W., P. J. Bartlein, A. M. Solomon, J. O. Holman, R. T. Busing, and S. L. Shafer, 2005: Climatic Controls of Fire in the Western United States: From the Atmosphere to Ecosystems. Report for Joint Fire Sciences Program Project 01-1-605, 28 pp.
[Available online at http://www.firescience.gov/projects/01-1-6-05/project/01-1-6-05_final_report.pdf.]
- Hughes, Kathryn K., 2001: Development of MOS Thunderstorm and Severe Thunderstorm Forecast Equations with Multiple Data Sources. *Preprints 18th Conference on Weather Analysis and Forecasting and the 14th Conference on Numerical Weather Prediction*, Ft. Lauderdale, FL, Amer. Meteor. Soc., [Available online at <http://www.nws.noaa.gov/mdl/synop/amspapers/18thwaftstm.pdf>.]
- Hughes, Kathryn K., 2004: Probabilistic lightning forecast guidance for aviation. *Preprints 11th Conference on Aviation, Range, and Aerospace*, Hyannis, MA, Amer. Meteor. Soc., [Available online at <http://www.nws.noaa.gov/mdl/synop/amspapers/hyannis.pdf>.]
- Huges, K. K., and R. A. Trimarco, 2004: Impacts of Resolution on Gridded Probability Thunderstorm Forecast Guidance. *Preprints 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc.
[Available online at http://www.nws.noaa.gov/mdl/synop/amspapers/seattlewaf_j6.5.pdf.]
- InciWeb, cited 2012. [Available online at inciweb.org].
- Kitagawa, N., M. Brook, and E. J. Workman (1962), Continuing currents in cloud-to-ground lightning discharges, *J. Geophys. Res.*, 67(2), 637–647, doi:[10.1029/JZ067i002p00637](https://doi.org/10.1029/JZ067i002p00637).

- Klein, William H., Billy M. Lewis, Isadore Enger, 1959: OBJECTIVE PREDICTION OF FIVE-DAY MEAN TEMPERATURES DURING WINTER. *J. Meteor.*, **16**, 672–682.
doi: [http://dx.doi.org/10.1175/1520-0469\(1959\)016<0672:OPOFDM>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1959)016<0672:OPOFDM>2.0.CO;2)
- Klein, William H., Joseph J. Charney, Morris H. McCutchan, John W. Benoit, 1996: Verification of Monthly Mean Forecasts for Fire Weather Elements in the Contiguous United States. *J. Climate*, **9**, 3317–3327.
doi: [http://dx.doi.org/10.1175/1520-0442\(1996\)009<3317:VOMMFF>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1996)009<3317:VOMMFF>2.0.CO;2)
- Lorenz, E. N., 1956: *Empirical orthogonal functions and statistical weather prediction*. Sci. Rep. No. 1, Stat. Fcst. Proj. Dept. Meteor., Mass. Inst. Tech, 40 pp.
- Marzban, Caren, Scott Sandgathe, Eugenia Kalnay, 2006: MOS, Perfect Prog, and Reanalysis. *Mon. Wea. Rev.*, **134**, 657–663.
doi: <http://dx.doi.org/10.1175/MWR3088.1>
- Mason, S.J., and N.E. Graham, 1999: Conditional Probabilities, Relative Operating Characteristics, and Relative Operating Levels. *Wea. Forecasting*, **14**, 713–725.
doi: [http://dx.doi.org/10.1175/1520-0434\(1999\)014<0713:CPROCA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(1999)014<0713:CPROCA>2.0.CO;2)
- Mitchell, et al., 2005: N. American Regional Reanalysis Land Surface: Configurations, Physics, Results, Cautions. Proceeding, *85th Annual Meeting*, San Diego, CA, Amer. Meteor. Soc., [Available online at http://www.emc.ncep.noaa.gov/mmb/rrean/land_sfc.ppt].
- Mesinger, Fedor, and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360.
doi: <http://dx.doi.org/10.1175/BAMS-87-3-343>
- National Interagency Fire Center, 2013: Lightning vs. human caused fires and acres. [Available online at http://www.nifc.gov/fireInfo/fireInfo_stats_lightng.html].
- Nelder, J.A. and Wedderburn, R. W. M. (1972) Generalized linear models. *J. R. Statist. Soc. A*, **135**, 370-384.

- North, Gerald R., Thomas L. Bell, Robert F. Cahalan, Fanthune J. Moeng, 1982: Sampling Errors in the Estimation of Empirical Orthogonal Functions. *Mon. Wea. Rev.*, **110**, 699–706.
doi: [http://dx.doi.org/10.1175/1520-0493\(1982\)110<0699:SEITEO>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2)
- O'Lenic, Edward A., Robert E. Livezey, 1988: Practical Considerations in the Use of Rotated Principal Component Analysis (RPCA) in Diagnostic Studies of Upper-Air Height Fields. *Mon. Wea. Rev.*, **116**, 1682–1689.
doi: [http://dx.doi.org/10.1175/1520-0493\(1988\)116<1682:PCITUO>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1988)116<1682:PCITUO>2.0.CO;2)
- Orville, Richard E., Gary R. Huffines, 2001: Cloud-to-Ground Lightning in the United States: NLDN Results in the First Decade, 1989–98. *Mon. Wea. Rev.*, **129**, 1179–1193.
doi: [http://dx.doi.org/10.1175/1520-0493\(2001\)129<1179:CTGLIT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(2001)129<1179:CTGLIT>2.0.CO;2)
- Orville, Richard E., 2008: Development of the National Lightning Detection Network. *Bull. Amer. Meteor. Soc.*, **89**, 180–190.
doi: <http://dx.doi.org/10.1175/BAMS-89-2-180>
- Reap, Ronald M., Donald S. Foster, 1979: Automated 12–36 Hour Probability Forecasts of Thunderstorms and Severe Local Storms. *J. Appl. Meteor.*, **18**, 1304–1315.
doi: [http://dx.doi.org/10.1175/1520-0450\(1979\)018<1304:AHPFOT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(1979)018<1304:AHPFOT>2.0.CO;2)
- Reap, Ronald M., 1991: Climatological Characteristics and Objective Prediction of Thunderstorms over Alaska. *Wea. Forecasting*, **6**, 309–319.
doi: [http://dx.doi.org/10.1175/1520-0434\(1991\)006<0309:CCAOP0>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(1991)006<0309:CCAOP0>2.0.CO;2)
- Reap, Ronald M., 1994: Analysis and Prediction of Lightning Strike Distributions Associated with Synoptic Map Types over Florida. *Mon. Wea. Rev.*, **122**, 1698–1715.
doi: [http://dx.doi.org/10.1175/1520-0493\(1994\)122<1698:AAPOLS>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1994)122<1698:AAPOLS>2.0.CO;2)
- Richman, Michael B., Peter J. Lamb, 1985: Climatic Pattern Analysis of Three- and Seven-Day Summer Rainfall in the Central United States: Some Methodological Considerations and a Regionalization. *J. Climate Appl. Meteor.*, **24**, 1325–1343.
doi: [http://dx.doi.org/10.1175/1520-0450\(1985\)024<1325:CPAOTA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(1985)024<1325:CPAOTA>2.0.CO;2)

- Richman, M. B. 1986: Rotation of principal components, *J. climatol.*, **6**, 293-335.
- Richman, Michael B., Xiaofeng Gong, 1999: Relationships between the Definition of the Hyperplane Width to the Fidelity of Principal Component Loading Patterns. *J. Climate*, **12**, 1557–1576.
doi: [http://dx.doi.org/10.1175/1520-0442\(1999\)012<1557:RBTDOT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1999)012<1557:RBTDOT>2.0.CO;2)
- Rochette, S. M., James T. Moore and Patrick S. Market, 1999: The importance of parcel choice in elevated CAPE computations. *Natl. Wea. Dig.*, 23:4, 20-32.
- Roebber, Paul J., 2009: Visualizing Multiple Measures of Forecast Quality. *Wea. Forecasting*, **24**, 601–608.
doi: <http://dx.doi.org/10.1175/2008WAF2222159.1>
- Rorig, Miriam L., Sue A. Ferguson, 1999: Characteristics of Lightning and Wildland Fire Ignition in the Pacific Northwest. *J. Appl. Meteor.*, **38**, 1565–1575.
doi: [http://dx.doi.org/10.1175/1520-0450\(1999\)038<1565:COLAWF>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(1999)038<1565:COLAWF>2.0.CO;2)
- Rorig, Miriam L., Sue A. Ferguson, 2002: The 2000 Fire Season: Lightning-Caused Fires. *J. Appl. Meteor.*, **41**, 786–791.
doi: [http://dx.doi.org/10.1175/1520-0450\(2002\)041<0786:TFSLCF>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(2002)041<0786:TFSLCF>2.0.CO;2)
- Rorig, Miriam L., Steven J. McKay, Sue A. Ferguson, Paul Werth, 2007: Model-Generated Predictions of Dry Thunderstorm Potential. *J. Appl. Meteor. Climatol.*, **46**, 605–614.
doi: <http://dx.doi.org/10.1175/JAM2482.1>
- Ruane, Alex C., 2010: NARR's Atmospheric Water Cycle Components. Part I: 20-Year Mean and Annual Interactions. *J. Hydrometeorol.*, **11**, 1205–1219.
doi: <http://dx.doi.org/10.1175/2010JHM1193.1>
- Ruane, Alex C., 2010: NARR's Atmospheric Water Cycle Components. Part II: Summertime Mean and Diurnal Interactions. *J. Hydrometeorol.*, **11**, 1220–1233.
doi: <http://dx.doi.org/10.1175/2010JHM1279.1>

- Shafer, P. E., and K. K. Gilbert, 2008: Developing GFS-based MOS thunderstorm guidance for Alaska. *Preprints, 3rd Conference on Meteorological Applications of Lightning Data*, New Orleans, Amer. Meteor. Soc., P2.9. [Available online at <http://www.nws.noaa.gov/mdl/synop/amspapers/Shofer-Gilbert-2008.pdf>.]
- Shafer, Phillip E., Henry E. Fuelberg, 2008: A Perfect Prognosis Scheme for Forecasting Warm-Season Lightning over Florida. *Mon. Wea. Rev.*, **136**, 1817–1846.
doi: <http://dx.doi.org/10.1175/2007MWR2222.1>
- Shafran, P., J. Woollen, W. Ebisuzaki, W. Shi, Y. Fan, R. Grumbine, and M. Fennessy, 2004: Observational data used for assimilation in the NCEP North American Regional Reanalysis. Preprints, *14th Conf. on Applied Climatology*, Seattle, WA, Amer. Meteor. Soc. [Available online at www.emc.ncep.noaa.gov/mmb/rreanl/RR_obs_12jun06.doc.]
- SULLIVAN, WILLIAM G., 1963: LOW-LEVEL CONVERGENCE AND THUNDERSTORMS IN ALASKA. *Mon. Wea. Rev.*, **91**, 89–92.
doi: [http://dx.doi.org/10.1175/1520-0493\(1963\)091<0089:LCATIA>2.3.CO;2](http://dx.doi.org/10.1175/1520-0493(1963)091<0089:LCATIA>2.3.CO;2)
- Tibco, 2011: *S-PLUS 8.2.0*
- Todd, Susan K., and Holly Ann Jewkes, 2006: Wildland fire in Alaska: a history of organized fire suppression and management in the last frontier. *Agricultural and Forestry Experiment Station Bulletin*, No. 114, University of Alaska, 64pp.
[Available online at <http://www.uaf.edu/files/snras/B114.pdf>.]
- Trenberth, Kevin E., Aiguo Dai, Roy M. Rasmussen, David B. Parsons, 2003: The Changing Character of Precipitation. *Bull. Amer. Meteor. Soc.*, **84**, 1205–1217.
doi: <http://dx.doi.org/10.1175/BAMS-84-9-1205>
- Valine, W. C., and E. P. Krider, Statistics and characteristics of cloud-to-ground lightning with multiple ground contacts, *J. Geophys. Res.*, 107(D20), 4441, doi:[10.1029/2001JD001360](https://doi.org/10.1029/2001JD001360), 2002.
- Vislocky, Robert L., George S. Young, 1989: The Use of Perfect Prog Forecasts to Improve Model Output Statistics Forecasts of Precipitation Probability. *Wea. Forecasting*, **4**, 202–209.
doi: [http://dx.doi.org/10.1175/1520-0434\(1989\)004<0202:TUOPPF>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(1989)004<0202:TUOPPF>2.0.CO;2)

- Wallmann, J., 2004: A procedure for forecasting dry thunderstorms in the Great Basin using the dynamic tropopause and alternate tools for assessing instability. NWS WR Tech. Attach. 04-04, 20 pp.
- Wallmann, James, Rhett Milne, Christopher Smallcomb, Matthew Mehle, 2010: Using the 21 June 2008 California Lightning Outbreak to Improve Dry Lightning Forecast Procedures. *Wea. Forecasting*, **25**, 1447–1462.
doi: <http://dx.doi.org/10.1175/2010WAF2222393.1>
- Walsh, John E., Michael B. Richman, David W. Allen, 1982: Spatial Coherence of Monthly Precipitation in the United States. *Mon. Wea. Rev.*, **110**, 272–286.
doi: [http://dx.doi.org/10.1175/1520-0493\(1982\)110<0272:SCOMPI>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1982)110<0272:SCOMPI>2.0.CO;2)
- Wang, Bin, Xihua Xu, 1997: Northern Hemisphere Summer Monsoon Singularities and Climatological Intraseasonal Oscillation. *J. Climate*, **10**, 1071–1085.
doi: [http://dx.doi.org/10.1175/1520-0442\(1997\)010<1071:NHMSA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1997)010<1071:NHMSA>2.0.CO;2)
- Watson, Andrew I., Ronald L. Holle, Raúl E. López, Robert Ortiz, James R. Nicholson, 1991: Surface Wind Convergence as a Short-Term Predictor of Cloud-to-Ground Lightning at Kennedy Space Center. *Wea. Forecasting*, **6**, 49–64.
doi: [http://dx.doi.org/10.1175/1520-0434\(1991\)006<0049:SWCAAS>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(1991)006<0049:SWCAAS>2.0.CO;2)
- Weiss, Steven J., Charles A. Doswell, Frederick P. Ostby, 1980: Comments on “Automated 12–36 Hour Probability Forecasts of Thunderstorms and Severe Local Storms”. *J. Appl. Meteor.*, **19**, 1328–1333.
doi: [http://dx.doi.org/10.1175/1520-0450\(1980\)019<1328:COHPFO>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(1980)019<1328:COHPFO>2.0.CO;2)
- West, Gregory L., W. James Steenburgh, William Y. Y. Cheng, 2007: Spurious Grid-Scale Precipitation in the North American Regional Reanalysis. *Mon. Wea. Rev.*, **135**, 2168–2184.
doi: <http://dx.doi.org/10.1175/MWR3375.1>
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd Edition. International Geophysics Series, Vol. 91, Academic Press, 627 pp.

Wilson, Laurence J., Marcel Vallée, 2003: The Canadian Updateable Model Output Statistics (UMOS) System: Validation against Perfect Prog. *Wea. Forecasting*, **18**, 288–302.

doi: [http://dx.doi.org/10.1175/1520-0434\(2003\)018<0288:TCUMOS>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(2003)018<0288:TCUMOS>2.0.CO;2)

Wilson, L. J. and Giles, A. (2013), A new index for the verification of accuracy and timeliness of weather warnings. *Met. Apps*, 20: 206–216.

doi: 10.1002/met.1404

Appendix A

Acronyms and Predictors

Lists of acronyms and the set of predictors used for this project.

Table A1: List of main acronyms used in the text.

Acronym	Name
AK	Alaska
AGL	Above Ground Level
ALDN	Alaska Lightning Detection Network
BLM	Bureau of Land Management
CAPE	Convective Available Potential Energy
CC	Congruence Coefficient
CG	Cloud-to-Ground (lightning)
CONUS	Contiguous United States
DF	Direction Finder
DTPI	Dry Thunder Potential Index
GEMPAK	General Meteorological Package
GFS	Global Forecast System
GLM	General Linear Model
IC	In-Cloud (lightning)
JFS	Joint Fire Sciences
MOS	Model Output Statistics
MUCAPE	Most Unstable (parcel) CAPE
NARR	North American Regional Reanalysis
NLDN	National Lightning Detection Network
PC	Principal Component
PCA	Principal Component Analysis
PP	Perfect Prognosis
PRISM	Parameter-Elevation Regressions on Independent Slopes
SGSP	Spurious Grid-Scale Precipitation
SPC	Storm Prediction Center
TOA	Time of Arrival

Table A2: List of predictors.

Name	Description
OMEG85	850 hPa Omega
UREL85	850 hPa U Wind Component
VREL85	850 hPa V Wind Component
RELH85	850 hPa Relative Humidity
MFCN85	850 hPa Moisture Flux Convergence
TWAD85	850 hPa Theta-W Advection
WDIV85	850 hPa Wind Divergence
RVRG85	850 hPa Relative Vorticity
FRNT85	850 hPa Frontogenesis
LHGT85	850 hPa Laplacian of the Geopotential Height
LTHA85	850 hPa Laplacian of Theta (Potential Temperature)
AWND85	850 hPa Ageostrophic Wind
GRDO85	850 hPa Gradient of Omega
THWC85	850 hPa Theta-W in Celsius
THTE85	850 hPa Theta-E
FRT8N85	850-1000 hPa Frontogenesis
SH8N85	850-1000 hPa Speed Shear
MT8N85	850-1000 hPa Max Thermal Wind
OMEG70	700 hPa Omega
UREL70	700 hPa U Wind Component
VREL70	700 hPa V Wind Component
RELH70	700 hPa Relative Humidity
MFCN70	700 hPa Moisture Flux Convergence
TWAD70	700 hPa Theta-W Advection
WDIV70	700 hPa Wind Divergence
RVRG70	700 hPa Relative Vorticity
FRNT70	700 hPa Frontogenesis
LHGT70	700 hPa Laplacian of the Geopotential Height
LTHA70	700 hPa Laplacian of Theta (Potential Temperature)
AWND70	700 hPa Ageostrophic Wind
GRDO70	700 hPa Gradient of Omega
THWC70	700 hPa Theta-W in Celsius
THTE70	700 hPa Theta-E
LP7870	700-850 hPa Theta Change/Pressure Difference 700-850
FRT7870	700-850 hPa Frontogenesis
SL7870	Sat. Theta-E Diff Divided by the 700-800 hPa Pressure Diff

Table A3: Predictors (Table A2) continued.

Name	Description
ML7870	700-850 hPa Theta-E Diff Divided by the Pressure Diff
SP7870	Saturated Geostrophic Potential Vorticity 700-850 hPa
SP5770	Saturated Geostrophic Potential Vorticity 500-700 hPa
SP3570	Saturated Geostrophic Potential Vorticity 300-500 hPa
MP7870	Moist Geostrophic Potential Vorticity 700-850 hPa
MP5770	Moist Geostrophic Potential Vorticity 500-700 hPa
MP3570	Moist Geostrophic Potential Vorticity 300-500 hPa
LC7870	Average Convergence of Theta Change/Pressure Difference 700-850 hPa
LC5770	Average Convergence of Theta Change/Pressure Difference 500-700 hPa
LC3570	Average Convergence of Theta Change/Pressure Difference 300-500 hPa
SH7870	700-850 hPa Speed Shear
MT7870	700-850 hPa Max Thermal Wind
SH7N70	700-1000 hPa Speed Shear
OMEG50	500 hPa Omega
UREL50	500 hPa U Wind Component
VREL50	500 hPa V Wind Component
RELH50	500 hPa Relative Humidity
LP5750	700-500 hPa Theta Change/Pressure Difference
LPS550	Surface-500 hPa Theta Change/Pressure Difference
FRT5750	500-700 hPa Frontogenesis
SL5750	Sat. Theta-E Diff Divided by the 500-700 hPa Pressure Diff
ML5750	500-700 hPa Theta-E Diff Divided by the Pressure Diff
AWND50	500 hPa Ageostrophic Wind
WDIV50	500 hPa Wind Divergence
FRNT50	500 hPa Frontogenesis
GRDO50	500 hPa Gradient of Omega
RVRG50	500 hPa Relative Vorticity
LHGT50	500 hPa Laplacian of the Geostrophic Height
LTHA50	500 hPa Laplacian of Theta (Potential Temperature)
SH5750	500-700 hPa Speed Shear
MT5750	500-700 hPa Max Thermal Wind
SH5850	500-850 hPa Speed Shear
SH5N50	500-1000 hPa Speed Shear
TL7550	700-500 hPa Temperature Difference
LR7550	700-500 hPa Lapse Rate
LR8550	850-500 hPa Lapse Rate
OMEG30	300 hPa Omega

Table A4: Predictors (Table A2) continued.

Name	Description
UREL30	300 hPa U Wind Component
VREL30	300 hPa V Wind Component
LP3530	300-500 hPa Theta Change/Pressure Difference
FRT3530	300-500 hPa Frontogenesis
SL3530	Sat. Theta-E Diff Divided by the 300-500 hPa Pressure Diff
ML3530	300-500 hPa Theta-E Diff Divided by the Pressure Diff
AWND30	300 hPa Ageostrophic Wind
WDIV30	300 hPa Wind Divergence
FRNT30	300 hPa Frontogenesis
GRDO30	300 hPa Gradient of Omega
RVRG30	300 hPa Relative Vorticity
LHGT30	300 hPa Laplacian of the Geopotential Height
LTHA30	300 hPa Laplacian of Theta (Potential Temperature)
SH3830	300-850 hPa Speed Shear
SH3530	300-500 hPa Speed Shear
MT3530	300-500 hPa Max Thermal Wind
SH3730	300-700 hPa Speed Shear
SH3N30	300-1000 hPa Speed Shear
AUAW30	300 hPa Magnitude of the Geostrophic Wind
PVOR30	300 hPa Potential Vorticity
CAPE30	MUCAPE between Surface-300 hPa
CINS30	Surface-300 hPa Convective Inhibition
PWTR0	Surface Precipitable Water
THTA0	Surface Theta (Potential Temperature)
CAPE0	Surface Based Cape
CINS0	Surface Based Convective Inhibition
LIFT0	Lifted Index based on MUCAPE Parcel
HLCY0	Helicity
CCTL0	Convective Cloud Top in meters
TMPC0	Surface Temperature in Celsius
MIXR0	Surface Mixing Ratio
UREL0	Surface U Wind Component
VREL0	Surface V Wind Component
PMSL0	Mean Sea-Level Pressure
LMSL0	Laplacian of the Mean Sea-Level Pressure
RELH0	Surface Relative Humidity
THTE0	Surface Theta-E
LTHA0	Surface Laplacian of Theta (Potential Temperature)

Table A5: Predictors (Table A2)continued.

Name	Description
ZMEG0	Terrain Upslope/Downslope
FRNT0	Surface Frontogenesis
MFCN0	Surface Moisture Flux Convergence
WDIV0	Surface Wind Divergence
TWAD0	Surface Theta-W Divergence
RVRG0	Surface Relative Vorticity
THWC0	Surface Theta-W in Celsius
AVETOTF	Total Lightning Flashes Climatology
AVECCFL	Cloud Contained (Low kAmp) Flash Climatology
AVEFONE	One or More Lightning Flashes Climatology
AVEFTHR	Three or More Lightning Flashes Climatology
AVEFTEN	Ten or More Lightning Flashes Climatology
AVEFTHT	Thirty or More Lightning Flashes Climatology
AVEFHUN	One Hundred or More Lightning Climatology
AVETSML	Three Hour Segments of 15-Minute Lightning Bins Climatology
BTOTF	CAPE30 * AVETOTF
BCCFL	CAPE30 * AVECCFL
BFONE	CAPE30 * AVEFONE
BFTHR	CAPE30 * AVEFTHR
BFTEN	CAPE30 * AVEFTEN
BFTHT	CAPE30 * AVEFTHT
BFHUN	CAPE30 * AVEFHUN
BTSML	CAPE30 * AVETSML
AVEPTOT	Total Precipitation Climatology (inches)
AVEPHDTH	Hundredth or More Climatology
AVEPTNTH	Tenth or More Climatology
AVEDRY1	Dry Thunder 1 Climatology (DRYTH1)
AVEDRY2	Dry Thunder 2 Climatology (DRYTH2)

Appendix B

Further Climatology Examples

Images of lightning and precipitation climatology are found in this Appendix. Following the pattern of DRYTH1 and DRYTH2 in Section 2.5, images represent values from every other pentad (10 days). For lightning, the full May-Sep values of AVETOTF and AVEFONE are shown. Precipitation values include the climatology for AVEPTOT and AVEPHDTH. The timestamp format here (e.g. 000707/0000) represents the year, month, day, and hour (yymmdd/hhhh); all climatology grids are placed at a year 2000 date.

As mentioned in Chapter 2, the units for these parameters are different between the fields. AVETOTF is measured in units of flashes, and AVEFONE values are measured in percentage. AVEPTOT fields are represented by hundredths of an inch, while AVEPHDTH is described by percentage. In other words, total amount parameters (AVETOTF and AVEPTOT) remain in their original units of measure. Binary fields described by the thresholds in Table 2.1 result in an average value described by percentage. Note that the magnitude of percentage may differ between the CONUS and AK domains.

Lightning fields (AVETOTF and AVEFONE) for the CONUS are shown first, followed by the precipitation climatologies (AVEPTOT and AVEPHDTH). Fields for Alaska follow the CONUS examples. Climatologies for other binary fields from Table 2.1 are not shown in this paper.

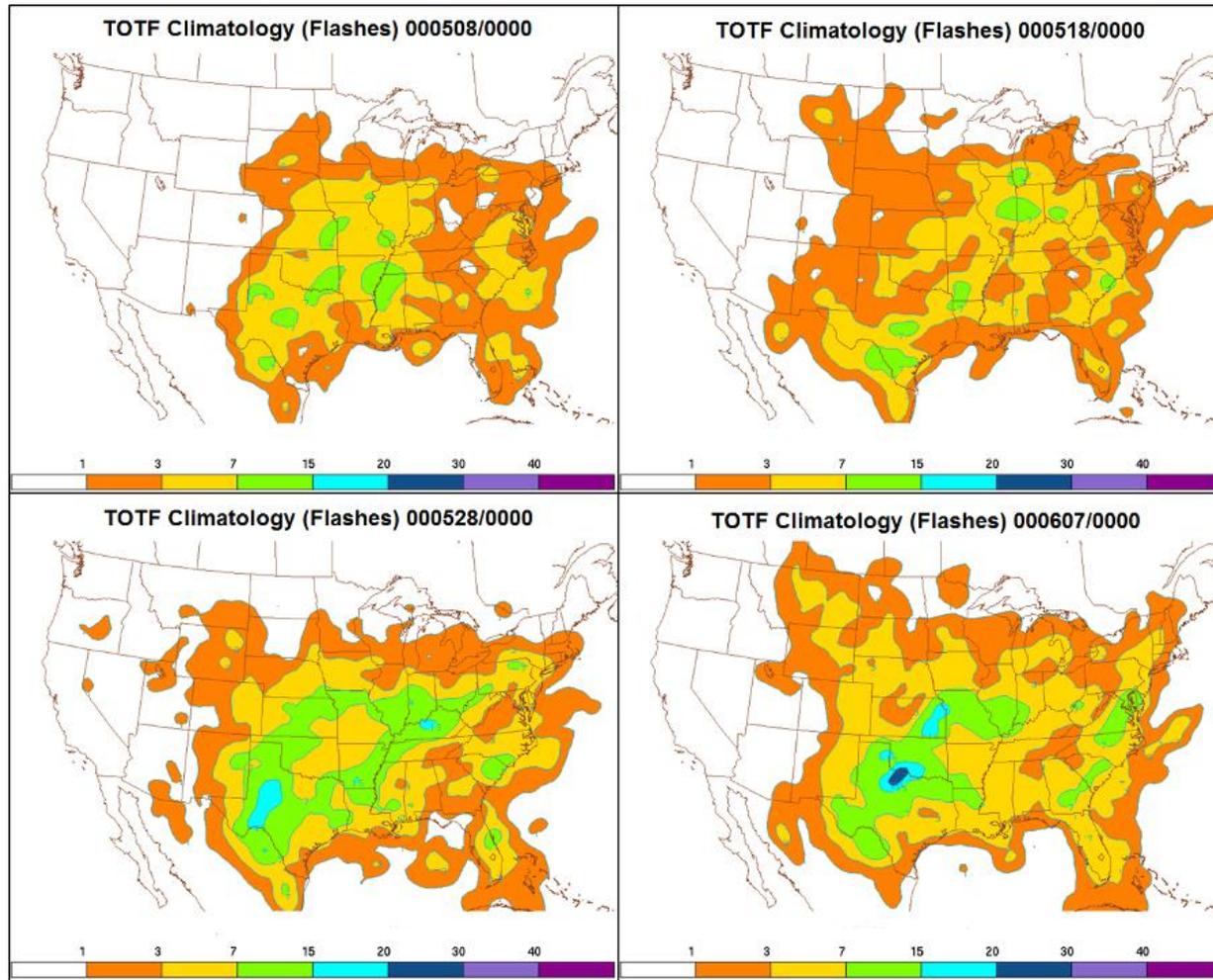


Figure B1: TOTF Climatology (AVETOTF) pentad examples for the CONUS from early May through early June – 00 UTC data.

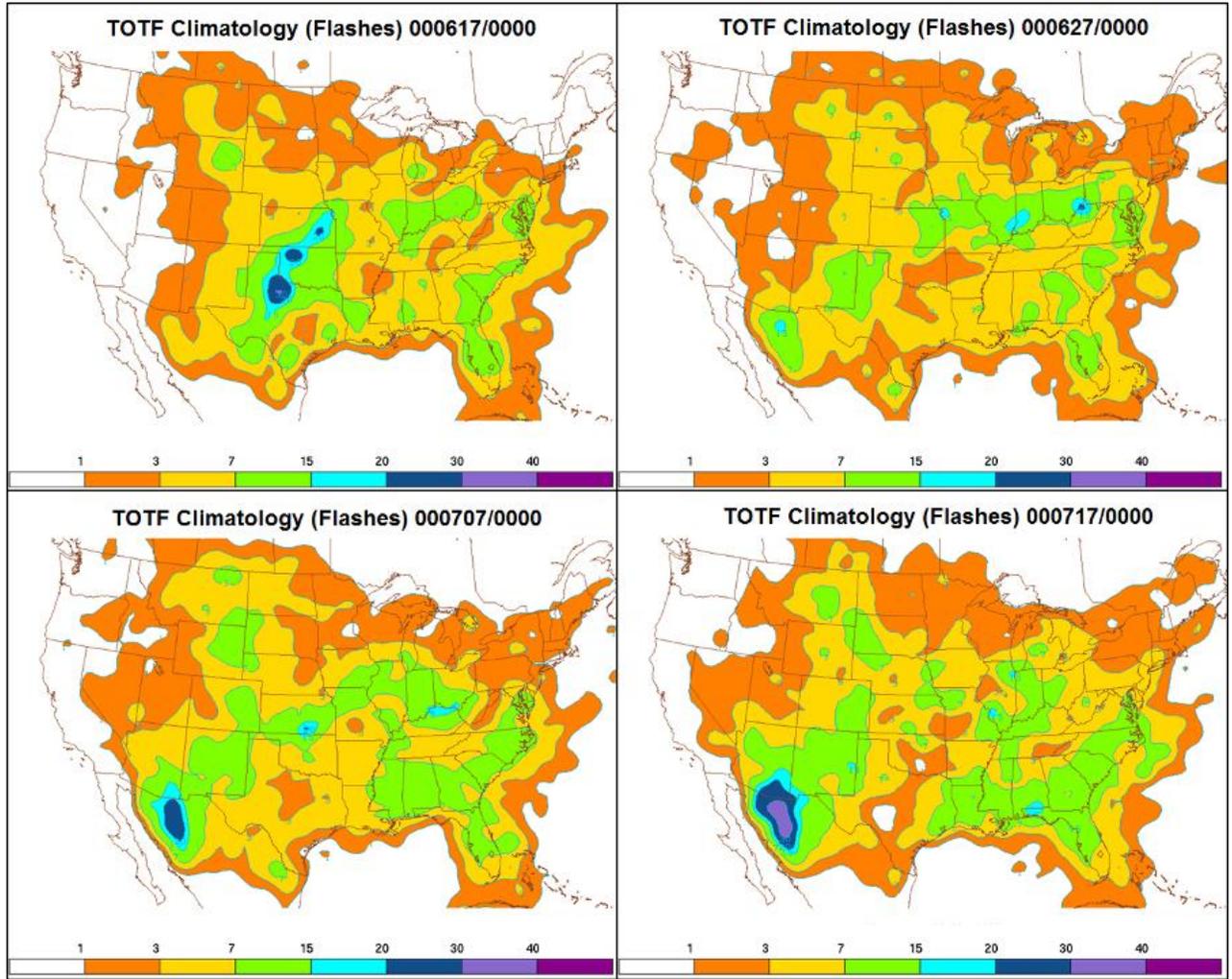


Figure B2: TOTF Climatology (AVETOTF) pentad examples for the CONUS from mid June through mid July – 00 UTC data.

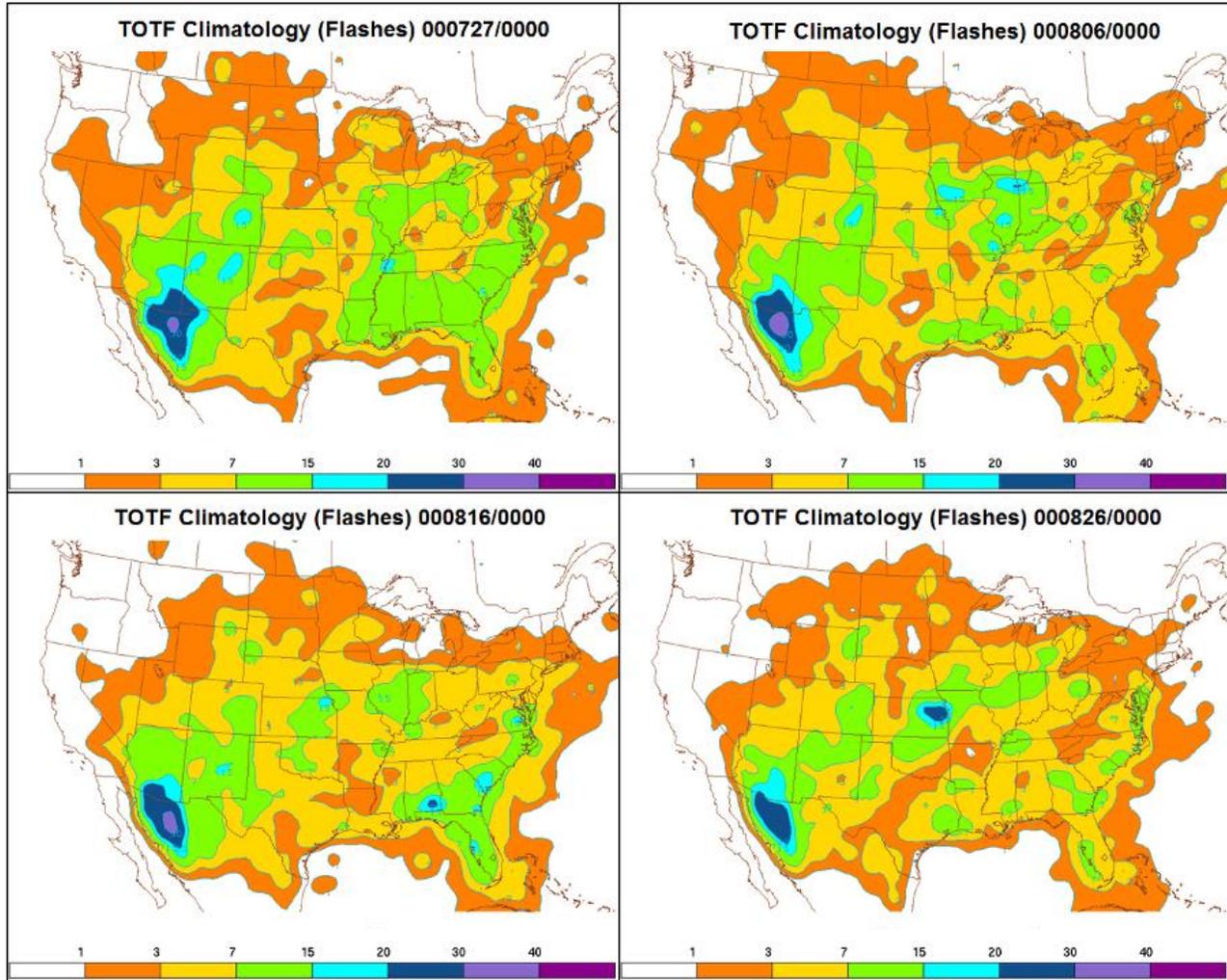


Figure B3: TOTF Climatology (AVETOTF) pentad examples for the CONUS from late July through late August – 00 UTC data.

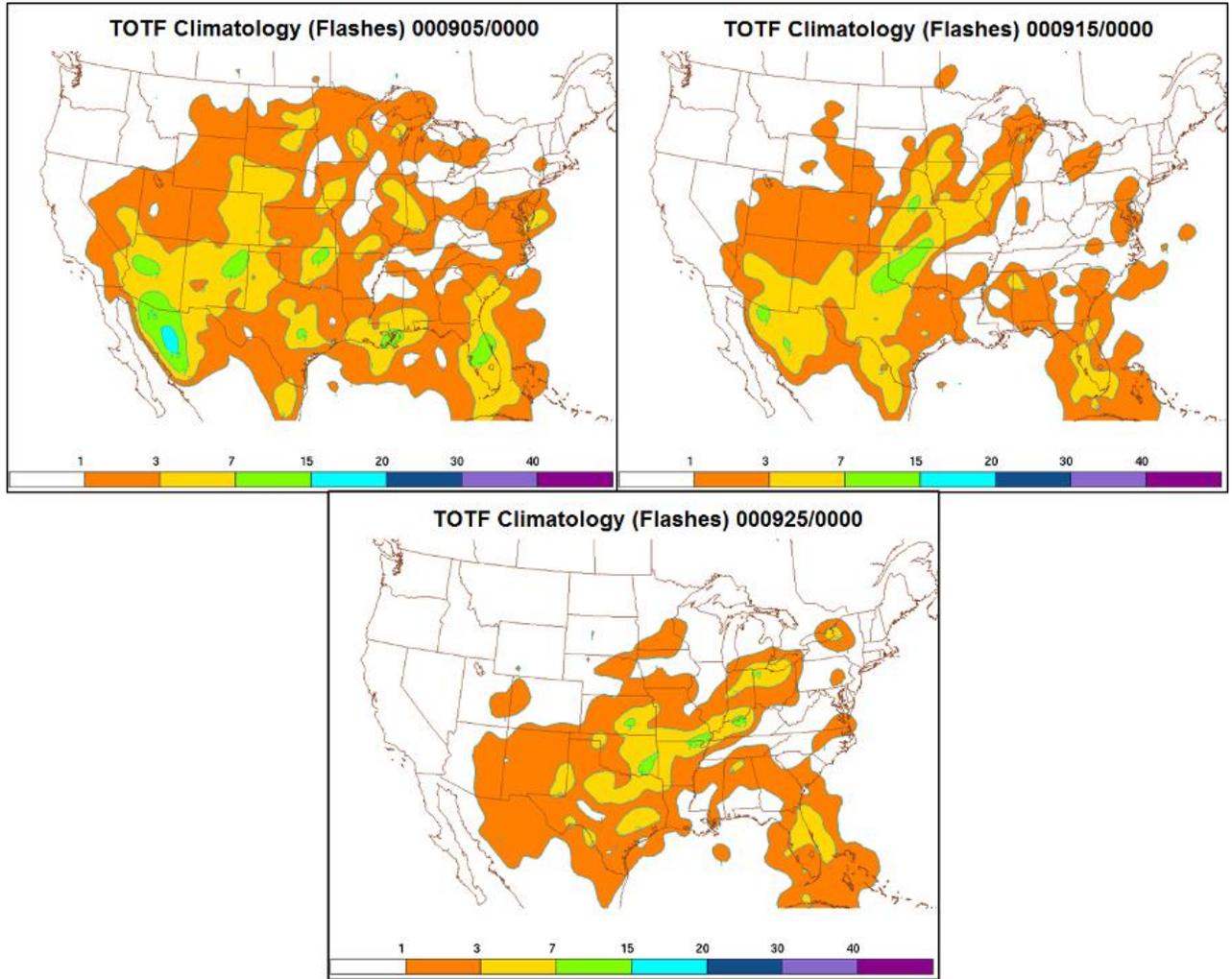


Figure B4: TOTF Climatology (AVETOTF) pentad examples for the CONUS from September – 00 UTC data.

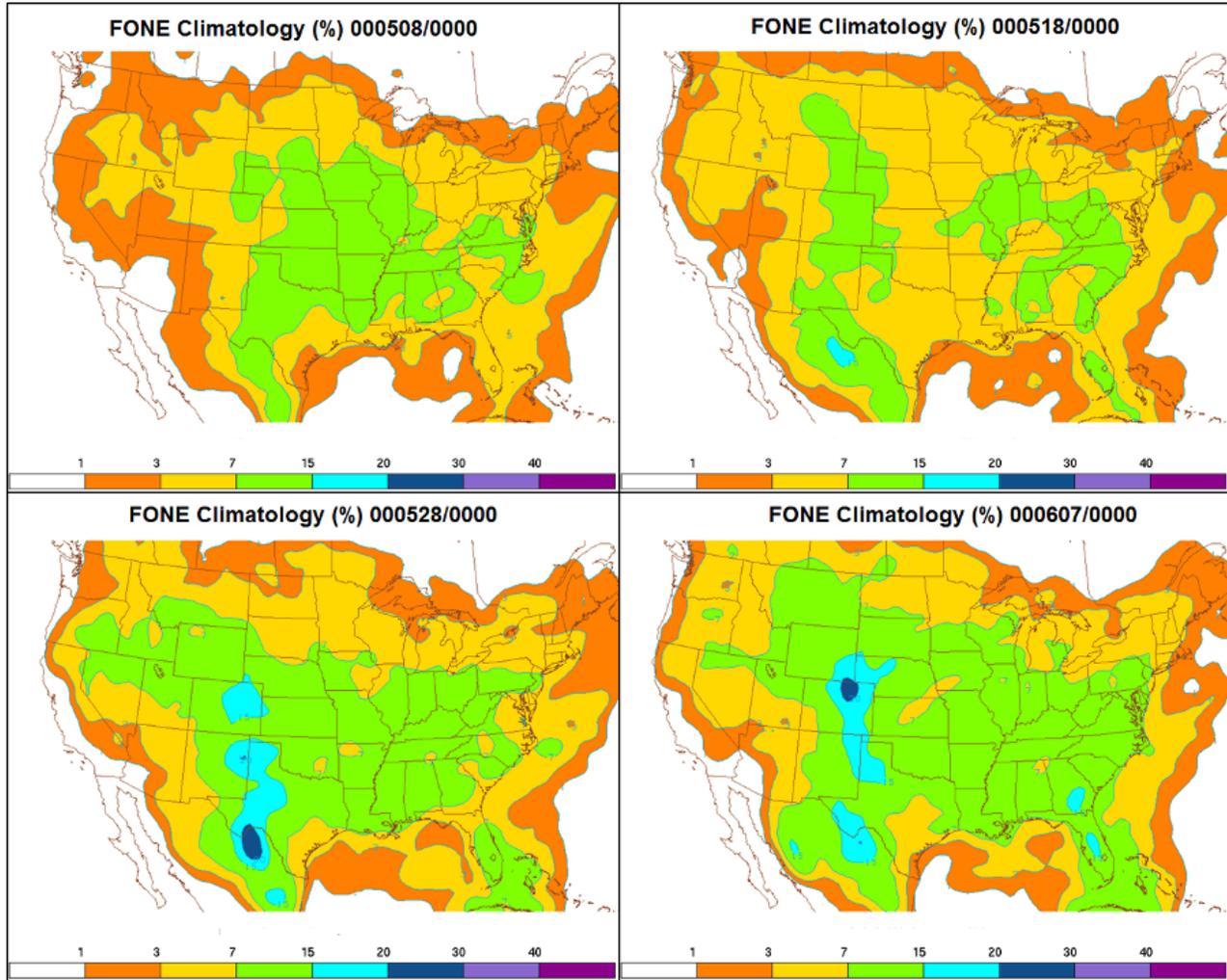


Figure B5: FONE Climatology (AVEFONE) pentad examples for the CONUS from early May through early June – 00 UTC data.

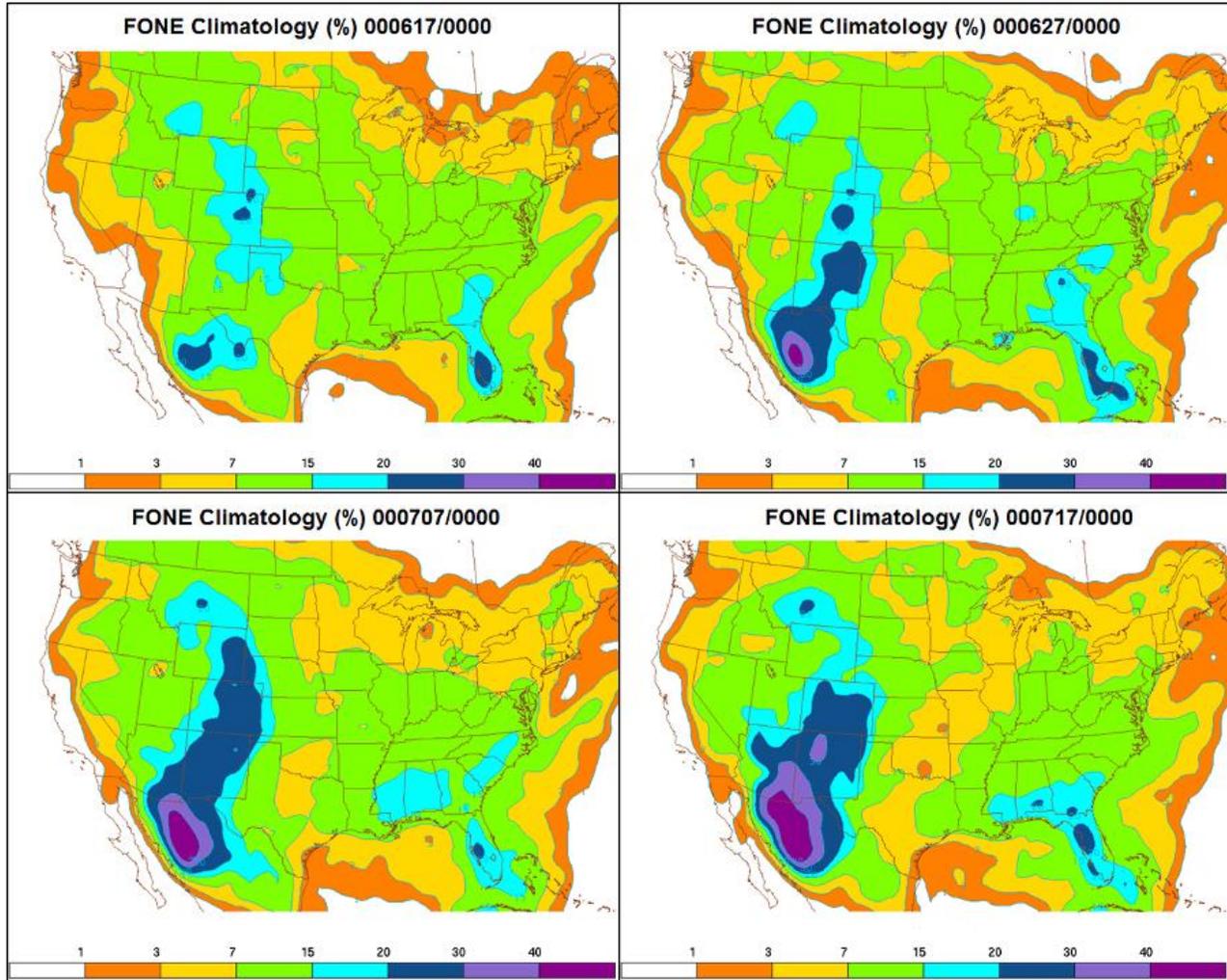


Figure B6: FONE Climatology (AVEFONE) pentad examples for the CONUS from mid June through mid July – 00 UTC data.

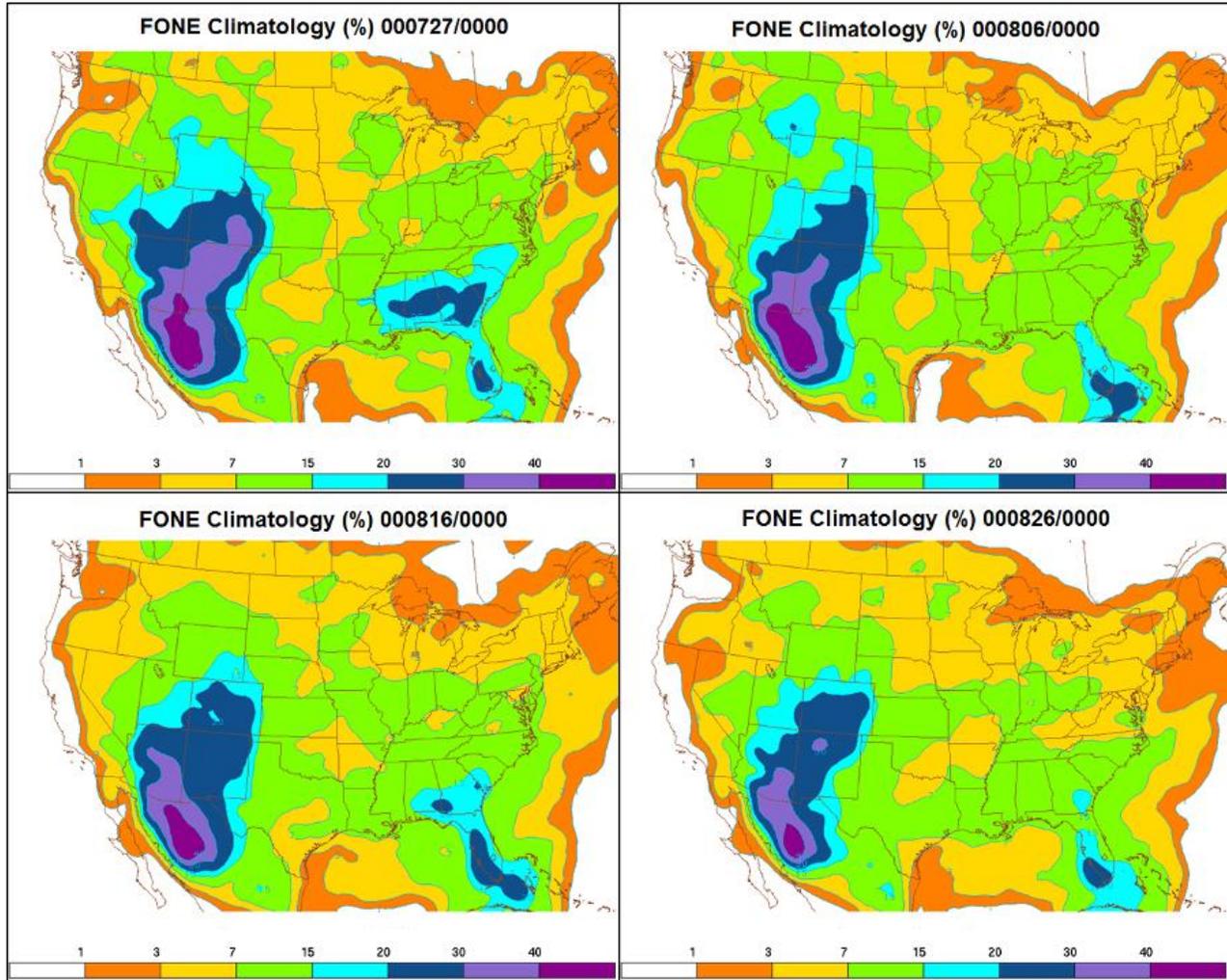


Figure B7: FONE Climatology (AVEFONE) pentad examples for the CONUS from late July through late August – 00 UTC data.

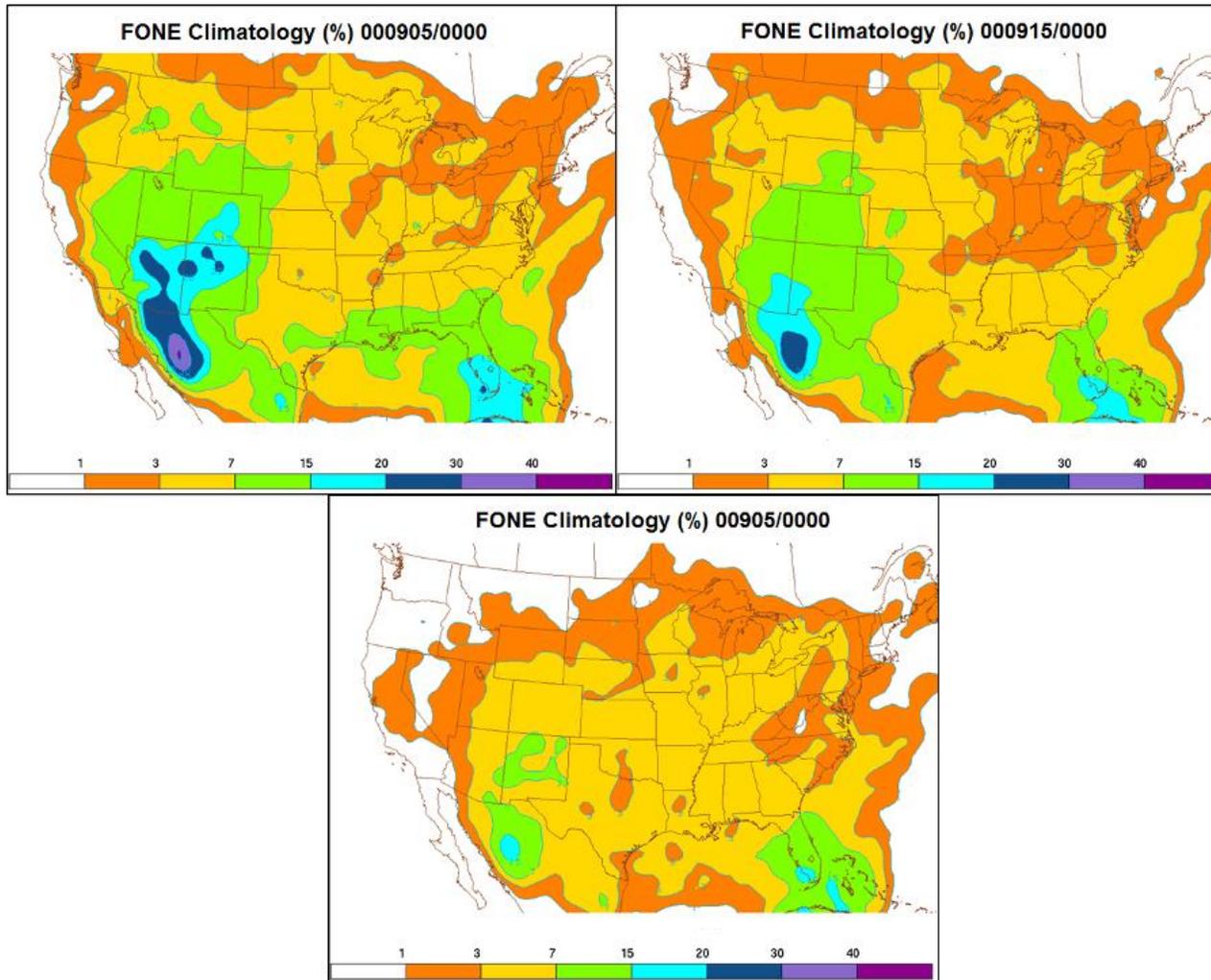


Figure B8: FONE Climatology (AVEFONE) pentad examples for the CONUS from September – 00 UTC data.

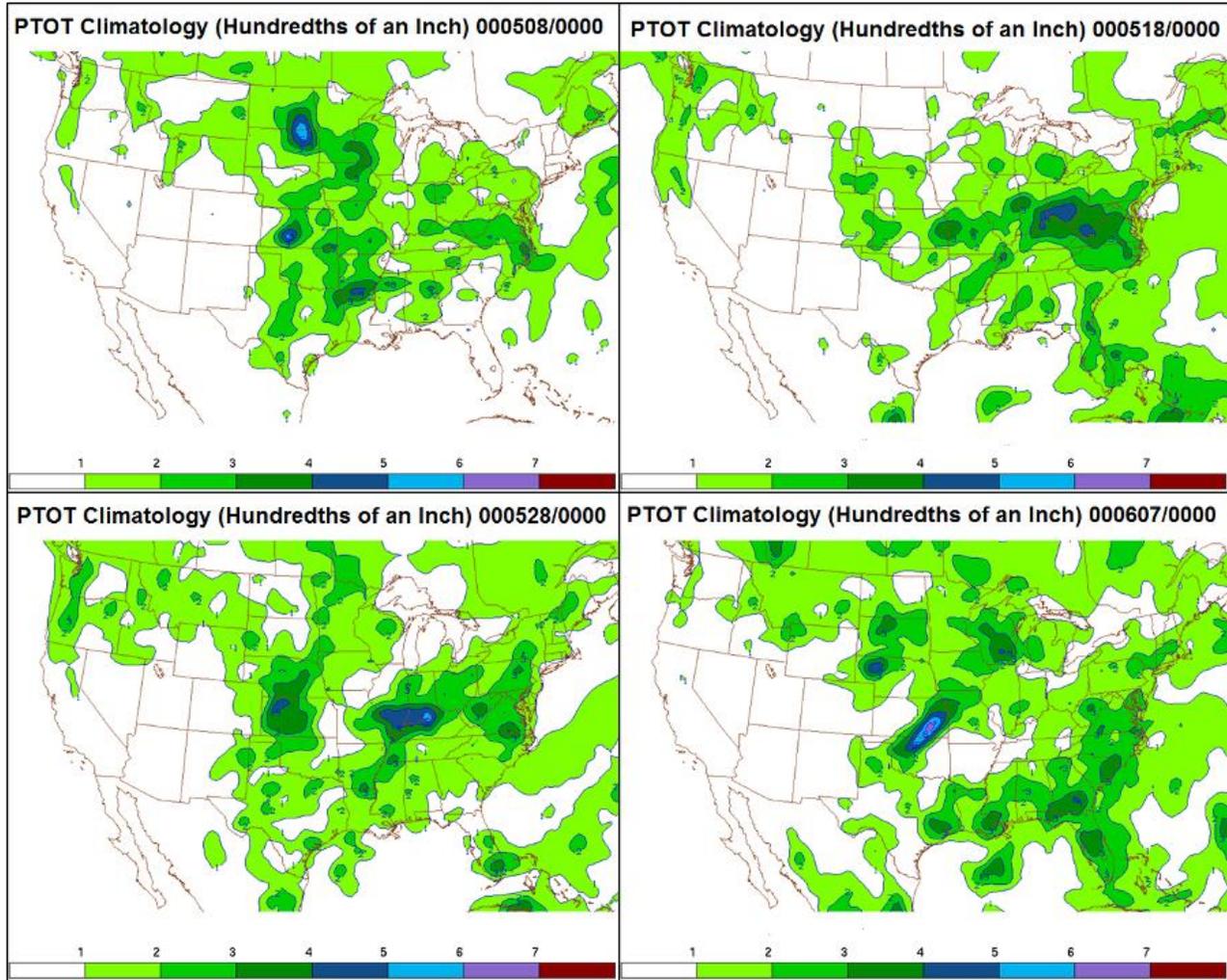


Figure B9: PTOT Climatology (AVEPTOT) pentad examples for the CONUS from early May through early June – 00 UTC data.

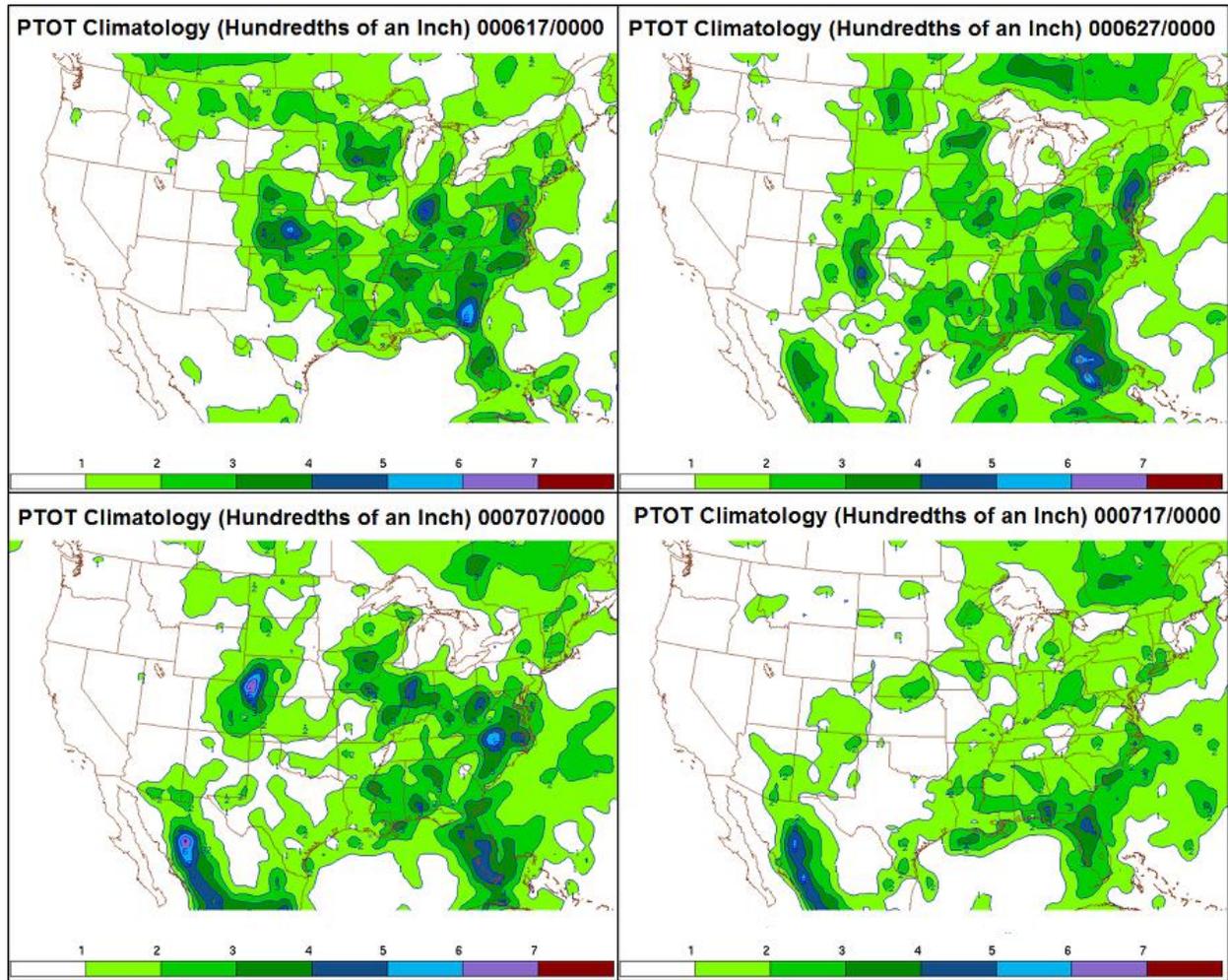


Figure B10: PTOT Climatology (AVEPTOT) pentad examples for the CONUS from mid June through mid July – 00 UTC data.

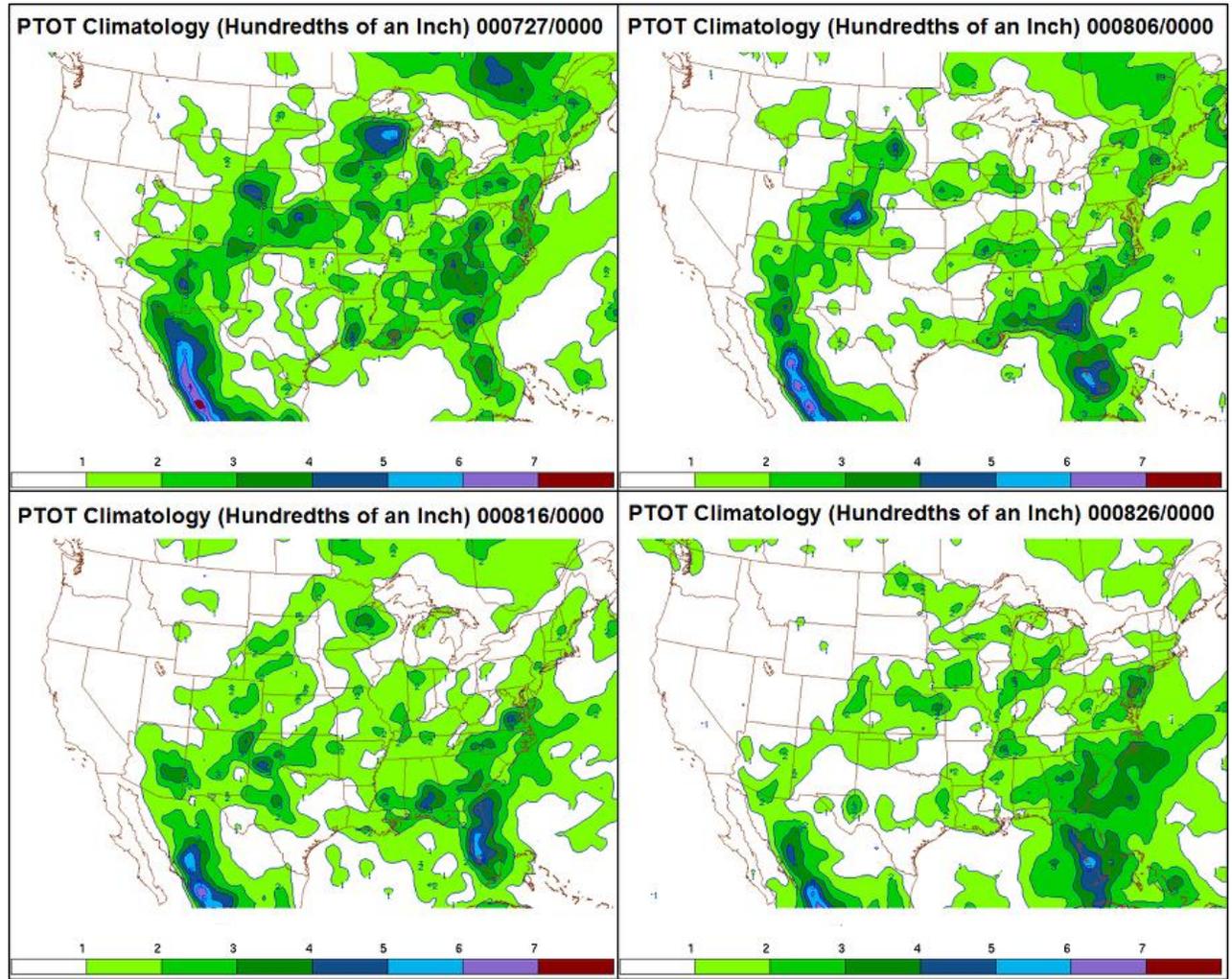


Figure B11: PTOT Climatology (AVEPTOT) pentad examples for the CONUS from late July through late August – 00 UTC data.

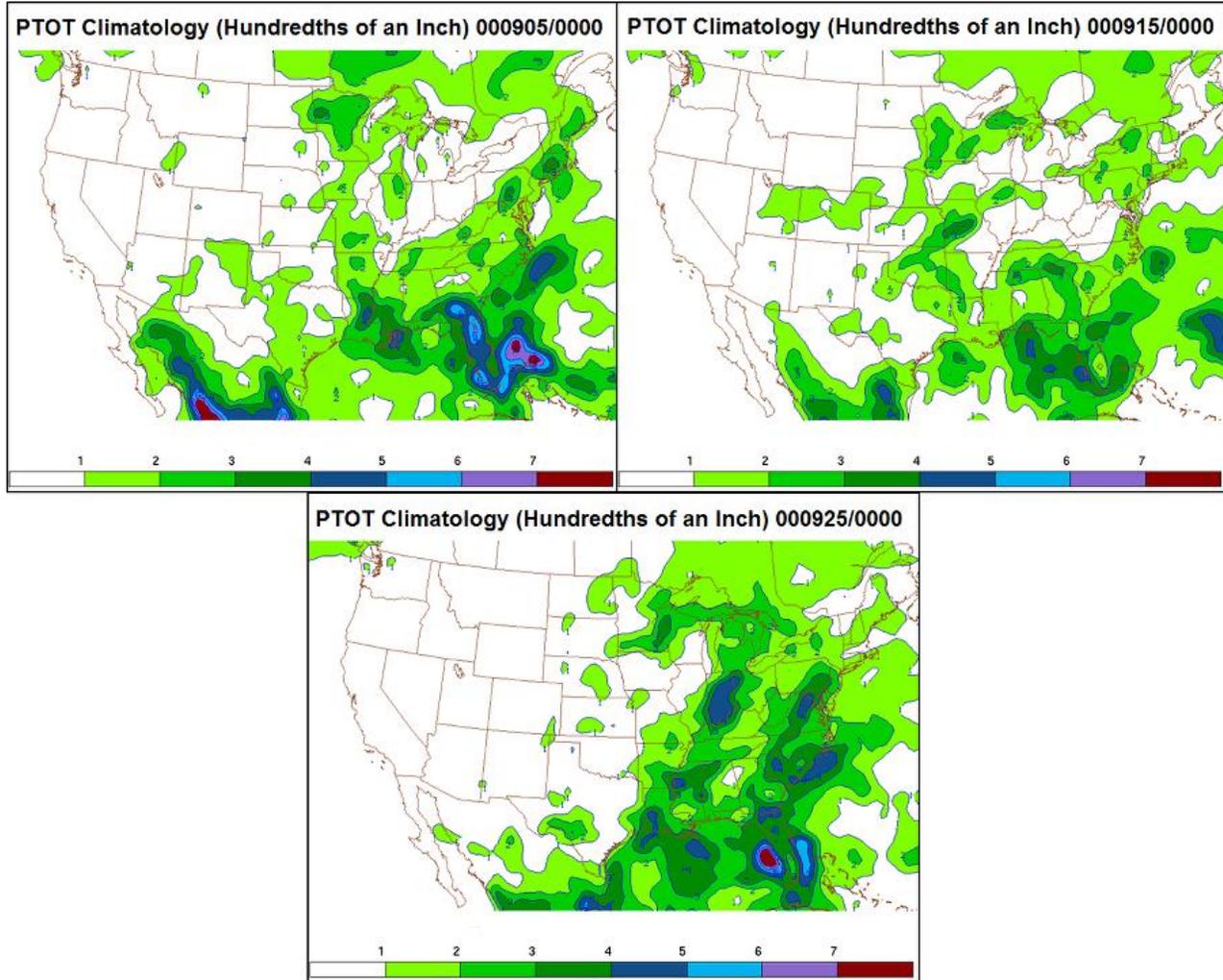


Figure B12: PTOT Climatology (AVEPTOT) pentad examples for the CONUS from September – 00 UTC data.

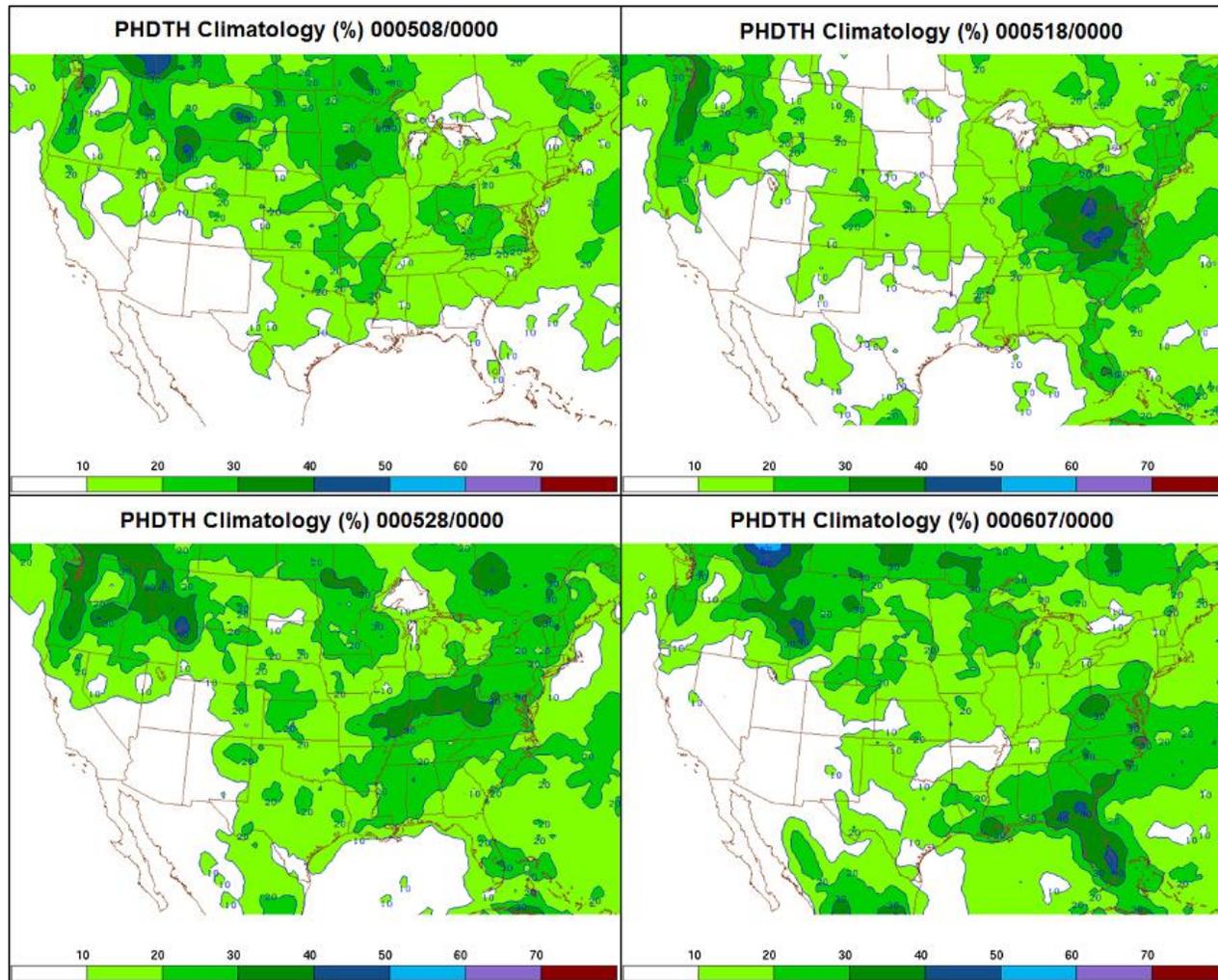


Figure B13: PHDTH Climatology (AVEPHDTH) pentad examples for the CONUS from early May through early June – 00 UTC data.

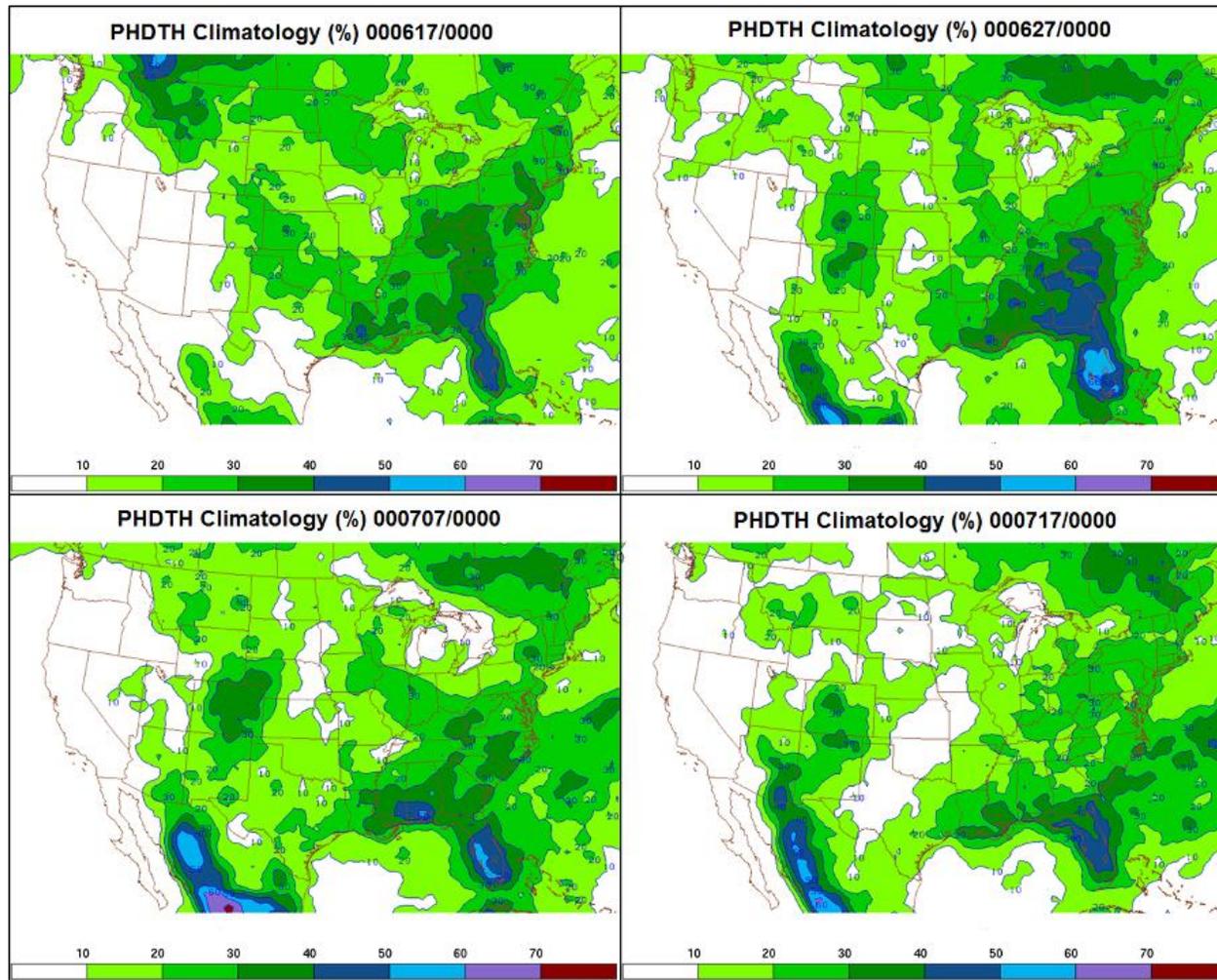


Figure B14: PHDTH Climatology (AVEPHDTH) pentad examples for the CONUS from mid June through mid July – 00 UTC data.

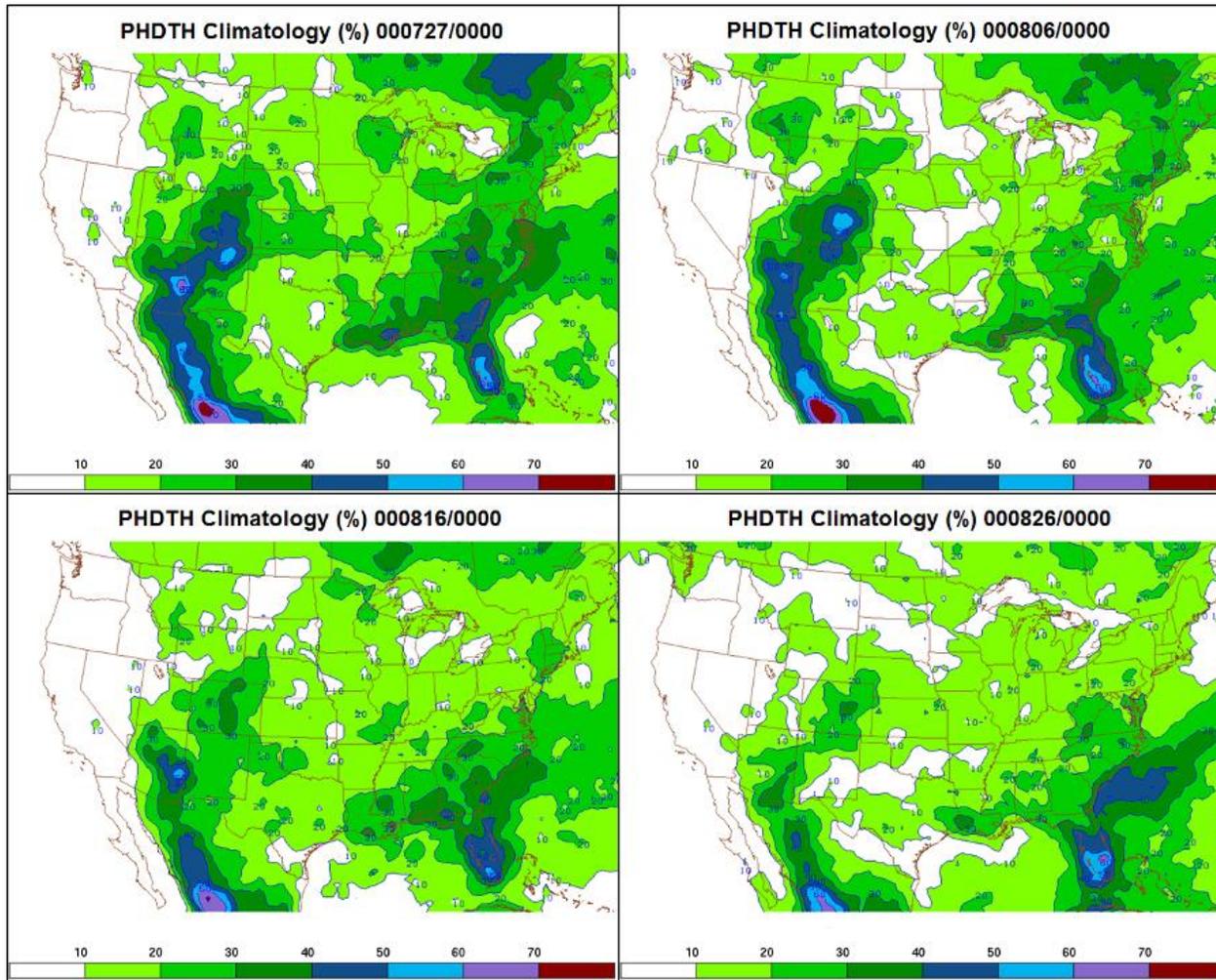


Figure B15: PHDTH Climatology (AVEPHDTH) pentad examples for the CONUS from late July through late August – 00 UTC data.

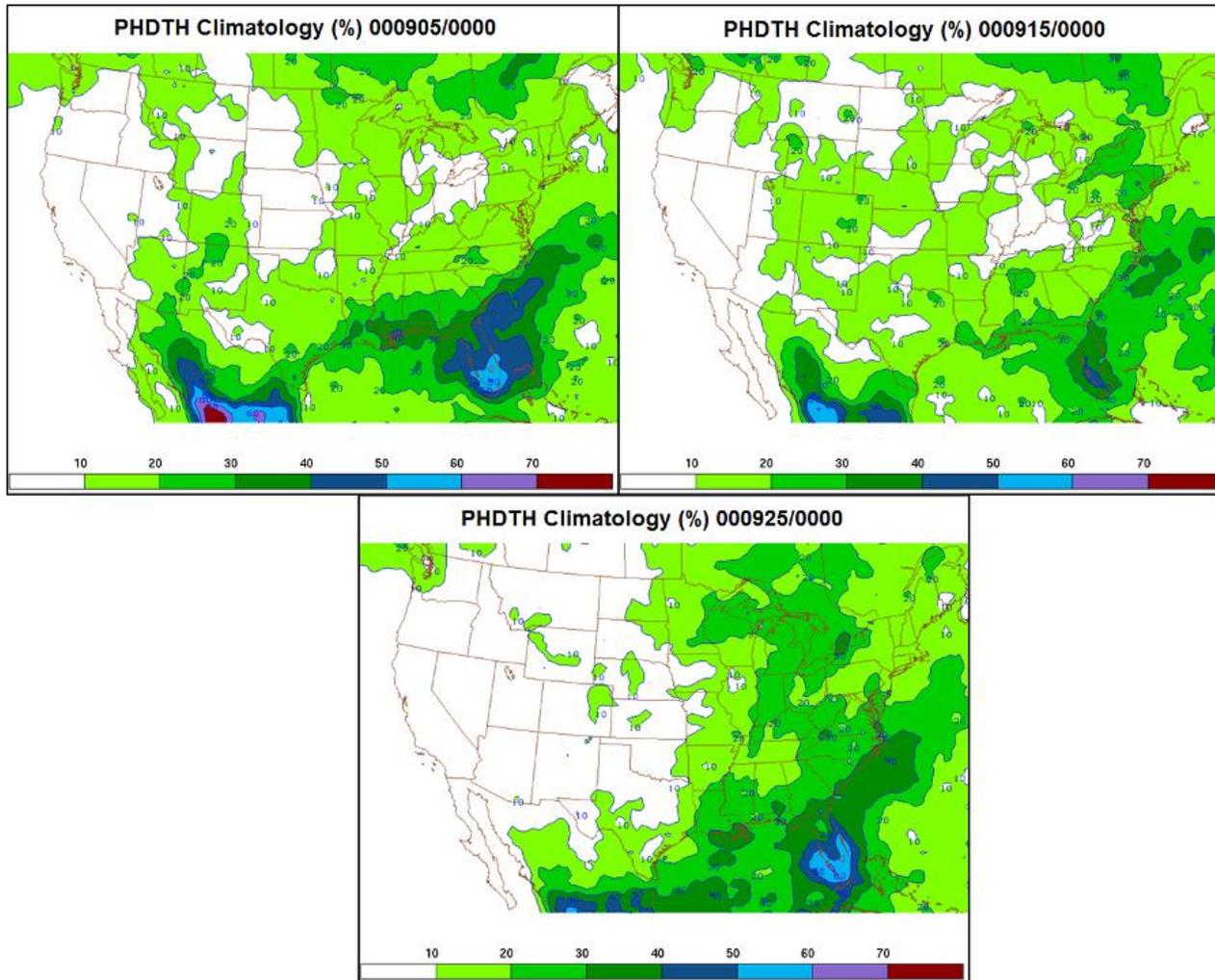


Figure B16: PHDTH Climatology (AVEPHDTH) pentad examples for the CONUS from September – 00 UTC data.

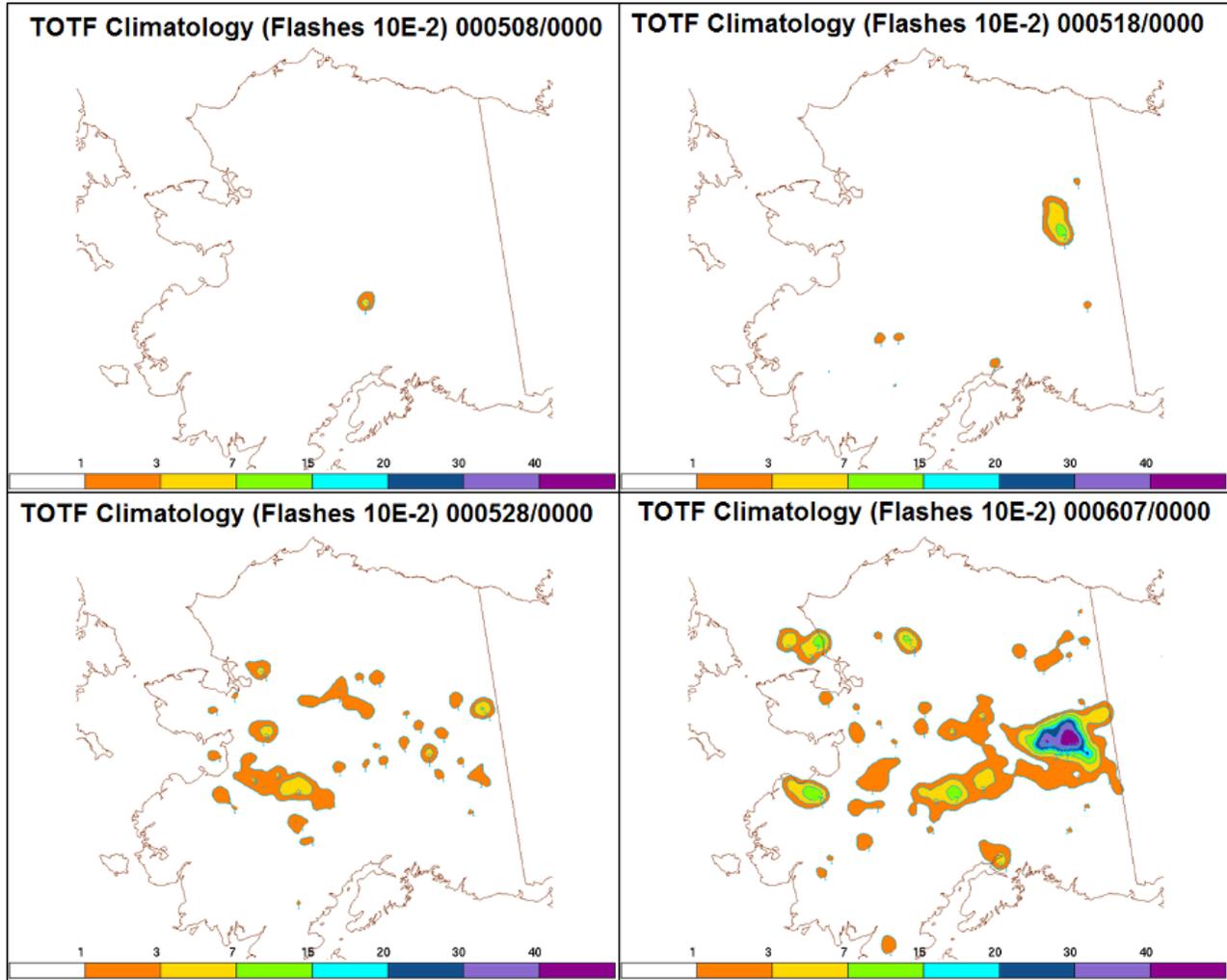


Figure B17: TOTF Climatology (AVETOTF) pentad examples for AK from early May through early June – 00 UTC data.

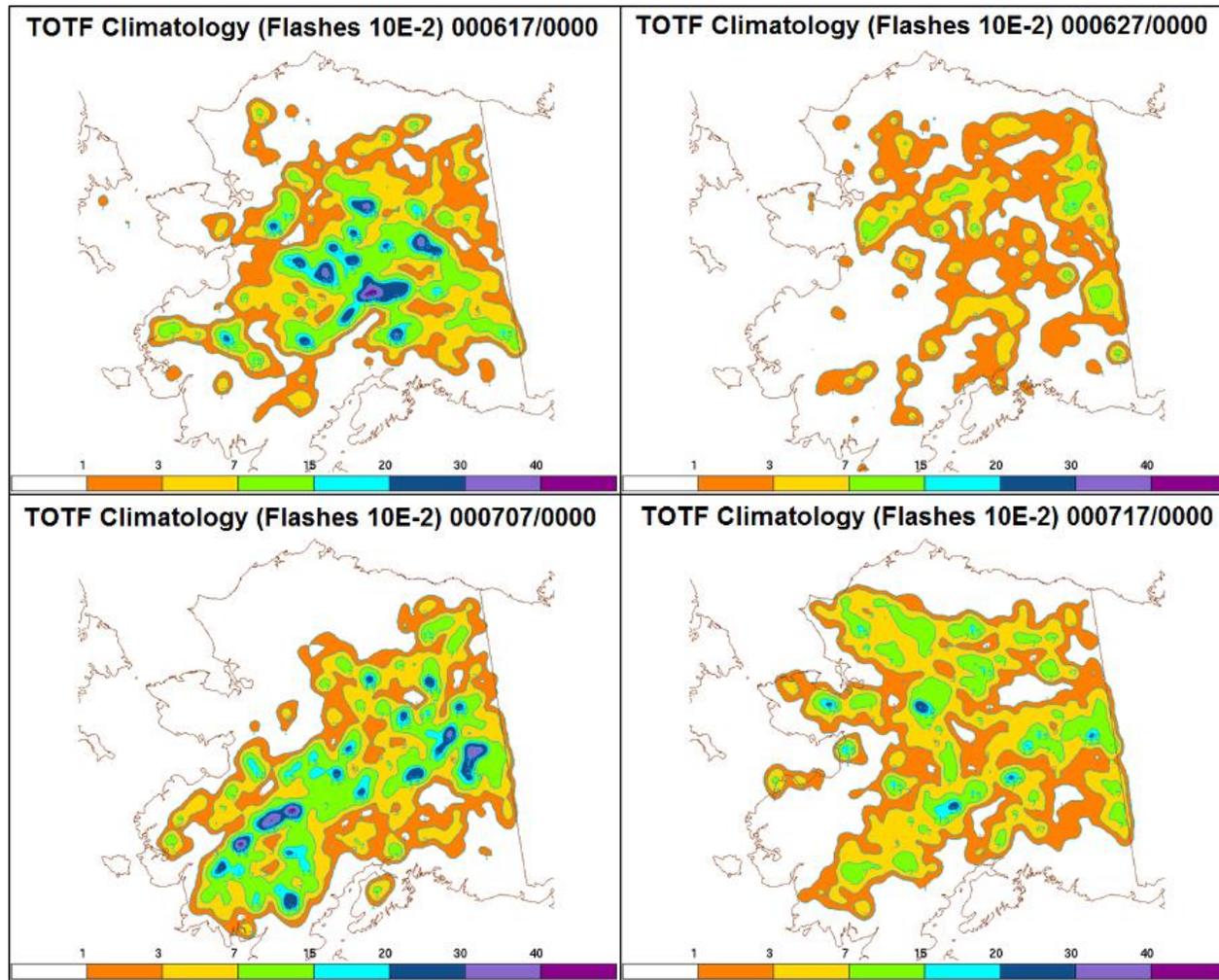


Figure B18: TOTF Climatology (AVETOTF) pentad examples for AK from mid June through mid July – 00 UTC data.

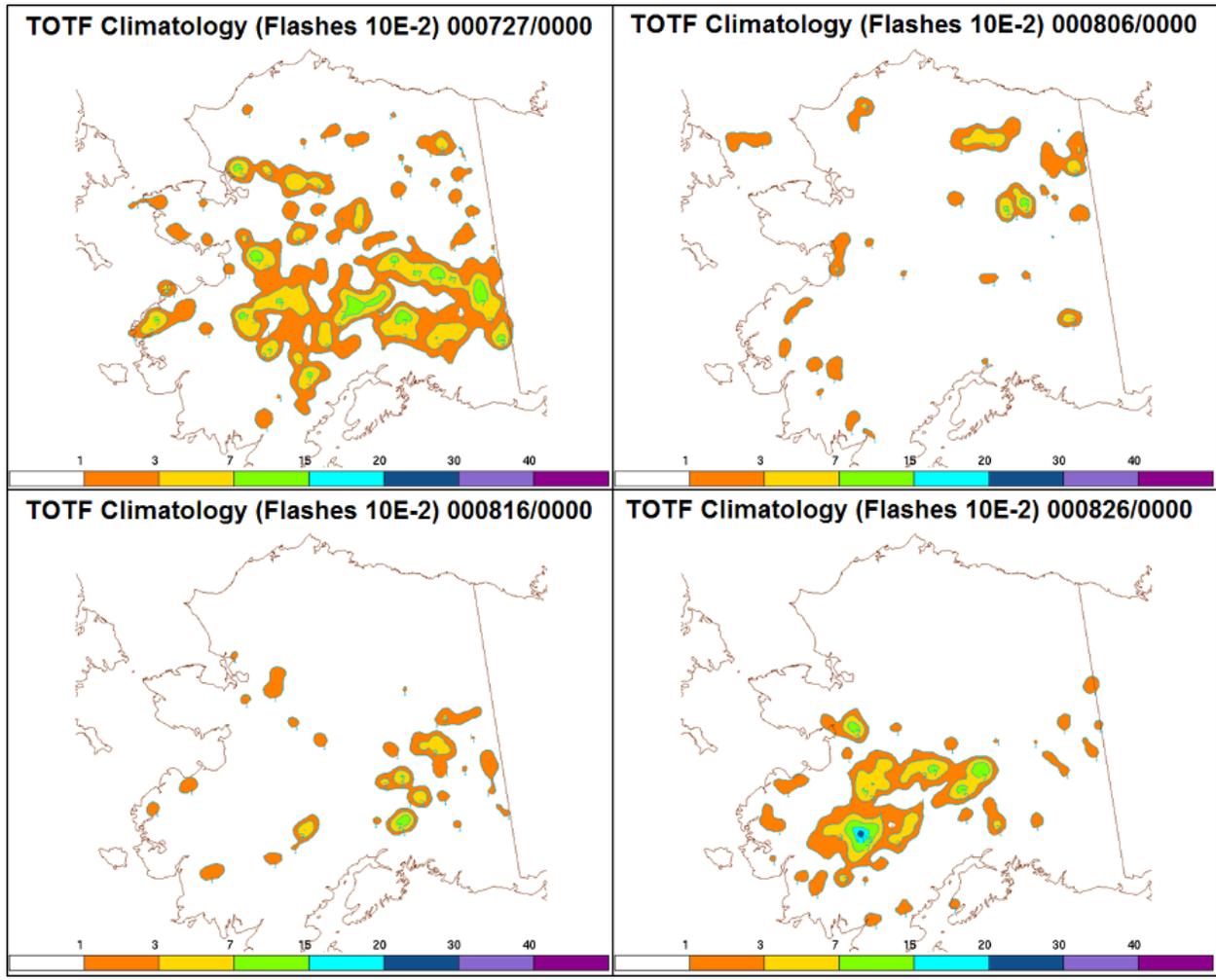


Figure B19: TOTF Climatology (AVETOTF) pentad examples for AK from late July through late August – 00 UTC data.

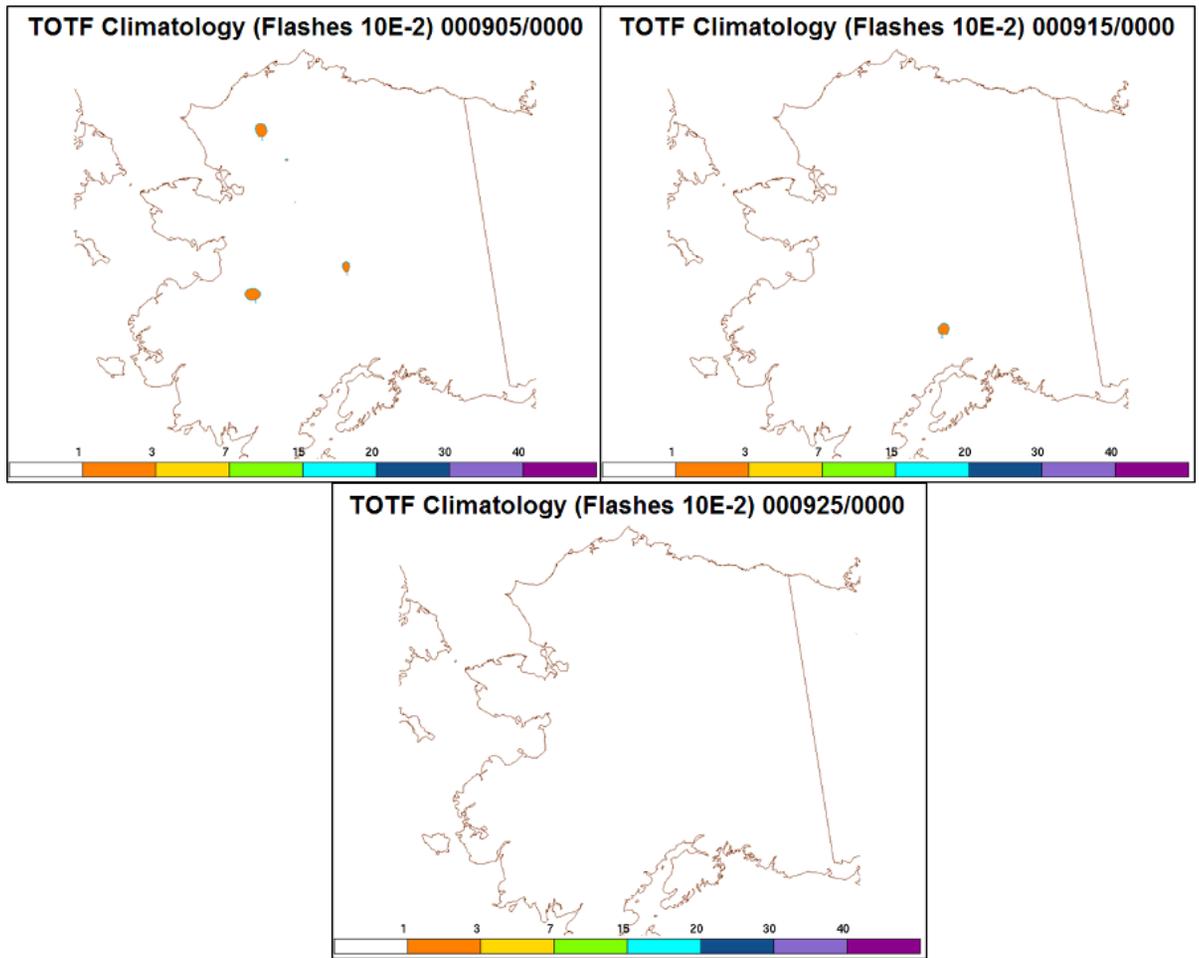


Figure B20: TOTF Climatology (AVETOTF) pentad examples for AK from September – 00 UTC data.

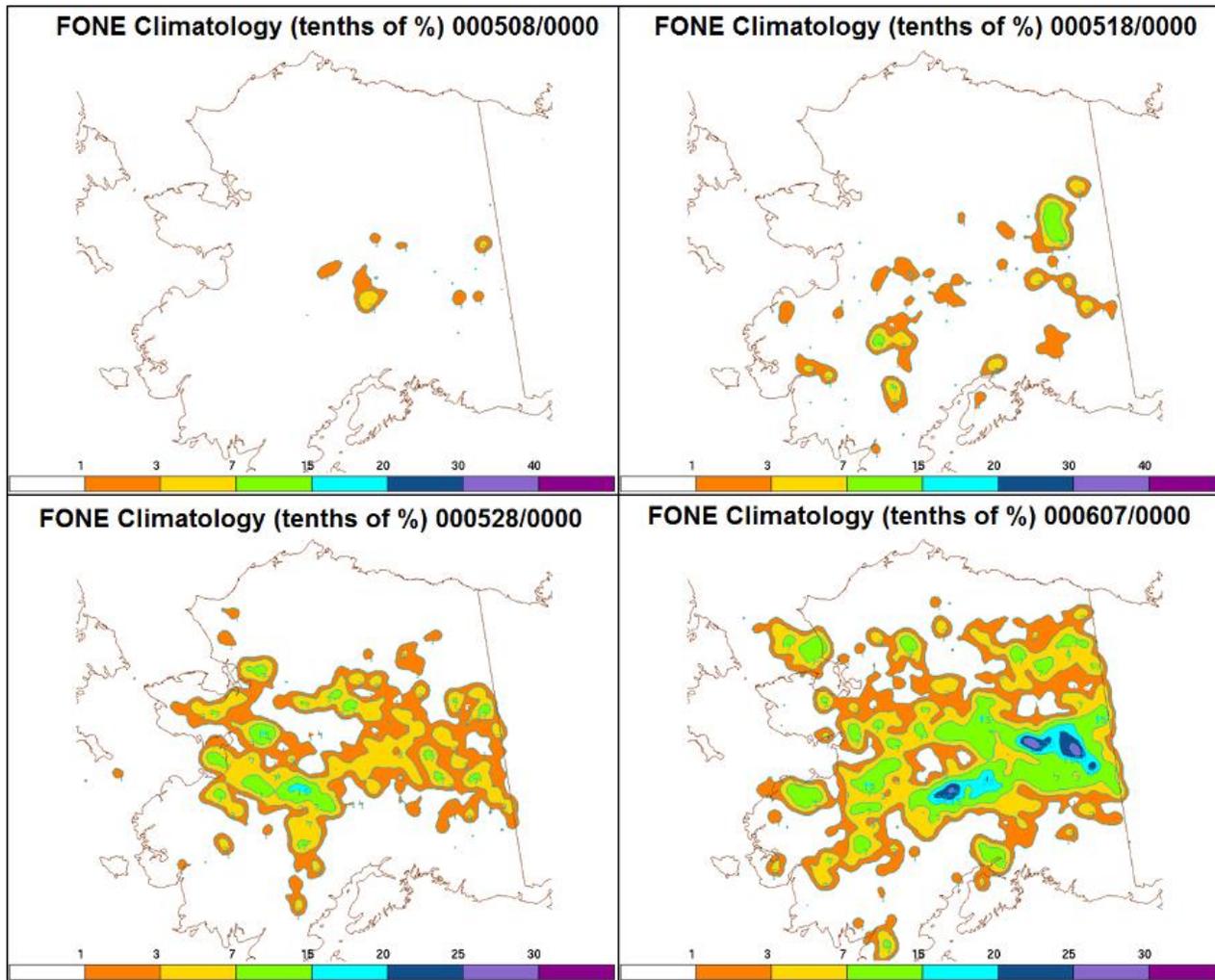


Figure B21: FONE Climatology (AVEFONE) pentad examples for AK from early May through early June – 00 UTC data.

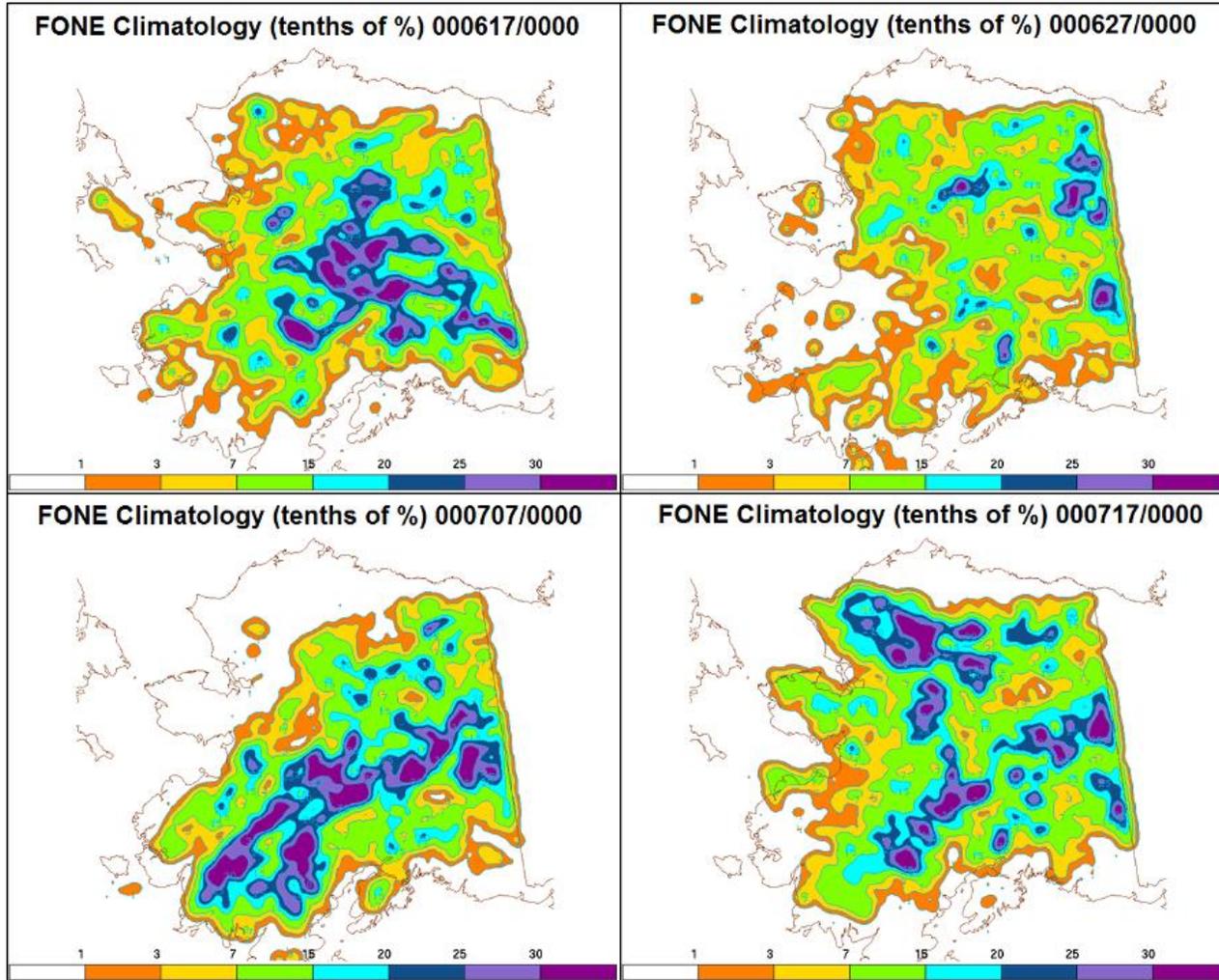


Figure B22: FONE Climatology (AVEFONE) pentad examples for AK from mid June through mid July – 00 UTC data.

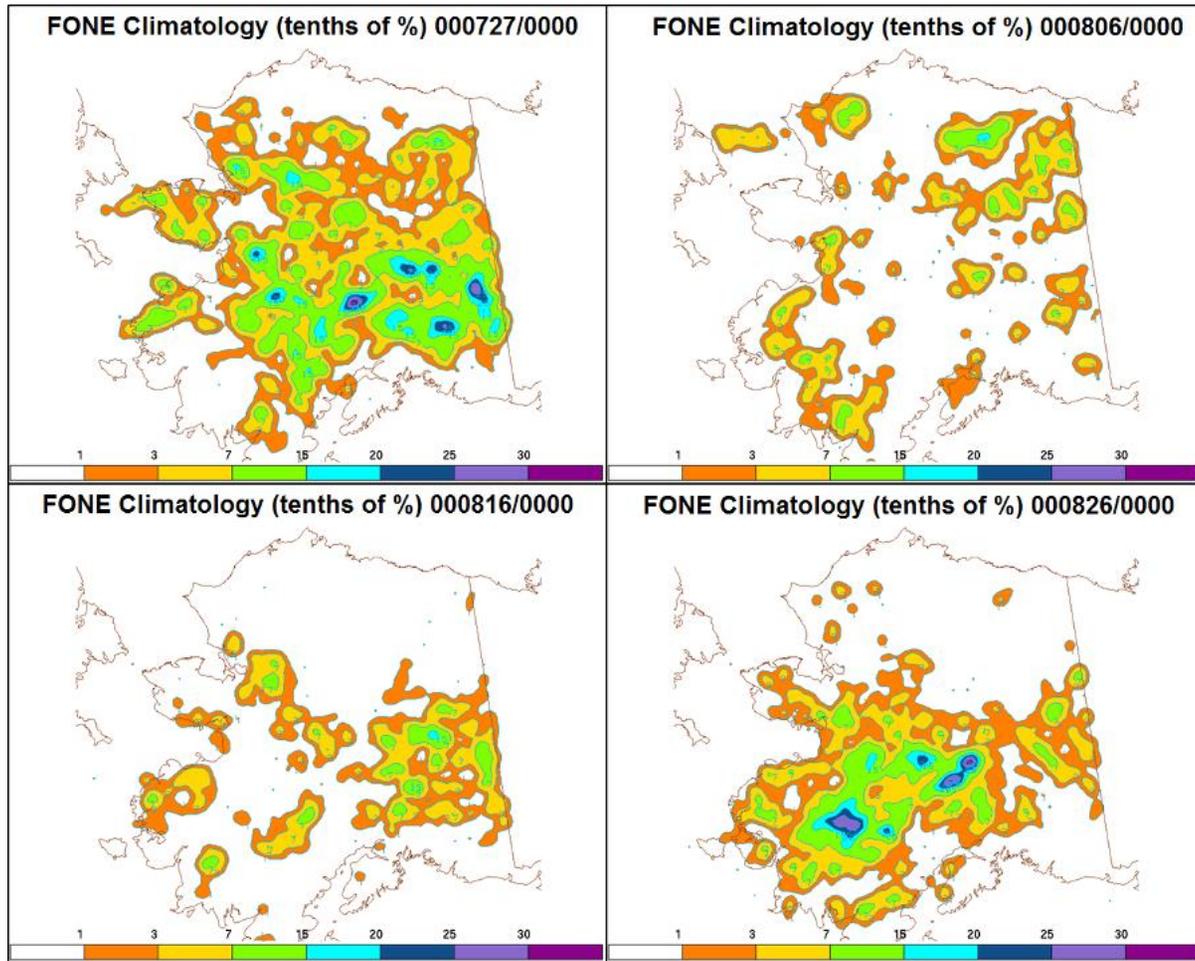


Figure B23: FONE Climatology (AVEFONE) pentad examples for AK from late July through late August – 00 UTC data.

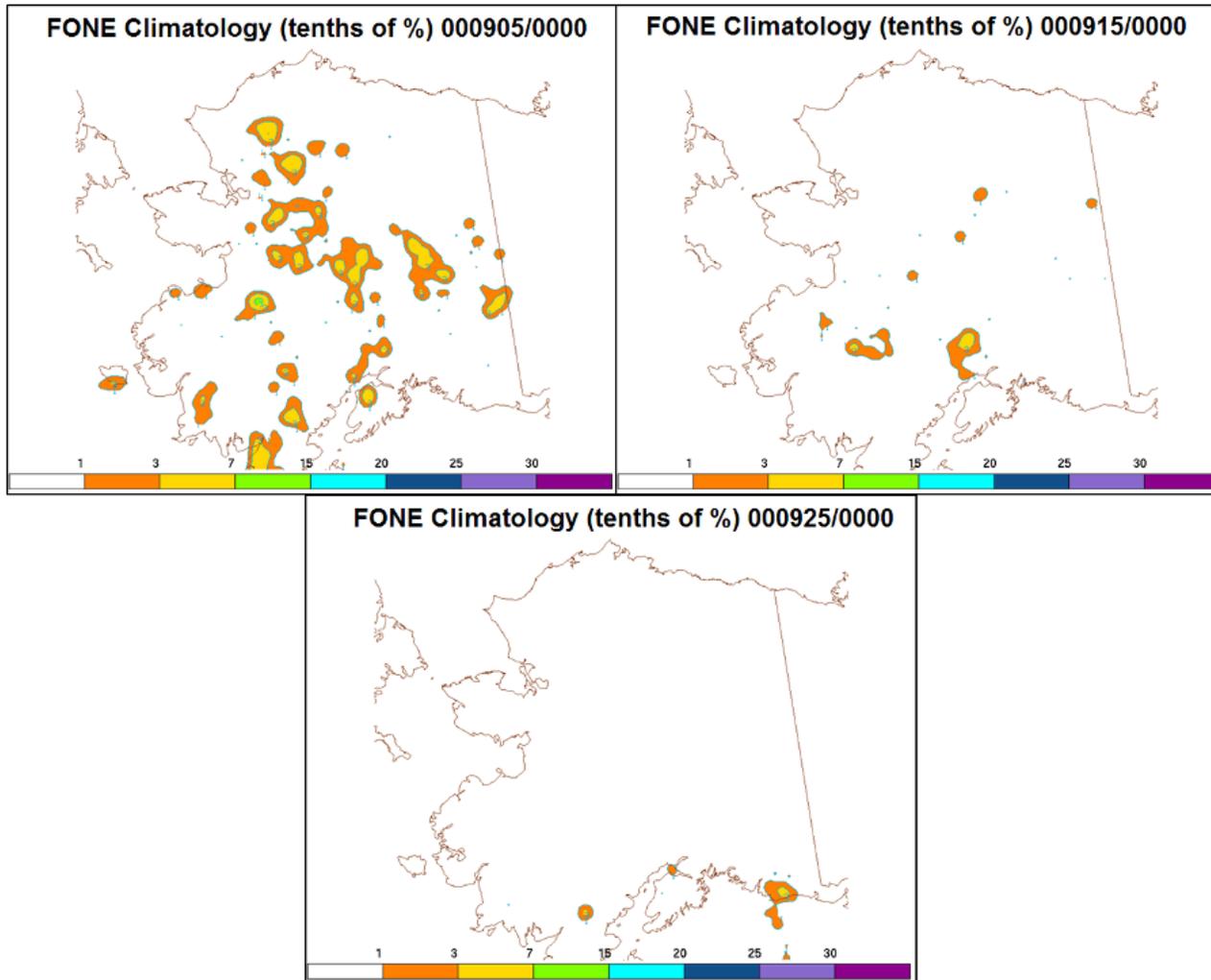


Figure B24: FONE Climatology (AVEFONE) pentad examples for AK from September – 00 UTC data.

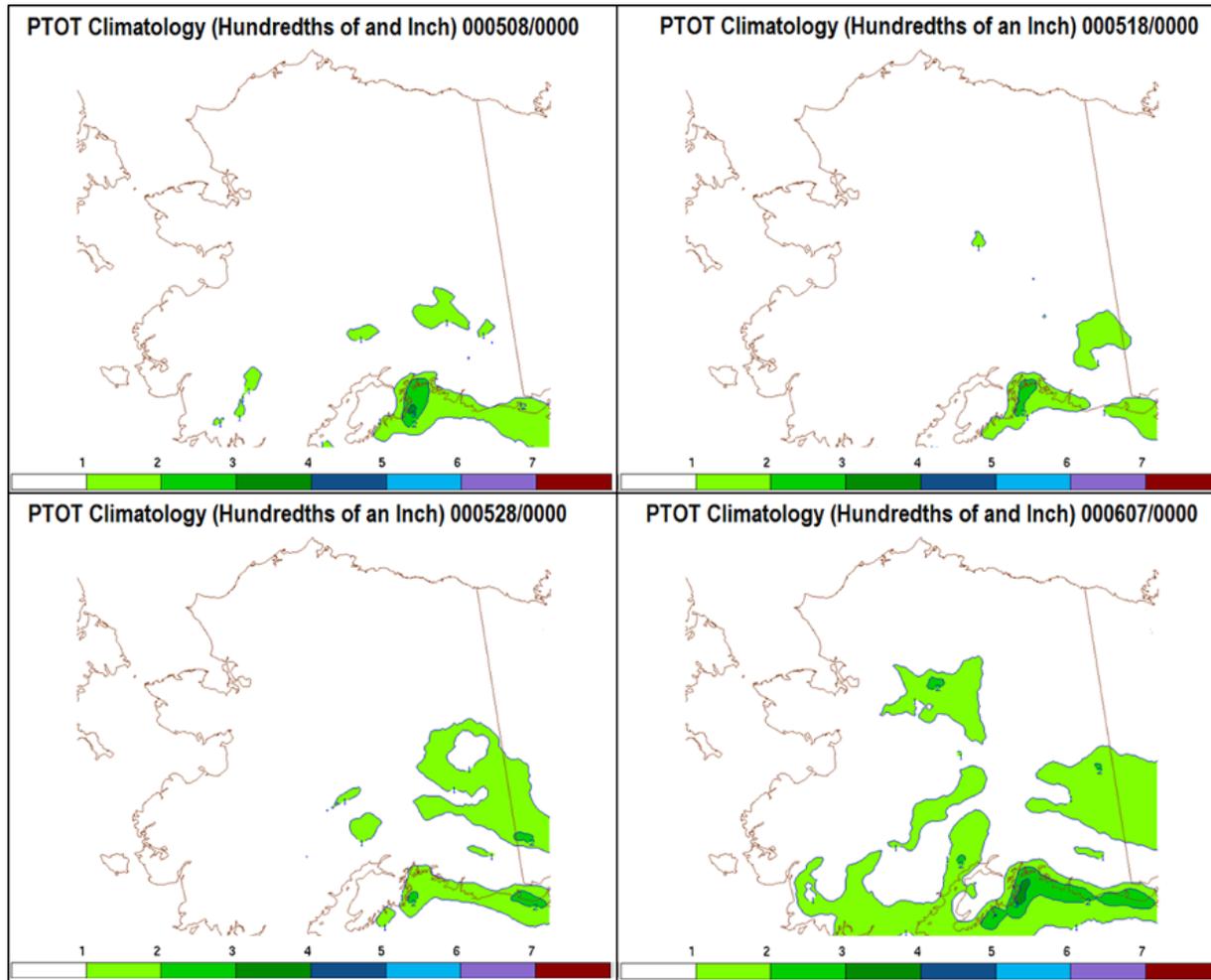


Figure B25: PTOT Climatology (AVEPTOT) pentad examples for AK from early May through early June – 00 UTC data.

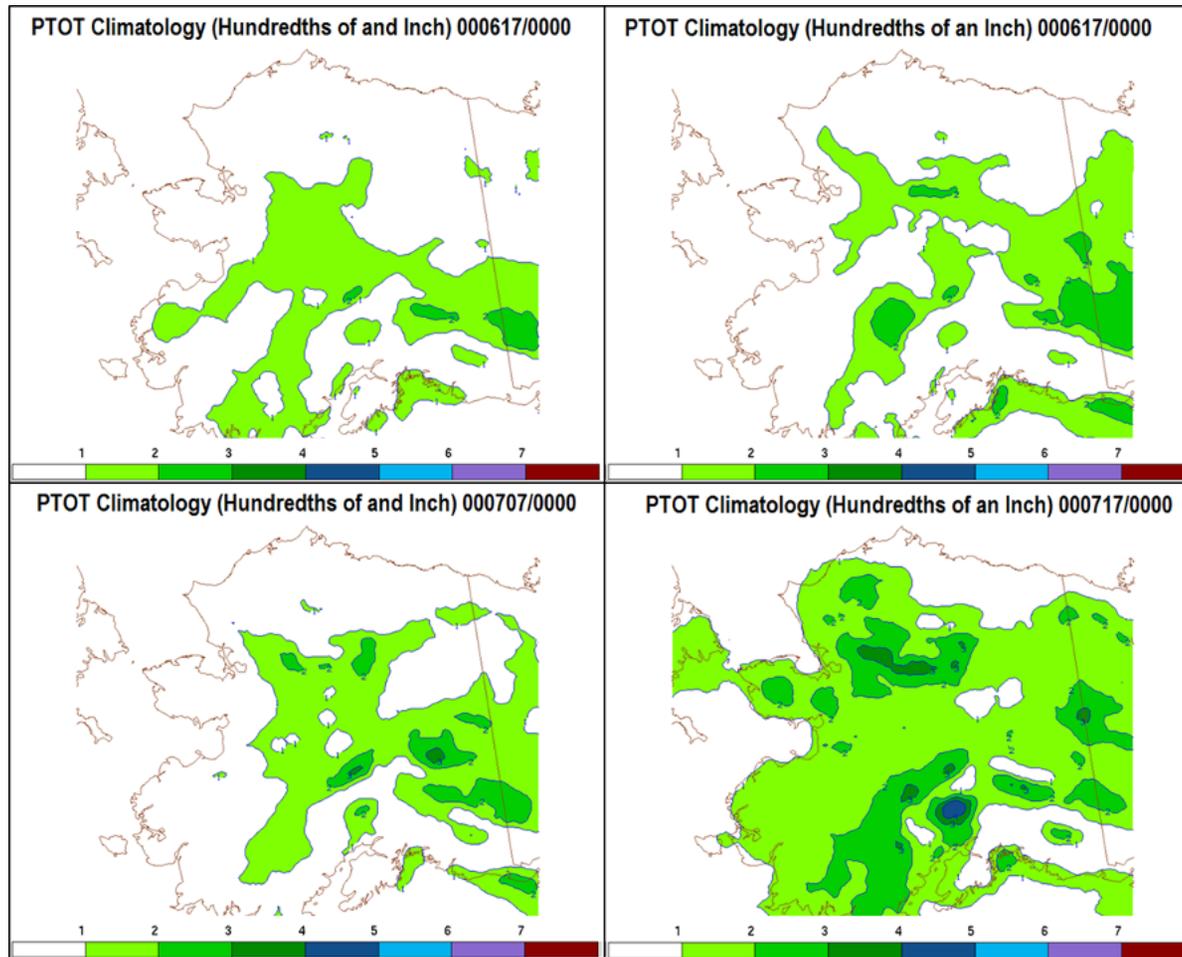


Figure B26: PTOT Climatology (AVEPTOT) pentad examples for AK from mid June through mid July – 00 UTC data.

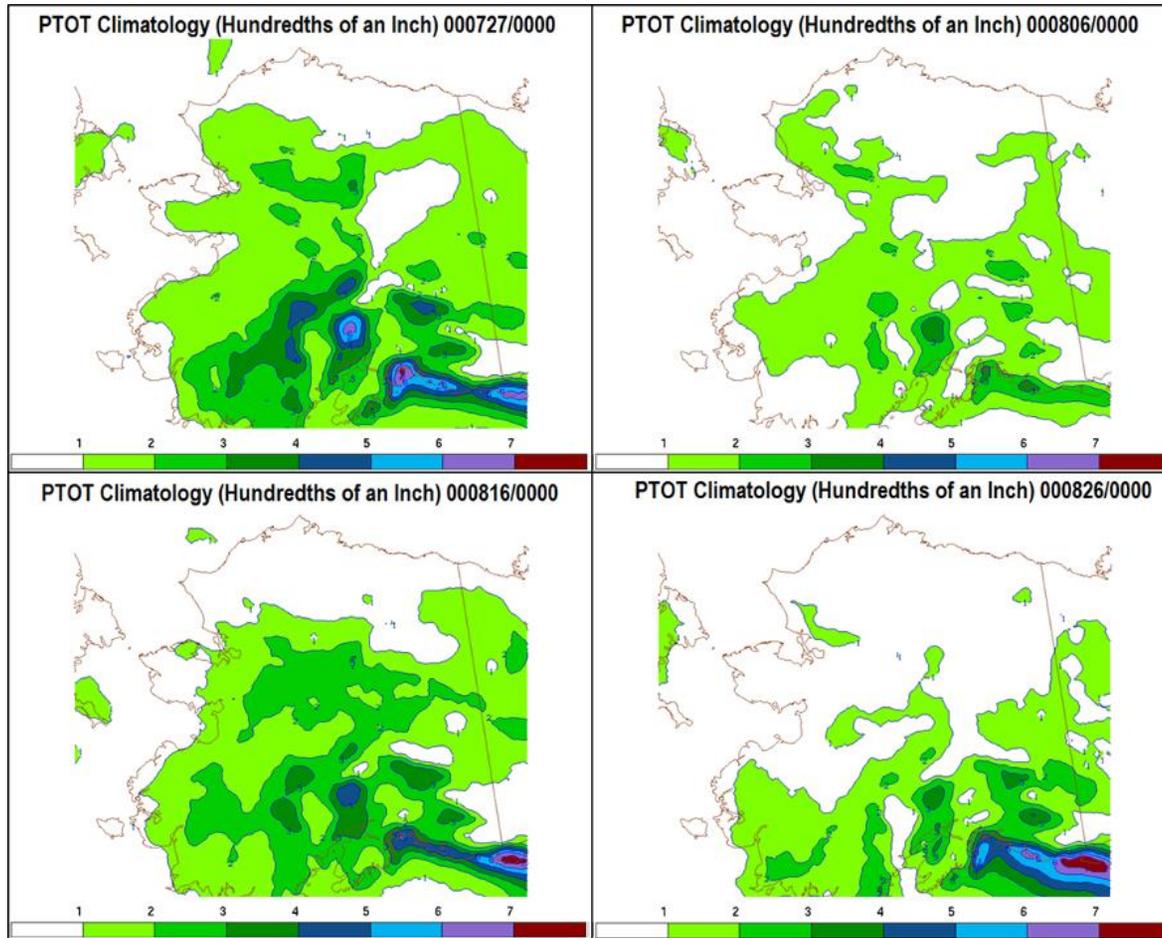


Figure B27: PTOT Climatology (AVEPTOT) pentad examples for AK from late July through late August – 00 UTC data.

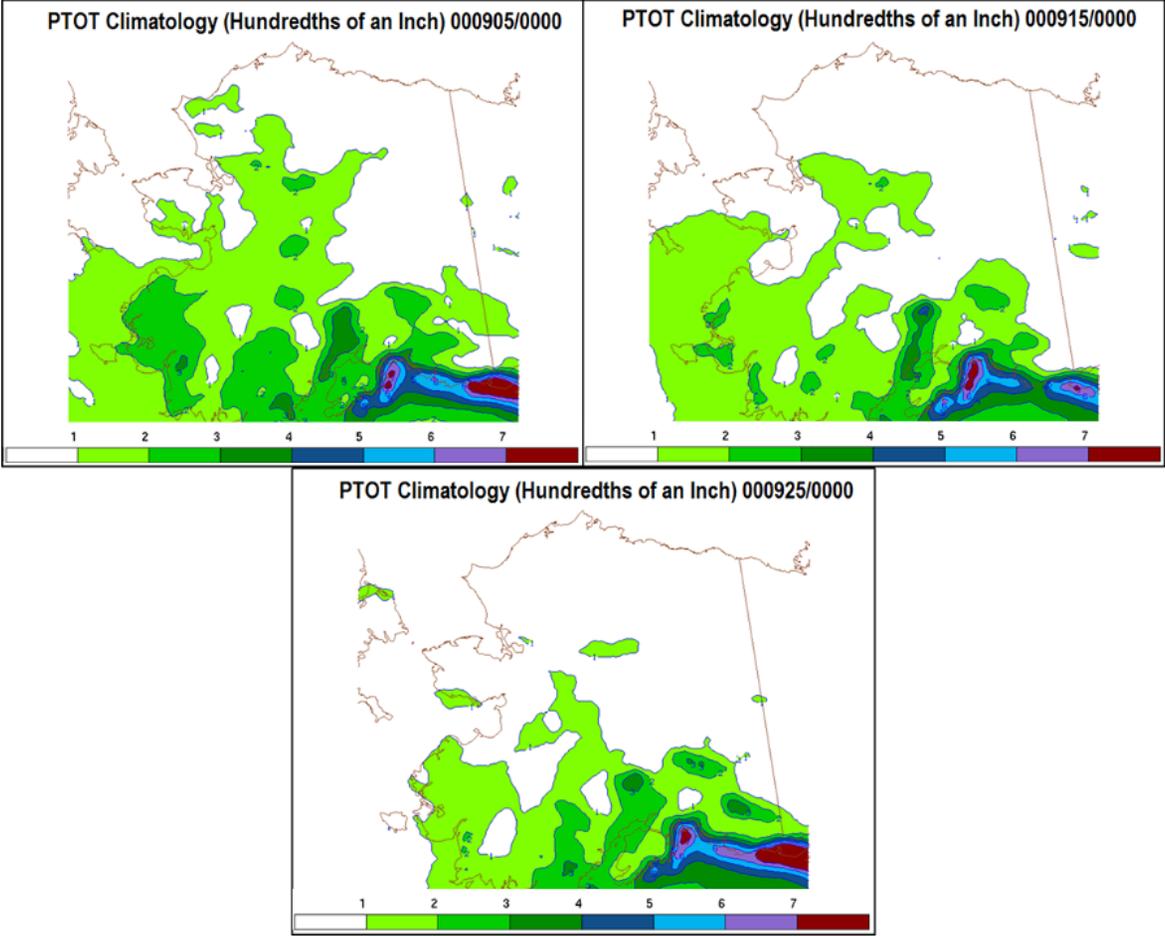


Figure B28: PTOT Climatology (AVEPTOT) pentad examples for AK from September – 00 UTC data.

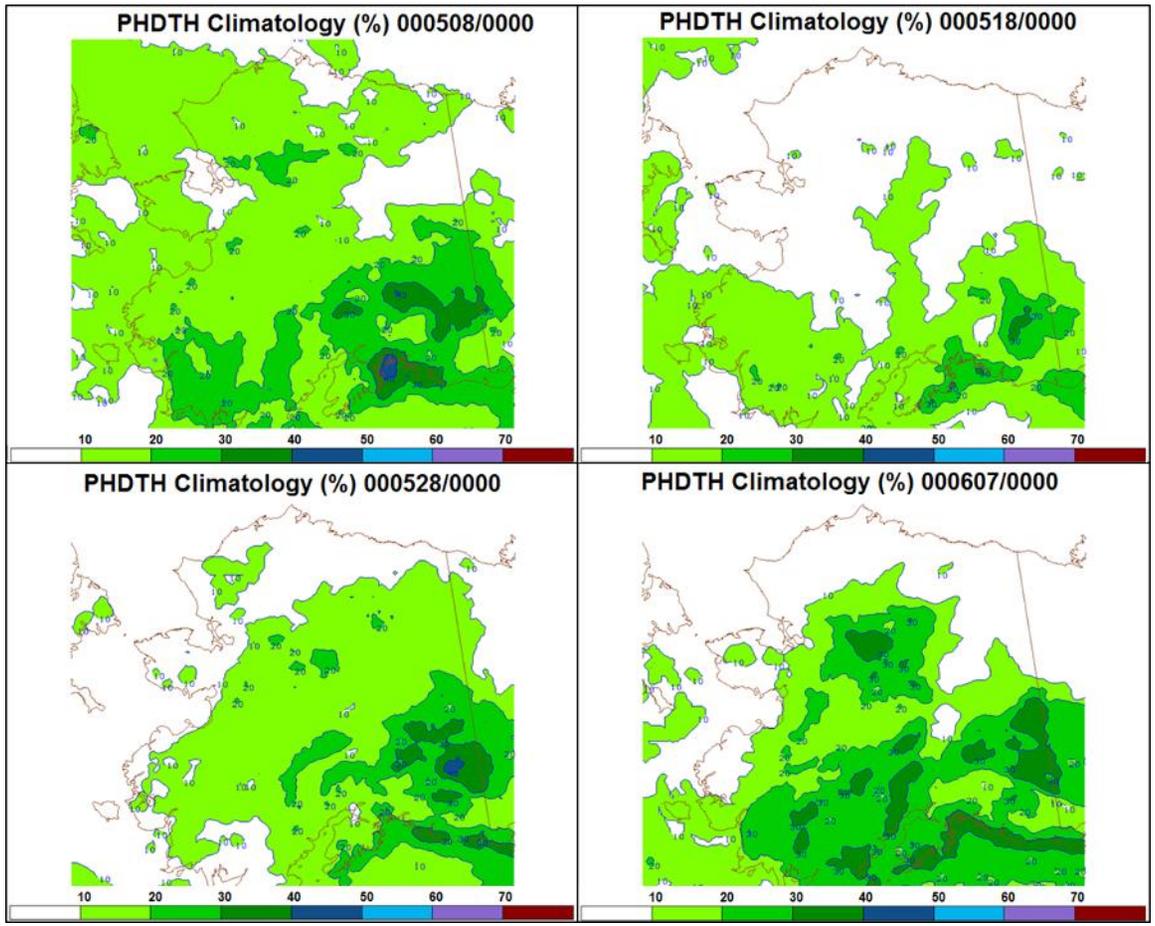


Figure B29: PHDTH Climatology (AVEPHDTH) pentad examples for AK from early May through early June – 00 UTC data.

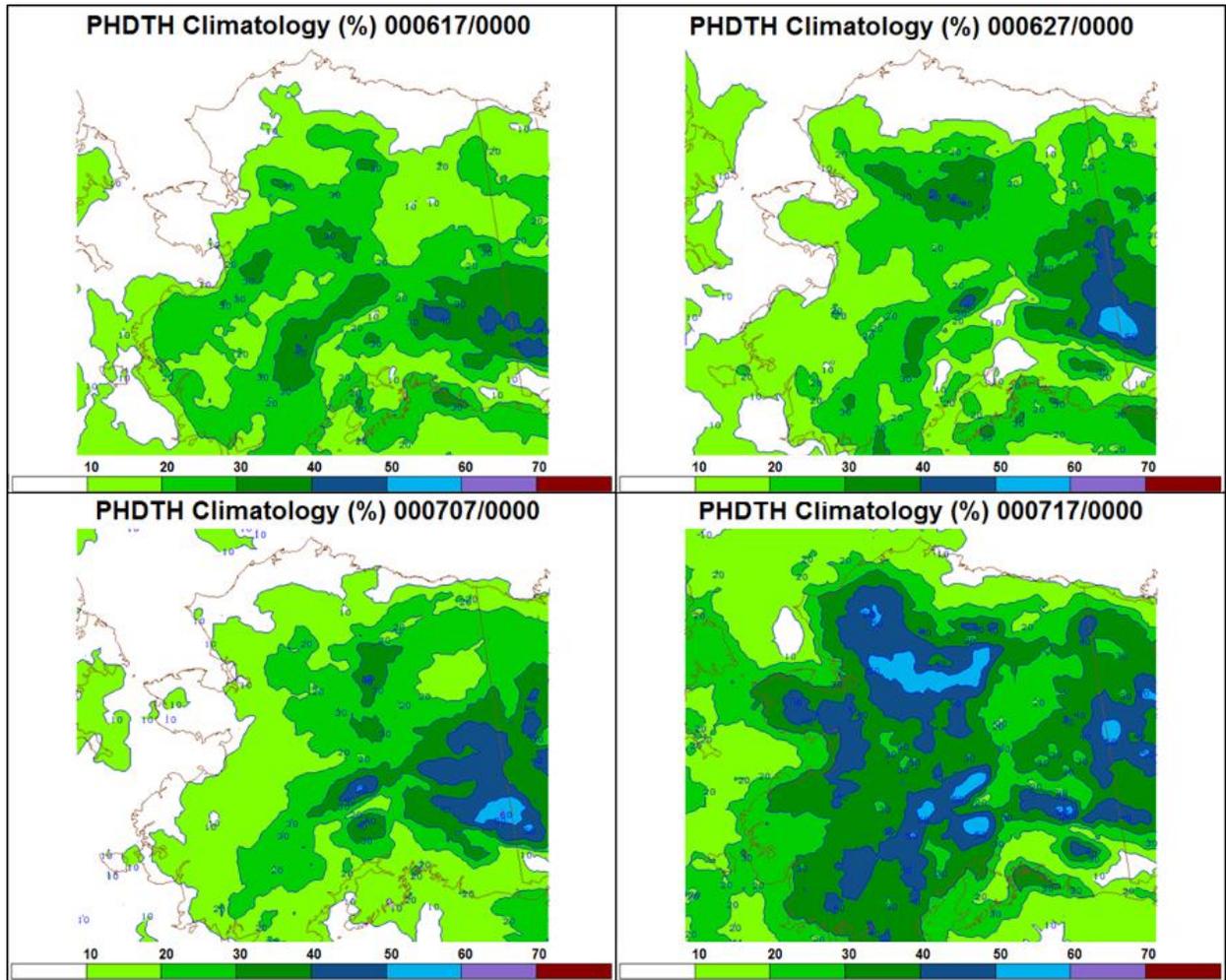


Figure B30: PHDTH Climatology (AVEPHDTH) pentad examples for AK from mid June through mid July – 00 UTC data.

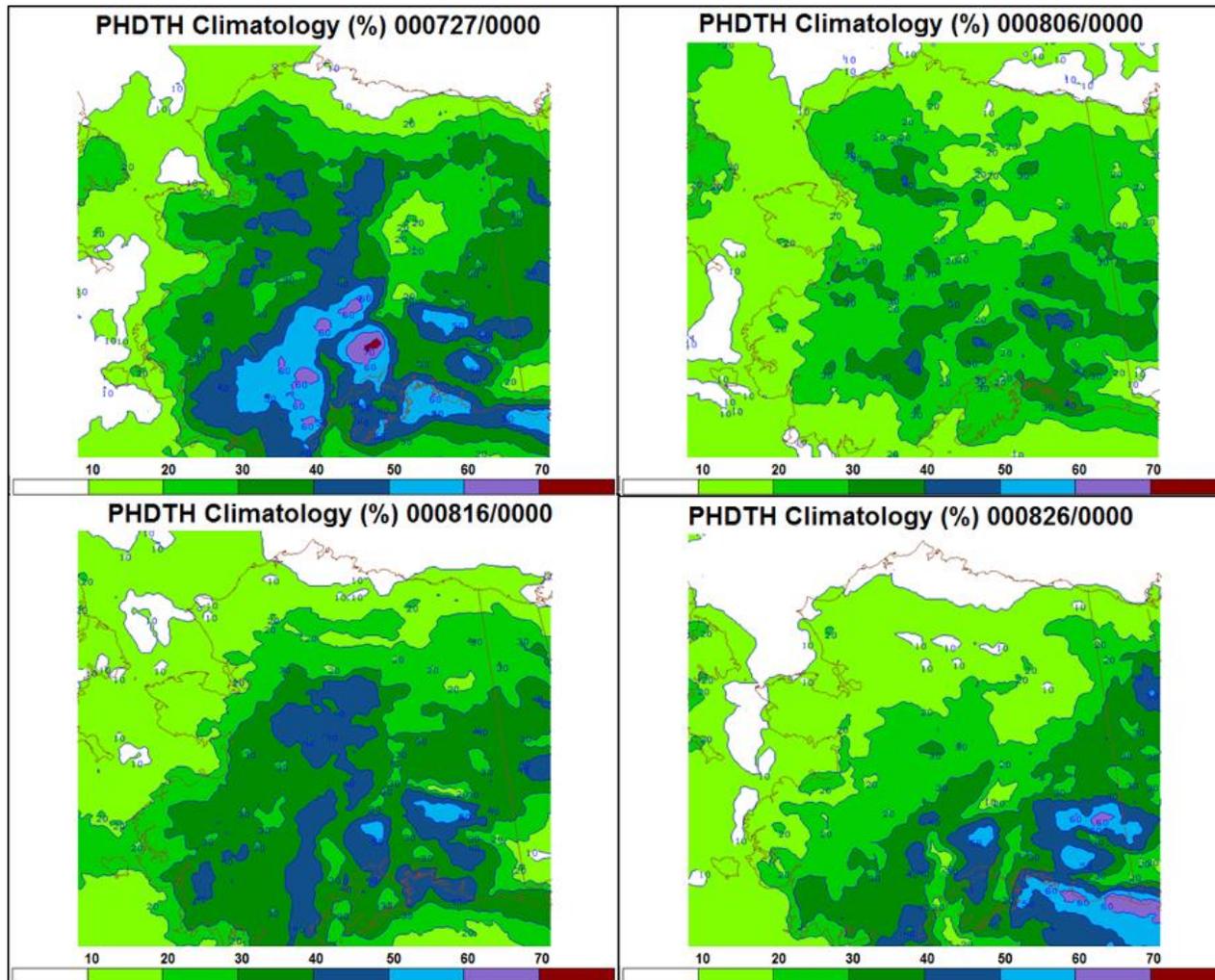


Figure B31: PHDTH Climatology (AVEPHDTH) pentad examples for AK from late July through late August – 00 UTC data.

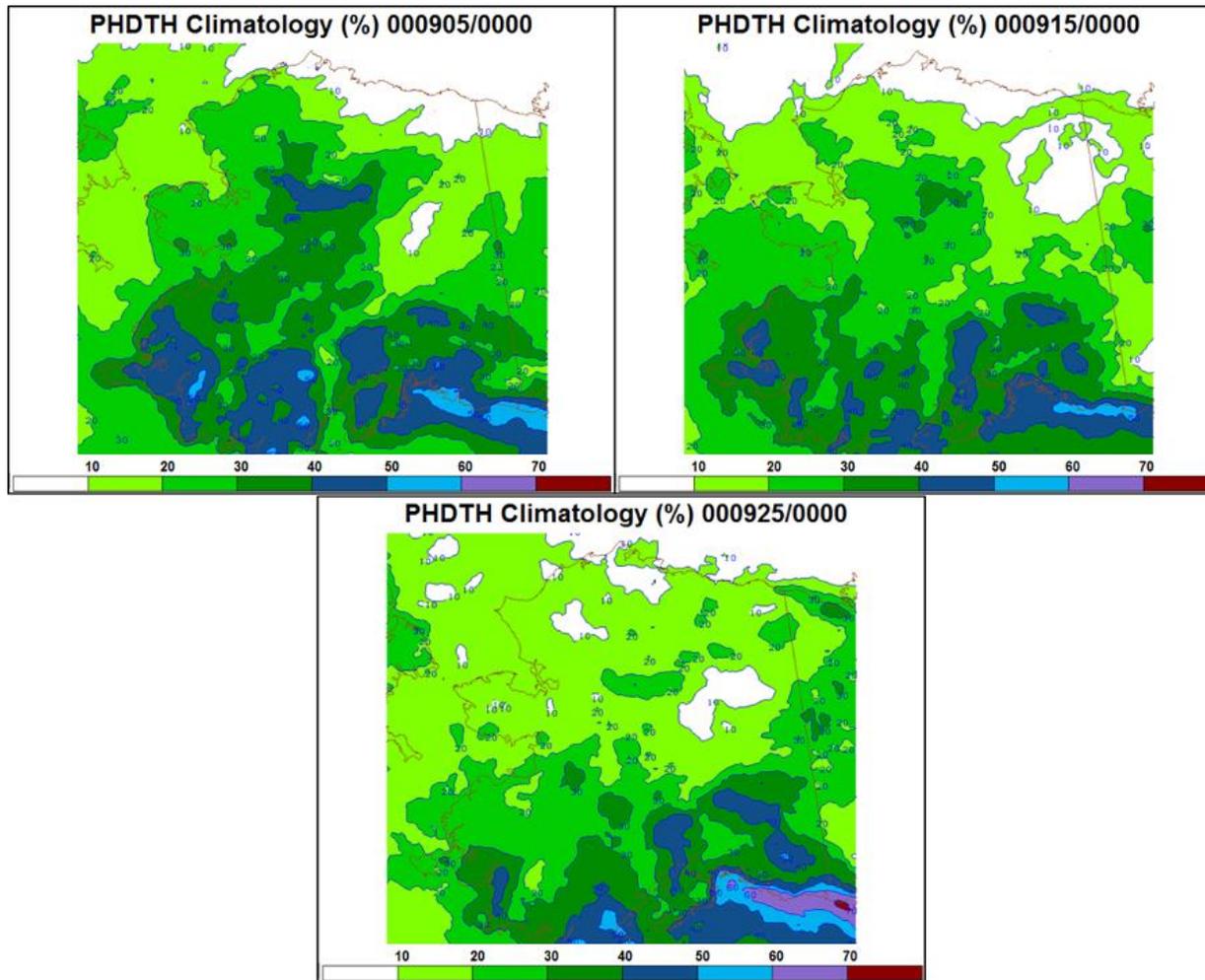


Figure B32: PHDTH Climatology (AVEPHDTH) pentad examples for AK September – 00 UTC data.

Appendix C

Further PCA Discussion and Examples

Principal Component Analysis (PCA) gathers Analysis information from Data using certain statistical techniques. In particular, most of the analysis involves matrices. Because a data set may have many variables and dimensions, the first step involves flattening information into a 2D matrix. For this study, the time dimension is flattened while the space and variables make up the rows and columns respectively. This “time-flattening” works by appending information from multiple days (a single month) as rows to the matrix. Figure C1 shows an example of this flattening process for this study. As mentioned in previous sections, the data matrix variables contain the predictors listed in Tables A2-A5.

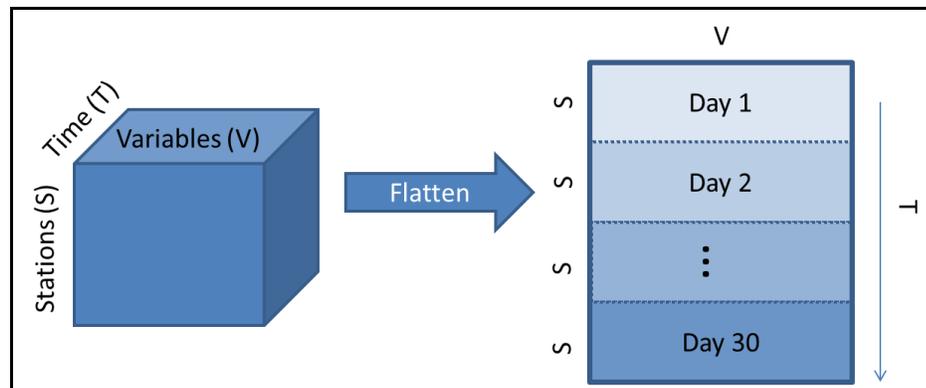


Figure C1: A 3-D data matrix is converted into a 2-D matrix by flattening with respect to time. This means pattern analysis of variable across stations (grid points) holds more interest than temporal changes – they are not necessarily more important. With this method, variables can be assessed for an entire month as preferred in this study. The setup shown here most closely resembles R-mode analysis.

Now the data matrix (\mathbf{Z}) can be explained by the Scores (\mathbf{F}) and Loadings (\mathbf{A}) (as seen in Equation 1 in Sec. 3.2 – repeated below). Using a

$$\mathbf{Z} = \mathbf{FA}^T \quad (1)$$

standardization method, like the one shown in Equation 2, allows the mean to remain the same while the variance changes. X represents the data,

$$Z = \frac{X - \bar{X}}{\sigma_x} \quad (2)$$

overbar denotes the mean, and σ_x is defined as shown in Equation 3 below.

$$\sigma_x = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}} \quad (3)$$

A correlation of the data can be taken to implicitly scale the data and give each variable equal weight for consideration (Richman 1986). This correlation matrix (**R**) can be described by Equation 4 or 5 where **V** represents the associated eigenvectors, and **D** is the associated diagonal matrix of eigenvalues. Note that **R** is a square matrix by this point.

$$\mathbf{R} = \mathbf{Z}^T \mathbf{Z} \quad (4)$$

$$\mathbf{R} = \mathbf{V} \mathbf{D} \mathbf{V}^T \quad (5)$$

Once eigenvalues/vectors are found, loadings can be computed (Equation 6). Loadings can be rotated using a specific transformation matrix (**T**). Only a certain number of PCs are rotated as mentioned in Section 3.2.

$$\mathbf{A} = \mathbf{V} \mathbf{D}^{1/2} \quad (6)$$

Note that a Varimax (orthogonal) rotation method is used here. That is, axes are rotated orthogonally to fit matches of PCs. Simple structure can be assessed following the rotation (Richman and Gong 1999) by assessing pair plots. Good simple structure is noted when most points from comparing two

PCs fall along the axes. Figure C2 shows a cartoon of rotation. Simple structure is achieved when most points fall along the axis within a given cutoff value – magnitude 0.4 here. Plots with several points outside of this cutoff and/or plots that portray diagonals do not count as simple structure. Only ~10% of the plots (C11-C22) do not contain simple structure (based on a simple count assessment), so Varimax is a valid rotation method for this study.

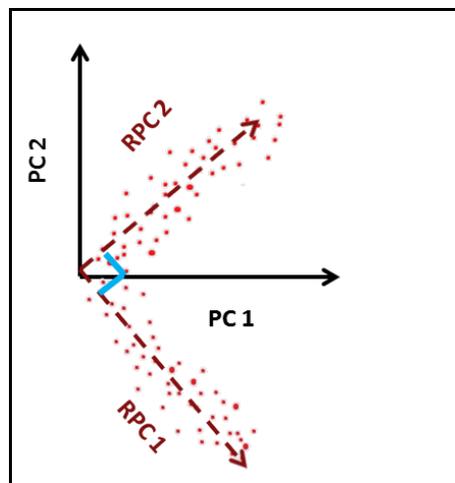


Figure C2: Example of how rotating the loadings aligns more data with the axes and increases simple structure (less points in main quadrants between axes).

One method of determining how many PCs to retain/rotate mentioned in the text comes from the CC Chart. The variance explained by keeping a certain number of PCs (the dashed red line in the charts) gets calculated by summing over the loading variance for a certain number of PCs over the total number of PCs/predictors (Equation 7). N represents the total number of predictors, k is the number of retained PCs, and a_{ij} is the loading for each index of a given PC. Results from Buckley (2009) explained about 54% of the variance when keeping 12 PCs on the 45-km grid. Charts from this study

$$\frac{\sum_{i=1}^k a_{ij}^2}{N} \quad (7)$$

show a similar amount of variance explained ~55% for each month on the 10-km AK domain when 12 PCs are retained. Keeping 12 PCs explains ~60% of the variance on the 40-km CONUS grid. These values may seem low, but can be attributed to the fact that it becomes difficult to distinguish signal from noise with rare event data.

The rotated loadings (**B**) shown in Equation 8 can be plugged back into Equation 1 in place of **A** (Equation 9). Rotation helps patterns look more

$$\mathbf{B} = \mathbf{A}\mathbf{T} \quad (8)$$

$$\mathbf{Z} = \mathbf{F}\mathbf{B}^{\mathbf{T}} \quad (9)$$

like meteorological patterns instead of the standard Buell patterns (North et al. 1982; Richman 1986) . With some matrix manipulation, scores can be found (Equation 10). Scores highlight deviations from the mean, which can then be associated with phenomena like dry thunderstorms.

$$\mathbf{F} = \mathbf{Z}(\mathbf{B})((\mathbf{B})^{\mathbf{T}}(\mathbf{B}))^{-1} \quad (10)$$

The remainder of this Appendix contains Tables of PC groups (Tables C1-C4), Jun CC Charts (Figs. C3-C4), Jun Score plots (Figs. C5-C10), and Jul Pair plots (Figs. C11-C22) for each domain.

Table C1: Predictors larger than the |0.4| cutoff that fall on each PC by month. Bold terms = 0.9, *italicized* terms are mixed across PCs. Purple (bottom) text = best guess of a process related to the group of terms.

CONUS			
PC #	May	Jun	Jul
1	VREL85, THWC85, THTE85, THWC70, THTE70, ML7870, MP7870, MP5770, ML5750, CAPE30, PWTR0, CAPE0, CCTLO, TMPC0, MIXR0, URELO, THTE0, THWC0 Low-Mid Lvl Moisture	THWC85, THTE85, ML7870, MP7870, CAPE30 , PWTR0, CAPE0, CCTLO, <i>TMPC0</i> , MIXR0, THTE0, THWC0, AVETOTF, AVECCFL, AVEFHUN, BTOTF , BCCFL, BFONE , BFTHR , BFTEN , BFTHT , BFHUN , BTSML LTG Climo * MUCAPE & Low Lvl Moisture	THWC85, THTE85, ML7870, MP7870, CAPE30, PWTR0, CAPE0, CCTLO, <i>TMPC0</i> , MIXR0, THTE0, THWC0, AVECCFL, BTOTF , BCCFL, BFONE , BFTHR , BFTEN , BFTHT , BFHUN , BTSML LTG Climo * MUCAPE & Low Lvl Moisture
2	RELH85, AWND85, SH8N85, LP7870, SL7870, SP7870, SP5770, <i>MP5770</i> , <i>SH7N70</i> , LP5750, LPS550 , SL5750, ML5750, TL7550, LR7550, LR8550, <i>PWTR0</i> , THTA0, <i>MIXR0</i> , PMSL0, RELH0 Low Lvl Moisture and Lapse Rate	RELH85, AWND85, <i>SH8N85</i> , RELH70, LP7870, SL7870, SP7870, SP5770, MP5770, <i>SH7N70</i> , LP5750 , LPS550 , SL5750 , ML5750, TL7550 , LR7550 , LR8550 , <i>PWTR0</i> , THTA0, <i>TMPC0</i> , <i>MIXR0</i> , PMSL0, RELH0 Low Lvl Moisture and Lapse Rate	RELH85, AWND85, SH8N85, THWC70, THTE70, LP7870, SL7870 , SP7870, SP5770 , MP5770, MT7870, LP5750 , LPS550 , SL5750 , ML5750, TL7550 , LR7550 , LR8550 , <i>PWTR0</i> , THTA0, <i>TMPC0</i> , MIXR0, PMSL0, RELH0, AVEDRY1 Low Lvl Moisture and Lapse Rate
3	MT7870, UREL50, UREL30, AWND30, SH5750, MT5750, SH5850, SH5N50, SH3830 , SH3530, SH3730 , SH3N30 , MT3530, AUAW30 Mid – Up Lvl Shear and Thermal Wind	MT7870, UREL50, UREL30, AWND30, SH5750, MT5750, <i>SH5850</i> , SH5N50, SH3830 , SH3530, SH3730 , SH3N30, <i>MT3530</i> , AUAW30 Mid – Up Lvl Shear and Thermal Wind	UREL70, UREL50, UREL30, THWC70, THTE70, SH5750, MT5750, <i>SH5850</i> , <i>SH5N50</i> , ML3530, AWND30, SH3830, SH3530, SH3730, SH3N30, MT3530, AUAW30 Mid-Up Lvl Shear and Thermal Wind
4	SP3570 , MP3570 , <i>RVRG50</i> , <i>LHGT50</i> , LTHA50, LP3530, SL3530, ML3530, <i>RVRG30</i> , <i>LHGT30</i> , <i>MT3530</i> , PVOR30 Up Lvl Vorticity	<i>OMEG85</i> , <i>OMEG70</i> , OMEG50, OMEG30, GRDO85, GRDO70, GRDO50, GRDO30, <i>VREL50</i> , RELH50, LC3570, WDIV30 OMEG & Up Lvl Divergence	<i>OMEG70</i> , OMEG50, OMEG30, GRDO85, GRDO70, GRDO50, GRDO30, <i>VREL50</i> , RELH50, <i>RVRG50</i> , WDIV30 OMEG & Up Lvl Divergence
5	<i>OMEG85</i> , <i>OMEG70</i> , OMEG50, OMEG30, GRDO85, GRDO70, GRDO50, GRDO30, <i>VREL70</i> , <i>VREL50</i> , <i>VREL30</i> , RELH70, RELH50, <i>LC7870</i> , LC3570, WDIV30 OMEG & Up Lvl Divergence	SP3570, MP3570 , <i>RVRG50</i> , <i>LHGT50</i> , LTHA50, SL3530, RVRG30 , LHGT30 , PVOR30 Up Lvl Vorticity	SP3570, MP3570, <i>RVRG50</i> , <i>LHGT50</i> , LTHA50, SL3530, <i>RVRG30</i> , LHGT30 , PVOR30 Up Lvl Vorticity
6	AVETOTF, AVECCFL, AVEFONE , AVEFTHR , AVEFTEN , AVEFTHT , AVEFHUN, AVETSML , AVEPHDTH, AVEPTNTH, AVEDRY1, AVEDRY2 Lightning Climatology	VREL85, VREL70, <i>VREL50</i> , VRELO, THWC85, THWC70, THWC0, THTE85, THTE70, THTE0, <i>PWTR0</i> , CINS0, LIFT0 Meridional Moisture Transport	VREL85, VREL70, VRELO, THWC85, THWC70, THWC0, THTE85, THTE70, THTE0, <i>PWTR0</i> , CINS0, LIFT0, CCTLO, <i>MIXR0</i> Meridional Moisture Transport
7	<i>OMEG85</i> , MFCN70, WDIV70, FRNT70, LTHA70, FRT7870, <i>LC7870</i> , <i>LC5770</i> , <i>FRT5750</i> Mid Lvl Forcing	MFCN85, WDIV85, <i>RVRG85</i> , FRNT85, <i>LHGT85</i> , LTHA85, FRT8N85, <i>OMEG70</i> , <i>LC7870</i> , WDIV50, FRNT0, MFCN0, WDIV0 Low Lvl Forcing	MFCN85, WDIV85, <i>RVRG85</i> , FRNT85, <i>LHGT85</i> , LTHA85, FRT8N85, <i>OMEG70</i> , <i>FRT7870</i> , <i>LC7870</i> , WDIV50, FRNT0, MFCN0, WDIV0 Low Lvl Forcing
8	MFCN85, WDIV85, <i>RVRG85</i> , <i>LHGT85</i> , LTHA85, FRT8N85, LMSL0, FRNT0, MFCN0, WDIV0 Low Lvl Forcing	<i>OMEG85</i> , MFCN70, WDIV70, FRNT70, FRT7870, <i>LC7870</i> , <i>LC5770</i> , FRT5750 Mid Lvl Forcing	<i>OMEG85</i> , MFCN70, WDIV70, FRNT70, <i>FRT7870</i> , <i>LC7870</i> , <i>LC5770</i> Mid Lvl Forcing
9	MFCN85, WDIV85, FRNT85, <i>OMEG70</i> , <i>LC5770</i> , FRT5750, WDIV50, FRNT50, HLCY0 Low – Mid Lvl Forcing	<i>RVRG85</i> , <i>RVRG70</i> , <i>RVRG50</i> , <i>LHGT85</i> , <i>LHGT70</i> , <i>LHGT50</i> Low – Mid Lvl Vorticity	AVETOTF, AVECCFL, AVEFONE , AVEFTHR , AVEFTEN , AVEFHUN, AVETSML , AVEPHDTH, AVEPTNTH, AVEDRY1, AVEDRY2 Lightning Climatology
10	<i>ML7870</i> , CAPE30, CAPE0, CCTLO, BTOTF , BCCFL, BFONE, BFTHR , BFTEN , BFTHT , BFHUN , BTSML LTG Climo * MUCAPE	<i>SH8N85</i> , <i>SH7870</i> , <i>SH7N70</i> , <i>SH5850</i> , <i>SH5N50</i> , HLCY0 Shear	<i>RVRG85</i> , <i>RVRG70</i> , <i>RVRG50</i> , <i>LHGT85</i> , <i>LHGT70</i> , <i>LHGT50</i> Low – Mid Lvl Vorticity
11	<i>RVRG85</i> , <i>RVRG70</i> , <i>RVRG50</i> , <i>LHGT85</i> , <i>LHGT70</i> , <i>LHGT50</i> Low – Mid Lvl Vorticity	AVETOTF, AVECCFL, AVEFONE , AVEFTHR , AVEFTHT , AVEFHUN, AVETSML , AVEPHDTH, AVEPTNTH, AVEDRY1, AVEDRY2 Lightning Climatology	SH7870, MT7870, SH7N70, <i>SH5850</i> , SH5N50, HLCY0 Shear
12	UREL85, UREL70, UREL50, UREL30, <i>VREL85</i> , <i>SH8N85</i> , <i>SH7N70</i> , SH5N50, HLCY0, VRELO Wind and Shear	UREL85, UREL70, <i>UREL50</i> , UREL30, URELO Zonal Wind	UREL85, <i>UREL70</i> , <i>UREL50</i> , URELO Zonal Wind

Table C2: Continuation of Table C1 for August and September.

CONUS		
PC #	Aug	Sep
1	<p><i>THWC85, THTE85, ML7870, MP7870, CAPE30, PWTR0, CAPE0, CCTLO, TMPC0, MIXR0, THTE0, THWC0, BTOTF, BCCFL, BPHONE, BFTHR, BFTEN, BFHT, BFHUN, BTSML</i> LTG Climo * MUCAPE & Low Lvl Moisture</p>	<p><i>THWC85, ML7870, CAPE30, PWTR0, CAPE0, CCTLO, TMPC0, MIXR0, THTE0, THWC0, BTOTF, BCCFL, BPHONE, BFTHR, BFTEN, BFHT, BFHUN, BTSML, AVEPHDTH, AVEPTNTH</i> LTG Climo * MUCAPE & Low Lvl Moisture</p>
2	<p><i>RELH85, AWND85, SH8N85, LP7870, SL7870, SP7870, SP5770, MP5770, LPS550, TL7550, LR7550, LR8550, PWTR0, THTA0, TMPC0, MIXR0, RELH0</i> Low Lvl Moisture and Lapse Rate</p>	<p><i>RELH85, SH8N85, LP7870, SL7870, SP7870, SP5770, SH7N70, LP5750, LPS550, SL5750, ML5750, TL7550, LR7550, LR8550, THTA0, TMPC0, MIXR0, PMSL, RELH0</i> Low Lvl Moisture and Lapse Rate</p>
3	<p><i>UREL70, UREL50, UREL30, THWC70, THTE70, SH5750, MT5750, SH5850, SH5N50, ML3530, AWND30, SH3830, SH3530, SH3730, SH3N30, MT3530, AUAW30</i> Mid – Up Lvl Shear Thermal Wind</p>	<p><i>UREL70, UREL50, UREL30, SH5750, MT5750, SH5850, SH7870, MT7870, SH5N50, AWND30, SH3830, SH3530, SH3730, SH3N30, MT3530, AUAW30</i> Mid – Up Lvl Shear and Thermal Wind</p>
4	<p><i>GRDO85, GRDO70, GRDO50, GRDO30, OMEG70, OMEG50, OMEG30, VREL50, RELH50, WDIV30</i> OMEG & Up Lvl Divergence</p>	<p><i>RVRG85, RVRG70, RVRG50, LHGT85, LHGT70, LHGT50</i> Low – Mid Lvl Vorticity</p>
5	<p><i>SP3570, MP3570, RVRG50, LHGT50, LTHA50, SL3530, RVRG30, LHGT30, PVOR30</i> Up Lvl Vorticity</p>	<p>SP3570, MP3570, RVRG50, LHGT50, LTHA50, LP3530, SL3530, ML3530, MT3530, RVRG30, LHGT30, PVOR30 Up Lvl Vorticity</p>
6	<p><i>VREL85, VREL70, VREL0, THWC85, THWC70, THWC0, THTE85, THTE70, RELH70, MP5770, ML5750, LP3530, PWTR0, CINSO, LIFT0, CCTLO</i> Meridional Moisture Transport</p>	<p><i>THWC85, THWC70, THTE85, THTE70, SL5750, TL7550, LR7550, LR8550</i> Low – Mid Lvl THTE and Lapse Rate</p>
7	<p><i>MFCN85, WDIV85, RVRG85, FRNT85, LHGT85, LTHA85, FRT8N85, OMEG70, FRT7870, LC7870, WDIV50, FRNT0, MFCN0, WDIV0</i> Low Lvl Forcing</p>	<p><i>OMEG85, MFCN70, WDIV70, FRNT70, FRT7870, LC7870, LC5770, FRT5750</i> Mid Lvl Forcing</p>
8	<p><i>OMEG85, MFCN70, WDIV70, FRNT70, FRT7870, LC7870, LC5770</i> Mid Lvl Forcing</p>	<p><i>MFCN85, WDIV85, RVRG85, FRNT85, LHGT85, LTHA85, FRT8N85, OMEG70, FRT7870, LC7870, FRNT0, MFCN0, WDIV0</i> Low Lvl Forcing</p>
9	<p><i>AVETOTF, AVECCFL, AVEFONE, AVEFTHR, AVEFTEN, AVEFTHT, AVEFHUN, AVETSML, BFTEN, BFHT, BFHUN, BTSML, AVEPHDTH, AVEPTNTH, AVEDRY1, AVEDRY2</i> Lightning Climatology</p>	<p><i>OMEG85, OMEG70, OMEG50, OMEG30, GRDO70, GRDO50, GRDO30, RELH70, RELH50, FRT3530, WDIV30</i> OMEG & Up Lvl Divergence</p>
10	<p><i>SH7870, MT7870, SH5850, SH5N50, HLCY0</i> Shear</p>	<p><i>UREL85, UREL70, VREL85, VREL70, VREL50, VREL30, VREL0, SH7N70, HLCY0</i> Meridional Wind</p>
11	<p><i>RVRG85, RVRG70, RVRG50, LHGT85, LHGT70, LHGT50</i> Low-Mid Lvl Vorticity</p>	<p>AVETOTF, AVECCFL, AVEFONE, AVEFTHR, AVEFTEN, AVEFTHT, AVEFHUN, AVETSML, BFHT, BFHUN, AVEDRY1, AVEDRY2 Lightning Climatology</p>
12	<p><i>UREL85</i> UREL85 (Summer Ridge)</p>	<p><i>THWC85, ML7870, LP3530, PWTR0, CINSO, LIFT0, THTE0, THWC0</i> Low Lvl Moisture</p>

Table C3: Same as Table C1 for Alaska during May-Jul.

AK			
PC #	May	Jun	Jul
1	AVETOTF, AVEFONE, AVEFTHR, AVEFTEN, AVETSML, BTOTF, BFONE, BFTHR, BFTEN, BTSML, AVEDRY1, AVEDRY2, CAPE30 Lightning Climatology	AVETOTF, AVEFONE, AVEFTHR, AVEFTEN, AVETSML, BTOTF, BFONE, BFTHR, BFTEN, BTSML, AVEDRY1, AVEDRY2, CAPE30, CCTLO Lightning Climatology	AVETOTF, AVEFONE, AVEFTHR, AVEFTEN, AVETSML, BTOTF, BFONE, BFTHR, BFTEN, BTSML, AVEDRY1, AVEDRY2, CAPE30, CCTLO Lightning Climatology
2	SH3530, AWND30, SH3830, MT3530, SH3730, SH3N30, AUAW30, SH5750, MT5750, SH5850, SH5N50 Mid – Up Lvl Shear and Thermal Wind	SH3530, AWND30, SH3830, MT3530, SH3730, SH3N30, AUAW30, SH5750, MT5750, SH5850, SH5N50 Mid – Up Lvl Shear and Thermal Wind	RVRG85, LHGT85, RVRG70, LHGT70, RVRG50, LHGT50, RVRG30, LHGT30, SP5770, SP3570, MP5770, MP3570, PVOR30 Mid Lvl Vorticity
3	THWC85, THWC70, MP3570, SP3570, THTE85, THTE70, LP3530, SL3530, ML3530, PVOR30, PWTR0, MIXR0, TMPC0, THTE0, THWC0 Low Lvl Moisture and Lapse Rate	RVRG85, LHGT85, RVRG70, LHGT70, RVRG50, LHGT50, RVRG30, LHGT30, SP5770, SP3570, MP5770, MP3570 Mid Lvl Vorticity	SH3530, AWND30, SH3830, MT3530, SH3830, SH3N30, AUAW30, SH5750, MT5750, SH5850, SH5N50, SH3730 Mid – Up Lvl Shear and Thermal Wind
4	LC5770, OMEG85, OMEG70, MFCN70, WDIV70, LC3570, FRT7570, WDIV50, FRT50 Mid Lvl Forcing	MFCN85, WDIV85, RVRG85, FRNT85, LHGT85, SH8N85, FRT7870, LC7870, ZMEG0, FRNT0, MFCN0, WDIV0 Low Lvl Forcing	OMEG85, OMEG70, OMEG50, OMEG30, GRDO85, GRDO70, GRDO50, RELH85, RELH70, RELH50, WDIV30, LC3570 Mid – Up Lvl OMEG and Divergence
5	MFCN85, WDIV85, RVRG85, FRNT85, LHGT85, SH8N85, FRT7870, LC7870, ZMEG0, FRNT0, MFCN0, WDIV0 Low Lvl Forcing	SP5770, MP5770, LP5750, SL5750, ML5750, TL7550, LR7550, LR8550, LIFT0 Mid Lvl Lapse Rate	MFCN85, WDIV85, RVRG85, FRNT85, LHGT85, LTHA85, FRT8N85, FRT7870, LC7870, ZMEG0, FRNT0, MFCN0, WDIV0 Low Lvl Forcing
6	THWC85, THTE85, SH8N85, LPS550, LP7870, SL7870, ML7870, SP7870, MP7870, LR8550, CAPE30, THTA0, LIFT0, TMPC0, RELH0, THTE0, THWC0 Low-Mid Lvl Moisture	OMEG85, OMEG70, OMEG50, OMEG30, GRDO85, GRDO70, GRDO50, RELH70, RELH50, WDIV30, LC3570 Mid – Up Lvl OMEG and Divergence	THWC85, THTE85, SH8N85, LPS550, LP7870, SL7870, ML7870, SP7870, MP7870, LR8550, CAPE30, THTA0, LIFT0, CCTLO, MIXR0, TMPC0, RELH0, THTE0, THWC0, AVEPHDTH Low-Mid Lvl Moisture
7	LP5750, SP5770, SL5750, ML5750, TL7550, LR7550, MP5770, LR8550, LIFT0 Mid Lvl Lapse Rate	THWC85, THTE85, SH8N85, LPS550, LP7870, SL7870, ML7870, SP7870, MP7870, LR8550, CAPE30, THTA0, LIFT0, CCTLO, MIXR0, RELH0, THTE0, THWC0, AVEPHDTH Low Lvl Moisture and Lapse Rate	SP5770, MP5770, LP5750, SL5750, ML5750, TL7550, LR7550, LR8550, LIFT0 Mid Lvl Lapse Rate
8	UREL85, UREL70, UREL50, UREL30, URELO Zonal Wind	UREL85, UREL70, UREL50, UREL30, URELO Zonal Wind	THWC85, THWC70, MP3570, SP3570, THTE85, THTE70, LP3530, SL3530, ML3530, PVOR30, PWTR0, MIXR0, THTE0, THWC0 Low Lvl Moisture and Lapse Rate
9	AWND85, GRDO85, AWND70, GRDO70, AWND50, GRDO50, GRDO30, AVEPHDTH Ageostrophic Wind	THWC85, THWC70, MP3570, SP3570, THTE85, THTE70, LP3530, SL3530, ML3530, PVOR30, PWTR0, MIXR0, TMPC0, THTE0, THWC0 Low-Mid Lvl Moisture	UREL85, UREL70, UREL50, UREL30, URELO Zonal Wind
10	VREL85, RELH85, VREL70, RELH70, VREL0, PWTR0 Low Lvl Moisture Transport ?	VREL85, VREL70, VREL50, VREL30, VREL0, RELH70, PWTR0 Meridional Wind	OMEG85, MFCN70, WDIV70, FRNT70, LC5770, FRT5750 Mid Lvl Forcing
11	RVRG85, LHGT85, RVRG70, LHGT70, RVRG50, LHGT50, RVRG30, LHGT30, SP5770, SP3570, MP5770, MP3570 Mid Lvl RVRG and LHGT	OMEG85, MFCN70, WDIV70, FRNT70, FRT7870, LC7870, LC5770, FRT5750 Mid Lvl Forcing	VREL85, VREL70, VREL50, VREL30, VREL0 Zonal Wind
12	SH7870, MT7870, SH7N70, SH5850, SH5N50, HLCY0, SH5750, MT5750, SH3N30 Low Lvl Shear and Helicity	MT8N85, MT7870, SH7870, SH7N70, SH5850, SH5N50, HLCY0 Low Lvl Shear and Helicity	MT8N85, MT7870, SH7870, SH7N70, SH5850, SH5N50, HLCY0, MIXR0 Low Lvl Shear and Helicity

Table C4: Continuation of Table C3 for Aug and Sep.

AK		
PC #	Aug	Sep
1	SH7870, SH5750, MT5750, SH5850, SH5N50, UREL30, LP3530, AWND30, SH3830 , SH3530, SH3730 , SH3N30, AUAW30 Mid – Up Lvl Shear and Thermal Wind	SH5750, MT5750, SH5850, SH5N50, UREL50, UREL30, LP3530, ML3530, AWND30, SH3830 , SH3530, SH3730 , MT3530, SH3N30 , AUAW30 Mid – Up Lvl Shear and Thermal Wind
2	RVRG70, LHGT70, SP5770, SP3570, MP3570 , SH7870, RVRG50 , LHGT50 , LP3530, SL3530, ML3530, RVRG30, LHGT30, PVOR30 Mid Lvl Vorticity	VREL85, VREL70, VREL50, VREL30, VREL0, RELH85, RELH70, RELH50, RELH0, SH8N85, OMEG70, OMEG50, OMEG30, PWTR0, HLCY0 Meridional Moisture Transport
3	AVETOTF, AVEFONE, AVEFTHR, AVEFTEN, AVETSML , BTOTF, Bfone, BFTHR, BFTEN, BTSML, AVEDRY1, AVEDRY2 Lightning Climatology	RVRG70, LHGT70, SP5770, MP5570, SP3570 , MP3570 , SH7870, RVRG50 , LHGT50 , LP3530, SL3530, ML3530, RVRG30, LHGT30, PVOR30 Mid Lvl Vorticity
4	VREL85, VREL70, VREL50, VREL30, VREL0, RELH70, PWTR0 Meridional Wind	MFCN85, WDIV85, RVRG85, FRNT85, LHGT85, FRT8N85, RVRG70, LHGT70, LTHA70, FRT7870, LC7870, ZMEG0, FRNT0, MFCN0, WDIV0 Low Lvl Forcing
5	MFCN85, WDIV85, RVRG85, FRNT85, LHGT85, FRT8N85, FRT7870, LC7870, ZMEG0, FRNT0, MFCN0, WDIV0 Low Lvl Forcing	AVETOTF, AVEFONE, AVEFTHR, AVETSML , BTOTF, Bfone, BFTHR, BTSML, AVEDRY1 , AVEDRY2 Lightning Climatology
6	SH8N85, LPS550, LP7870, SL7870, ML7870, SP7870, MP7870, LR8550, SH7N70, CAPE30, THTA0, LIFT0, CCTLO, TMPC0, RELH0, THWC0, AVEPHDTH Low Lvl Moisture and Lapse Rate	THWC85, THWC70, SL3530, ML3530, PWTR0, THTA0, TMPC0, MIXR0, THTE0, THWC0, AVEPHDTH Low-Mid Lvl Moisture
7	SP5770, MP5770, LP5750, SL5750, ML5750, TL7550, LR7550, LR8550, LIFT0 Mid Lvl Lapse Rate	SP5770, MP5770, LP5750, LPS550, SL5750, ML5750, TL7550, LR7550, LR8550, LIFT0 Mid Lvl Lapse Rate
8	THWC85, THWC70, THTE85, THTE70, SL3530, ML3530, PWTR0, TMPC0, MIXR0, THTE0, THWC0 Low Lvl Moisture	AWND85, AWND70, AWND50, GRDO85, GRDO70, GRDO50, GRDO30, SH7870, MT7870, SH5850, HLCY0 Ageostrophic Wind
9	UREL85, UREL70, UREL50, UREL30, UREL0 Zonal Wind	THTE85, THTE70, TL7550, LR7550, LR8550 Mid Lvl THTE and Lapse Rate
10	OMEG85, MFCN70, WDIV70, FRNT70, LC5770, FRT5750, FRNT50 Mid Lvl Forcing	OMEG85, OMEG70, MFCN70, WDIV70, RVRG70, LC5770, FRT5750, WDIV50, FRNT50 Mid Lvl Forcing
11	OMEG85, OMEG70, OMEG50, OMEG30, RELH70, RELH50, WDIV30, LC3570 OMEG and Up Lvl Divergence	UREL85, UREL70, UREL50, UREL30, UREL0 Zonal Wind
12	AWND85, AWND70, AWND50, GRDO85, GRDO70, GRDO50, SH7870, MT7870, HLCY0 Ageostrophic Wind	LP7870, SL7870, ML7870, SP7870, MP7870, LPS550, LR8550, THTA0, LIFT0 Low Lvl Lapse Rate

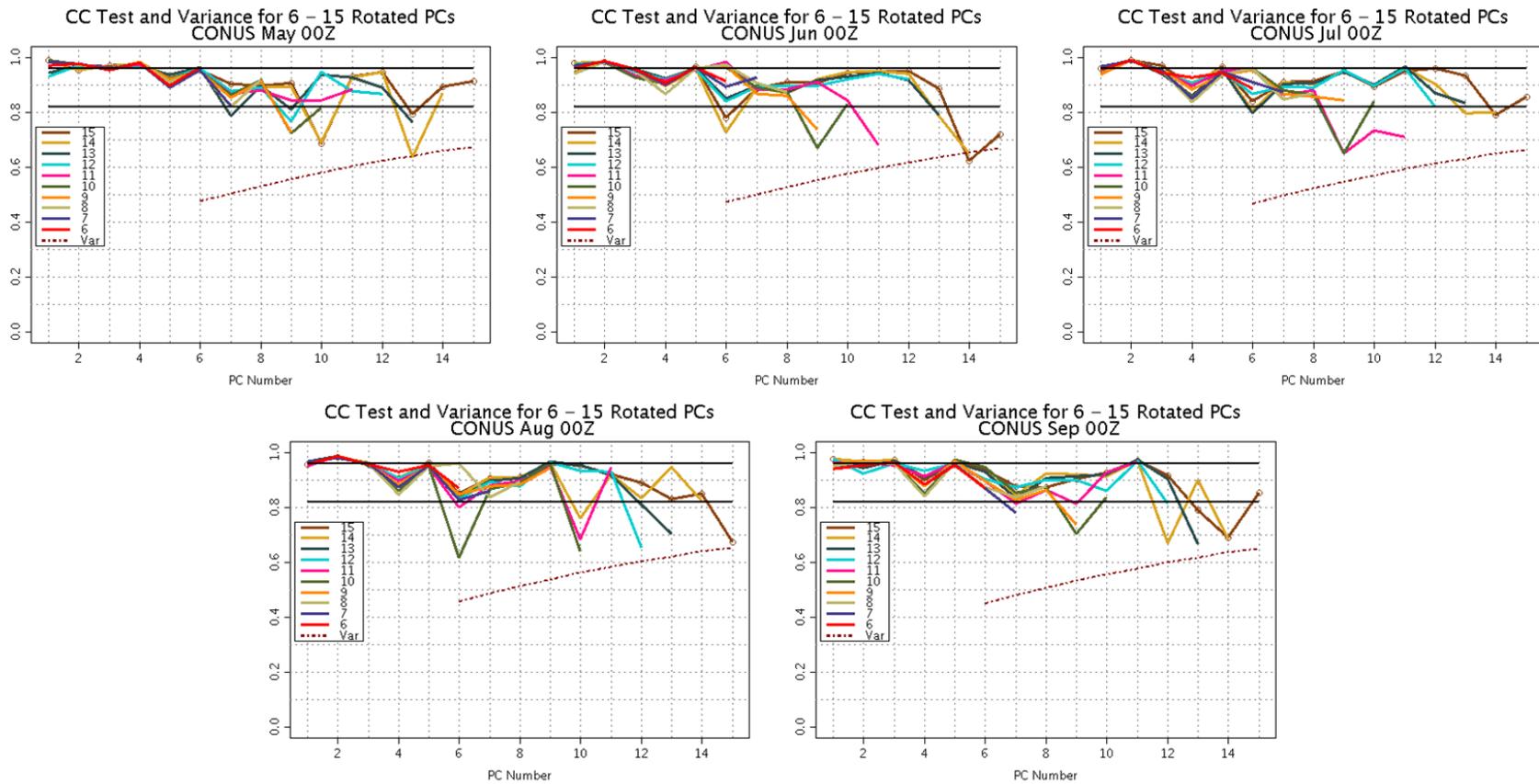


Figure C3: CC Chart for CONUS May - Sep. The majority of the line appears to be a good match when keeping 12 PCs (cyan). Note: 00Z = 00 UTC. August may be more erratic overall due to the warm summer ridge that often dominates weather patterns.

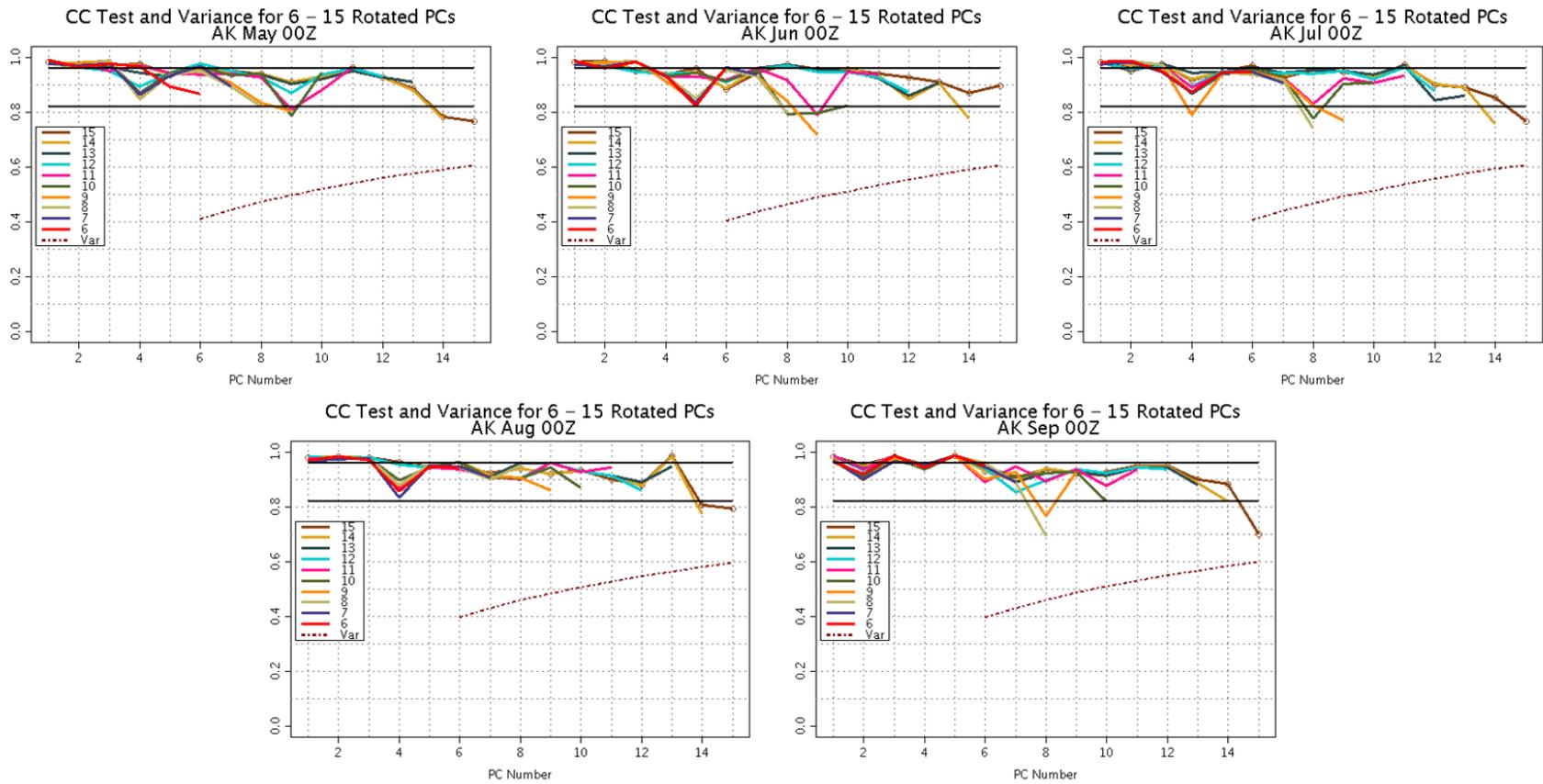


Figure C4: CC Charts for AK May - Sep. The majority of the line appears to be a good match when keeping 12 PCs (cyan). Note: 00Z = 00 UTC.

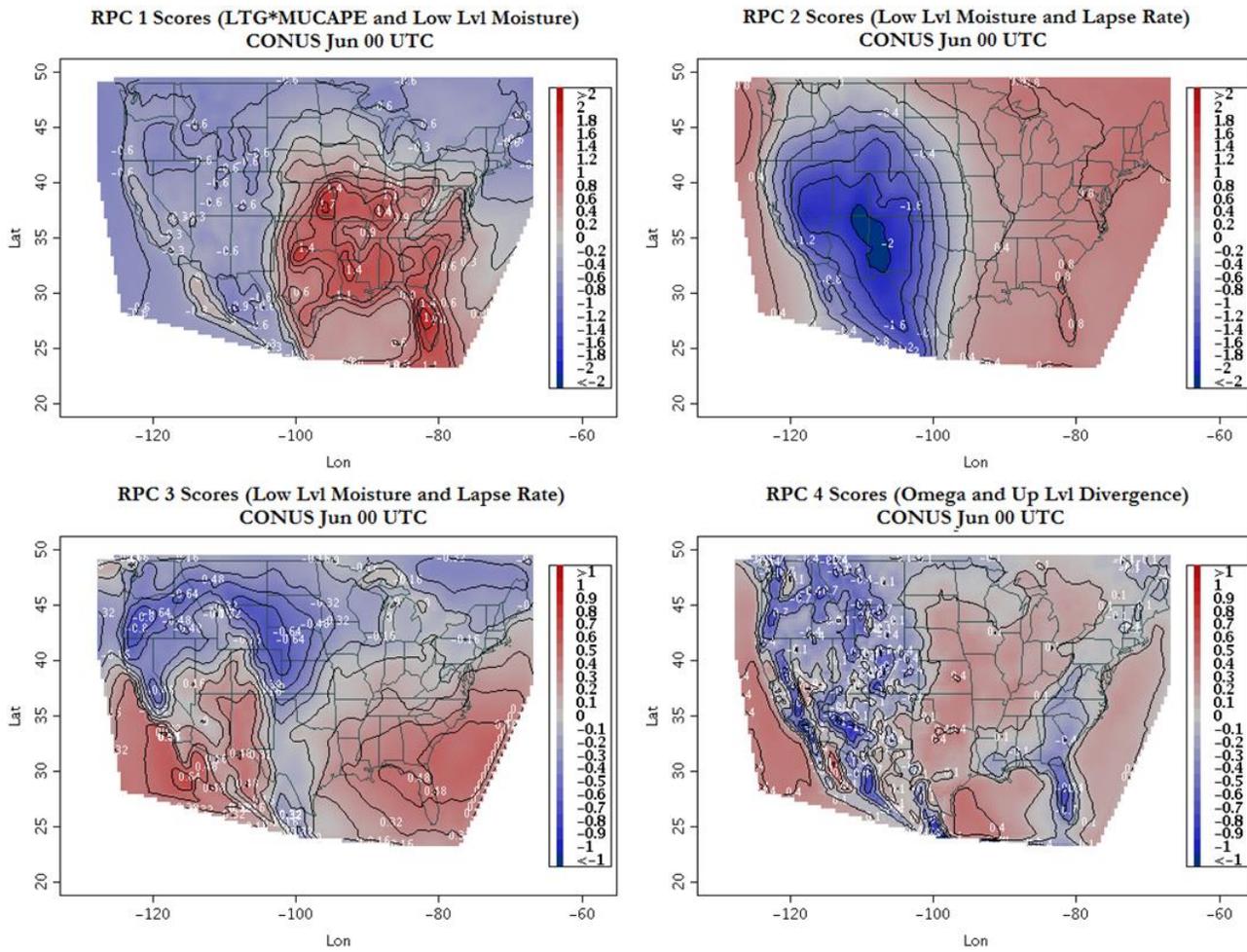


Figure C5: CONUS Score plots for RPC 1-4. Note that the scales vary, and spatial variation is important compared to intensity.

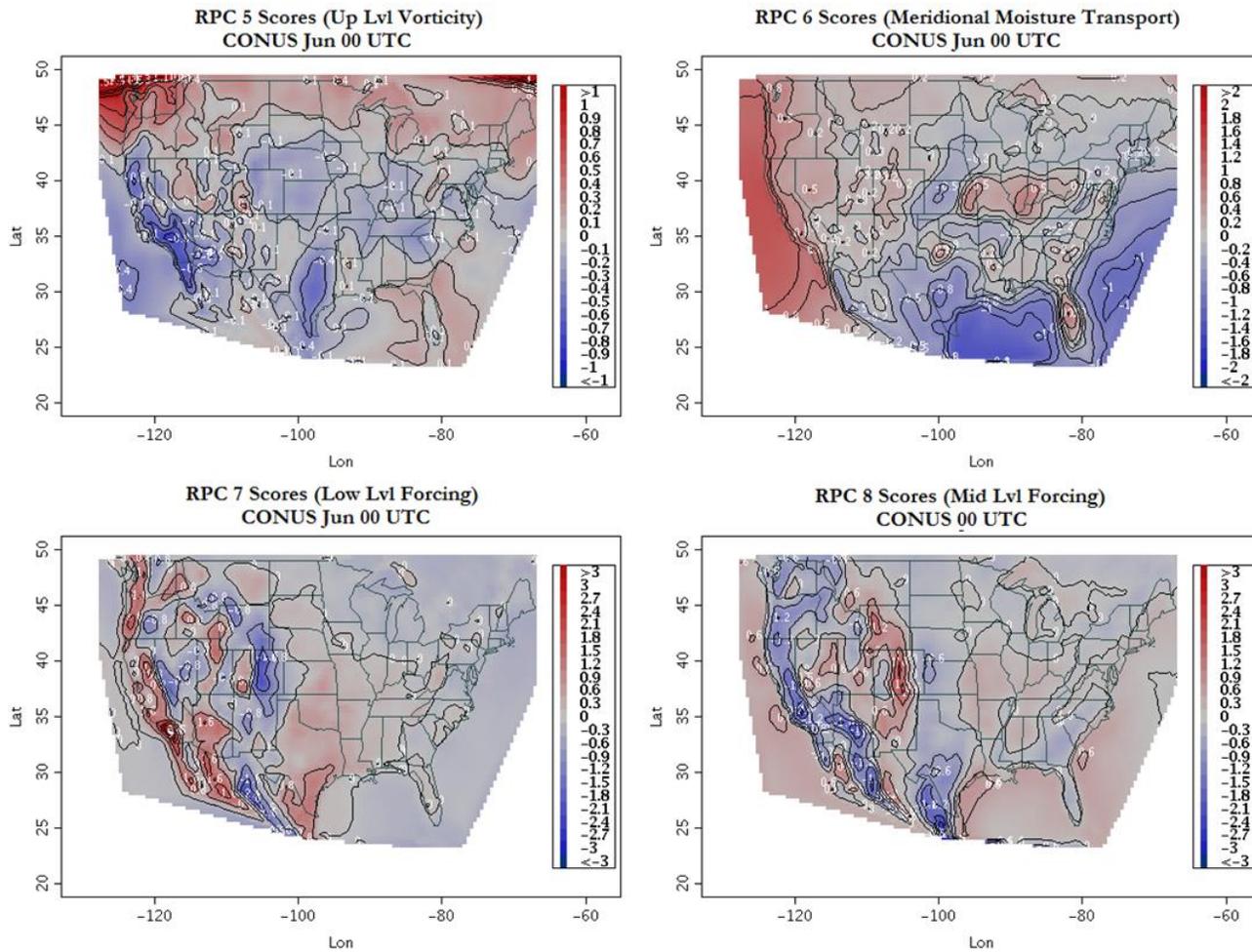


Figure C6: CONUS Score Plots for RPC 5-8. Note that the scales vary, and spatial variation is important compared to intensity.

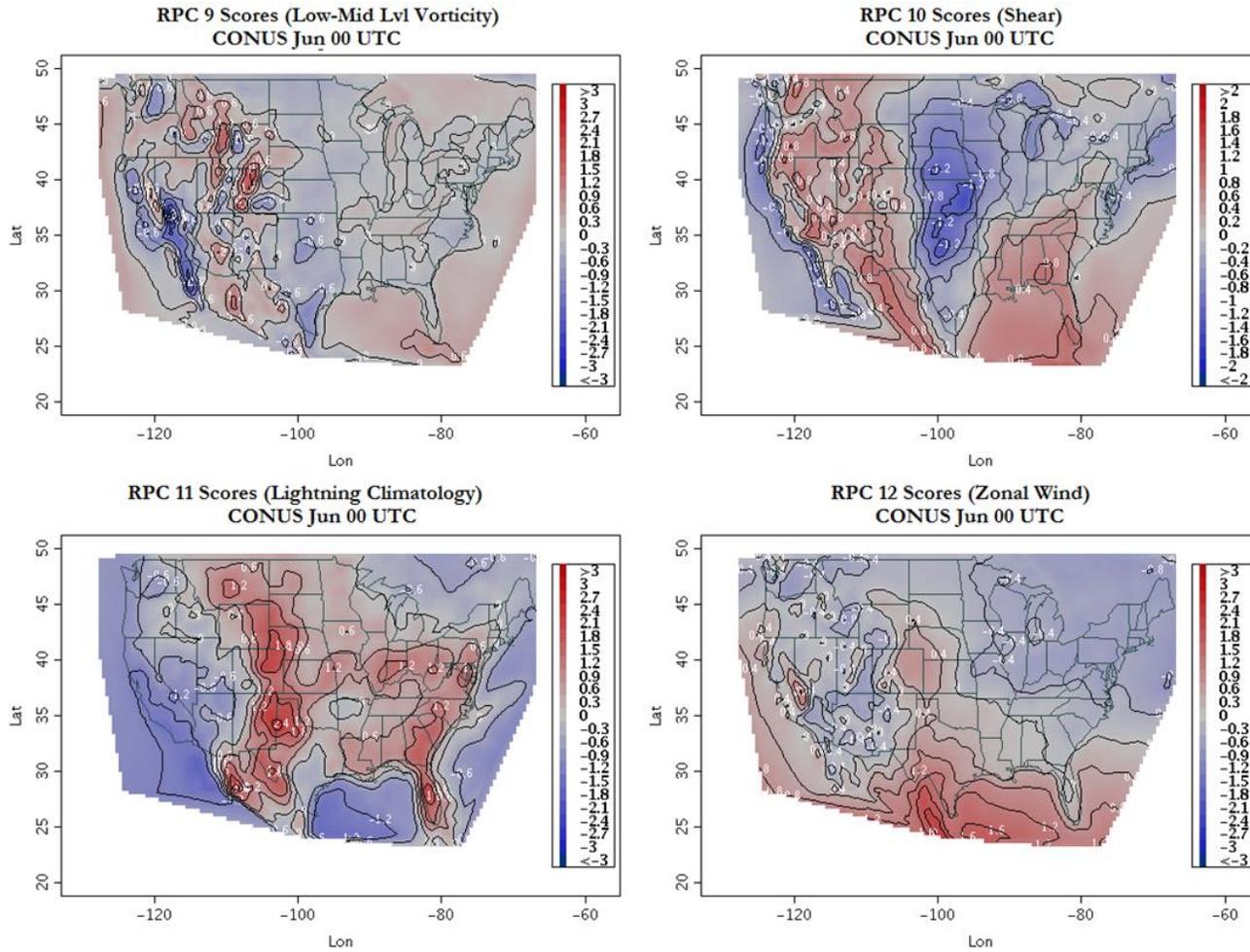
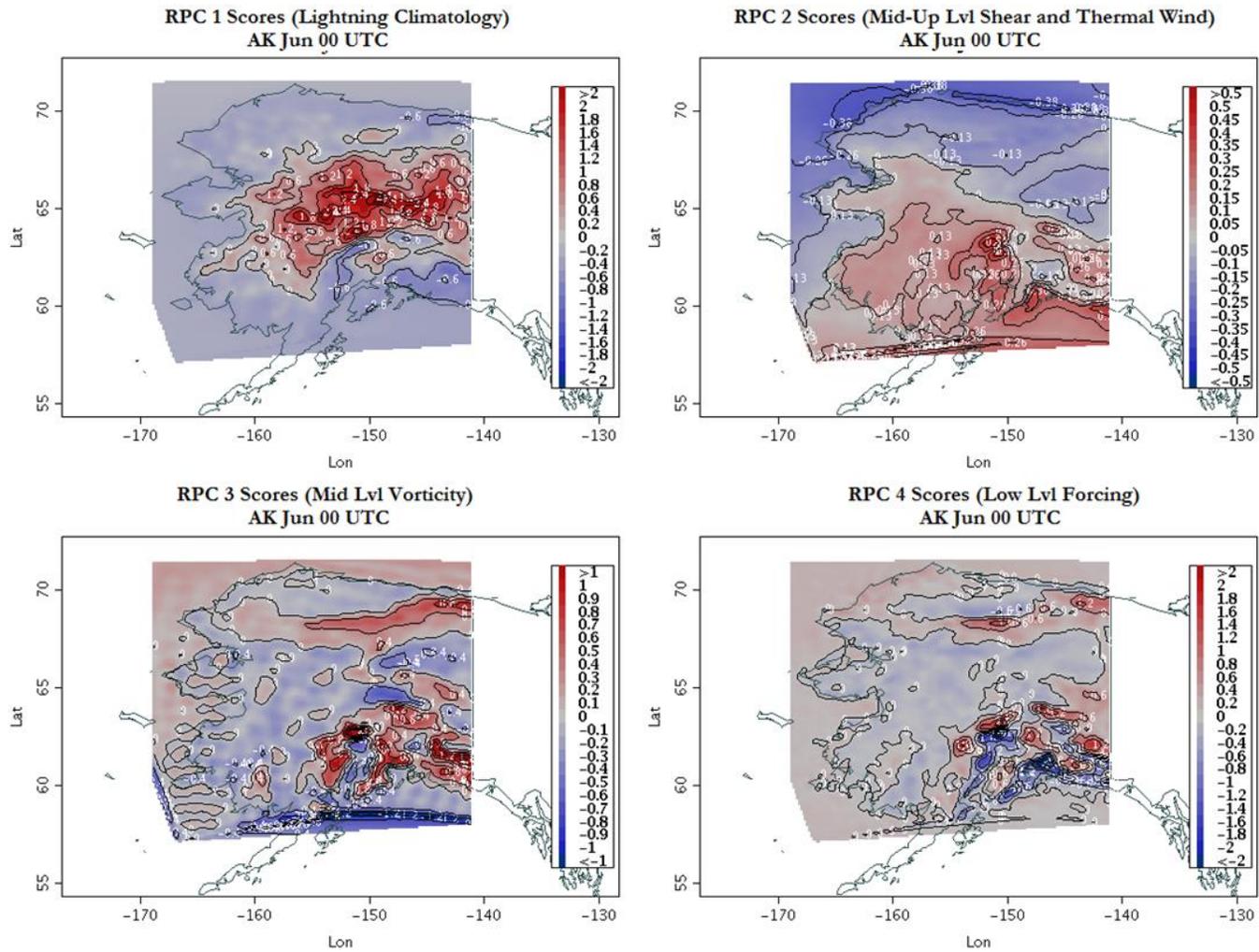


Figure C7: CONUS Score Plots for RPC 9-12. Note that the scales vary, and spatial variation is important compared to intensity.



171 **Figure C8: AK Score Plots for RPC 1-4. Note that the scales vary, and spatial variation is important compared to intensity.**

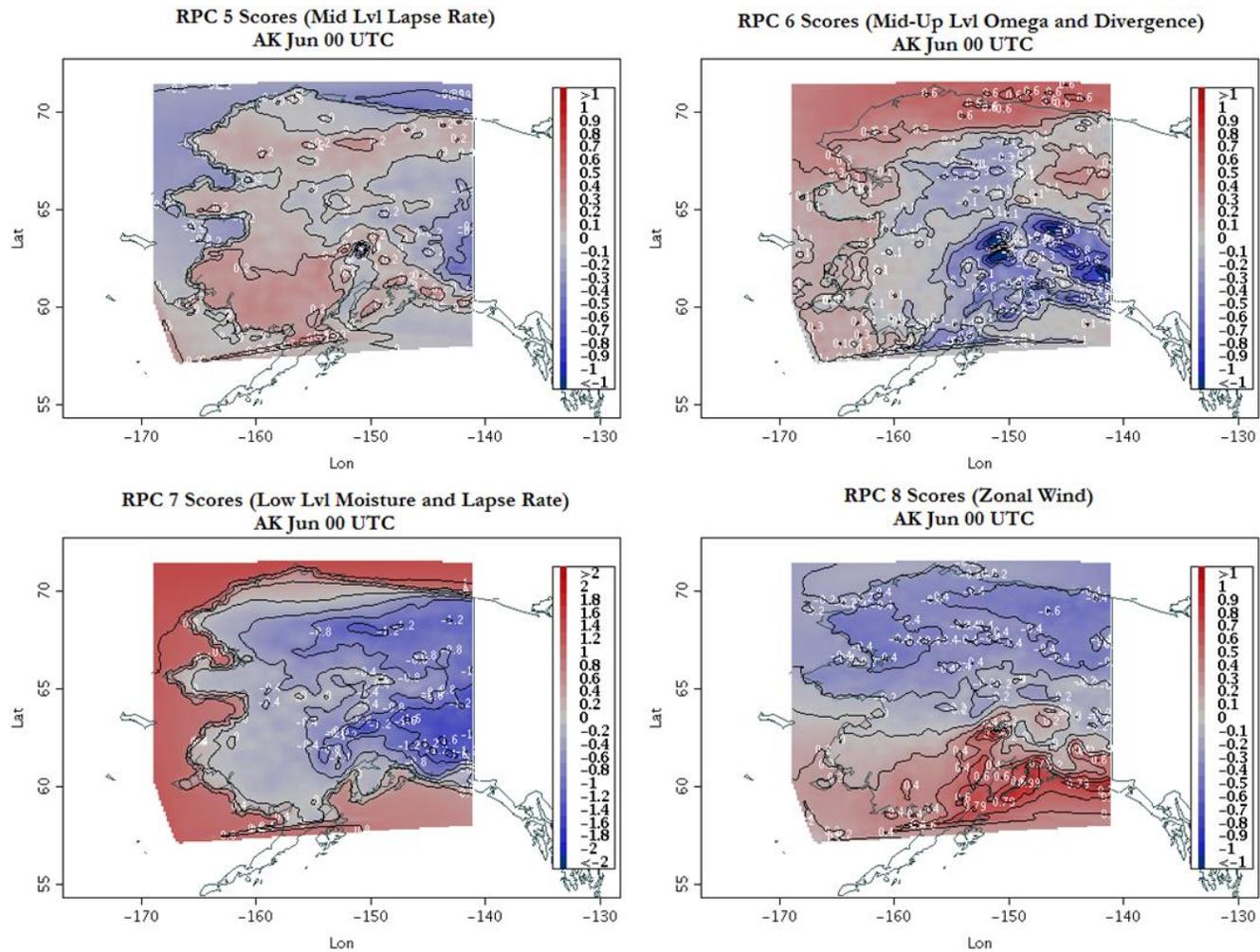


Figure C9: AK Score Plots for RPC 5-8. Note that the scales vary, and spatial variation is important compared to intensity.

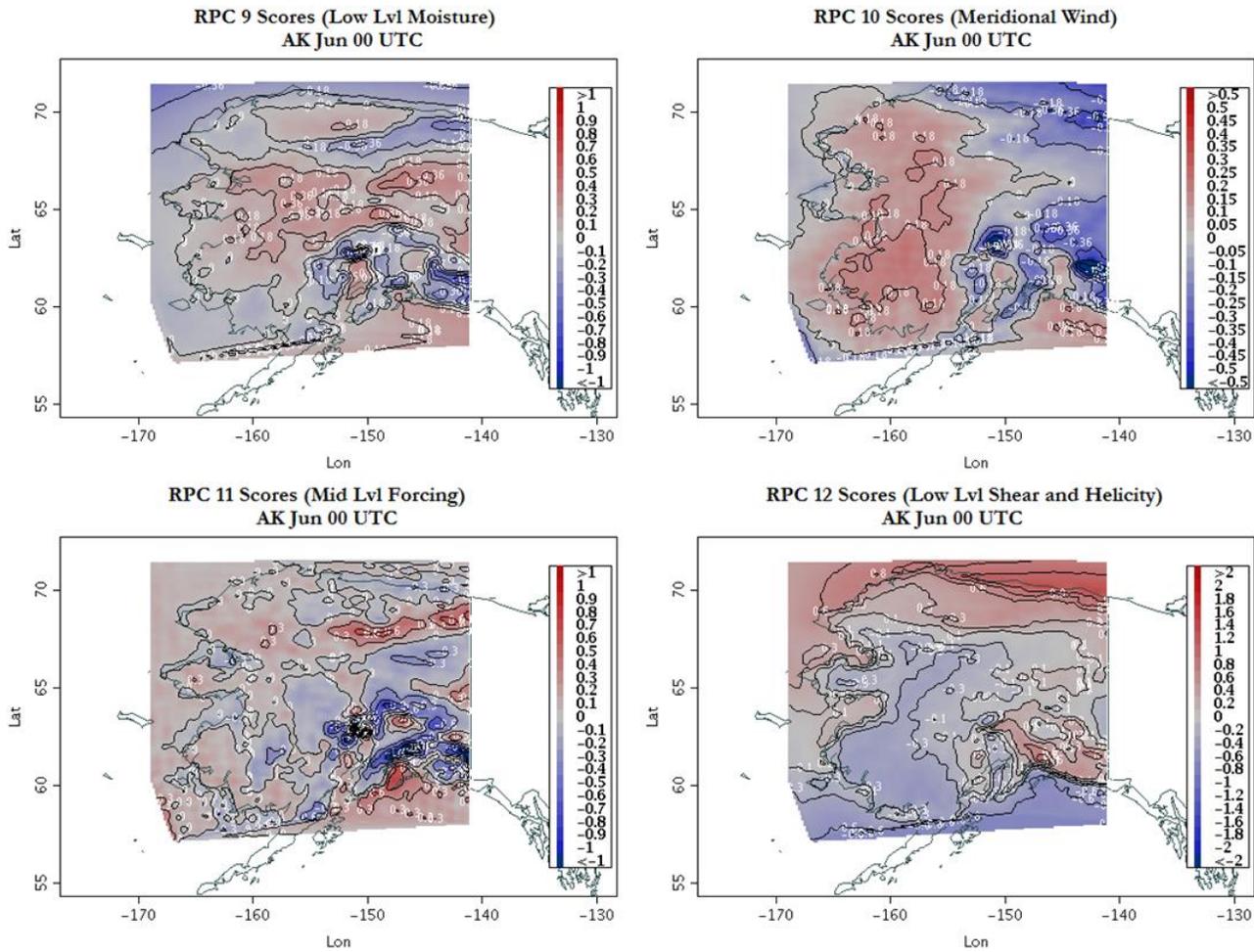


Figure C10: AK Score Plots for RPC 9-12. Note that the scales vary, and spatial variation is important compared to intensity.

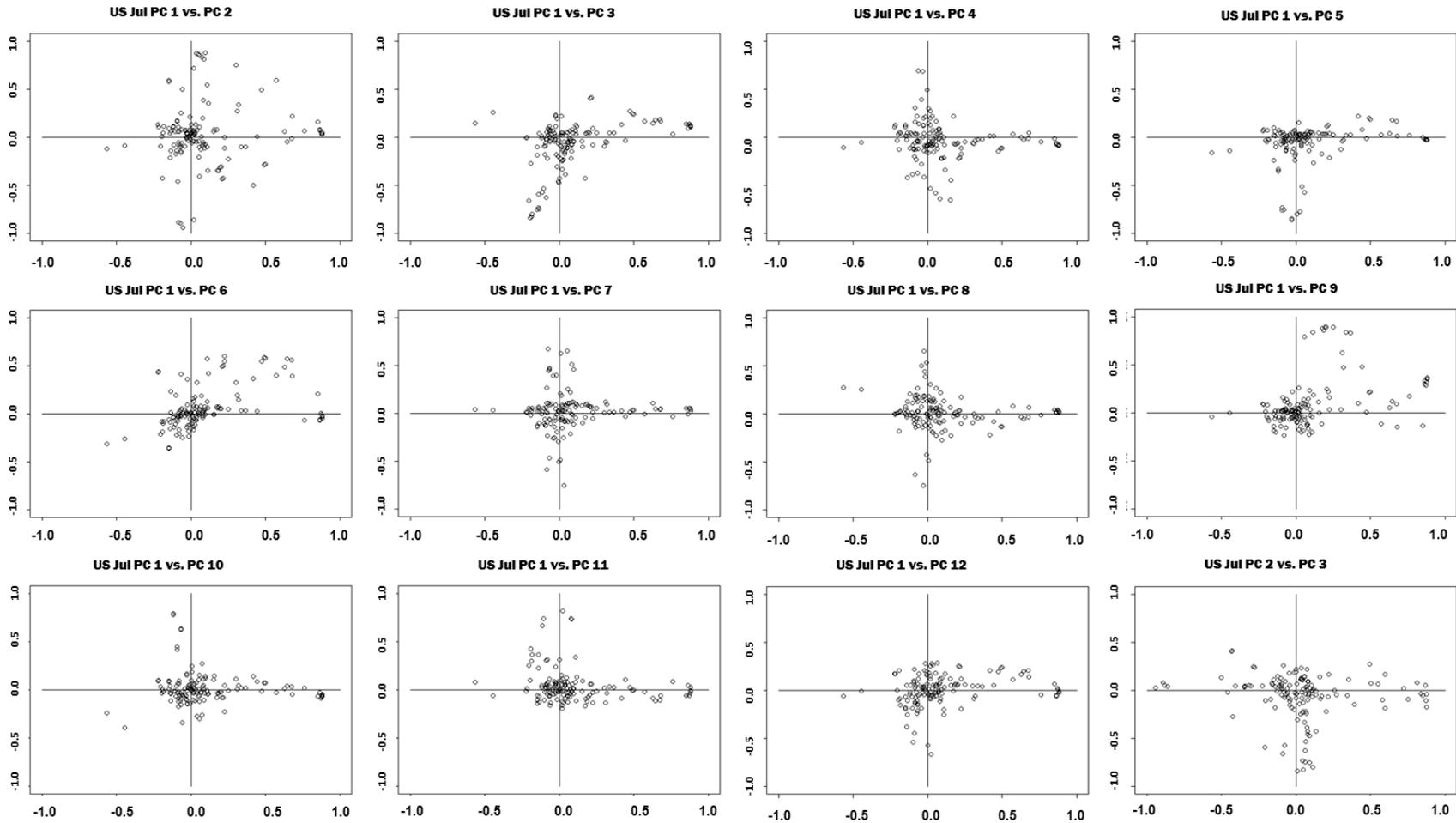


Figure C11: CONUS Pair plots

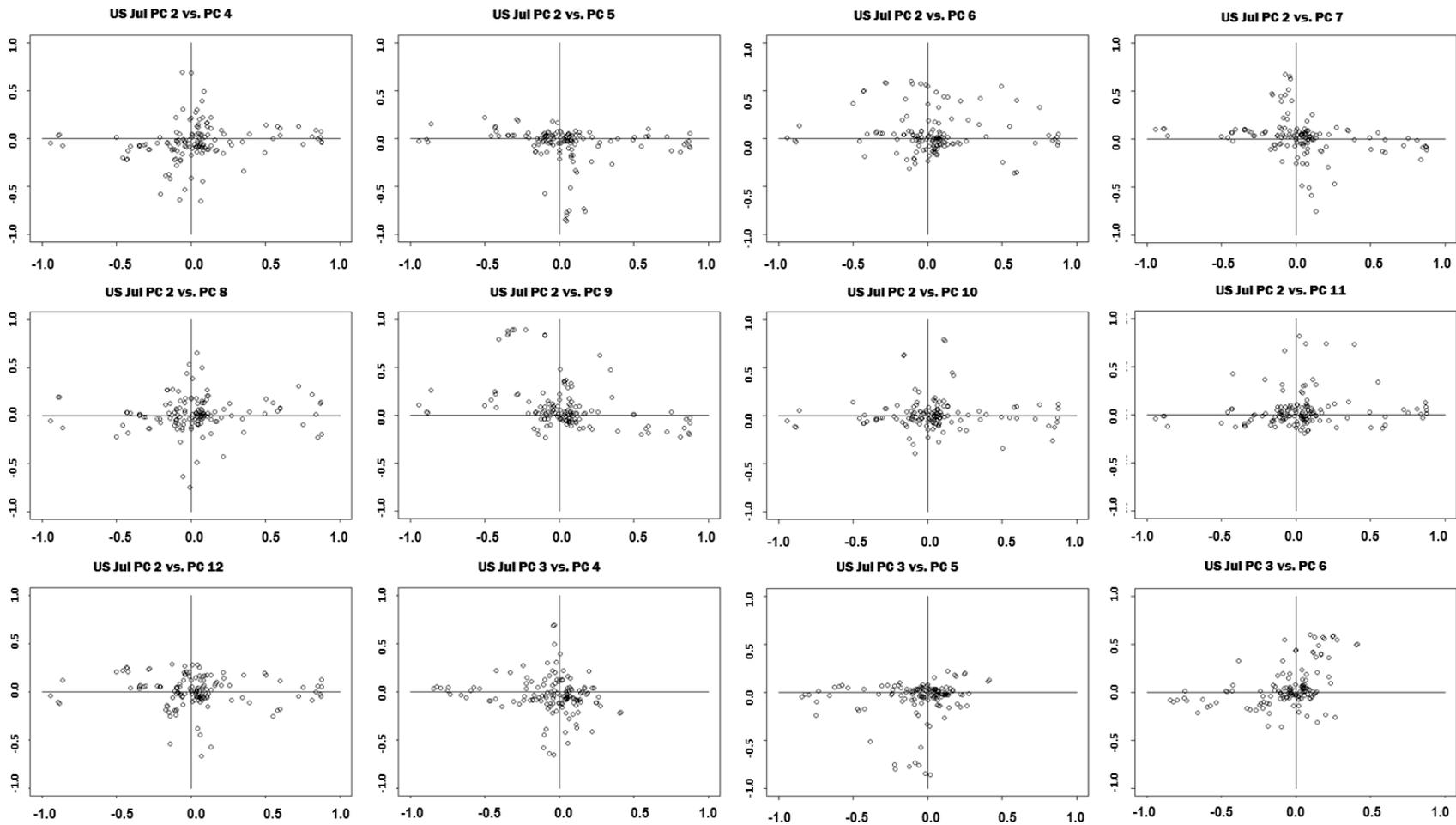


Figure C12: CONUS Pair plots continued.

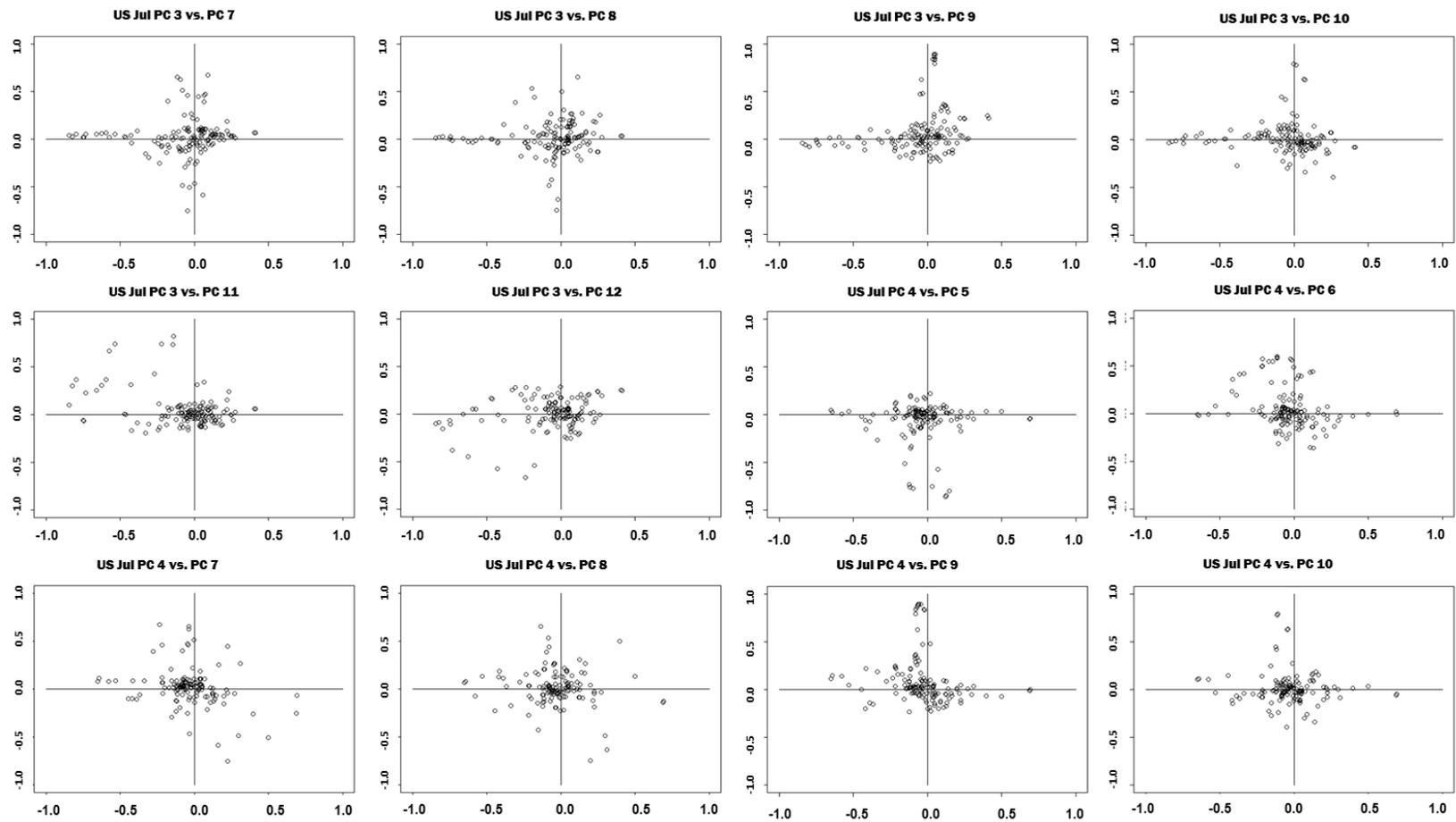


Figure C13: CONUS Pair plots continued.

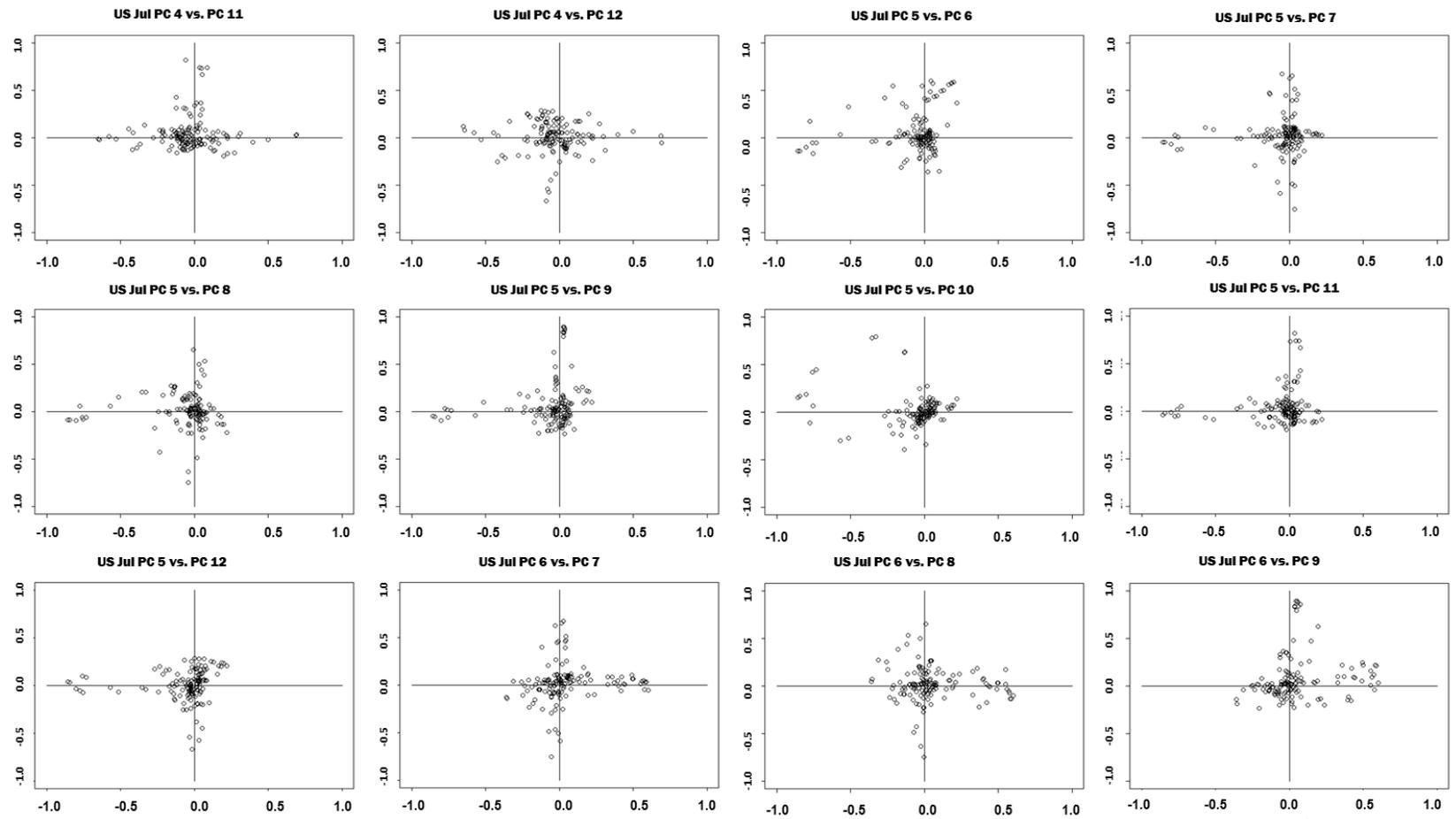


Figure C14: CONUS Pair plots continued.

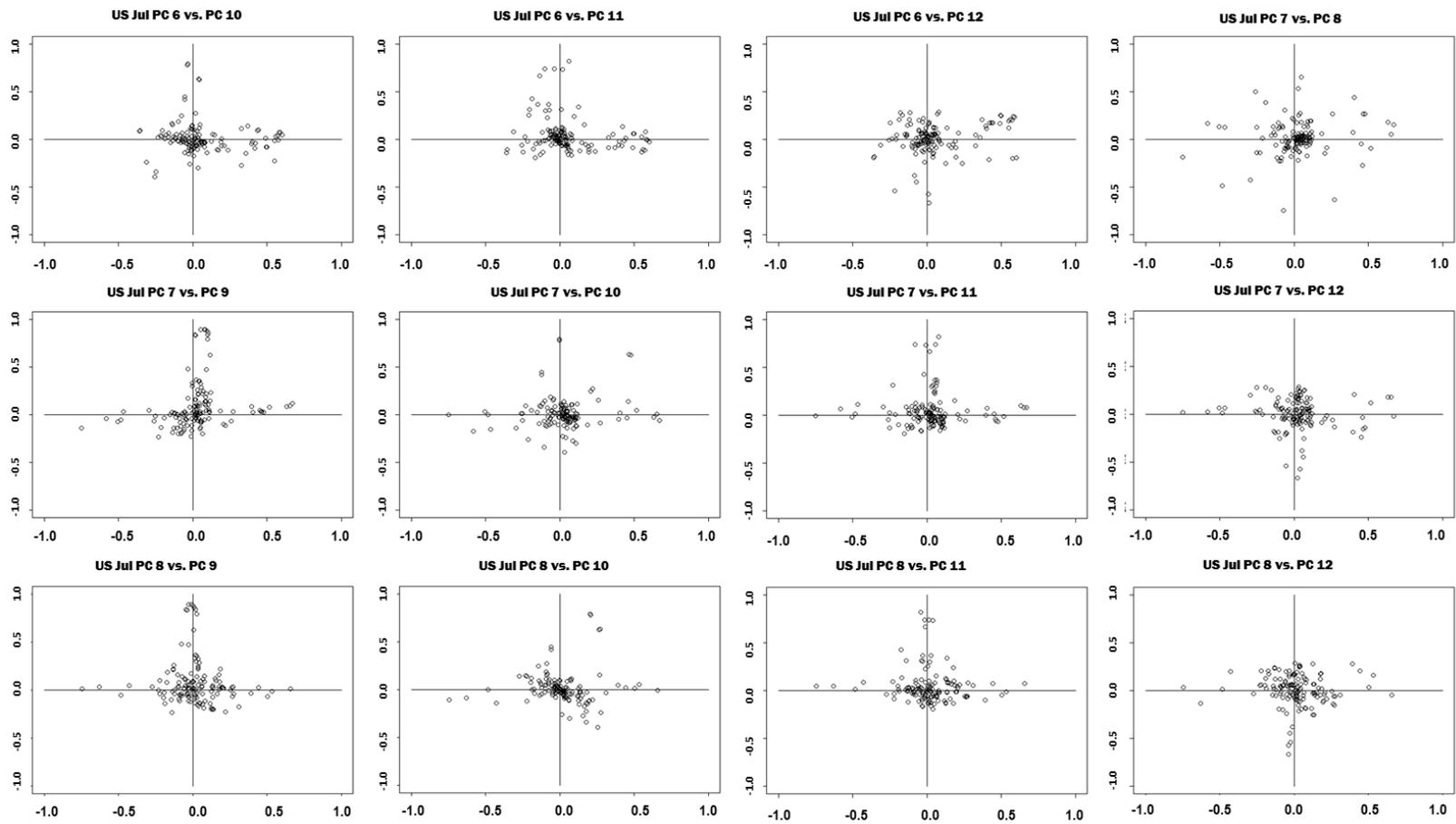


Figure C15: CONUS Pair plots continued.

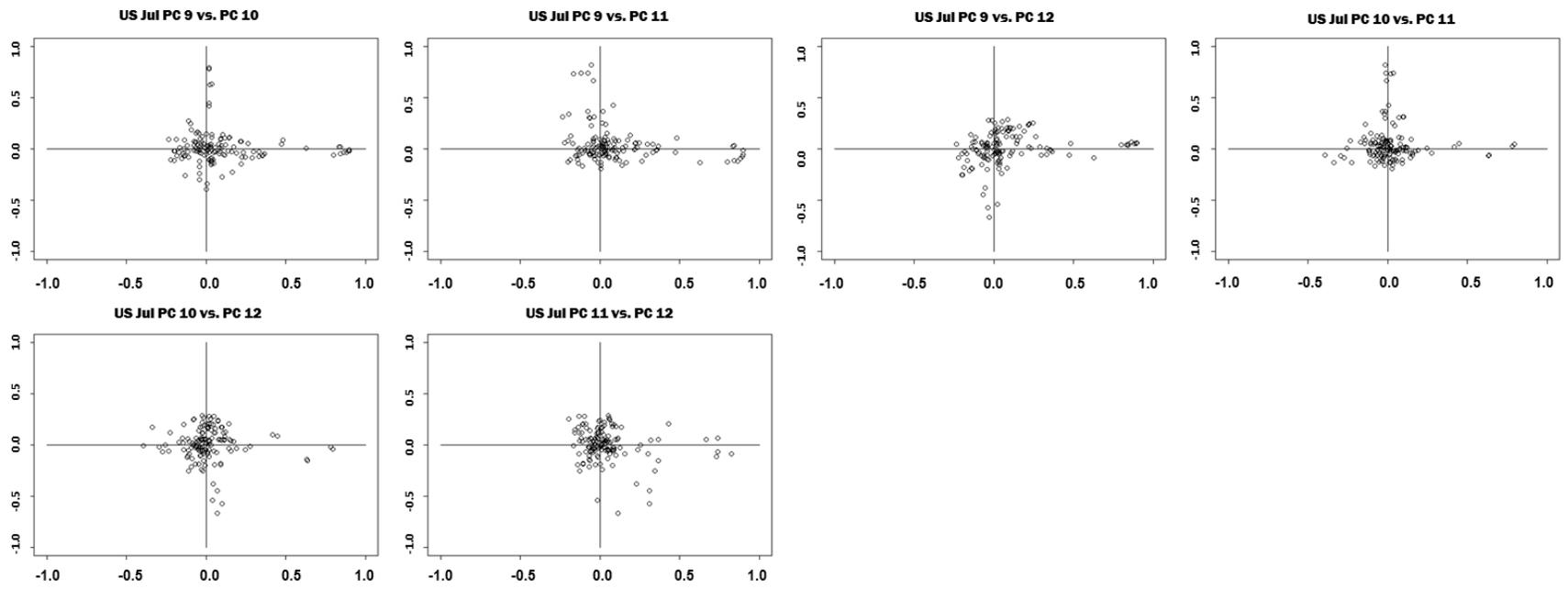


Figure C16: CONUS Pair plots continued.

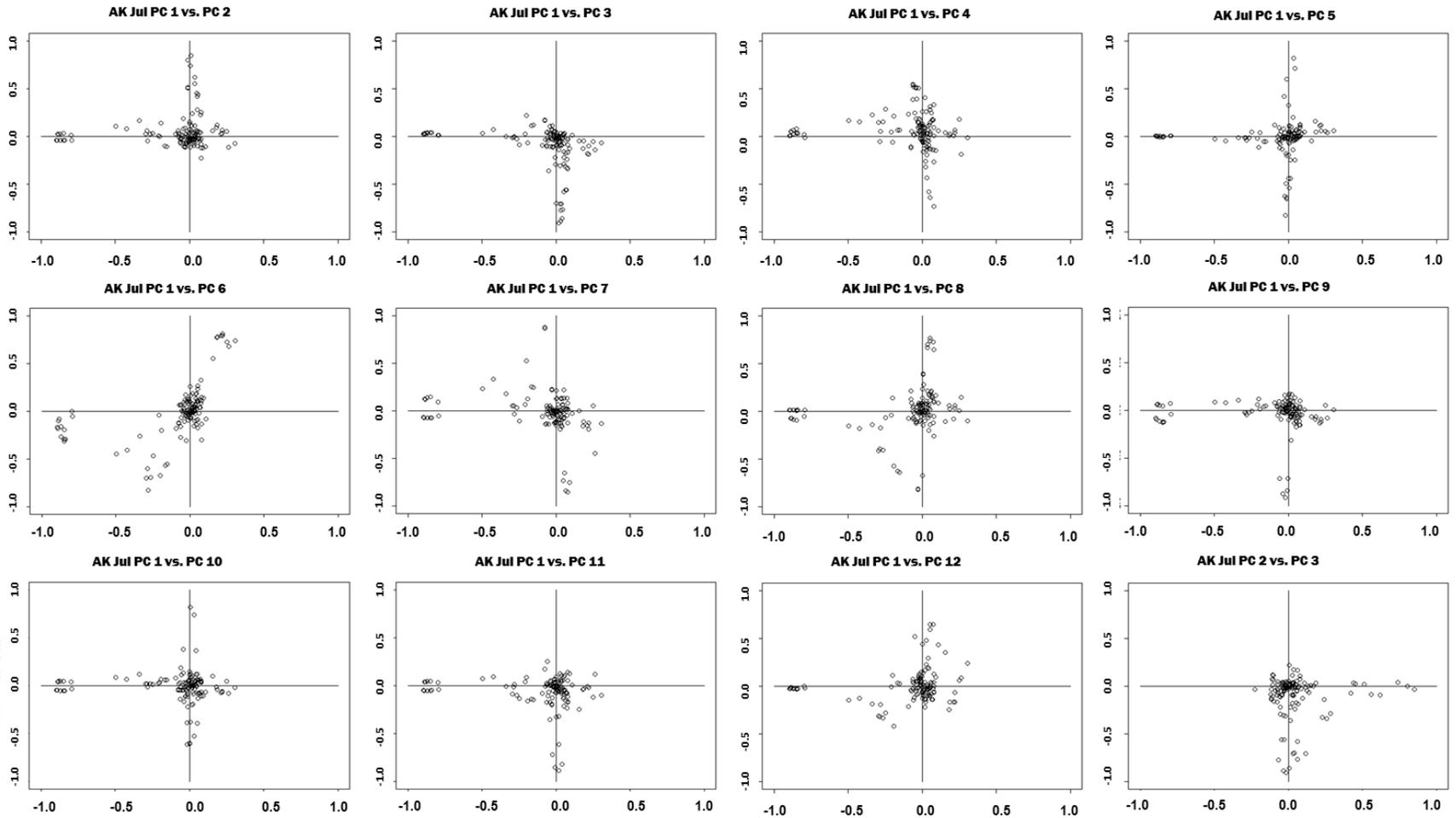


Figure C17: AK Pair plots.

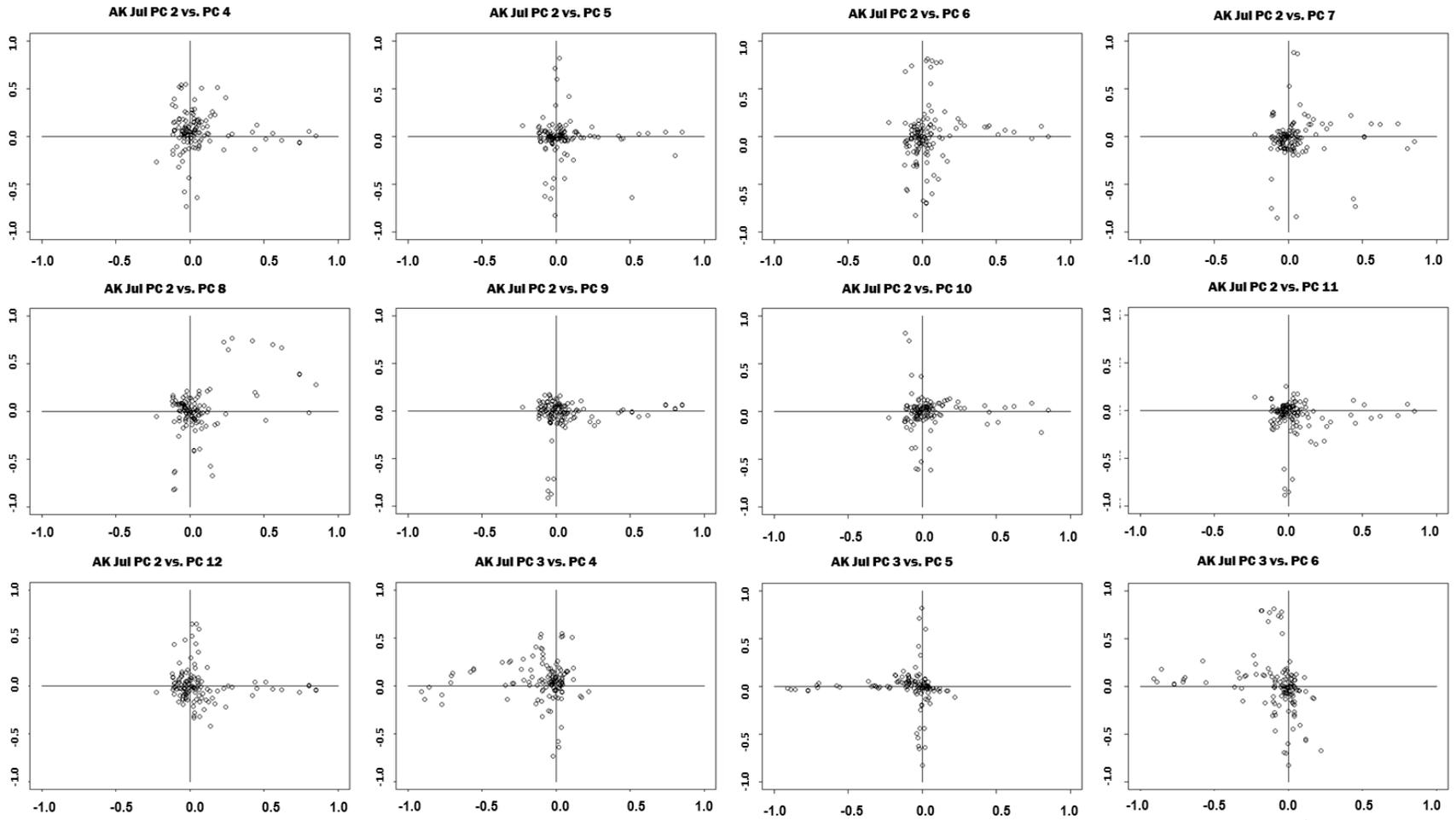


Figure C18: AK Pair plots continued.

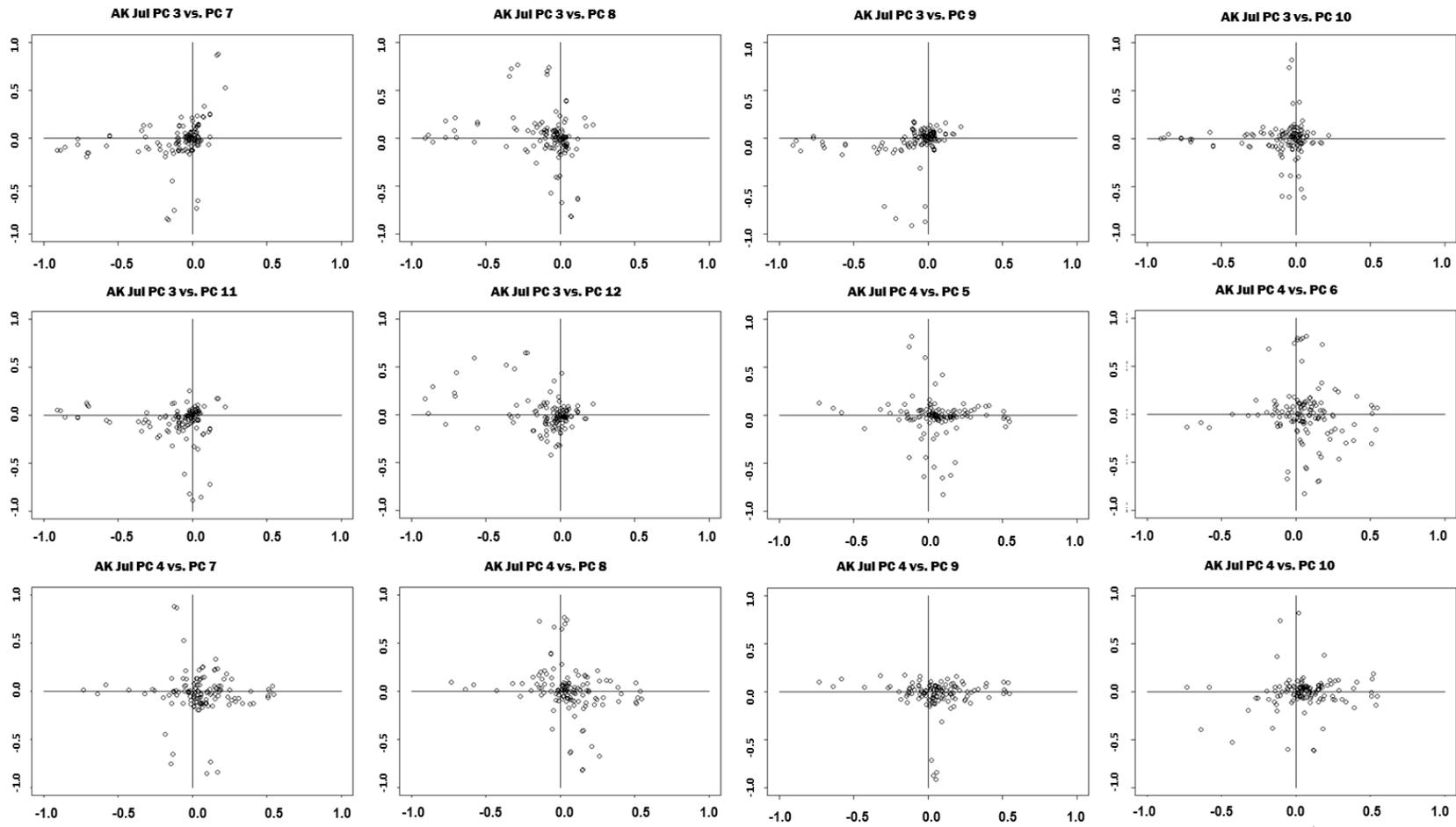


Figure C19: AK Pair plots continued.

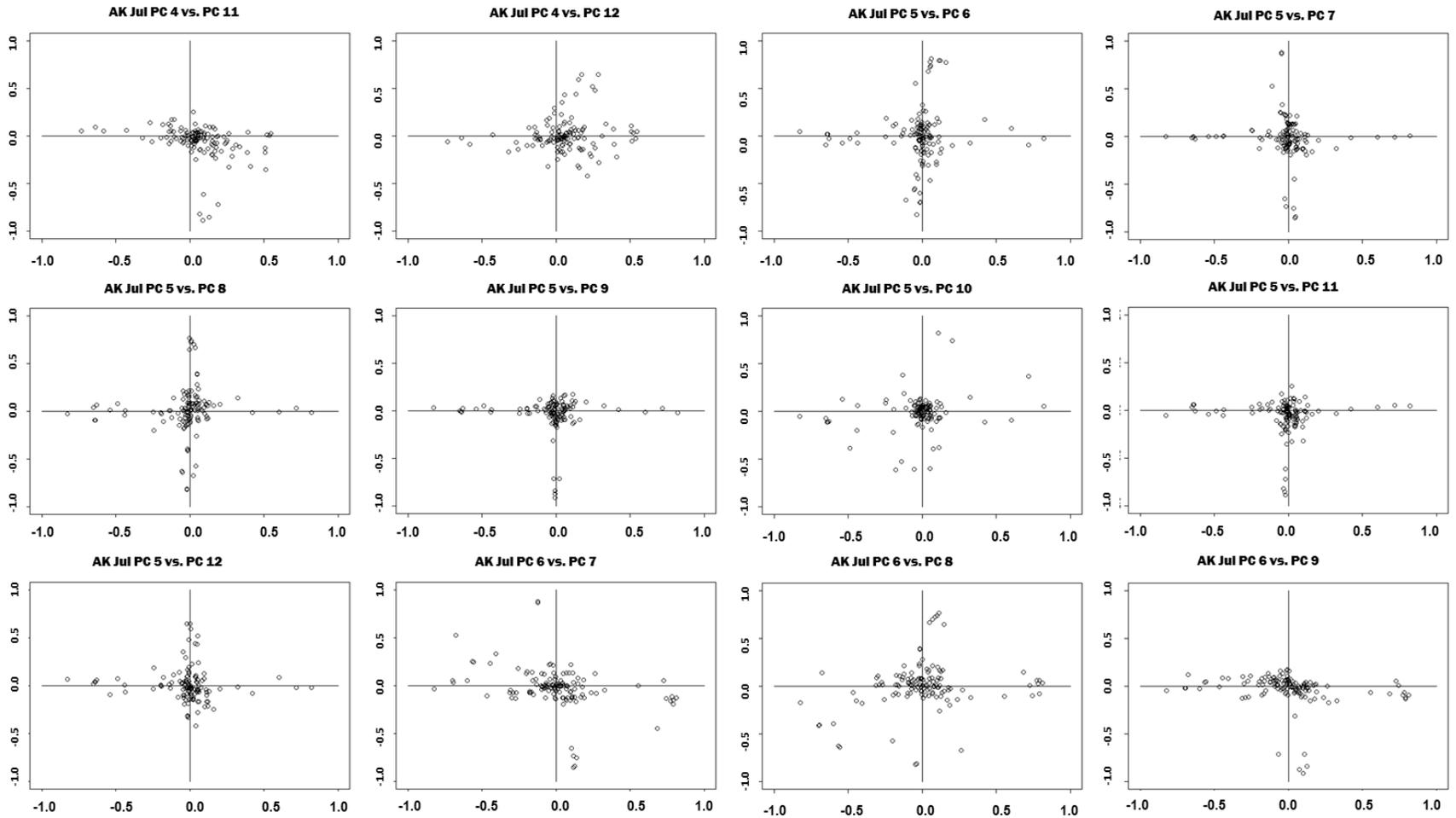


Figure C20: AK Pair plots continued.

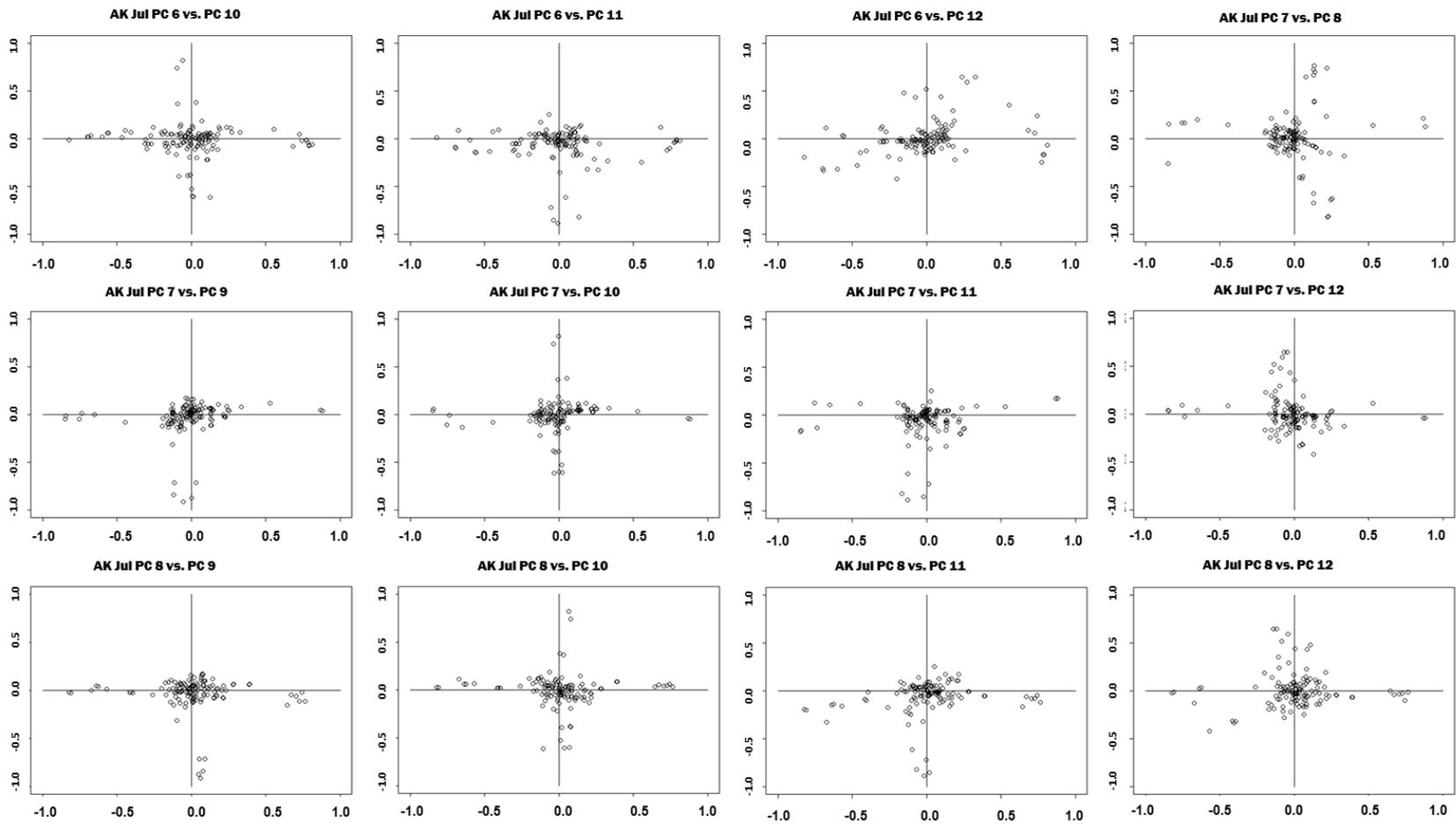


Figure C21: AK Pair plots continued.

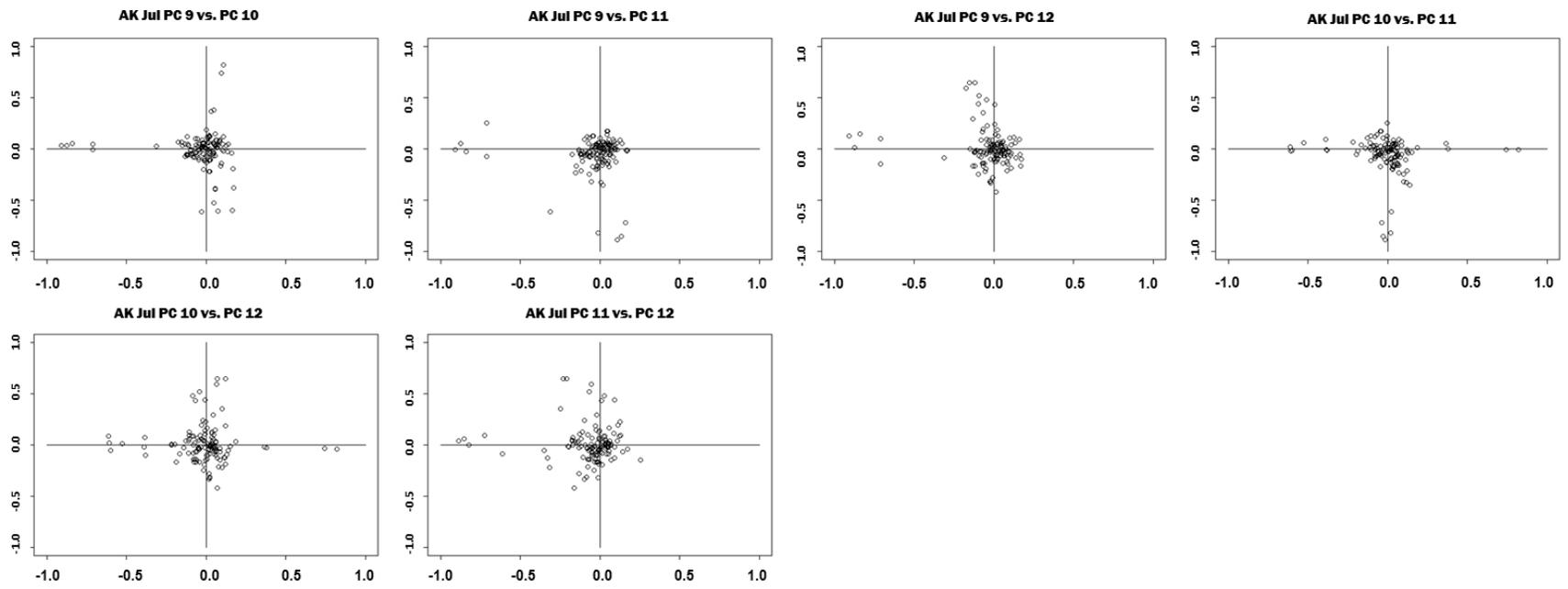


Figure C22: AK Pair plots continued.