UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

# Verification of the Tornado and Lightning Plumes and Evaluation of a New Kernel for the Tornado and Lightning Plumes

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN METEOROLOGY

By

IAN WILLIAM GESELL
Norman, Oklahoma
2020

Verification of the Tornado and Lightning Plumes and Evaluation of a New Kernel for the Tornado and Lightning Plumes

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Harold Brooks, Chair

Dr. Kristin Calhoun

Dr. Jason Furtado

Dr. Michael Biggerstaff

## Acknowledgments

This work would not have been possible with the support and advise from countless individuals. First, I would like to thank my advisors Dr. Harold Brooks and Dr. Kristin Calhoun for their guidance and assistance throughout this project. I would also like to thank them for their assistance in the writing process.

I would like to thank Adrian Campbell for his help with understanding the code for the PHI Tool and for providing me with the lightning and tornado data files. I would like to thank Greg Stumpf for his assistance and guidance with creating the tornado verification. I would also like to thank my committee Dr. Jason Furtado and Dr. Michael Biggerstaff for their guidance. I would like all the professors I have had at the University of Oklahoma for helping me to enhance my meteorological knowledge and grow as a scientist.

I would also like to thank you to my friends and family who supported me during my time at the University of Oklahoma. I would like to thank my parents for my supporting my move to Oklahoma. I would like Kevin Thiel for being able to bounce ideas off of in our office and putting up with in our office for the past two years.

# Table of Contents

## List of Tables

## List of Figures

## Abstract

Though tornado forecasts have been of interest much longer, the National Weather Service (NWS) has been issuing tornado alerts and warnings for more than 50 years. On 1 October 2007, the NWS began issuing storm-based warnings. However, surveys, assessments, and evaluations following the 26-28 April 2011 southeast tornado outbreak and the 22 May 2011 Joplin, MO tornado noted the desire for additional information and communication outside of the watch/warning paradigm. Due to this, and other issues with storm-based warnings, Forecasting a Continuum of Environmental Threats (FACETs) was proposed. FACETs is a framework that could modify the current watch/warning system to communicate hazardous weather information to serve the public. A key part of FACETs is Probabilistic Hazard Information (PHI). Currently warnings are yes/no (binary); they have a start and end time, but do not provide a likelihood of the event or a time of arrival or departure. PHI changes this by putting threat probabilities onto grids that are updated as the storm moves.

PHI plumes for tornado, lightning, wind, and hail were created by NWS forecasters as part of experiments in the NOAA Hazardous Weather Testbed (HWT) in 2016 and 2017. These probabilistic plumes were produced using the storm size and speed to determine the width and length of the plumes. A Gaussian smoother from the center point was then used to provide a visualization of the forecast threat area and region of uncertainty in the HWT.

Since these forecasts are probabilistic, a method of forecast verification is needed to determine the accuracy of these plumes instead of binary verification. Currently, the NWS verifies warnings in this binary nature if a severe local storm report is in the warning. However, most probabilistic forecasts are not verified in this manner. For example, the Storm Prediction Center verifies outlooks with a correct forecast corresponding to an event occurring within 25 mi (40 km) of a forecast point. To move away from the deterministic verification of warnings, this study first examines event distances for both lightning and

tornado hazards, and then applies verification methods used for other probabilistic forecasts to define an appropriate reference class.

Lightning and tornado hazards were chosen for this study since the verification for each of these are the most reliable. The verification data used for the lightning plumes was the National Lightning Detection Network. The tornado plumes were verified by tracking the mesocyclone coordinates using start and end time from the Storm Events Database or using tornado warnings.

This study uses four different tornado and lightning cases from the 2017 HWT PHI experiment. First, for a case the PHI tool code was run to regenerate all the plumes for tornado and lightning hazards from the forecasters. Then, all the plumes for a given time step were merged together spatially. Next, the maximum probability of the plumes at all time steps for a case were merged together and saved producing an accumulation plot. Attributes and ROC diagrams were created for the forecaster data. Practically perfect plumes were then created for comparison to the forecaster PHI to better understand what an ideal forecast could be following the same forecasting constraints and/or rules.

Multiple thresholds of distance (i.e, 0.5 km - 7.5 km) were used to determine the appropriate neighborhood or definition of "an event" for the probabilistic hazard. For each of the tornado and lightning cases, it was found that using a 7.5 km radial distance away from the mesocyclone coordinates or the NLDN flashes, as a neighborhood where the event was defined, was where the attributes diagrams indicated the plumes were most accurate. Practically perfect plumes were used in addition to the forecaster plumes to evaluate the tornado and lightning plumes for an ideal forecast.

Epanechnikov, Quartic, and Triangular kernels were applied to the forecaster and practically perfect created probabilities to see if there was a better kernel than Gaussian for either of the tornado or lightning plumes. It was found that for the tornado plumes the Gaussian kernel provided the best results. However, for lightning the Epanechnikov kernel was the best kernel to use to apply to the probabilities.

# Chapter 1

# Introduction

## 1.1 FACETS

Severe convective warnings issued by the National Weather Service (NWS) help protect life and property. Sgt. John P. Finely, of the U.S. Signal Corps, began to perform research on tornadoes in the 1880s. Based off his research, he created what today would be considered a tornado watch. This tornado watch was referred to as a tornado alert and was created for a large area and long duration (Coleman et al. 2011). Based on Finely's work, Edward S. Holden created a plan for a short term tornado warning system using telegraph wires (Coleman et al. 2011). At this time, it was thought by many scientists that issuing tornado warnings would cause panic which would lead to more harm than good, so the official issuance of tornado warnings ended in 1887 (Coleman et al. 2011).

The U.S. Weather Bureau (USWB), the precursor to the NWS, was formed in 1890. The USWB banned the use of the word tornado in warnings until 1938 (Coleman et al. 2011; Bates 1962; Doswell et al. 1999). Major E. J. Fawbush and Captain R. C. Miller created and issued a tornado forecast and warning for Tinker Air Force Base on 25 March 1948 which were verified (Coleman et al. 2011). In 1950, the USWB removed the ban on tornado forecasting and the era of modern tornado warnings began. The modern severe convective warning was officially adopted and the USWB began issuing warnings in 1965 (Coleman et al. 2011; Rothfusz et al. 2018). Over time with the development of weather radar and more research on tornadoes, tornado warnings improved.

In 1970, the USWB was renamed the NWS (Uccellini and Ten Hoeve 2019). The NWS has been issuing warnings for severe convective weather for about 50 years. Based on research and testing the NWS decided to issue storm-based or polygon warnings on 1 October 2007 (Rothfusz et al. 2018; Coleman et al. 2011; Ferree et al. 2007; Harrison

and Karstens 2017). Storm-based warnings replaced warnings which were issued based on geopolitical boundaries, usually counties (Coleman et al. 2011; Harrison and Karstens 2017). Two of the goals with issuing storm-based warnings were to provide more specific information about the hazard and to alert only those in immediate danger by issuing warnings that covered less area (smaller warnings; Harrison and Karstens 2017).

However, there are several issues that arise with storm-based warnings. Often, there are overlapping warnings which can create confusion for emergency managers and the general public. Also some of the "legacy" (county based alert) systems, such as the National Oceanic and Atmospheric Association (NOAA) Weather Radio, have not been upgraded to disseminate storm-based warnings. Some of the issues with county-based warnings, such as undesirable public response, still exist with the storm-based warning system (Harrison and Karstens 2017). Another issue that can occur when storm-based warnings are issued is county clipping, where portions of the warning are removed prior to issuance to accommodate the legacy system (Karstens et al. 2015).

Due to these challenges and communication issues within the current NWS watch/warning system, new methods of creating and disseminating hazardous weather information are being developed and tested. A new framework and hazard forecasting paradigm called Forecasting a Continuum of Environmental Threats (FACETs) has been proposed (Rothfusz et al. 2018; Harrison and Karstens 2017) to provide a framework for updating the current watch/warning system. FACETs was proposed following internal assessments performed by the NWS after the 26-28 April 2011 southeast tornado outbreak (National Weather Service 2011a) and 22 May 2011 Joplin, MO tornado (National Weather Service 2011b), in addition to Weather-Ready Nation, National Research Council, and National Institute of Standards and Technology recommendations (Rothfusz et al. 2018). FACETs is a framework that will modify the current deterministic watch/warning system to a continuum of information with the use of probabilistic hazard information (PHI; Karstens et al. 2015).

FACETs consists of seven facets. This work involves four of the facets. It involves method and manner of the hazard communication (FACET 1), observations and guidance for forecasters (FACET 2), tools for forecasts (FACET 4), and forecast verification (FACET 7). Method and manner refers to how the hazards are forecasted and communicated (Rothfusz et al. 2018). Observations and guidance refers to the tools and data that are used to create, diagnose, and predict the hazard (s; Rothfusz et al. 2018). Tools refers to the software, hardware, and systems used to create and disseminate the forecasts (Rothfusz et al. 2018). Verification refers to measures used to validate the effectiveness of the forecasts (Rothfusz et al. 2018). PHI is defined as the probability of a weather hazard occurring within a given spatial and temporal domain (Rothfusz et al. 2018). Currently warnings are yes/no (binary), but PHI changes this by putting threat probabilities onto grids. The PHI Tool was developed for forecasters as a way to create PHI forecasts in the NOAA Hazardous Weather Testbed (HWT).

### 1.1.1 PHI Tool

The PHI tool is a geospatial web application designed to allow scientists to generate PHI for severe convective weather. An image of the PHI tool interface is shown in Figure 1.1. Forecasters come to the NOAA Hazardous Weather Testbed (HWT) and participate in a variety of experiments involving the PHI tool to test the feasibility of use in an operational environment. For the lightning hazard, a forecaster can choose to create their own object or modify an object created through automated guidance, ProbLightning. ProbLightning is a model that uses machine-learning in addition to current cloud and cloud-to-ground (CG) lightning strikes, radar data, and the environmental conditions near the storm (Calhoun et al. 2017; Cartier 2017). During the 2017 experiment in the HWT, forecasters had access to ProbLightning and ProbSevere (Cintineo et al. 2018), which provided a storm-based probability of hail, wind, and tornado likelihood, but there was not yet automated guidance available for creating tornado objects (Karstens et al. 2018).

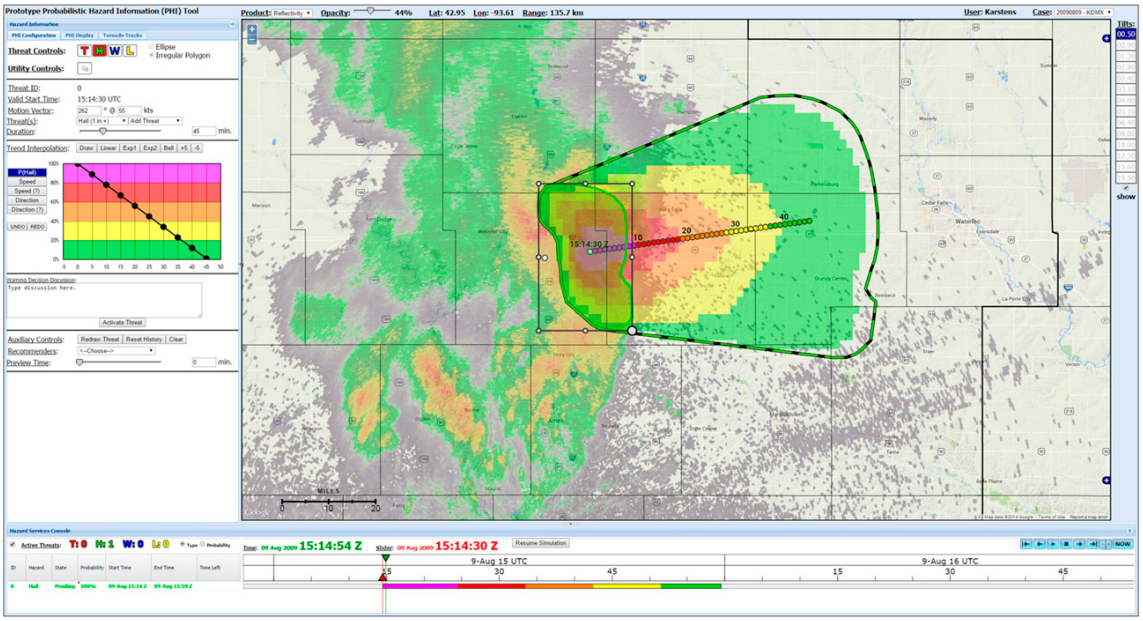



Figure 1.1: An image of the PHI tool interface adapted from Karstens et al. (2015).

To manually create an object, the forecaster draws a polygon or irregular shape to represent the current broad threat area. The forecaster, working within the PHI tool, provides the area swept out by the moving object through its duration. The PHI tool computes the mean motion vector and object's motion uncertainty in the previous step, but the forecaster can choose to override it (Karstens et al. 2015). Next, the forecaster interactively draws a probability trend and specifies a duration for the object (Karstens et al. 2015). To modify a PHI object created using the automated guidance, similar steps are followed. The steps for creating PHI objects changed slightly from 2014 through 2017 as mentioned in Karstens et al. (2018). The PHI tool takes the coordinates of the PHI object and determines the storm size. The storm size along with the speed and direction is used by the PHI tool to determine the width and length of the plumes. The PHI tool then uses a Gaussian smoother from the center point of the object to provide a visualization of the forecast threat area and region of uncertainty using the speed and direction uncertainty. The plumes that are generated from these objects were saved to files.

## 1.2 Current Warning/SPC Verification

Currently, NWS warnings are verified based on an event in the warning. If there is a local storm report (LSR), that exceeds the thresholds for a severe event (severe LSR), that occurs in the warning polygon during the duration of the warning then that warning is verified. If there is a severe LSR that occurs either not during the warning duration or outside of warning polygon, then that LSR corresponds to a missed event. Some work has been done with verifying the PHI plumes by considering a plume to have been verified if a severe LSR occurs somewhere in the plume during the duration associated with the plume (e.g., Harrison 2018). Karstens et al. (2015) determined that due to certain limitations in the current verification metrics/techniques, such as the fact that one severe LSR verifies a warning, different and/or new verification methods/techniques need to be used to verify the PHI plumes. The Storm Prediction Center (SPC) Outlooks are verified if an event occurs within 40 km (25 mi) of a forecast point. SPC watches are verified if a severe LSR occurs in the watch.

## 1.3 Grid Based Verification

Stumpf et al. (2015) developed a method of gridded verification to verify tornado and severe thunderstorm warnings. In this method, the observations of an event were placed on a 1 $km^2$ grid and then each grid box of the warning was given a value of a one (the grid boxes outside the warning were assigned a value of zero) and placed on the same grid. In the creation of the observation grid, a sphere of influence around the points was used to account for location uncertainties (Stumpf et al. 2015). The sphere of influence was objectively determined based on user opinion surveys about the region around the observation where a recipient of the warning thought the warning was justified by the event occurring (Stumpf et al. 2015). Additionally, Stumpf et al. (2015) used radar data in combination with damage

surveys to determine the tornado track. Stumpf et al. (2015) choose to use a sphere of influence radius of 5 km around the tornado track for verification.

Annual Spring Forecasting Experiments (SFEs) have been run in the NOAA HWT for the past 20 years as way to test new tools and techniques for forecasting severe weather (Gallo et al. 2017). In the 2014 SFE, participants created hourly probabilistic forecasts which were verified using a 40 $km^2$ grid and the forecasts were found to be reliable (Gallo et al. 2017). But when the forecasts were verified using a 20 $km^2$ grid, the forecasts were found to be overforecasting the threat of severe weather (Gallo et al. 2017). This verification was performed by placing the LSRs and a radar derived hail product on a grid (Gallo et al. 2017). This shows that changing the distance away from an event that is considered a correct forecast affects the reliability of the forecasts. This research shows that changing the definition of an event changes the reliability of the forecasts.

As numerical weather prediction models' grid spacing has decreased and with the development of convective resolving/allowing models there have been numerous grid-based/gridded forecast verification studies to try to quantify the improvement that the new models offer. This has been most often done for precipitation forecasts. However, outside of Stumpf et al. (2015), there has not been any research completed on performing grid-based verification of severe weather on the warning scale. Gilleland et al. (2009) compared neighborhood, scale separation, feature-based, and field deformation approaches to spatial forecast verification. The methods are compared using gridded forecasts with an observation field that is on the same grid. The forecasts and observations are described in Ahijevych et al. (2009). Features-based approaches identify individual features within a field and analyzes them separately while the field deformation methods apply to the entire field as a whole. There are three items that scale separation approaches aim to accomplish. These three are: "assess the scale dependency of the error, determine the skill-no-skill transition, and assess the capability of the forecast to reproduce the observed scale structure

in the observations" (Gilleland et al. 2009). The neighborhood approach smooths the forecasts over a range of increasing scales and the filtered fields broadly resemble the original field while the scale-separation approach treats each scale independently (Gilleland et al. 2009). Additionally, Ahijevych et al. (2009) found that for realistic precipitation cases using forecasts created with Numerical Weather Prediction models the new spatial verification methods (neighborhood, scale separation, feature-based, and field deformation) can give credit for close forecasts or forecasts that resemble the observation field. Rewarding close forecasts seems reasonable for the tornado and lightning plumes because of flying debris for tornadoes and the fact that lightning can strike a distance away from the storm it is associated with. None of these methods have been used to perform forecast verification at the warning scale.

Ebert (2008) stated that neighborhood verification can be referred to as fuzzy verification due to neighborhood verification methods allowing a forecast to be both partially correct and partially incorrect. These techniques require that the forecast and observations are in approximate agreement (Ebert 2008). This should be true for the PHI plumes. An assumption of fuzzy verification is that it is okay for the forecast to be slightly displaced and still be useful. The area that contains the maximum allowable displacement is referred to as a neighborhood (Ebert 2008). Since lightning usually covers a large area in a storm, and if lightning is seen within 10 km sporting events have to be stopped for 15 minutes per strike this seems to be a reasonable approach to verifying the lightning plumes. Additionally, due to flying debris from tornadoes this seems to be a reasonable approach to verify the tornado plumes.

By verifying model forecasts of precipitation, Ebert (2009) found that the strength of neighborhood verification is in showing the scales at which the forecast has useful skill. Neighborhood verification methods address different decision models about what makes a useful forecast (Ebert 2009). Roberts and Lean (2008) developed a method of spatial verification called the Fraction Skill Score (FSS) and used it to evaluate the United Kingdom

Met Office Unified Model. The forecast and observed rain fractional coverage are compared in Roberts and Lean (2008) and it was determined that using a 1 km grid spacing resulted in a better FSS value than using 4 km or 12 km grid spacing. The question for this research is how close to a lightning strike or a tornado counts as an event. A certain neighborhood around a tornado or lightning strike and/or the FSS can be used to determine at what scale the probabilistic tornado and lightning plumes are useful forecasts. The scale at which the tornado and lightning plumes are useful forecasts would be the scale at which an event can be defined.

## 1.4 Study Purpose

Currently, there has been limited research performed on the verification of PHI plumes. The PHI tool uses a Gaussian kernel applied to the forecaster and automated (for the lightning hazard only) probabilities to create the plumes. There are other kernels that could be used, but have not been tested yet. For this research an event definition is needed to determine at what scale the tornado and lightning plumes can be considered useful. The event definition will have a substantial impact on the verification results found and discussed in this study. Using a different kernel than the Gaussian will change how the probability values are used in the tornado and lightning plumes. This means that there potentially could be a more useful kernel for the tornado and lightning plumes in some or all forecast environments/scenarios. The main goals of this research are to:

1. Provide an event definition for each hazard, tornado and lightning.

2. Create a method to verify the tornado and lightning plumes that were created by forecasters and the automated guidance (lightning only) in the NOAA HWT.

3. Determine if the Gaussian kernel or a different kernel should be applied to the forecaster and automated (lightning only) probabilities to create the plumes for both the tornado and lightning hazards.

## 1.5 Summary

- The concept of tornado forecasts have been around for more than 100 years.
- The NWS has been issuing warnings for about 50 years.
- The NWS began issuing storm-based warnings on 1 October 2007.
- There are numerous issues currently present with storm-based warnings.
- Due to the 26-28 April 2011 tornado outbreak and the 22 May 2011 Joplin, MO tornado FACETS was proposed as a new watch/warning paradigm.
- FACETs is a framework to modify the current deterministic watch and warning system to a continuum of information with the use of PHI.
- The PHI tool was developed to allow forecasters to test the creation of PHI in an operational environment.
- Forecasters in the NOAA HWT used the PHI tool to create probabilistic plumes.
- A Gaussian smoother was incorportated into the PHI tool to produce probabilistic risk spatially from the center of the storm track. Other kernals have not been tested.
- Currently watches and warning are verified if there is a severe LSR in the watch or warning.
- Previous grid based verification research has been primarily focused on verification of precipitation forecasts from numerical weather prediction models.
- One way to perform spatial grid based verification is by using a neighborhood approach (or event definition).
- This study will determine an appropriate event definition for both tornado and lightning hazards.

- This study will create a method to verify the tornado and lightning plumes that were created by forecasters and the automated guidance (lightning only) in the NOAA HWT.
- This study will determine if the Gaussian kernel or a different kernel should be used to be applied to the probabilities to create the plumes for both the tornado and lightning hazards.

# Chapter 2

# Data and Methods

## 2.1 Cases

PHI plumes for both tornado and lightning were created by forecasters in the 2017 HWT experiment. Four cases for each tornado (Table 2.1) and lightning (Table 2.2) were chosen from this experiment to examine event definition and kernel options through a variety of verification methods. These cases will be discussed in more detail in Chapter 3.

Table 2.1: Four cases from the 2017 HWT that were evaluated for the tornado plumes. The time for each case is in Universal Time Coordinate (UTC).

| Date | Time (UTC) | Location | Radar |
|---|---|---|---|
| 25-26 May 2015 | 22:45-00:45 | Bennington, KS | KTWX |
| 24 May 2016 | 22:00-23:30 | Dodge City, KS | KDDC |
| 8 May 2017 | 20:40-21:20 | Denver CO | KFTG |
| 24 May 2017 | 18:40-20:30 | Columbia, SC | KCAE |

Table 2.2: Four cases from the 2017 HWT that were evaluated for the lightning plumes. The time for each case is in Universal Time Coordinate (UTC).

| Date | Time (UTC) | Location | Radar |
|---|---|---|---|
| 25-26 May 2016 | 22:45-00:45 | Bennington KS | KDDC |
| 22 July 2016 | 17:20 - 18:55 | Grand Junction, CO | KGJX |
| 9 September 2016 | 17:55-19:55 | Melbourne, FL | KMLB |
| 25-26 May 2017 | 22:45-00:00 | Goodland, KS | KGLD |

## 2.2 Verification Methods

New code was developed for generating plumes using tornado and lightning forecaster data files which were created using the PHI tool during the 2017 HWT experiment. For lightning PHI objects, the forecaster and automated plumes were merged together at each individual two-minute time step. For tornado PHI objects, the forecaster plumes were merged together at each individual two-minute time step. For each case, an accumulated maximum probability of each hazard plume was created created using the maximum value at any grid point (one km by one km) during each case. The probability that corresponded to the plumes ranged from 0 to 100. For each case a grid of all the probability values from all plumes (either tornado or lightning) was created to evaluate forecaster performance.

### 2.2.1 Tornado Verification

A combination of the tornado tracks provided by the NWS in the Storm Events Database (start point and end point) and the radar-indicated mesocyclone track were used to verify the tornado hazard plumes. This verification method was used because the data in the Storm Events Database includes only the start and end time of tornadoes and tornadoes usually do not move in a straight line. Similar methods were used in recent studies including Stumpf et al. (2015) and Flora et al. (2019) to verify tornado warnings/forecasts. Using radar derived mesocyclone tracks removed some of limitations associated with the storm reports in the Storm Events Database such as non-meteorological bias (Brooks et al. 2003; Flora et al. 2019) and population bias (Brooks et al. 2003).

For each tornado case, either the initial point of the tornado track from the Storm Events Database was used or, if there was only a warning (no confirmed tornado) on the storm, the location of the mesocyclone associated with a warning was used as the beginning point in the mesocyclone track. The mesocyclone was determined using base velocity (0.5°) from the nearest WSR-88D radar and looking for the strongest velocity couplet. If it was difficult to find the velocity couplet then normalized rotation (NROT) was used for assistance.

NROT is azimuthal shear divided by the area of the bins used to calculate it (Lemon and Umscheid 2008). Each of the coordinates of the tornado location were recorded in the center of the mesocyclone found as described above. If there was a Tornado Vortex Signature (TVS) or rotation signature that matched the current path, the coordinates of the center of the signature were recorded. The coordinates were recorded and interpolated so that there were coordinates at every minute assuming a constant speed between coordinates.

For an individual storm, recording of coordinates stopped at the time of the end of the tornado path from the Storm Events Database, at the end of the warning, or at the end time for the case. Then a grid of zeroes and ones was created for each tornado case in a similar manner as with the lightning.

### 2.2.2 Lightning Verification

The National Lightning Detection Network (NLDN) by Vaisala was used to verify CG lightning activity within and around the lightning plumes. The NLDN has been providing real-time lightning information for the U.S. since the 1980s with an accuracy within 2 km and a few microseconds (Cummins and Murphy 2009). For each lightning case the latitude and longitude bounds of the plume were used to create spatial bounds for a grid of coordinates for the NLDN data. The time of the first and last plumes for a case were used to create temporal bounds. The NLDN data with times outside of the temporal and spatial bounds were removed.

Using the technique of gridded verification from Stumpf et al. (2015) it was determined that since the probability plumes represents a one km by one km by two minute grid point, the NLDN coordinate grid should have the same resolution. Since the NLDN contained data down to millisecond, every other minutes' worth of NLDN coordinates for a case was removed beginning with the start time for the first plume of a case. This new set of data was then used to make a NLDN grid of zeros where there were not NLDN coordinates and ones where there were NLDN coordinates.

The NLDN grid for a case was created by rounding the NLDN coordinates for that case to the hundredths place. A grid of zeros was created for the NLDN grid. The latitude and longitude coordinates for each position in the NLDN grid was calculated. Each of the NLDN coordinates was compared to the latitude and longitude coordinates of the grid. If a NLDN coordinate matched a latitude and longitude coordinate of the NLDN grid then a one was placed in that location in the grid. The NLDN grid was saved as a binary file. These steps were repeated to a make a NLDN grid of zeros and ones for all four lightning cases.

## 2.3 Kernels

The PHI tool uses a Gaussian (Equation 2.1) kernel (distribution) to smooth the forecaster created probabilities. In order to determine if the Gaussian kernel is the most appropriate kernel for both lightning and tornado hazards, for each tornado and lightning case an Epanechnikov (Equation 2.2), Quartic (Equation 2.3), and Triangular (Equation 2.4) kernel (distribution) was applied to the forecaster and practically perfect plumes created probabilities. The equation for the Gaussian kernel is shown below:

$$Gaussian = \frac{1}{\sqrt{2\pi}} \exp \frac{-1}{2} u^2 \tag{2.1}$$

where u was the outer boundary of the PHI object (plume). The equation for the Epanechnikov kernel is shown below:

$$Epanechnikov = \frac{3}{4}(1-u^2) \tag{2.2}$$

The equation for the Quartic kernel is shown below:

$$Quartic = \frac{15}{16}(1-u^2)^2 \tag{2.3}$$

The equation for the Triangular kernel is shown below:

$$Triangular = (1-|u|) \tag{2.4}$$

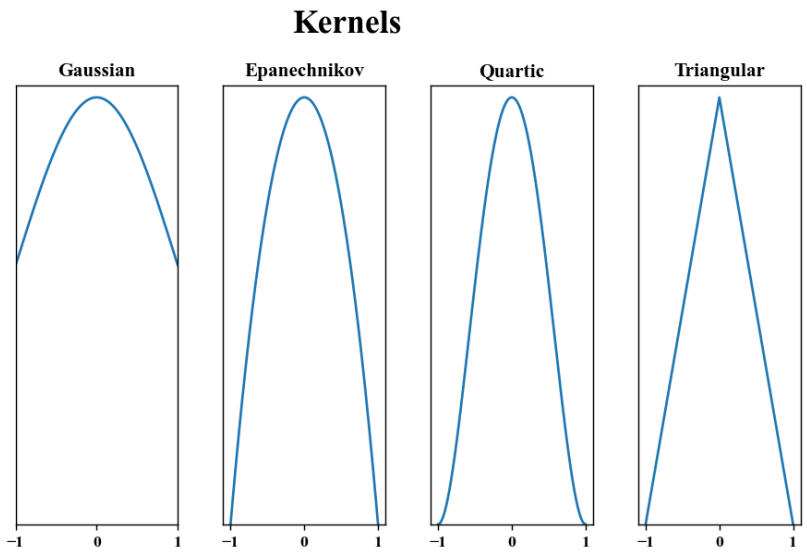


Figure 2.1: The four kernels used in this study are plotted. For this study the peak of each kernel was at one.

The four kernels are shown in Figure 2.1. Equations 2.2, 2.3, and 2.4 were adjusted slightly to have approximately the same area under the curve and peak height as the Gaussian kernel. The Gaussian kernel has a wider peak than the other kernels and would go to infinity but is truncated by the PHI tool. This means that the Gaussian kernel is going to use more higher probability values than the other kernels. All of the other kernels go to zero at one as opposed to at infinity. The Epanechnikov kernel has a rounded narrow peak and so it will use the least number of the higher probability values out of the four kernels. This means that forecasts created using this kernel are going to have the appearance that the forecaster has more confidence than forecasts created using any of the other three kernels. The triangular kernel has a pointed peak and this means that it is going to utilize more of the higher probability values than the Epanechnikov kernel. In some instances this could be a good way to indicate confidence (a different amount of confidence than the Epanechnikov kernel) in the probabilistic plumes. The Quartic kernel has a wider peak than the Triangular and Epanechnikov kernels and this means that it is going to be more similar to the Gaussian kernel in terms of the higher probability values.

## 2.4 Verification Diagrams

### 2.4.1 Brier Score

The Brier Score (BS) is a common way to verify probabilistic forecasts of a yes/no event. Mathematically the BS is defined as

$$BS = \frac{1}{N}\sum_{i=1}^{N}(f_i - x_i)^2 \tag{2.5}$$

with $N$ as the number of forecasts, $f_i$ is the forecast probability, and $x_i$ is the actual outcome of the event at instance $i$. $x_i$ is either zero or one. The BS represents the mean squared error (mse) of the forecast. The BS ranges from zero to one. Since it is a measure of error, zero is the best possible score. With the use of algebra the BS (equation 2.5) can be decomposed into the following

$$BS = \frac{1}{N}\sum_{i=1}^{I} n_i(f_i - \bar{x_i})^2 - \frac{1}{N}\sum_{i=1}^{I} n_i(\bar{x_i} - \bar{x})^2 + \bar{x}(1-\bar{x}) \tag{2.6}$$

with $N$ defined as the total number of forecasts, $I$ the number of unique forecasts issued, $\bar{x}$ the base rate of the event, $x_i$ the number of forecasts at each probability, $\bar{x_i}$ the observed frequency of the event at each probability, $n_i$ being the number of forecasts at each probability, and $f_i$ being the forecast probabilities. The base rate is the mean of the event divided by the mean of the forecasts of the event. The first term on the right side of equation 2.6 is reliability. Perfect reliability results in a value of zero. Reliability refers to how close the forecast probabilities are to the observed frequency at each forecast probability. The second term on the right side of equation 2.6 is resolution. Perfect resolution results in a value of one. Resolution refers to how well the forecast probabilities are distinguishable from the base rate of the event. The third term on the right hand side of equation 2.6 is uncertainty. Uncertainty is proportional to the square of the base rate. This means that the BS can be thought of as reliability minus resolution plus uncertainty.

#### 2.4.1.1 Attributes Diagram

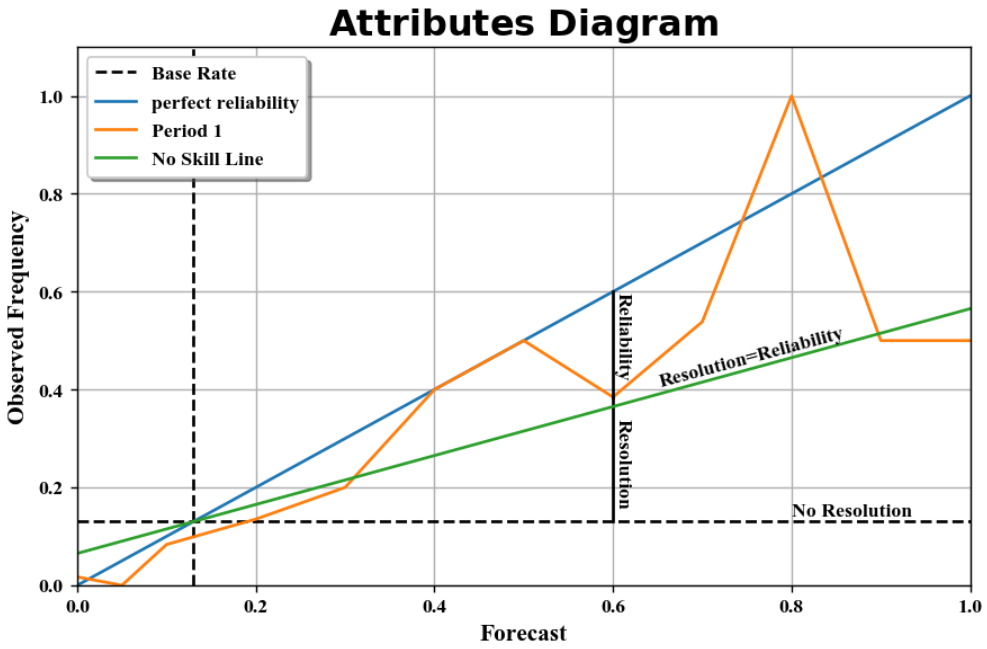



Figure 2.2: An example attributes diagram for precipitation events over the course of a year.

The attributes diagram is a visual representation of the terms of the BS. Figure 2.2 shows an attributes diagram for precipitation forecasts over the course of a year. On these diagrams observed frequency is on the vertical axis and forecast probability is on the horizontal axis. The line at a 45° angle is perfect reliability. The vertical distance from this line to the Period 1 line is the reliability. The dashed blacked lines are the base rate. The horizontal base rate line is also the no resolution line.

The distance from this line to the Period 1 line is the resolution. This means that to have good resolution you might not have good reliability and vice versa. This is especially true at forecast value of 0.2 because if there is good reliability there will be bad resolution. So there is a trade-off between having good resolution versus having good reliability. Halfway between the perfect reliability line and the no resolution line is the no skill line. This refers to no skill relative to the BS. The no skill line is where resolution equals reliability or where term 1 on the right hand side equals term 2 on the right hand side of equation 2.6. A forecast can be useful and have no skill and a forecast can be skillful and not be useful. When there

is a section or point of the Period 1 line that is above the perfect reliability line, that means there is underforecasting and when there is a section or point of the Period 1 line below the perfect reliability line, that means there is overforecasting. A perfect forecast is at the points 0,0 and 1,1.

### 2.4.2 ROC Diagram

Another way to verify forecasts and calculate different forecast evaluation scores or metrics is with the use of a 2x2 Contingency Table (Table 2.3). A hit means that a yes forecast was made and the event occurred. A false alarm means that a yes forecast was made but the event did not occur. A miss means that a no forecast was made but the event occurred. A null means that a no forecast was made and the event did not occur. There are 8 different scores or metrics that can be calculated using this table. Two of them are probability of detection (POD) and probability of false detection (POFD). These are defined below using the letters representing the different boxes from Table 2.3.

Table 2.3: 2x2 Contingency Table

| | | Observations | |
|---|---|---|---|
| | | Yes | No |
| Forecasts | Yes | a. hits | b. false alarms |
| | No | c. misses | d. nulls |

$$POD = \frac{a}{a+c} \tag{2.7}$$

$$POFD = \frac{b}{b+d} \tag{2.8}$$

POD and POFD vary between zero and one. The best possible score for POD is one while the best possible score for POFD is zero. POFD is sometimes also referred to as false alarm rate. A way to visually show a comparison between POD and POFD is a Relative Operating Curve (ROC) diagram (Mason 1982). An example ROC diagram is shown in Figure 2.3. This ROC diagram was created from the method developed by Mason (1982). First, POD and POFD were calculated using the technique of Mason (1982).

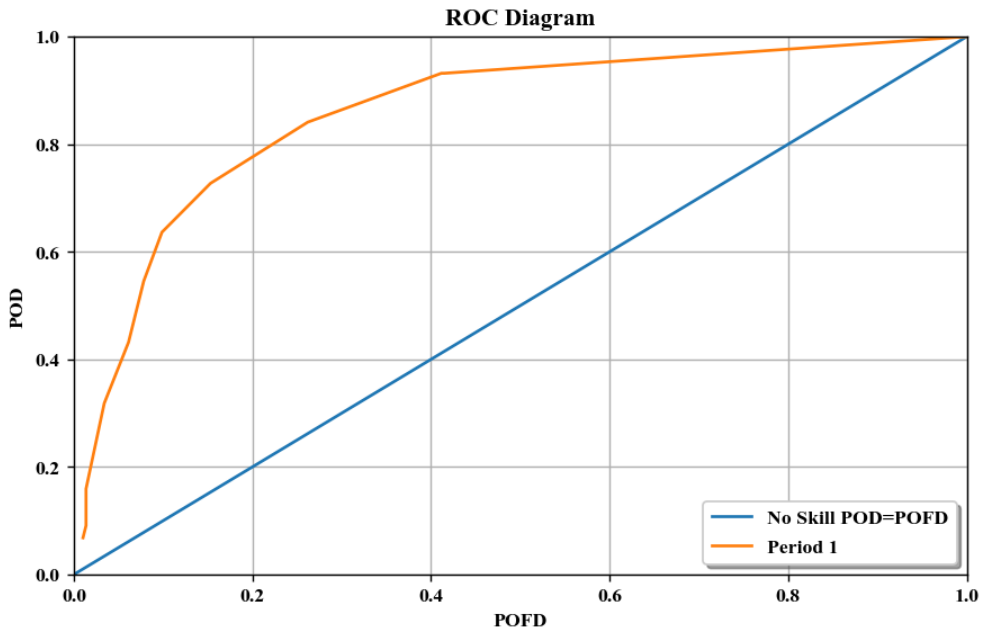



Figure 2.3: An example ROC diagram for precipitation events over the course of a year.

This was done by creating n-1 2x2 contingency tables where n is the number of forecast probabilities which for Figure 2.3 was 11. The POD and POFD were calculated for all tables assuming a yes forecast was every forecast greater than the given threshold for the 2x2 table. For example, for the 1st 2x2 contingency table every forecast greater than a probability of 0 was treated as a yes forecast. For the 2nd 2x2 table every forecast greater than a probability of 10 was treated as a yes forecast. Area Under the Curve (AUC) is calculated from ROC diagrams. An AUC greater than 0.7 means that a yes forecast can be determined. POD is on the vertical axis and POFD is on the horizontal axis. Additionally, a line at a 45° angle is plotted as the no skill line. This is where POD=POFD as shown in Figure 2.3.

### 2.4.3 Attributes and ROC Diagrams

Attributes and ROC diagrams were used to evaluate multiple event definitions used with the four kernels. Numerous studies have used attributes diagrams for verifying forecasts (e.g., Stumpf et al. 2015; Karstens et al. 2015; Cintineo et al. 2018). A few studies have used both attributes diagrams and ROC diagrams for verifying forecasts (e.g., Gallo et al. 2017; Lu et al. 2007).

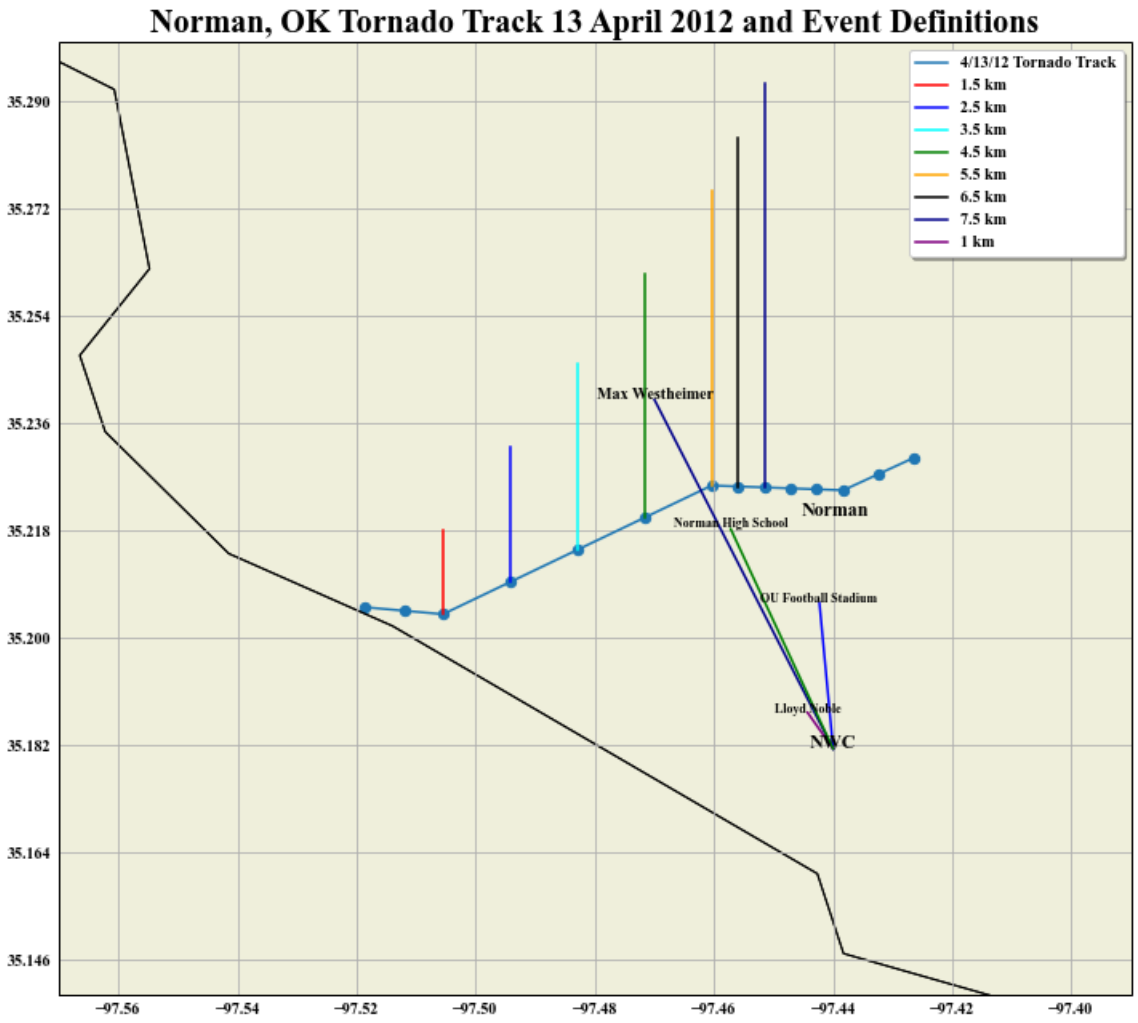



Figure 2.4: Map of Norman showing the 13 April 2012 tornado track and distances 0.5 to 7.5 km away from it. Additionally, a few different distances away from the National Weather Center are shown for perspective.

These diagrams were created using event definitions from 0.5 km to 7.5 km. The question that arises is how close does a person need to be to consider a lightning flash or a tornado an event. For the NCAA, when lightning is within 10 km to 13 km then play must be stopped. Based on surveys that have been completed, people felt that they experienced a tornado when it was about 8 km away from them (Krocak 2017). Based on the two different distances used for tornado and lightning events, for this study a 4.5 km and 7.5 km event definition were evaluated. A 4.5 km event definition represents the distance from the National Weather Center (NWC) to Norman High school (Figure 2.4). A 7.5 km event definition represents the distance between the NWC and Max Westheimer Airport (Figure 2.4).

Additionally, combined ROC and Attributes diagrams were created for all four kernels for both tornado and lightning to provide a summary of all activity in the HWT for a 4.5 km and 7.5 km event definition.

## 2.5 Practically Perfect Forecasts and Plumes

To determine how good the ROC and attributes diagrams can be expected to be, the concept of practically perfect forecasts (Hitchens et al. 2013) can be used. Additionally, Karstens et al. (2015) suggested that this concept could be used as a method to help verify the probabilistic PHI plumes. A practically perfect forecast is the best possible one that a forecaster could make if the observations were known ahead of time, using the same forecasting constraints and/or rules. Using practically perfect forecasts is especially useful when verifying forecasts of rare events. Rare events include severe weather hazards such as tornadoes, lightning, and hail. Practically perfect forecasts can determine the max (min) possible expected value for the resolution and reliability terms of the BS. They can also be used to determine the maximum possible expected value of AUC for a ROC diagram. There have been numerous studies and experiments involving creating practically perfect forecasts to assist in verifying forecasts of rare events. On example occurred in SFE 2015 where practically perfect hail forecasts were created by applying a two-dimensional Gaussian smoother to hail LSRs which were then verified using a 40 $km^2$ grid (Gallo et al. 2017). The participants in SFE 2015 evaluated hail forecasts created in SFE 2015 compared to those practically perfect forecasts (Gallo et al. 2017). Therefore, practically perfect forecasts provide a way to evaluate and verify forecasts of rare events.

Following the method of Hitchens et al. (2013), practically perfect plumes were created for each of the tornado and lightning cases. These were developed to create practically perfect attributes and ROC diagrams to compare to the forecaster attributes and ROC diagrams. The maximum possible improvement in the forecaster plumes could then be determined.

### 2.5.1 Tornado

For each of the tornado cases a circle of 4.5 km and 7.5 km radius around the mesocyclone coordinates at every two minutes from the start time of the case were used to generate a PHI object. The coordinates of the PHI object were saved to a file. The mesocyclone coordinates were then used to determine a distance and a speed at each two-minute time step for the duration of the case. The direction was calculated as the average of the three previous directions, provided a mesocyclone track consisted of at least 8 coordinates. Default speed uncertainty values, direction uncertainty values, duration, and probability values from the forecaster tornado PHI plumes were saved to the same file for each two-minute time step of each case. New "practically perfect" tornado plumes were created using this information in the same way the forecaster did in the HWT. These practically perfect plumes were compared to the forecaster plumes using multiple event definitions and kernels to better understand the best possible scenario for each event definition and kernel choice.

### 2.5.2 Lightning

For each of the lightning cases a circle of 4.5 km and 7.5 km radius around each of the NLDN coordinates were used to create a PHI object. Average speed, direction, speed uncertainty values, direction uncertainty values, and duration values from the forecaster lightning PHI plumes were saved to the same file for every time step of each case. The same probability values as the forecaster lighting PHI plumes were used. The modified PHI Tool code was run in the same way as the forecaster using a similar method to the tornado practically perfect plumes as described above. These practically perfect plumes were compared to the forecaster plumes using multiple event definitions and kernels to better understand the best possible scenario for each event definition and kernel choice.

Combined practically perfect ROC and attributes diagrams were created for all four kernels for both tornado and lightning for a 4.5 km and 7.5 km event definition to provide a summary of the maximum possible improvement of the plumes created in the HWT.

## 2.6 Summary

- Four tornado and for lightning cases from the 2017 HWT were used.
- CG locations and times from NLDN were used for lightning verification.
- Tornado paths were created using a combination of the Storm Events Database start and end times with the mesocyclone track from the 0.5° velocity from the closest WSR-88D radar.
- BS is a common way to verify probabilistic forecasts of a yes/no event.
- The BS is the mean squared error of the forecasts.
- The BS can be decomposed into three terms: reliability, resolution, and uncertainty.
- An attributes diagram visually shows two components (reliability and resolution) of the BS.
- There is a trade-off between having good resolution and reliability and having a skillful and useful forecast as shown on a attributes diagram.
- Another way to verify forecasts is with the ROC diagram which is a plot of POFD versus POD and shows the AUC.
- In addition to the Gaussian kernel, Epanechnikov, Quartic and Triangular kernels were applied to the forecaster created probabilities for all four tornado and lightning cases.
- Gridded verification was used to create attributes diagrams for all four tornado and lightning cases.
- Practically perfect forecasts can be created to determine how good ROC and attributes diagrams can be expected to be for comparison.

- Practically perfect plumes were created for all four tornado and lightning cases as a way to determine what the best possible attributes and ROC diagrams could be given forecaster constraints.
- Combined attributes and ROC diagrams were created for both tornado and lightning cases and for the practically perfect plumes.

# Chapter 3

# Case Description

During the 2017 HWT PHI experiment, forecasters issued PHI plumes for a variety of real-time events and archived events set in displaced-real-time. The goal of these cases was to test the feasibility of the PHI concept in NWS operations (Karstens et al. 2018). For the tornado events used in this study, two of the cases were used in a displaced real-time setting in the HWT (25-26 May 2016, Bennington, KS and 24 May 2016, Dodge City, KS) and two were handled by forecasters in real-time as the event unfolded (8 May 2017, Denver, CO and 25 May 2017, Columbia, SC). For the lightning events, three of these events were archived cases set in displaced real-time (25-26 May 2016, Bennington, KS, overlapping the tornado event, 22 July 2016, Grand Junction, CO, and 1 September 2016 Melbourne, FL) and one occurred real-time during the experiment (25 May 2017, Goodland, KS). The synoptic and near-storm environments of each of these cases as well as the storm evolution and tornado or lightning activity are described below to provide a general understanding of each case chosen for this study.

## 3.1 Tornado

### 3.1.1 Bennington, KS

The 2000 UTC 25 May 2016 SPC Convective Outlook (Categorical, Tornado, Hail, and Wind) showed (Figure 3.1) that there was a slight categorical risk, very slight risk of tornadoes (2%), and slightly larger risk of severe hail and wind (15%) over the domain of the case. Additionally, the SPC issued a tornado watch for the domain of the case that lasted from 2045 UTC 25 May 2016 - 0200 UTC 26 May 2016. The SPC issued a Mesoscale Discussion (MD) at 1947 UTC 25 May 2016 which showed a dryline to the west and a stationary/cold front to the north of the domain of the case.

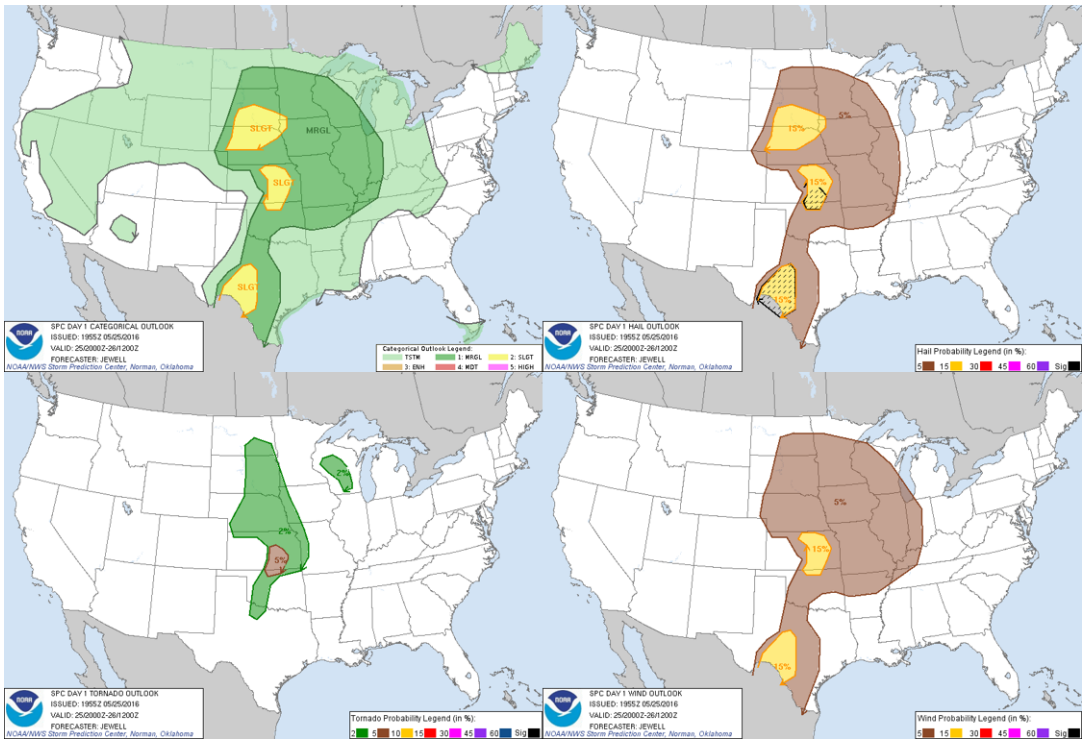
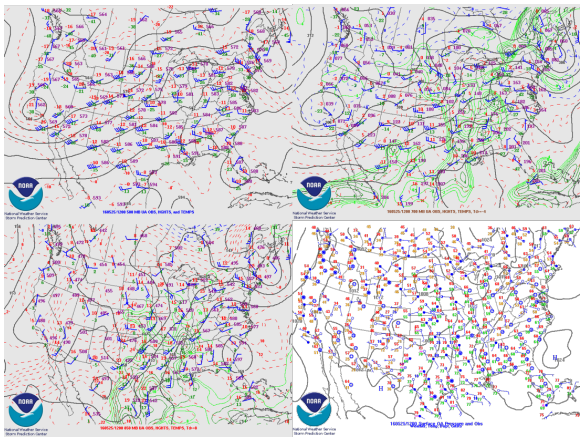
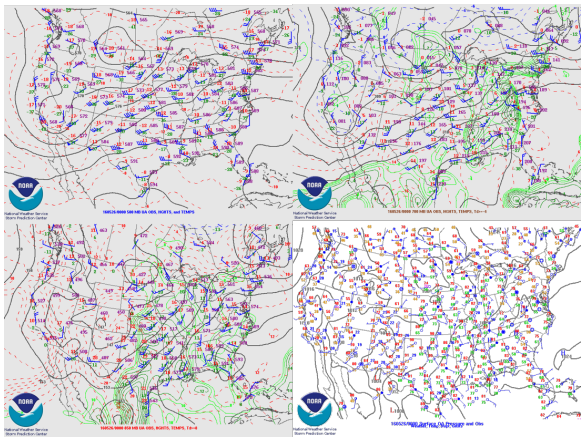



Figure 3.1: The Categorical Outlook (Top Left), Hail Outlook (Top Right), Tornado Outlook (Bottom Left), and Wind Outlook (Bottom Right) for 2000 UTC 25 May 2016.

A second MD was issued at 0005 UTC 26 May 2016 which showed a dryline to the southwest, a stationary front to the north, and a low to the west of the the domain of the case.



(a) Upper Level and Surface Maps for 1200 UTC 25 May 2015

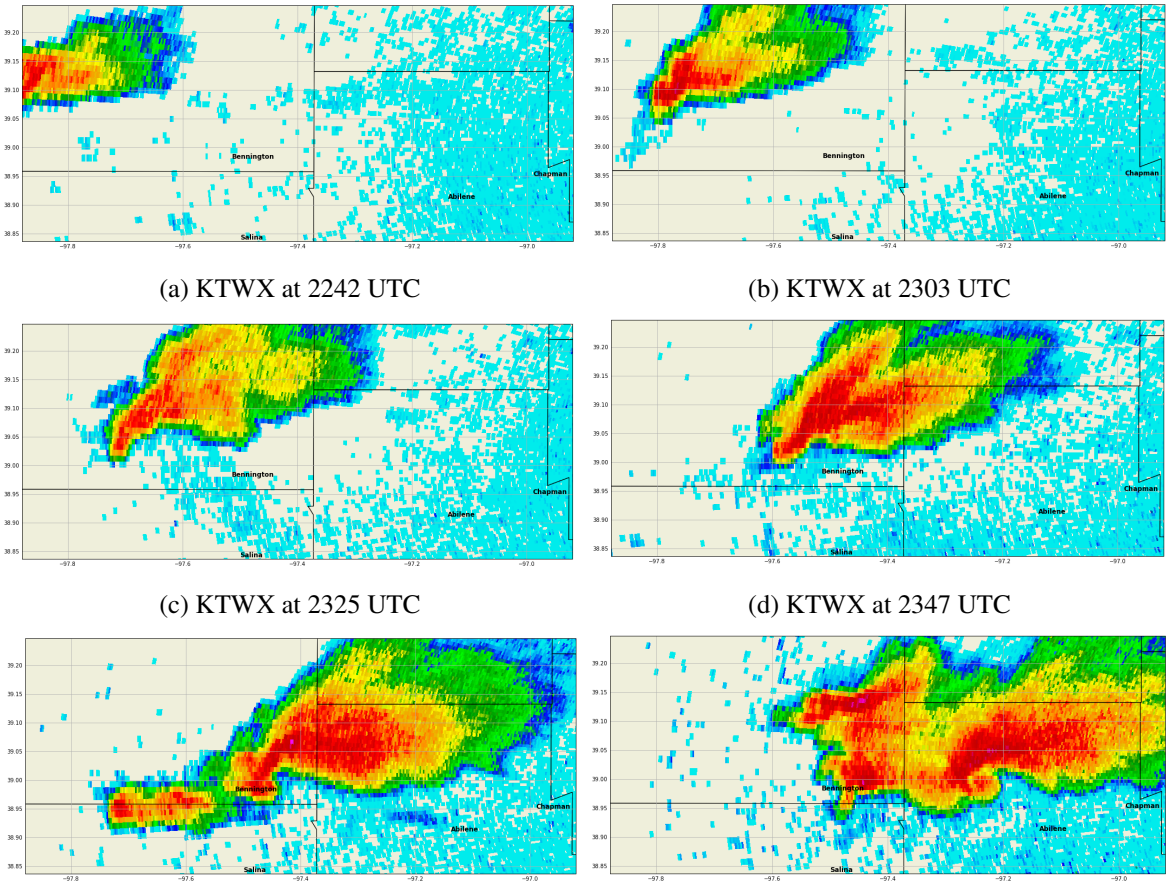


(b) Upper Level and Surface Maps for 0000 UTC 26 May 2016

Figure 3.2: The 500 mb (Top Left), 750 mb (Top Right), 800 mb (Bottom Left), and Surface map (Bottom Right) for 1200 UTC 25 May 2016 (a) and for 0000 UTC 26 May 2016 (b).

The 500 mb map at 0000 UTC 26 May 2016 (Figure 3.2 (b)) showed that there was an upper level low over southern California and a ridge over Kansas. Additionally, it showed that there was slight positive vorticity advection (PVA) in western Kansas. Upper level and surface maps at 1200 UTC 25 May 2016 (Figure 3.2 (a)) and 0000 UTC 26 May 2016 indicated that the there was essentially no veering with height and weak upper level winds.

From 2300 UTC 25 May 2016 to 0100 UTC 26 May 2016 there were two tornadoes that occurred from one storm that tracked near Bennington, KS.




(a) KTWX at 2242 UTC

(b) KTWX at 2303 UTC

(c) KTWX at 2325 UTC

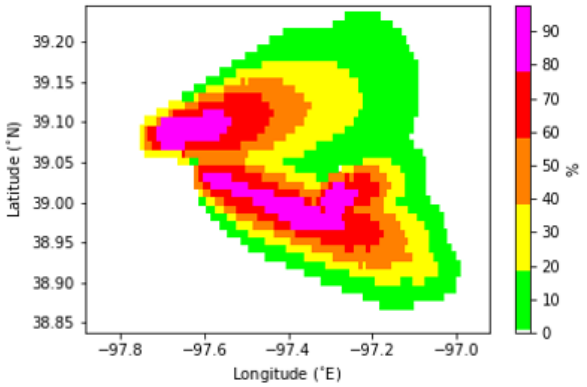
(d) KTWX at 2347 UTC

(e) KTWX at 0010 UTC

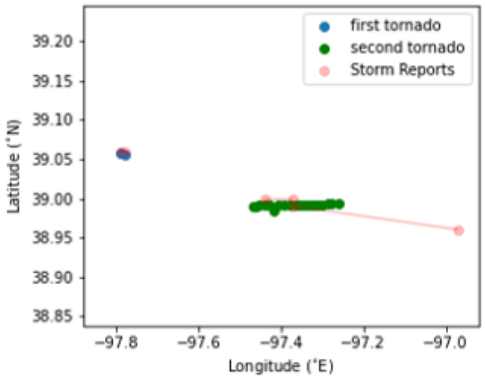
(f) KTWX at 0046 UTC

Figure 3.3: The radar images from the KTWX radar for the Bennington, KS case at 2242 UTC 25 May 2016 (a), at 2303 UTC 25 May 2016 (b), at 2325 UTC 25 May 2016 (c), at 2347 UTC 25 May 2016 (d), at 0010 UTC 26 May 2016 (e), and at 0046 UTC 26 May 2016 (f).

This storm formed in Ottawa County, Kansas. The first tornado touched down at 2308 UTC 25 May 2016 approximately 5 km south of Minneapolis, KS and 10 km northwest of Bennington, KS and lasted for one minute. (National Weather Service ndc). This tornado was rated as an EF-0. The second tornado touched down north of Niles, KS at 0007 UTC 26 May 2016. This tornado lasted for approximately 1 hour and 33 minutes and was rated as an EF-4. This tornado moved east southeast and passed about 3 km north of Abilene, KS. This tornado then moved northeast of Abilene, KS and weakened slightly. It moved southeast as it was approaching Interstate 70 and intensified. Six time steps of the KTWX radar for this case were shown (Figure 3.3).




(a) Forecaster Created Plume




(b) Mesocyclone Track with Storm Reports

Figure 3.4: The merged maximum forecasted probability at any time step over the duration of the case (a) and the storm reports from the Storm Events Database plotted on top of the radar derived mesocyclone tracks (b) for the Bennington, KS case 25-26 May 2016.

The storm reports from the Storm Events Database were plotted on top of the radar derived mesocyclone tracks for the Bennington, KS tornado case (Figure 3.4 b). The second mesocyclone track corresponded to two tornado reports from Storm Data. The second tornado of those tornadoes lasted longer than the duration of the Bennington, KS tornado case in the HWT. Additionally, the first storm report lined up well with the first tornado, while the second storm report was east of where the second tornado was found to be using radar.

The forecaster merged maximum plume at any time step over the duration

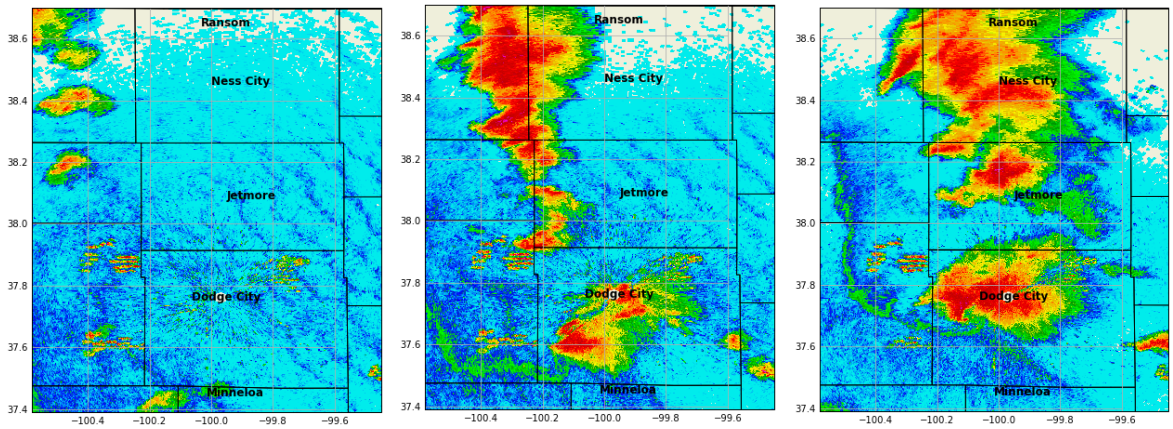
of the Bennington, KS tornado case is shown in Figure 3.4 (a). The forecaster created three PHI objects for this case. It appears that the forecaster thought that the supercell would split into two storms with both producing a tornado. At this time a southern storm moving north collided with the Bennington, KS storm. The forecaster did a fairly good job when making the plume compared to the mesocyclone track (Figure 3.4 (b)). The forecaster missed the first tornado but captured where the second tornado was fairly well. The forecaster was off slightly on the storm's direction.

### 3.1.2 Dodge City, KS

The 2000 UTC 24 May 2016 SPC Convective Outlook showed that there was an enhanced categorical risk, enhanced risk of strong tornadoes (10%), enhanced risk of large severe hail (30%), and a slight risk of wind (15%) over the southern portion of the domain of the case. Additionally, the SPC issued a tornado watch for the domain of the case. The SPC issued a MD at 1931 UTC 24 May 2016 which showed a surface low at the Kansas - Colorado border west of the domain of the case. Additionally, there was a dryline extending south from the low. This MD also showed an outflow boundary going through the center of the domain of the case and a stationary/cold front to the north of the domain of the case. The SPC issued a MD at 2314 UTC 24 May 2016 which showed a dryline and an outflow boundary intersecting in the center of the domain of the case.

The 500 mb map at 1200 UTC 24 May 2016 showed that there was an upper level low off the coast of southern California and a ridge over Kansas. Additionally, it showed that there was slight PVA in western Kansas. Upper level and surface maps at 1200 UTC 24 May 2016 indicated that the wind profile had weak upper level winds and some veering with height. Additionally, there was a surface low near the southwestern corner of Kansas and southeastern corner of Colorado at 1200 UTC 24 May 2016.

On 24 May 2016 near and north of Dodge City, KS a series of supercells formed during the late afternoon and early evening and produced more than 10 tornadoes.

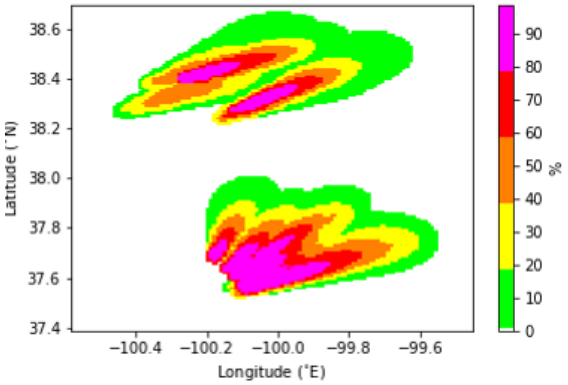



(a) KDDC at 2200 UTC (b) KDDC at 2249 UTC (c) KDDC at 2235 UTC

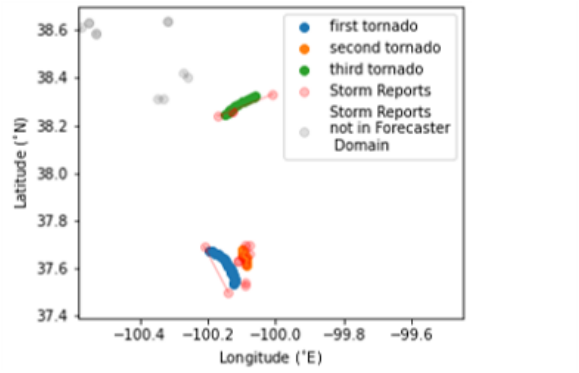
Figure 3.5: The radar images from the KDDC radar for the Dodge City, KS case at 2200 UTC (a), at 2249 UTC (b), and at 2235 UTC (c) 24 May 2016.

There was one supercell that produced numerous cyclic tornadoes and had a mesocyclone that kept reforming. The first tornado from this supercell began a few km north of Minneola, KS and the last tornado from this supercell dissipated ten km southwest of Jetmore, KS (National Weather Service ndb). Tornadoes also occurred in Hodge County, KS, Ness County, KS, and Edwards County, KS. Three time steps of the KDDC radar for this case were shown (Figure 3.5).

The storm reports from Storm Data were plotted on top of the radar derived mesocyclone tracks for the Dodge City, KS case (Figure 3.6 (b)). The southern tornadoes were challenging to track because the mesocyclone kept wrapping up and reforming. The third tornado track lined up fairly well with the storm reports. The forecaster merged maximum plume at any time step over the duration of the Dodge City, KS case is shown in Figure 3.6 (a). In addition to the north moving supercell that formed south of and moved over Dodge City, KS there were a few storms north of Dodge City, KS up through the northern part of the domain that moved east and one of the storms produced a tornado. Due to having storms moving different directions in the domain of this case and the mesocyclone reforming in a supercell, this case was challenging for the forecaster.




(a) Forecaster Created Plume




(b) Mesocyclone Track with Storm Reports

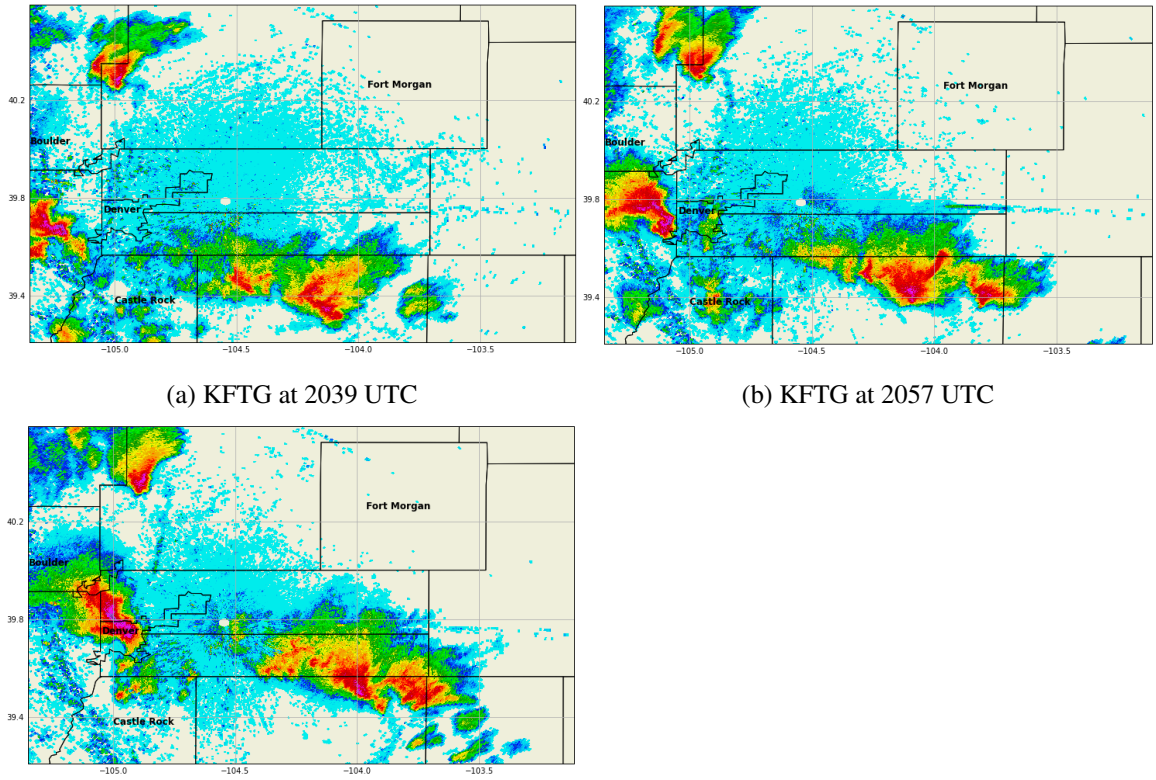
Figure 3.6: The merged maximum forecasted probability at any time step over the duration of the case (a) and the storm reports from the Storm Events Database plotted on top of the radar derived mesocyclone tracks (b) for the Dodge City, KS case 24 May 2016. The storm reports in gray are storm reports that were outside of the forecaster domain.

However, the forecaster still captured the overarching trends of the location of tornadoes based on the mesocyclone track (Figure 3.6).

### 3.1.3 Denver, CO

The 2000 UTC 8 May 2017 SPC Convective Outlook showed that there was a slight categorical risk, very slight risk of tornadoes (2%), and slightly larger risk of severe hail and wind (15%) over the domain of the case. Additionally, the SPC issued a severe thunderstorm watch for the domain of the case. The SPC issued a MD at 1801 UTC 8 May 2017 which showed a cool front/wind shift south of the domain of the case.

The 500 mb map at 1200 UTC 8 May 2017 showed that there was an upper level low off the coast of northern Baja California and a ridge over eastern Colorado. Additionally, it showed that there was PVA in eastern Colorado. Upper level and surface maps at 1200 UTC 8 May 2017 indicated that the wind profile had no veering with height and weak upper level winds.



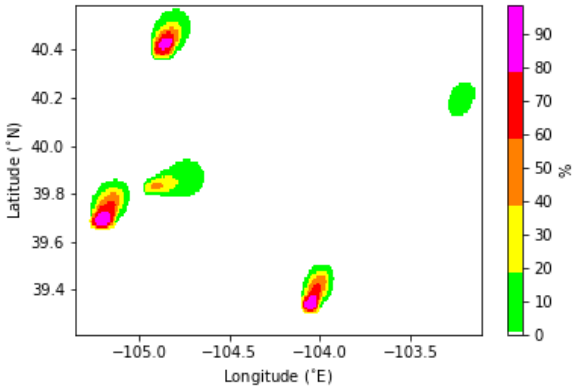

(a) KFTG at 2039 UTC

(b) KFTG at 2057 UTC

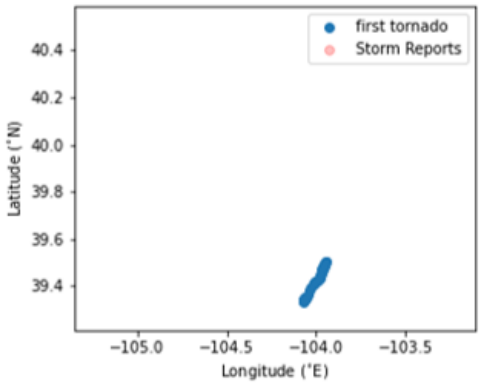
(c) KFTG at 2120 UTC

Figure 3.7: The radar images from the KFTG radar for the Denver, CO case at 2039 UTC (a), at 2057 UTC (b), and at 2120 UTC (c) 8 May 2017.

For the Denver, CO case there was a tornadic warned storm which formed in Elbert County, CO about 55 km east of Castle Rock, CO at 2020 UTC 8 May 2017. This storm intensified and moved northwest over the course of the next 20 minutes. Then at 2050 UTC 8 May 2017 a tornado warning was issued for this storm. At this time the storm was moving north. This storm continued to move north and began to slightly weaken. The warning was cancelled at 2118 UTC 8 May 2017. Three time steps of the KFTG radar for this case were shown (Figure 3.7).




(a) Forecaster Created Plume




(b) Mesocyclone Track with Storm Reports

Figure 3.8: The merged maximum forecasted probability at any time step over the duration of the case (a) and the storm reports from the Storm Events Database plotted on top of the radar derived mesocyclone track (b) for the Denver, CO case 8 May 2017.

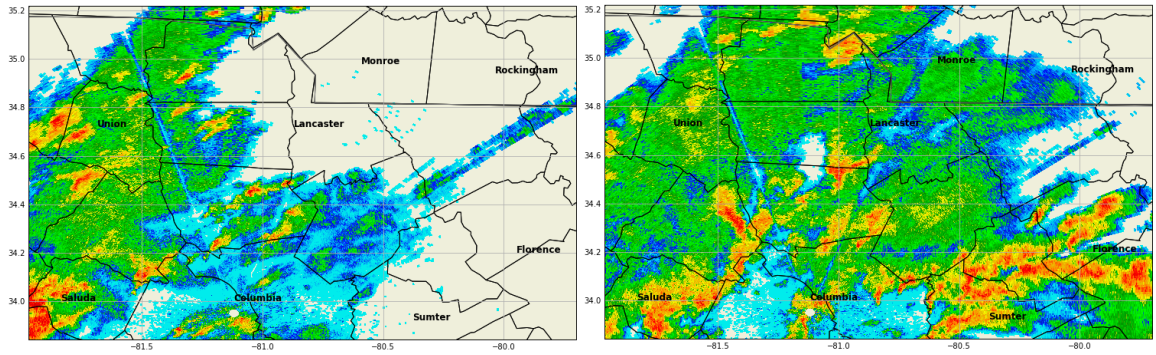
The storm reports from Storm Data were plotted on top of the radar derived mesocyclone track for the Denver, CO case (Figure 3.8 (b)). Due to there only being a tornado warning for this case there were no Storm Reports in the temporal and spatial domain of the case. The forecaster merged maximum plume at any time step over the duration of the Denver, CO case is shown in Figure 3.8 (a). Due to the storms forming near and in the mountains and a lack of good radar coverage, this case was challenging for the forecaster. However, the forecaster still correctly forecasted the location (based on the tornado warning) of one of the five forecasted tornadoes based on the mesocyclone track (Figure 3.8 (b)).

### 3.1.4 Columbia, SC

The 1630 UTC 24 May 2017 SPC Convective Outlook showed that there was an enhanced categorical risk, enhanced risk of tornadoes (10%), and enhanced risk of severe hail (30%), and slight risk of severe wind (15%) over the domain of the case. The SPC issued three tornado watches for different portions of the domain of this case. The SPC issued a MD at 1229 UTC 24 May 2017 which showed that the dewpoints in the domain were in the mid-70s. The SPC issued a MD at 1810 UTC 24 May 2017 which showed that the winds were backed from the southeast.

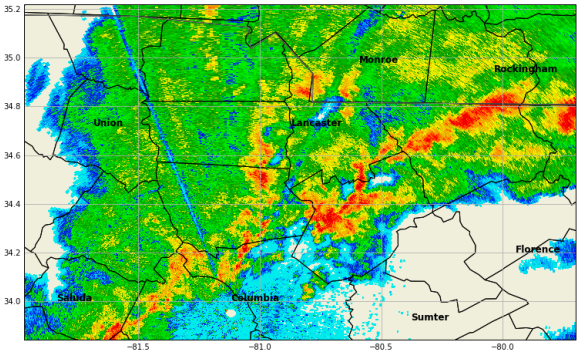
The 500 mb map at 1200 UTC 24 May 2017 showed that there was an upper level low over Illinois and a ridge over South Carolina. Additionally, it showed that there was PVA over South Carolina. Upper level and surface maps at 1200 UTC 24 May 2017 indicated that the wind profile had veering with height and strong upper level winds. At the surface at 1200 UTC 24 May 2017 there was a low in central Tennessee. This low tracked east throughout the morning. There was an occluded front that extended south from this low through central Alabama. Extending east from the end of the occluded front through central South Carolina there was a warm front. Extending southwest from the occluded front there was a cold front (National Weather Service nda).

On 24 May 2017, a Quasi-Linear Convective System (QLCS) formed over east central Georgia. This storm produced three tornadoes and there was an additional tornado warning for a portion of this storm that did not produce a tornado. The first tornado touched down near Fruit Hill road just north of US Highway 378 approximately 10 km west of Saluda, SC (National Weather Service nda). It continued northeast crossing Henley Road and Old Charleston Road. This tornado traveled for five km and dissipated at approximately the intersection of Old Chappell Ferry Road and Simmons Road six km northwest of Saluda, SC. The second tornado touched down near Denny Highway in Northern Saluda County, SC 16 km northeast of Saluda, SC. The tornado moved northeast across the Saluda River into Newberry County, SC.




(a) KCAE at 1836 UTC

(b) KCAE at 1930 UTC

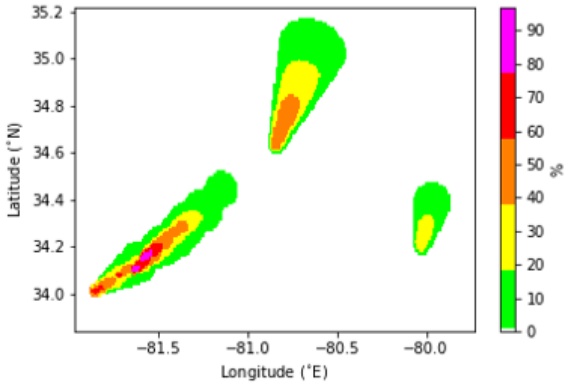



(c) KCAE at 2025 UTC

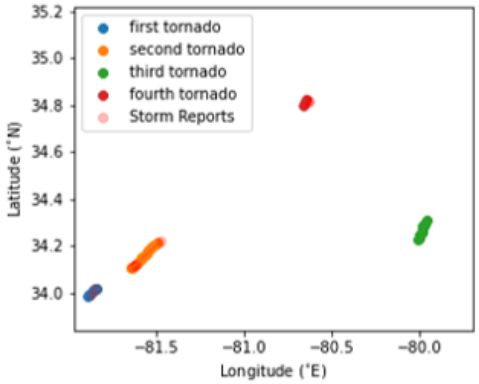
Figure 3.9: The radar images from the KCAE radar for the Columbia, SC case at 1836 UTC (a), at 1930 UTC (b), and at 2025 UTC (c) 24 May 2017.

It eventually passed about three km south of Prosperity, SC and dissipated about five km east of Prosperity, SC (National Weather Service nda). This tornado had a track of about 19 km. The third tornado touched down northeast of Lancaster, SC and close to the border with North Carolina. This tornado continued northward and moved across the state line and tracked northeast for around 11 km. A tornado warning was issued for a portion of the storm 22 km west of Florence, SC. This portion of the storm moved northeast and had a tornado warning associated with it for about 15 minutes. Three time steps of the KCAE radar for this case were shown (Figure 3.9).

The storm reports from Storm Data were plotted on top of the radar derived mesocyclone tracks for the Columbia, SC case (Figure 3.10). The mesocyclone tracks for tornadoes one, two, and four lined up well with the storm reports.




(a) Forecaster Created Plume




(b) Mesocyclone Track with Storm Reports

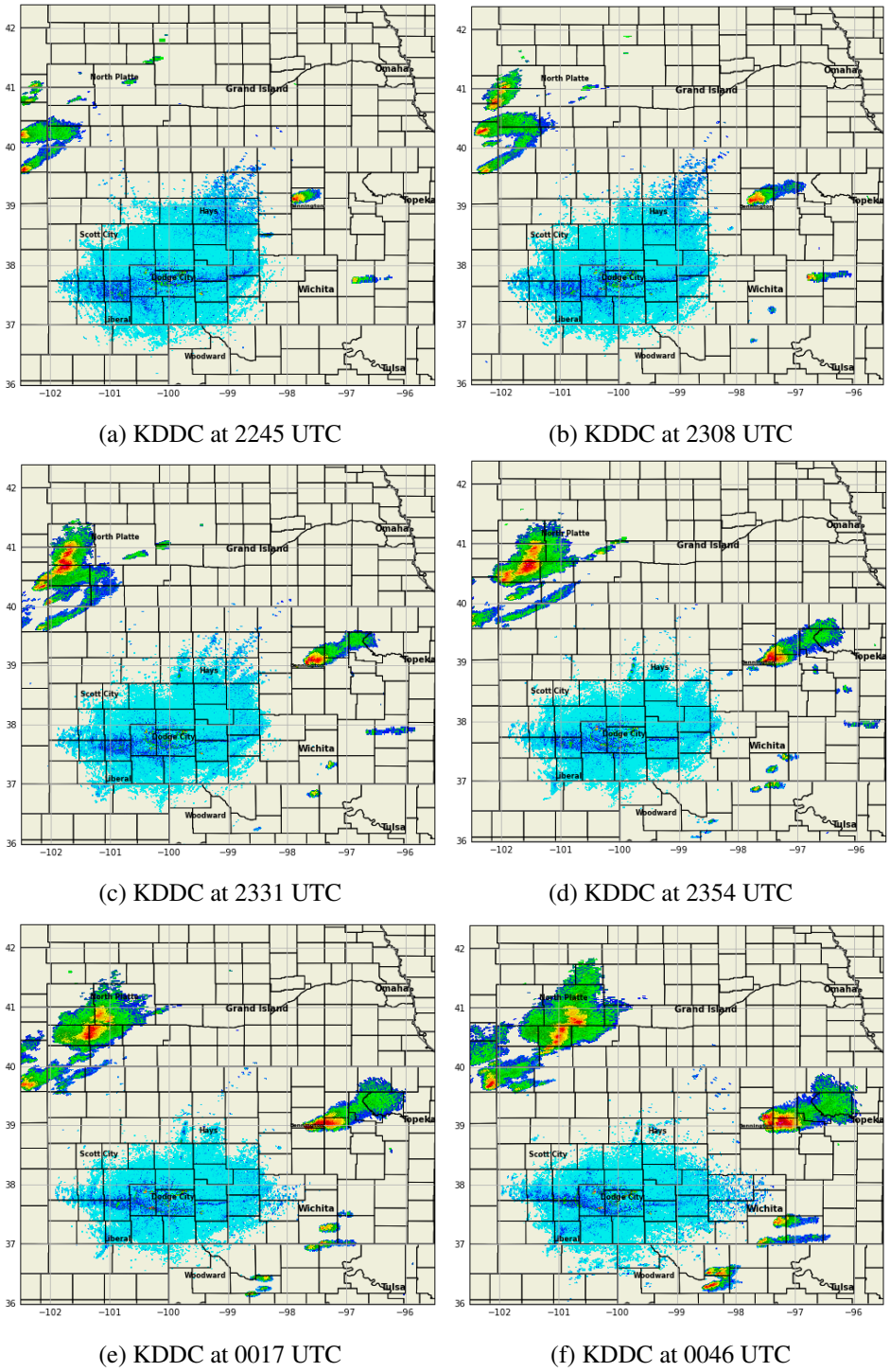
Figure 3.10: The merged maximum forecasted probability at any time step over the duration of the case (a) and the storm reports from the Storm Events Database plotted on top of the radar derived mesocyclone track (b) for the Columbia, SC case 24 May 2017.

The third tornado was from a warning so there was not an associated storm report for it. The forecaster merged maximum plume at any time step over the duration of the Columbia, SC case is shown in Figure 3.10. Forecasting for QLCS tornadoes is challenging. The forecaster created four PHI objects and the associated plumes corresponded to a mesocyclone track (Figure 3.10 (b)). The northern plume did not correspond to the mesocyclone track as well as the other plumes. The forecaster did an excellent job of forecasting for this case. However, the forecaster was not overly confident in the plumes that were created as shown by the lack of high probability values in the three plumes (Figure 3.10 (a)).

## 3.2 Lightning

There were many differences between the tornado and lightning cases. The primary difference was that the spatial domain for the lightning cases was much larger than for the tornado cases. Additionally, as opposed to a line for the tornado verification (mesocyclone tracks) the lightning verification (NLDN) was a cluster and/or a scattering of points.

### 3.2.1 Bennington, KS




(a) KDDC at 2245 UTC

(b) KDDC at 2308 UTC

(c) KDDC at 2331 UTC

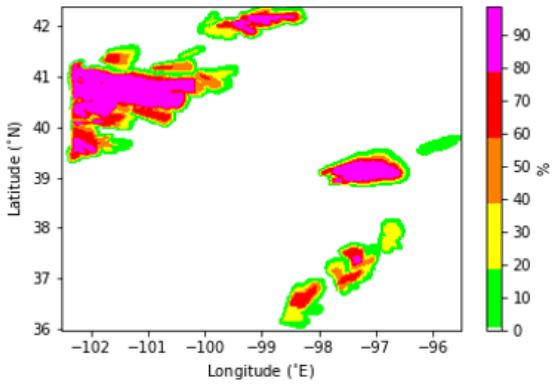
(d) KDDC at 2354 UTC

(e) KDDC at 0017 UTC

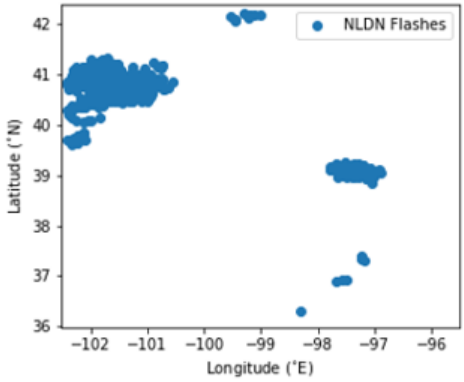
(f) KDDC at 0046 UTC

Figure 3.11: The radar images from the KDDC radar for the Bennington, KS lightning case at 2245 UTC 25 May 2016 (a), at 2308 UTC 25 May 2016 (b), at 2331 UTC 25 May 2016 (c), at 2354 UTC 25 May 2016 (d), at 0017 UTC 26 May 2016 (e), and at 0046 UTC 26 May 2016 (f).

The SPC Convective Outlook (Figure 3.1), watches, and MDs were discussed and described previously in subsection 3.1.1. The synoptic setup for this case was described previously in subsection 3.1.1, using Figure 3.2. Note that the lightning forecaster created lightning PHI objects over the same region with a larger domain simultaneously to the tornado activity.




(a) Forecaster Created Plume




(b) NLDN Flashes

Figure 3.12: The merged maximum forecasted probability at any time step over the duration of the case (a) and NLDN flashes (b) for the Bennington, KS lightning case 25-26 May 2016.

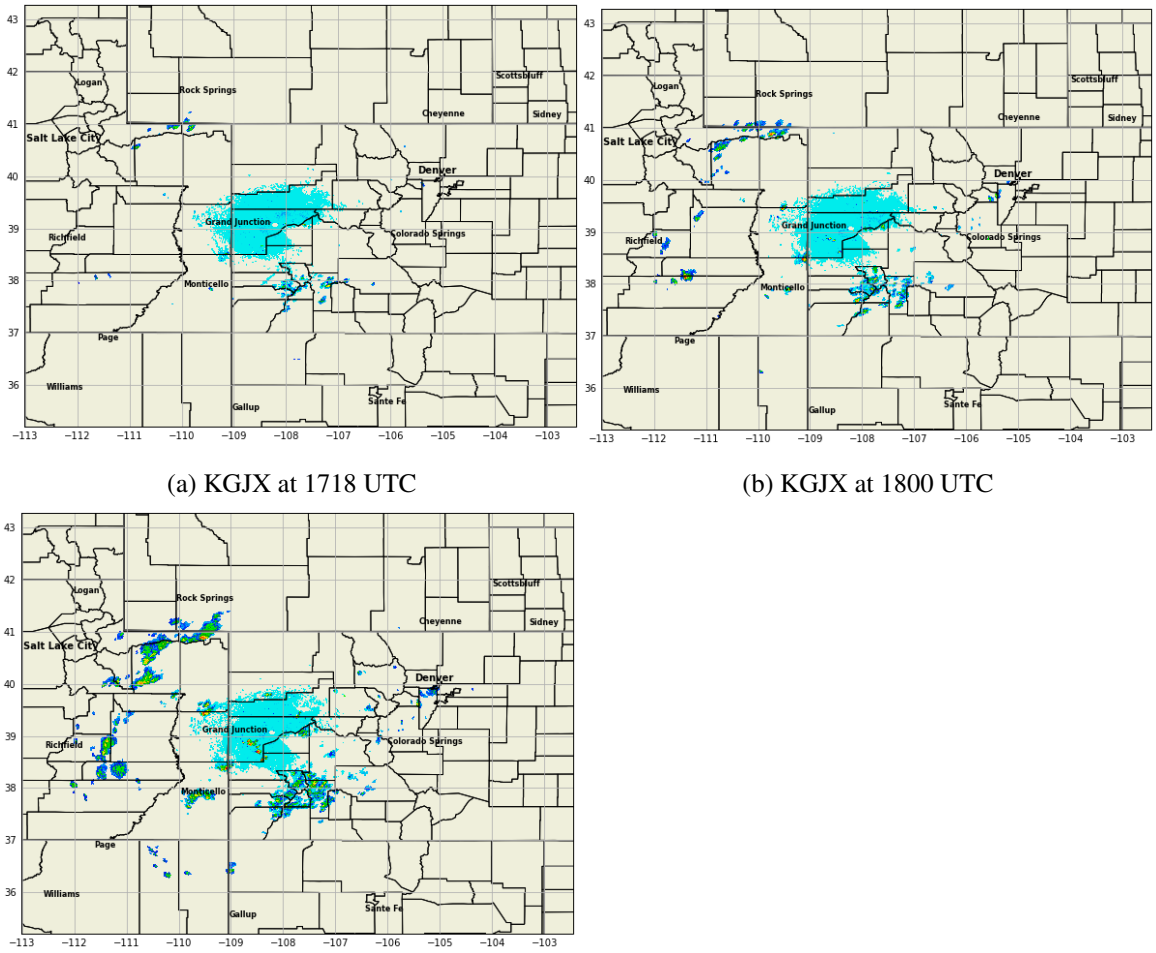
For the Bennington, KS lightning case, in the west-central portion of the domain of the case there was a supercell thunderstorm that formed as was described previously in subsection 3.1.1. Additionally, there was a line of storms that developed south of the supercell in southern Kansas and Oklahoma that moved northeast. There were also a few storms that originated in northeastern Colorado and moved into the southwestern corner of Nebraska. These storms moved east. Six time steps of the KDDC radar for this case were shown (Figure 3.11).

The forecaster merged maximum plume at any time step over the duration of the Bennington, KS lightning case was shown in Figure 3.12 (a). In the location of the path of the supercell the lightning plume had primarily one probability value compared to the many different probability values for the tornado hazard for this

case. Based on the NLDN Flashes (Figure 3.12 (b)) the forecaster did a fairly good job of making this forecast. However, there were a few plumes that did not verify.

### 3.2.2 Grand Junction, CO

The 1630 UTC 22 July 2016 SPC Convective Outlook showed that there was a marginal categorical risk, no risk of tornadoes, and a marginal risk of severe hail and wind (5%) over the extreme northeast-central portion of the domain of the case. The SPC issued no watches and no MDs prior to or during the duration of the case for the domain of this case.



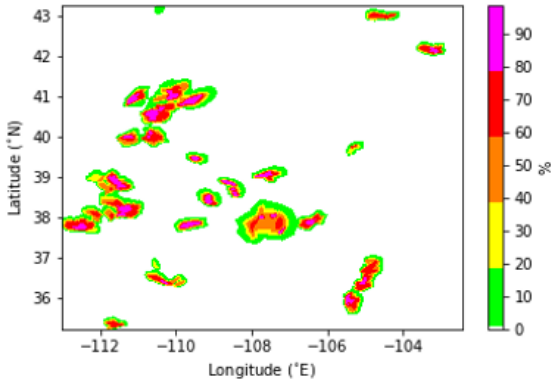

(a) KGJX at 1718 UTC

(b) KGJX at 1800 UTC

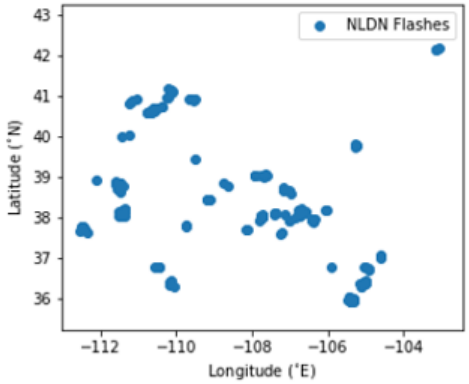
(c) KGJX at 1855 UTC

Figure 3.13: The radar images from the KGJX radar for the Grand Junction, CO case at 1718 UTC (a), at 1800 UTC (b), and at 1855 UTC (c) 22 July 2016

The 500 mb map at 1200 UTC 22 July 2016 showed that there was an upper level low over eastern Washington and a ridge over Colorado. Upper level and surface maps at 1200 UTC 22 July 2016 indicated that the wind profile had no veering with height and weak upper level winds. Additionally, there was a surface high in northwestern Colorado.




(a) Forecaster Created Plume




(b) NLDN Flashes

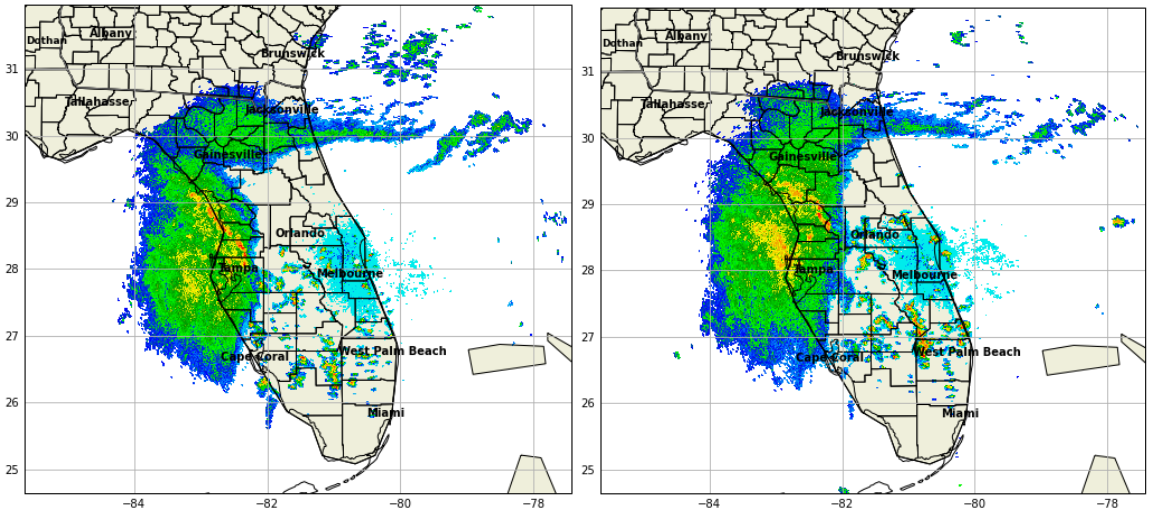
Figure 3.14: The merged maximum forecasted probability at any time step over the duration of the case (a) and NLDN flashes (b) for the Grand Junction, CO case 22 July 2016.

There were many slow moving, short duration storms that formed around Grand Junction, CO in the mountains over the course of the early afternoon on 22 July 2016. None of these storms were tornadic. Additionally, almost all of these storms were not severe. Three time steps of the KGJX radar for this case were shown (Figure 3.13).

The forecaster merged maximum plume at any time step over the duration of the Grand Junction, CO case was shown in Figure 3.14 (a). Due to the short duration nature of the storms in this case their were a lot of small plumes. Additionally, due to this case being in the mountains the radar coverage was poor and caused this case to be challenging to forecast. However, based on the NLDN flashes (Figure 3.14 (b)) the forecaster did a fairly good job of making this challenging forecast.

### 3.2.3 Melbourne, FL

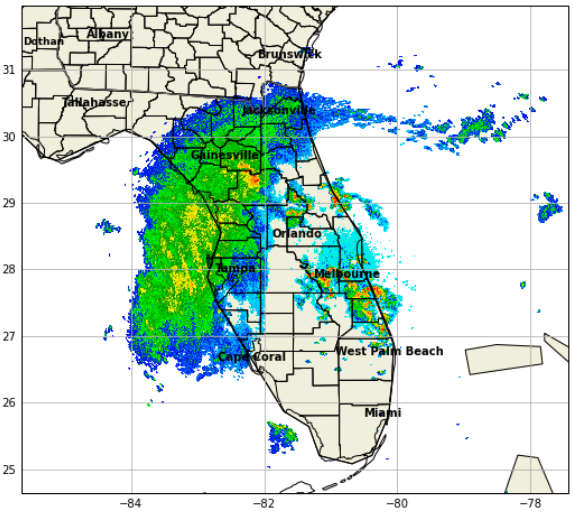
The 1630 UTC 1 September 2016 SPC Convective Outlook showed that there was a slight categorical risk, marginal risk of wind (5%), and a slight risk of tornadoes (5%) over the north-central portion of the domain of the case. There was no risk for severe hail over the whole domain of the case. There was also a marginal categorical risk and very slight risk of tornadoes (2%) in a portion of the domain of the case as well. Additionally, the SPC issued a tornado watch for the east-central portion of the domain of the case.




(a) KMLB at 1755 UTC

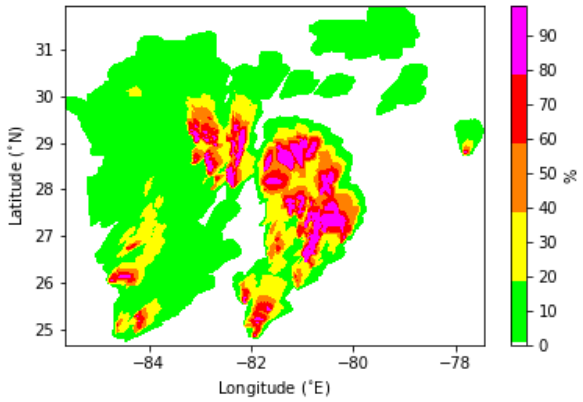



(b) KMLB at 1852 UTC

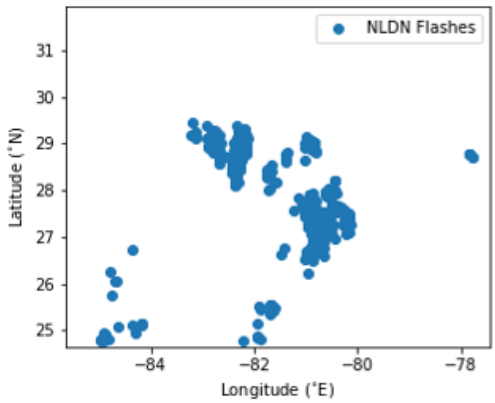



(c) KMLB at 1954 UTC

Figure 3.15: The radar images from the KMLB radar for the Melbourne, FL case at 1755 UTC (a), at 1852 UTC (b), and at 1954 UTC (c) 1 September 2016.




(a) Forecaster Created Plume




(b) NLDN Flashes

Figure 3.16: The merged maximum forecasted probability at any time step over the duration of the case (a) and NLDN flashes (b) for the Melbourne, FL case 1 September 2016.

The SPC issued a MD at 1549 UTC 1 September 2016 which showed that temperatures were in the low-80s. The SPC issued a MD at 1812 UTC 1 September 2016 which showed Hurricane Hermine off the coast of Florida slightly east of the central edge of the domain.

The Melbourne, FL case involves Tropical Storm Hermine which strengthened to become Hurricane Hermine during the case. Hurricane Hermine was a category one Hurricane that made landfall along the Big Bend coast of Florida just east of St. Marks. A tropical disturbance moved off the coast of Africa on 17 August 2016 that caused the formation of Hermine (Berg 2017). On 28 August 2016 the system developed a well-defined center. Later on this day the tropical depression formed approximately 80 km south southeast of the Florida Keys. The depression moved west across the southeastern Gulf of Mexico until 30 August 2016 (Berg 2017). On 31 August 2016 the depression began to move north northeast and strengthened into a tropical storm.

The tropical storm then continued to move north northeast and on 1 September 2016 at 1800 UTC it became a hurricane. At this time and through the duration of the case the outer rain bands of the hurricane were on the western coast of Florida. There were many

isolated storms to the east of the outer rain bands over much of southern Florida. These storms moved north during the course of the case. A few of the storms became severe. Three time steps of the KMLB radar for this case were shown (Figure 3.15).

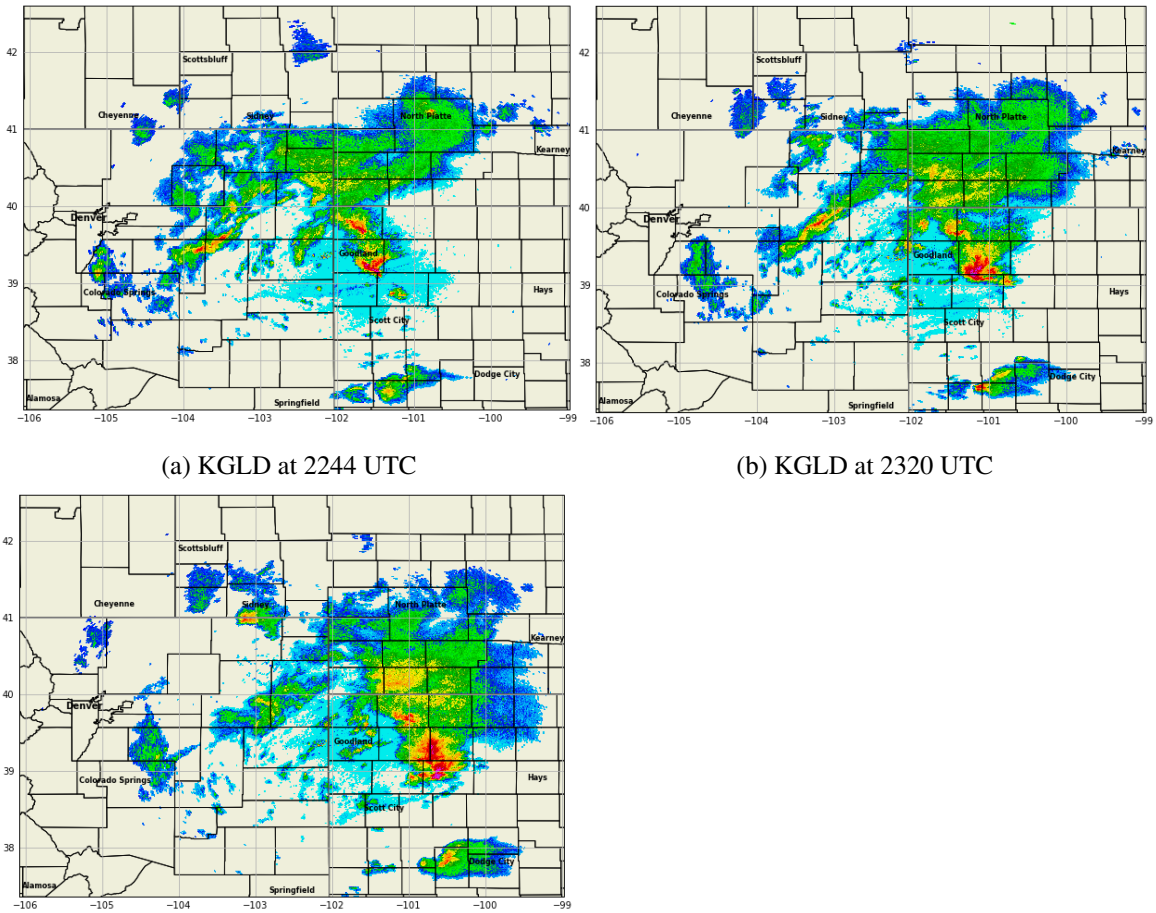
The forecaster merged maximum plume at any time step of the duration of the Melbourne, FL case was shown in Figure 3.16 (a). The hurricane caused there to be some challenges for the forecaster for this case. However, based on the NLDN flashes (Figure 3.16 (b)) the forecaster did a fairly good job of making this challenging forecast. The higher probabilities correlated very well with the location of the NLDN flashes.

### 3.2.4 Goodland, KS

The 2000 UTC 25 May 2017 SPC Convective Outlook showed that there was an enhanced categorical risk, an enhanced risk of wind (30%), a slight risk of severe hail (15%), and a marginal risk for tornadoes (2%) over the north-central portion of the domain of the case. There was a marginal and slight categorical risk and a marginal (5%) and slight (15%) risk of wind over a portion of the domain. Also, there was a marginal risk of severe hail (5%) over a portion of the domain. Additionally, the SPC issued a severe thunderstorm watch for the north-central portion of the domain. The SPC issued a MD at 1835 UTC 25 May 2017 which showed dewpoints ranging from the mid 40s to lower 50s.

The 500 mb map at 1200 UTC 25 May 2017 showed that there was an upper level low over the south-central portion of the Saskatchewan Province of Canada (north of Montana) and a ridge over Kansas and Nebraska. Additionally, it showed that there was slight PVA in western Nebraska. Upper level and surface maps at 1200 UTC 25 May 2017 indicated that the wind profile had no veering with height and weak upper level winds. Additionally, there was a surface low near where the borders of Nebraska, Wyoming, and Colorado meet.

For the Goodland, KS case, two severe thunderstorms formed in the northeastern corner of Colorado and moved east into Kansas at 2245 UTC 25 May 2017. The southern storm was 18.59 km southeast of Goodland, KS.




(a) KGLD at 2244 UTC

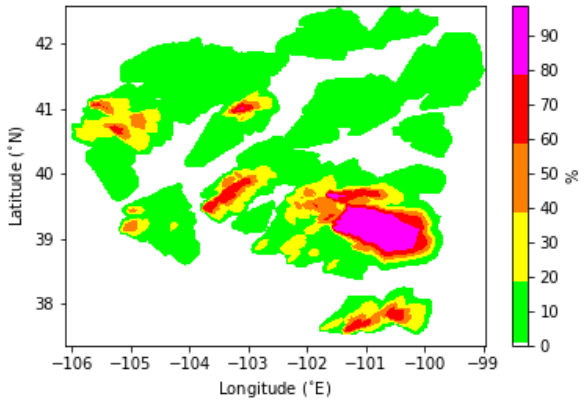
(b) KGLD at 2320 UTC

(c) KGLD at 2357 UTC

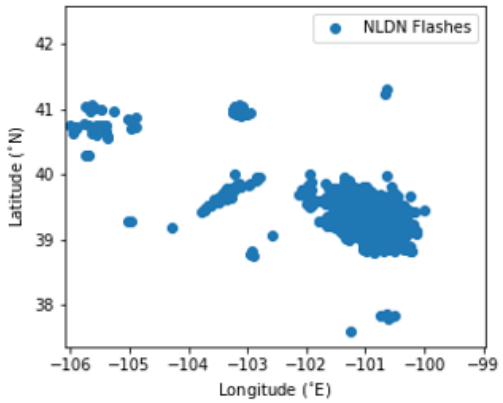
Figure 3.17: The radar images from the KGLD radar for the Goodland, KS case at 2244 UTC (a), at 2320 UTC (b), and at 2357 UTC (c) 25 May 2017.

Additionally, there were several small weak storms in the northeastern corner of Colorado and one in the southwestern corner of Kansas 115.66 km east of Springfield, CO at this time.

The southern severe storm in Kansas turned south and began to strengthen. At 2308 UTC 25 May 2017 it started to rotate and have a hook signature. At 2321 UTC 25 May 2017 it moved east and a tornado warning was issued for it. This storm continued to move east and have a tornado warning associated with it for the rest of the duration of the case. The northern severe storm moved east and began to weaken. The storm in the southwestern




(a) Forecaster Created Plume




(b) NLDN Flashes

Figure 3.18: The merged maximum forecasted probability at any time step over the duration of the case (a) and NLDN flashes (b) for the Goodland, KS case 25-26 May 2017.

corner of Kansas moved east northeast and began to strengthen. At 2314 UTC 25 May 2017 a severe thunderstorm warning was issued for it. Towards the end of the duration of the case a storm formed south of Sidney, NE and moved east. Three time steps of the KGLD radar for this case were shown (Figure 3.17). The forecaster merged maximum plume at any time step over the duration of the Goodland, KS case is shown in Figure 3.18 (a)). Based on the NLDN flashes (Figure 3.18 (b)) the forecaster did a fairly good job of making this forecast. The higher probabilities correlated very well with the location of the NLDN flashes.

The spatial domain of all four tornado and lightning cases was shown (Figure 3.19). The domain of any lightning case was somewhat larger than for any tornado case. Additionally, the domain of three of the lightning cases overlap in Colorado, Nebraska, and Kansas. The domain size of all eight cases vary drastically.

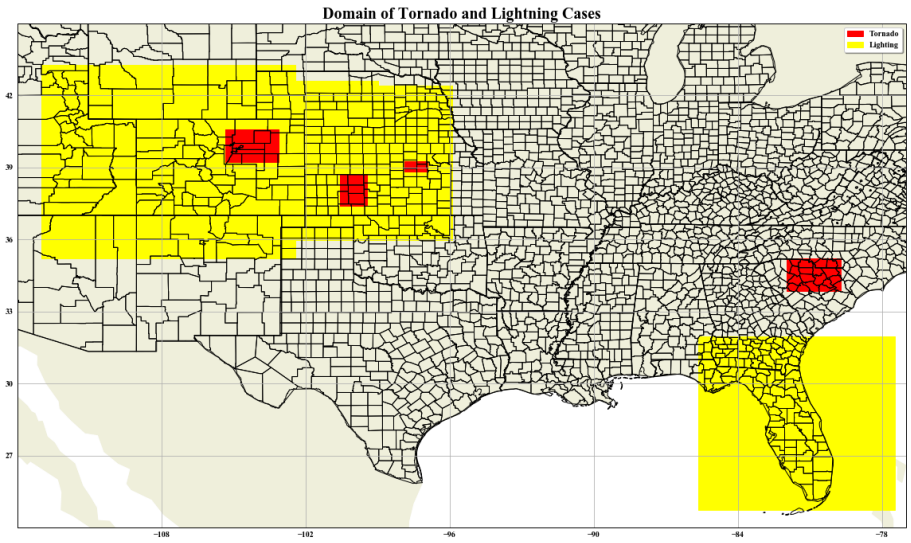



Figure 3.19: The spatial domain for all the cases with the lightning in yellow and tornado in red.

## 3.3 Summary

- The SPC issued a 2% risk of tornadoes in the 2000 UTC 25 May 2016 Convective Outlook and a tornado watch that lasted through the duration of the case.
- Two supercell tornadoes formed near Bennington, KS on 25 May 2016.
- The SPC issued a 10% risk of tornadoes in the 2000 UTC 24 May 2016 Convective Outlook and a tornado watch that lasted through the duration of the case.
- Several tornadoes formed near Dodge City, KS on 24 May 2016.
- The SPC issued a 2% risk of tornadoes in the 2000 UTC 8 May 2017 Convective Outlook and a severe thunderstorm watch that lasted through the duration of the case.
- There was one tornadic warned storm near Denver, CO on 8 May 2017.

- The SPC issued a 10% risk of tornadoes in the 2000 UTC 24 May 2017 Convective Outlook and three tornado watches for different portions of the domain of the case that lasted through the duration of the case.
- There were three tornadoes that formed from and a tornado warning was issued for a portion of a QLCS storm near Columbia, SC on 24 May 2017.
- On 25 May 2016 in addition, to a supercell thunderstorm there were a few lines of storms that moved through the area around Bennington, KS.
- The SPC issued a 5% risk of severe hail and wind in the 1630 UTC 22 July 2016 Convective Outlook and zero watches.
- Many slow moving, short duration storms formed in the mountains surrounding Grand Junction, CO on 22 July 2016.
- The SPC issued a 5% risk of wind and tornadoes in the 1630 UTC 1 September 2016 Convective Outlook and a tornado watch for the eastern portion of the domain.
- On 1 September 2016 there was a band of storms from Hurricane Hermine on the western coast of Florida and to the east of this band there were thunderstorms some of which became severe.
- The SPC issued a 15% risk of severe hail and a 30% risk of wind in the 2000 UTC 25 May 2017 Convective Outlook and a severe thunderstorm watch for the north-central portion of the domain that lasted through the duration of the case.
- On 25 May 2017 a few severe thunderstorms, one of which there was a tornado warning issued for moved through the region around Goodland, KS.
- The domain size of the lightning cases was much larger than for the tornado cases and the domain size between individual cases varies drastically.

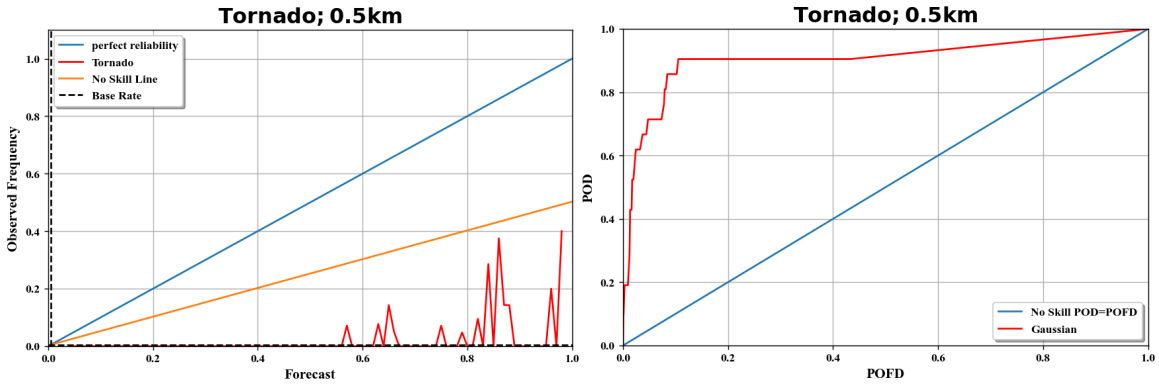
# Chapter 4

## Event Definition and Practically Perfect Plumes

The first question this research is trying to address is at what distance away from the event should the event be defined. In other words, at what distance away from or at what neighborhood around the event does the forecast become useful. This distance is then referred to as the distance at which the event is defined (event definition). The event definition could be different between the tornado and lightning hazards.

## 4.1 Tornado

First, a 0.5 km radius (1 km$^2$ grid) around the mesocyclone coordinates was tested as the event definition for the Bennington, KS tornado case. Using the Bennington, KS tornado case as an example, it is difficult to get meaningful information from the attributes diagram using a 0.5 km event definition (Figure 4.1 (a)). While the resolution is 0.001 and reliability is 0.078, this does not adequately represent the skill of the forecast. Based on Figure 3.4 the forecast was objectively good. However, the AUC for the ROC diagram (Figure 4.1 (b)) is 0.90.



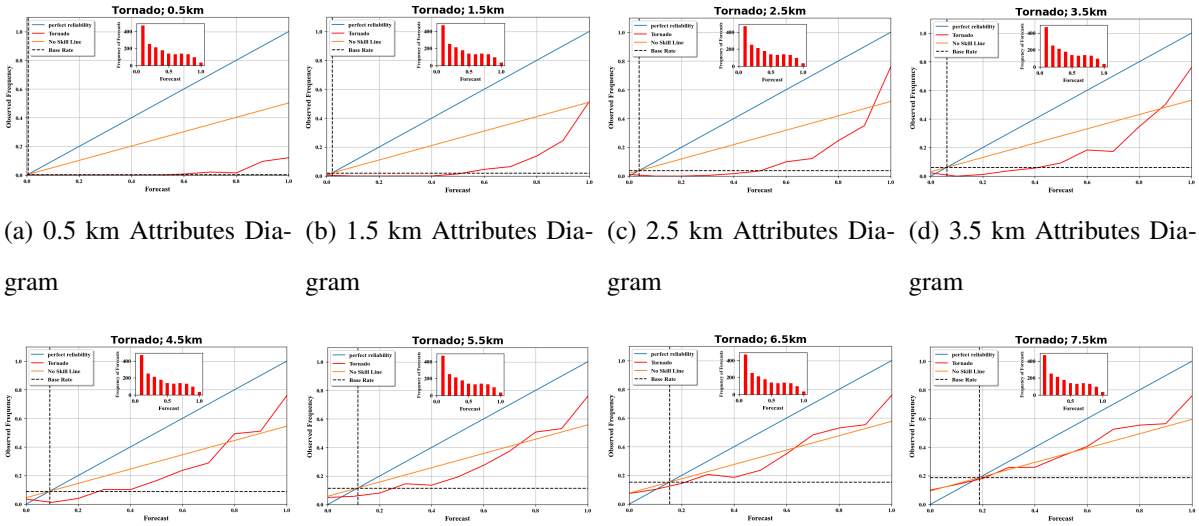

(a) Attributes Diagram (b) ROC Diagram

Figure 4.1: The attributes diagram (a) and ROC diagram (b) for the Bennington, KS tornado case 25-26 May 2016 at a 0.5 km radial distance away from the mesocyclone track.

Since the ROC diagram had an AUC of greater than 0.70, a yes forecast was able to be discriminated.

Using a single case such as this, also limits the sample size and produces the irregular nature of the plots (Figure 4.1). To reduce the impact of the sampling, probabilities are binned in groups of 10 (i.e. 1-10, 11-20, 21-30, etc) for all the following plots. Forecast frequency histograms, excluding the zero bin, are shown to indicate the usage of the forecasts at the different bins. The zero bin was excluded because the zero forecast bin contains many more points than any of the other forecast bins since the majority of the area is covered by a null forecast.

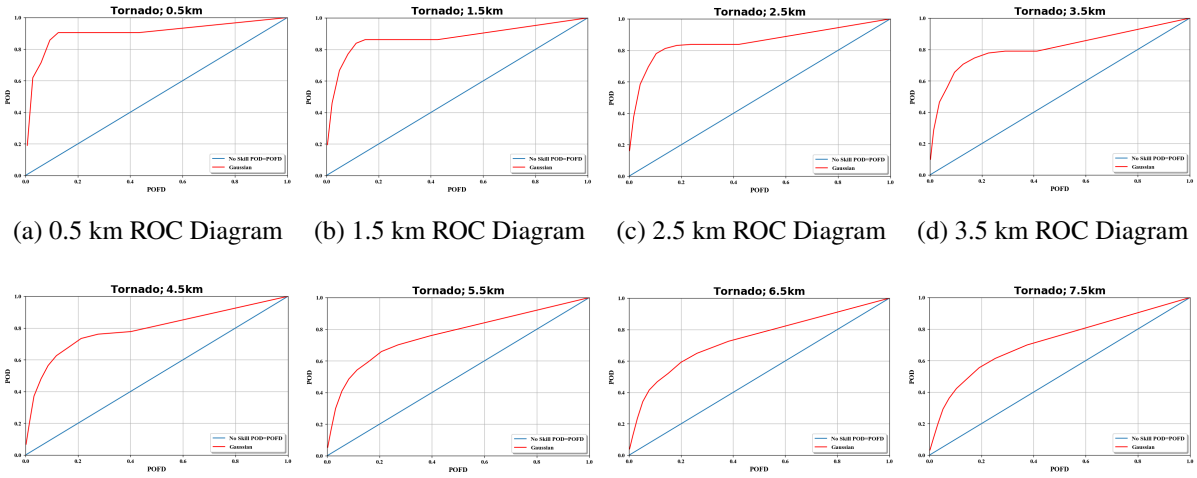
The effect of changing the event definition for the Bennington, KS tornado plume is shown in both the the attributes (Figure 4.2) and ROC diagrams (Figure 4.3).




(a) 0.5 km Attributes Diagram (b) 1.5 km Attributes Diagram (c) 2.5 km Attributes Diagram (d) 3.5 km Attributes Diagram

(e) 4.5 km Attributes Diagram (f) 5.5 km Attributes Diagram (g) 6.5 km Attributes Diagram (h) 7.5 km Attributes Diagram.

Figure 4.2: The attributes diagrams for the Bennington, KS tornado case 25-26 May 2016 at a 0.5 km (a), 1.5 km (b), 2.5 km (c), 3.5 km (d), 4.5 km (e), 5.5 km (f), 6.5 km (g), and 7.5 km (h) radial distance away from the mesocyclone track using a 10 probability (forecast) bin. In the top middle of each attributes diagram there is a histogram of the frequency of the forecasts.




(a) 0.5 km ROC Diagram (b) 1.5 km ROC Diagram (c) 2.5 km ROC Diagram (d) 3.5 km ROC Diagram

(e) 4.5 km ROC Diagram (f) 5.5 km ROC Diagram (g) 6.5 km ROC Diagram (h) 7.5 km ROC Diagram

Figure 4.3: The ROC diagrams for the Bennington, KS tornado case 25-26 May 2016 at a 0.5 km (a), 1.5 km (b), 2.5 km (c), 3.5 km (d), 4.5 km (e), 5.5 km (f), 6.5 km (g), and 7.5 km (h) radial distance away from the mesocyclone track using a 10 probability (forecast) bin.

Each subsequent plot in Figure 4.2 and Figure 4.3 shows the effect of increasing the radius of the event definition by 1 km from 0.5 km to 7.5 km. Since the forecast was the same as the event definition changed, the forecast frequency histograms in all the plots are identical.

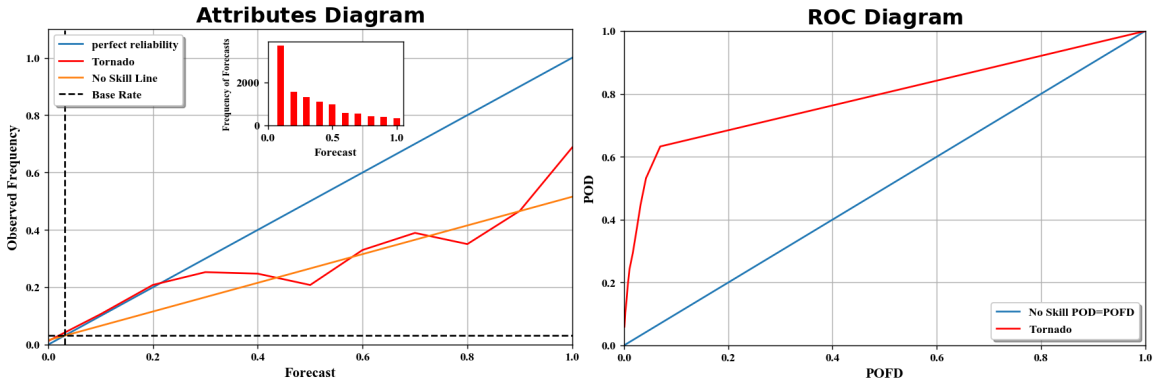
As the distance away was increased, there was a greater observed frequency at lower probabilities. This was especially true at a forecast of zero. As the distance away increased toward 7.5 km, the observed frequency at higher probabilities changed less compared to at the lower probabilities. At the 7.5 km event definition, the line on the attributes diagrams lies about on the no skill line. As the distance away increased toward 7.5 km, the AUC value decreased.

Based on Figures 4.2 and 4.3, the appropriate event definition uses a 7.5 km radius around the mesocyclone coordinates. For the 7.5 km attributes diagram the resolution was 0.021 and reliability was 0.015. For the 7.5 km ROC diagram the AUC was 0.72. Using a 7.5 km event definition compared to using a 0.5 km event definition improved the reliability

and resolution values and the AUC was still above 0.70 for this case. Additionally, further increasing the event definition up to 14.5 km was tested.

Due to the Columbia, SC and Denver, CO cases having a plume verify using a warning and not a storm report, terminology was developed to refer to the diagrams, resolution values, reliability values, and AUC values. The first terminology included the warning mesocyclone verification tracks and was referred to as Meso. The other one excluded the warning mesocyclone verification tracks and was referred to as Tor Official.

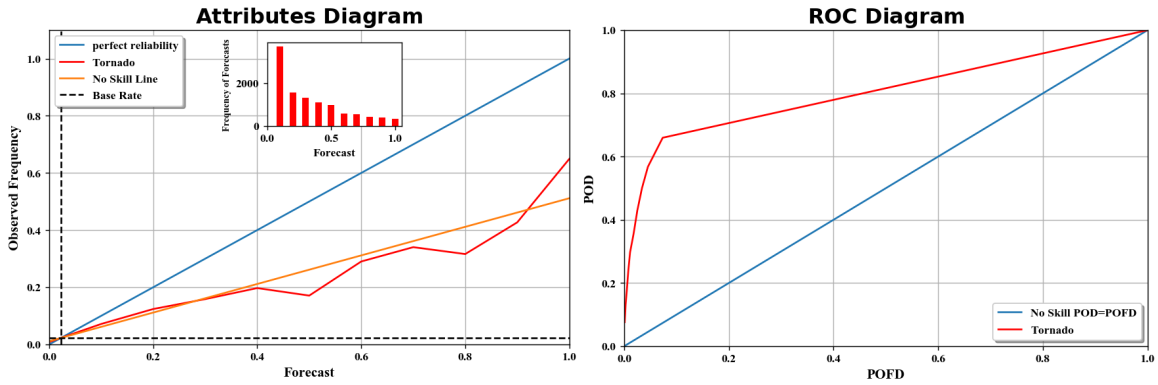
For the combined Meso attributes diagram (Figure 4.4 (a)) the resolution was 0.005 and reliability was 0.002. For the combined Meso ROC diagram (Figure 4.4 (b)) the AUC was 0.79. For the combined Tor Official attributes diagram (Figure 4.5 (a)) the resolution was 0.004 and reliability was 0.004. For the combined Tor Official ROC diagram (Figure 4.5 (b)) the AUC was 0.80.




(a) Attributes Diagram (b) ROC Diagram

Figure 4.4: The Meso attributes diagram (a) and Meso ROC diagram (b) for all tornado cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. In the top middle of the attributes diagram there is a histogram of the frequency of the forecasts.

Generally, for all cases and the combined cases when going from a 4.5 km to a 7.5 km event definition improves the resolution and reliability values (Figure 4.6). However, there was not a noticeable improvement for the Denver, CO case.




(a) Attributes Diagram (b) ROC Diagram

Figure 4.5: The Tor Official attributes diagram (a) and Tor Official ROC diagram (b) for all tornado cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. In the top middle of the attributes diagram there is a histogram of the frequency of the forecasts.

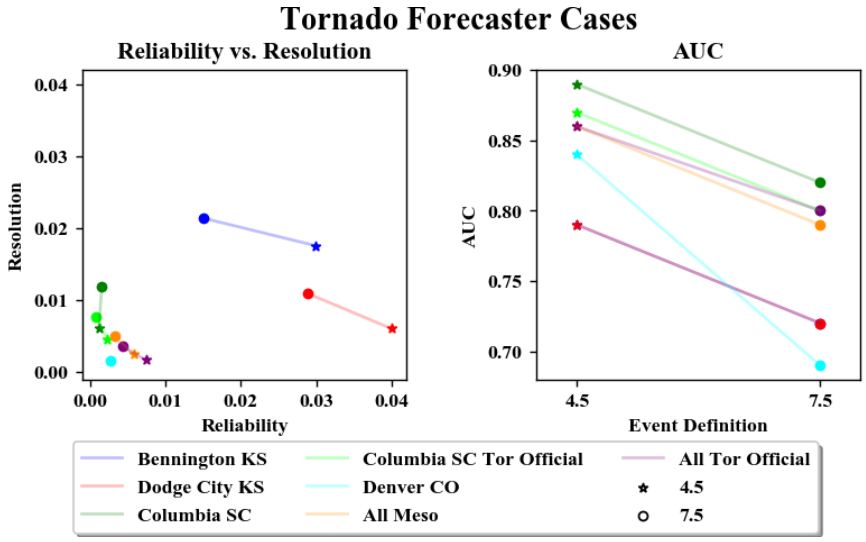
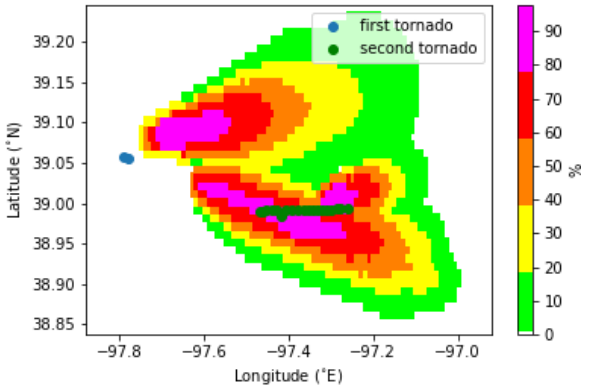


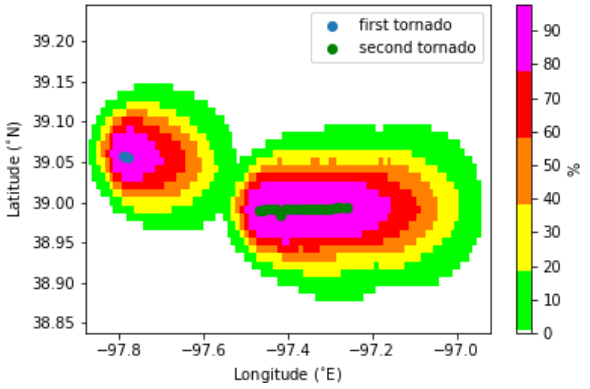

Figure 4.6: The forecaster reliability vs resolution terms (left) and the forecaster AUC values (right) for all tornado cases and the combined cases for a 4.5 km and 7.5 km event definition.

For all of the other cases except the Columbia, SC Meso and Tor Official cases there was a greater improvement in the reliability values than for the resolution values when going from a 4.5 km to a 7.5 km event definition. For the Columbia, SC Tor Official case there was not any improvement in the reliability values when going from a 4.5 km to a 7.5 km event definition. For the combined

Meso and Tor Official cases there was almost an equal improvement in the reliability and resolution values when going from a 4.5 km to a 7.5 km event definition. For all cases and the combined cases when going from a 4.5 km to a 7.5 km event definition the AUC values dropped. For all cases but the Denver, CO case the AUC values with a 7.5 km event definition was still at or above 0.70. While the AUC dropped when going from a 4.5 km to a 7.5 km event definition, but stayed above 0.70 (for all but one case), the improvement in the resolution and reliability terms indicates that a 7.5 km event definition should be used.




(a) Forecaster Created Plume

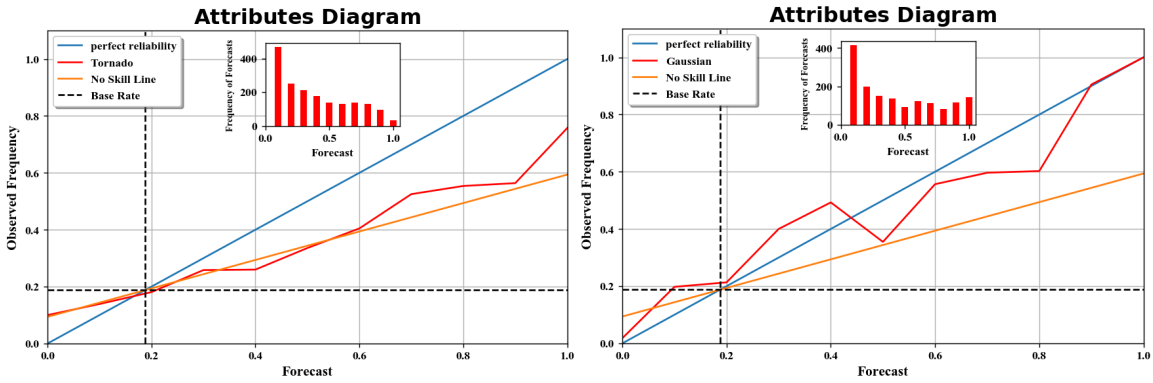



(b) Practically Perfect Plume

Figure 4.7: The mesocyclone track overlaid on top of the forecaster (a) and practically perfect (b) merged maximum plume at any time step over the duration of the Bennington, KS tornado case 25-26 May 2016.

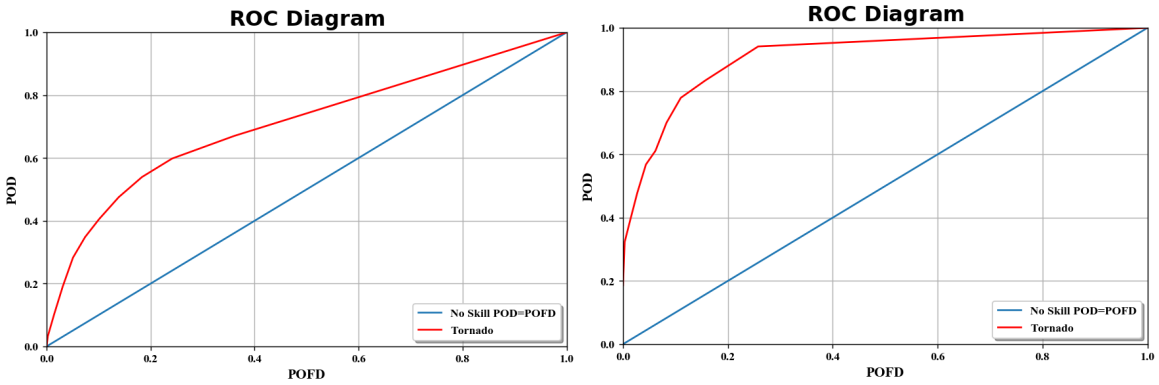
### 4.1.1 Practically Perfect Plumes

The concept of practically perfect forecasts was applied to the plumes to provide guidance as to what the "best" forecast will look like. The mesocyclone track overlaid on top of the practically perfect merged maximum plume at any time step over the duration of the Bennington, KS tornado case was shown in Figure 4.7 (b). Due to the forecaster forecasting the plume (Figure 4.7 (a)) well, there was not much difference between the practically perfect and the forecaster created plume. However, the first tornado was forecasted with the practically perfect plume and not the forecaster created plume.




(a) Forecaster (b) Practically Perfect

Figure 4.8: The forecaster (a) and practically perfect (b) attributes diagram for the Bennington, KS tornado case 25-26 May 2016 at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. In the top middle of the attributes diagrams there are histograms of the frequency of the forecasts.




(a) Forecaster (b) Practically Perfect

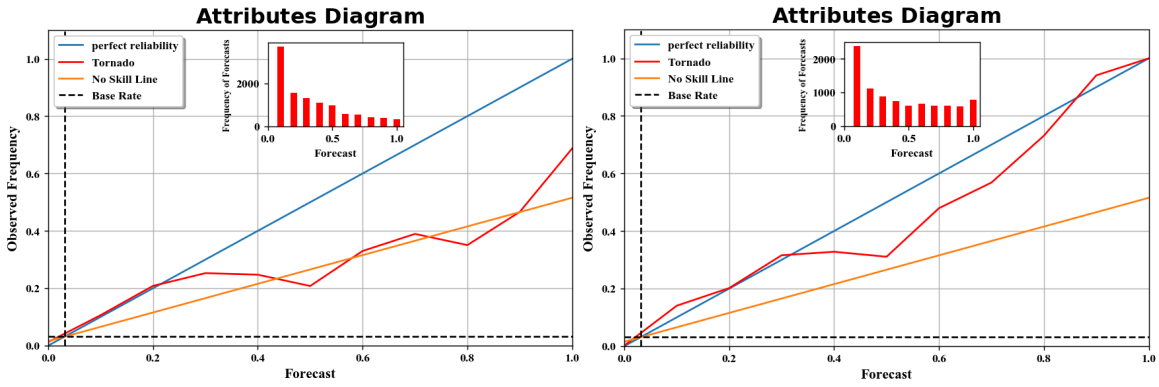
Figure 4.9: The forecaster (a) and practically perfect (b) ROC diagram for the Bennington, KS tornado case 25-26 May 2016 at a 7.5 km radial distance away from the mesocyclone track using a 10 probability(forecast) bin.

For the Bennington, KS tornado case practically perfect attributes diagram (Figure 4.8 (b)) the resolution was 0.073 and the reliability was 0.003. The reliability and resolution values for the practically perfect attributes diagram indicate that the forecaster plume for this case could be improved. Additionally the attributes diagrams for the forecaster (Figure 4.8 (a)) and practically perfect plumes are quite different, especially at the extremely high probability values. The histogram of forecast frequency is quite different between

the two attributes diagrams. For the practically perfect attributes diagram compared to the forecaster attributes diagram, there is a higher frequency of forecasts at the higher 10 bin forecast values.

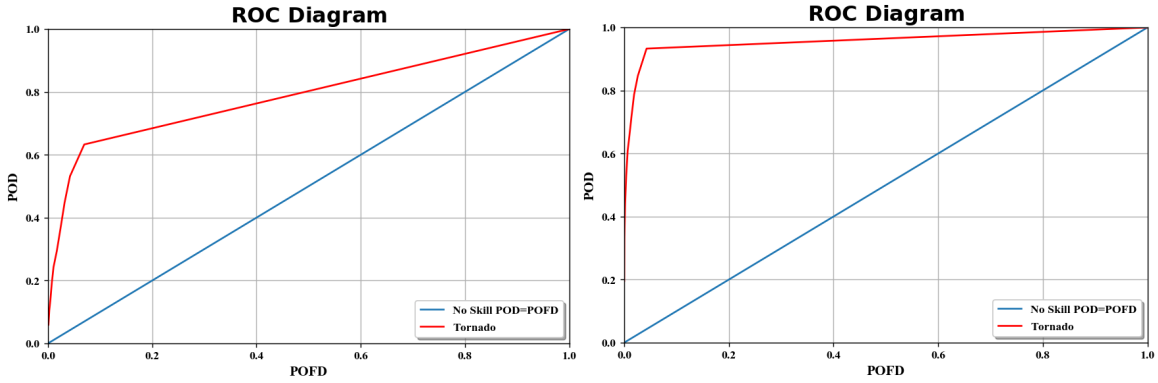
For the Bennington, KS tornado case practically perfect ROC diagram (Figure 4.9 (b)) the AUC was 0.91. The AUC value for the practically perfect ROC diagram indicates that the forecaster plume for this case could be improved. Additionally the ROC diagrams for the forecaster (Figure 4.9 (a)) and practically perfect plumes are quite different.

For the combined practically perfect Meso attributes diagram (Figure 4.10 (b)) the resolution was 0.017 and reliability was 0.000. The resolution and reliability values for the combined practically perfect Meso attributes diagram indicate that the forecaster plumes for all cases could be better. Additionally, the combined Meso attributes diagrams for forecaster (Figure 4.10 (a)) and practically perfect plumes are drastically different due to the challenging nature of some of the tornado cases. The histogram of forecast frequency is quite different between the combined practically perfect Meso and the combined forecaster Meso attributes diagrams. For the combined practically perfect Meso attributes diagram compared to the combined forecaster Meso attributes diagram, there is a higher frequency of forecasts at the higher 10 bin forecast values.




(a) Forecaster (b) Practically Perfect

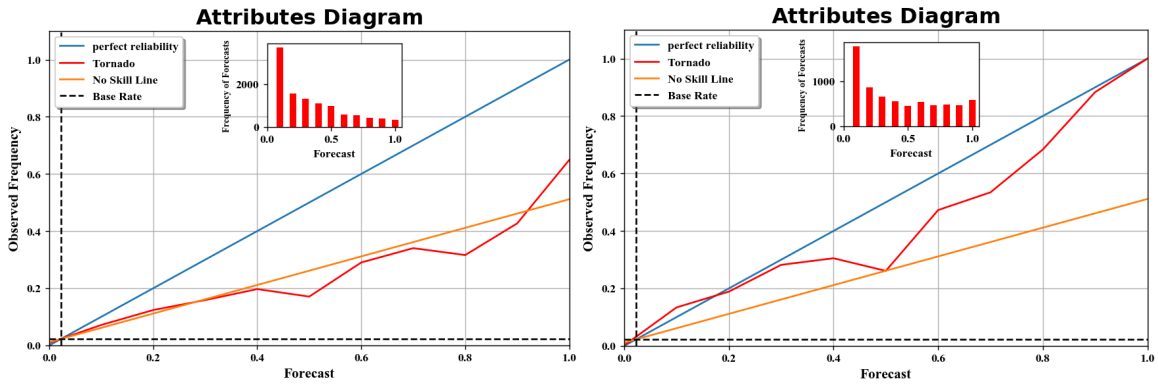
Figure 4.10: The forecaster (a) and practically perfect (b) Meso attributes diagram for all cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. In the top middle of the attributes diagrams there are histograms of the frequency of the forecasts.




(a) Forecaster (b) Practically Perfect

Figure 4.11: The forecaster (a) and practically perfect (b) Meso ROC diagrams for all cases at a radial distance 7.5 km away from the mesocyclone track using a 10 probability (forecast) bin.

For the combined practically perfect Meso ROC diagram (Figure 4.11 (b)) the AUC was 0.96. The AUC value for the combined practically perfect Meso ROC diagram indicates that the forecaster plumes for all cases could be improved. Additionally, the Meso ROC diagram for the forecaster (Figure 4.11 (a)) and practically perfect plumes are quite different.

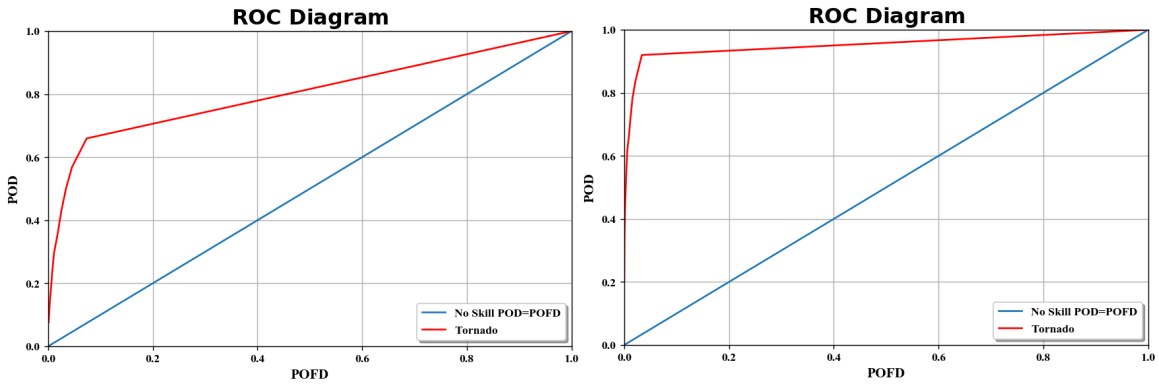



(a) Forecaster (b) Practically Perfect

Figure 4.12: The forecaster (a) and practically perfect (b) Tor Official attributes diagram for all cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. In the top middle of the attributes diagrams there are histograms of the frequency of the forecasts.

For the combined practically perfect Tor Official attributes diagram (Figure 4.12 b)) the resolution was 0.012 and reliability was 0.000. The resolution and reliability values for the combined practically perfect Tor Official attributes diagram indicate that the forecaster plumes for all cases could be better. Additionally, the Tor Official attributes diagrams forecaster (Figure 4.12 (a)) and practically perfect plumes are drastically different due to the challenging natures of some of the tornado cases. The histogram of forecast frequency is quite different between the combined practically perfect and forecaster Tor Official attributes diagrams. For the combined practically perfect Tor Official attributes diagram compared to the combined forecaster Tor Official attributes diagram, there is a higher frequency of forecasts at the higher 10 bin forecast values.

For the combined practically perfect Tor Official ROC diagram (Figure 4.13 (b)) the AUC was 0.95. The AUC value for the combined practically perfect Tor Official ROC diagram indicates that the forecaster plume for all cases could be improved. Additionally the forecaster Tor Official ROC diagram for the forecaster (Figure 4.13 (a)) and practically perfect plume are quite different.




(a) Forecaster (b) Practically Perfect

Figure 4.13: The forecaster (a) and practically perfect (b) Tor Official ROC diagrams for all cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin.

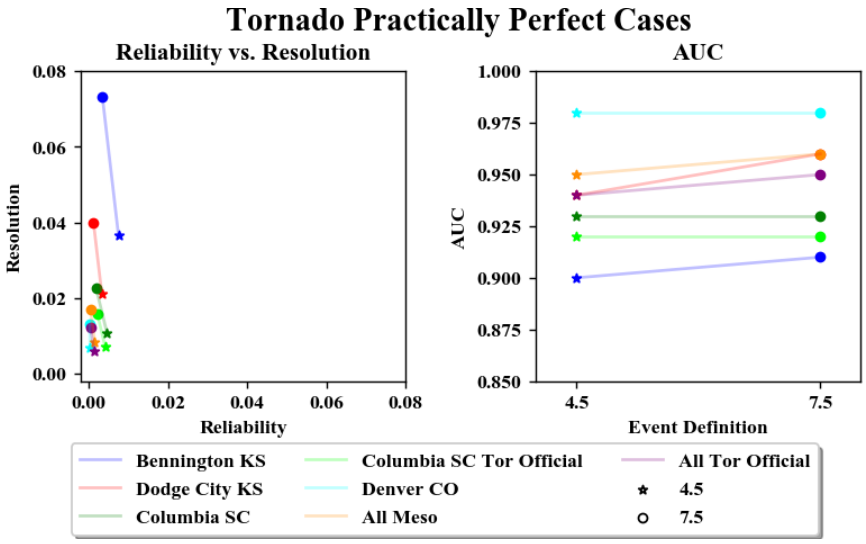



Figure 4.14: The practically perfect reliability vs resolution terms (left) and the practically perfect AUC values (right) for all tornado cases and the combined cases for a 4.5 km and 7.5 km event definition.

For all practically perfect cases and the practically perfect combined cases when going from a 4.5 km to a 7.5 km event definition improves the resolution and reliability values (Figure 4.14). For all of the practically perfect cases there was a greater improvement in the resolution values than for the reliability values when going from a 4.5 km to a 7.5 km event definition. For all practically perfect cases and the combined cases when going from a 4.5 km to a 7.5 km event definition the AUC values stayed the same or slightly increased.

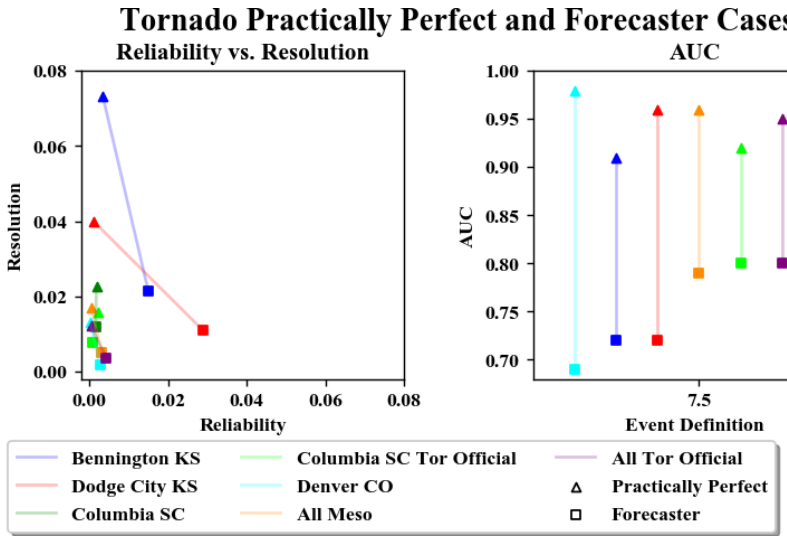
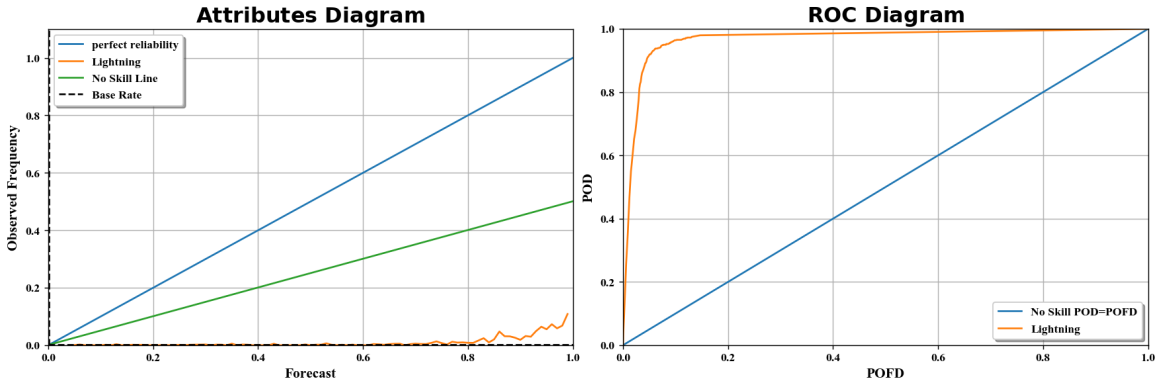



Figure 4.15: The practically perfect and forecaster reliability vs resolution terms (left) and the practically perfect and forecaster AUC values (right) for all tornado cases and the combined cases for a 7.5 km event definition.

For all forecaster cases and the forecaster combined cases the practically perfect plumes at a 7.5 km event definition improved the resolution values (Figure 4.15). For all of the forecaster cases except Columbia, SC Meso and Tor Official the practically perfect plumes at a 7.5 km event definition improved the reliability values. The practically perfect plumes improved the forecaster Bennington, KS tornado and Dodge City, KS cases' resolution and reliability values the most. For all forecaster cases and the combined cases the practically perfect plume at a 7.5 km event definition improved the AUC value.

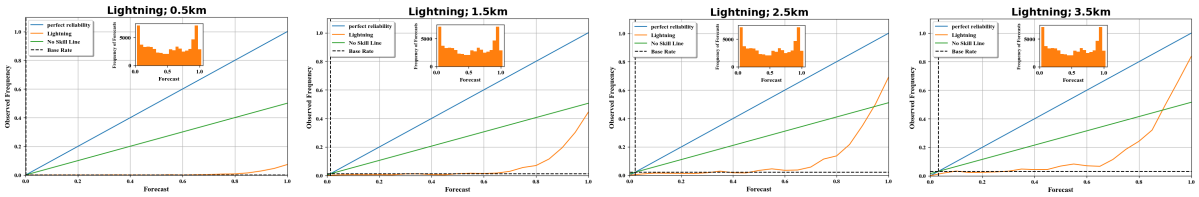
## 4.2 Lightning

First, a 0.5 km radius (1 km$^2$ grid) around the NLDN flashes was tested as the event definition for the Bennington, KS lightning case. Using the Bennington, KS lightning case as an example, it is difficult to get meaningful information from the attributes diagram using a 0.5 km event definition (Figure 4.16 (a)). While the resolution is 0.000 and reliability is 0.050 this does not adequately represent the skill of the forecast, just as with the Bennington KS, tornado case. Based on Figure 3.12 the forecast was objectively good. However, the AUC for the ROC diagram (Figure 4.16 (b)) is 0.97. Since the ROC diagram had an AUC of greater than 0.70 a yes forecast was able to be discriminated.

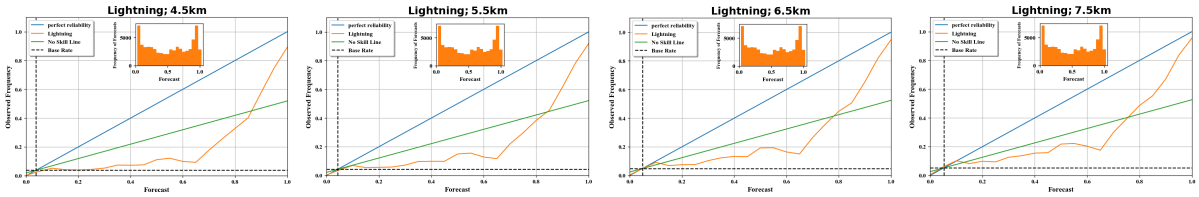



(a) Attributes Diagram (b) ROC Diagram

Figure 4.16: The attributes diagram (a) and ROC diagram (b) for the Bennington, KS lightning case 25-26 May 2016 at a 0.5 km radial distance away from the NLDN flashes.

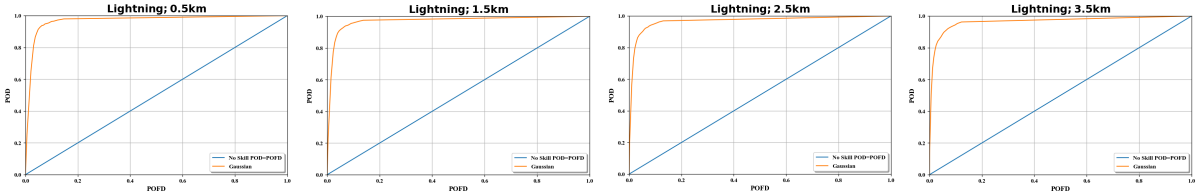



(a) 0.5 km Attributes Diagram (b) 1.5 km Attributes Diagram (c) 2.5 km Attributes Diagram (d) 3.5 km Attributes Diagram

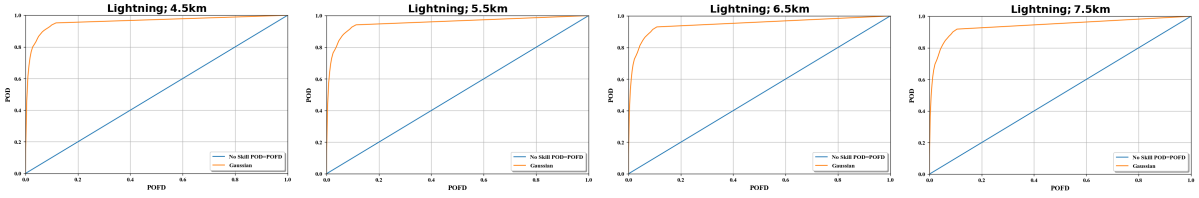



(e) 4.5 km Attributes Diagram (f) 5.5 km Attributes Diagram (g) 6.5 km Attributes Diagram (h) 7.5 km Attributes Diagram

Figure 4.17: The attributes diagrams for the Bennington, KS lightning case 25-26 May 2016 at a 0.5 km (a), 1.5 km (b), 2.5 km (c), 3.5 km (d), 4.5 km (e), 5.5 km (f), 6.5 km (g), and 7.5 km (h) radial distance away from the NLDN flashes using a five probability (forecast) bin. In the top middle of each attributes diagram there is a histogram of the frequency of the forecasts.




(a) 0.5 km ROC Diagram (b) 1.5 km ROC Diagram (c) 2.5 km ROC Diagram (d) 3.5 km ROC Diagram




(e) 4.5 km ROC Diagram (f) 5.5 km ROC Diagram (g) 6.5 km ROC Diagram (h) 7.5 km ROC Diagram

Figure 4.18: The ROC diagrams for the Bennington, KS lightning case 25-26 May 2016 at a 0.5 km (a), 1.5 km (b), 2.5 km (c), 3.5 km (d), 4.5 km (e), 5.5 km (f), 6.5 km (g), and 7.5 km (h) radial distance away from the NLDN flashes using a five probability (forecast) bin.

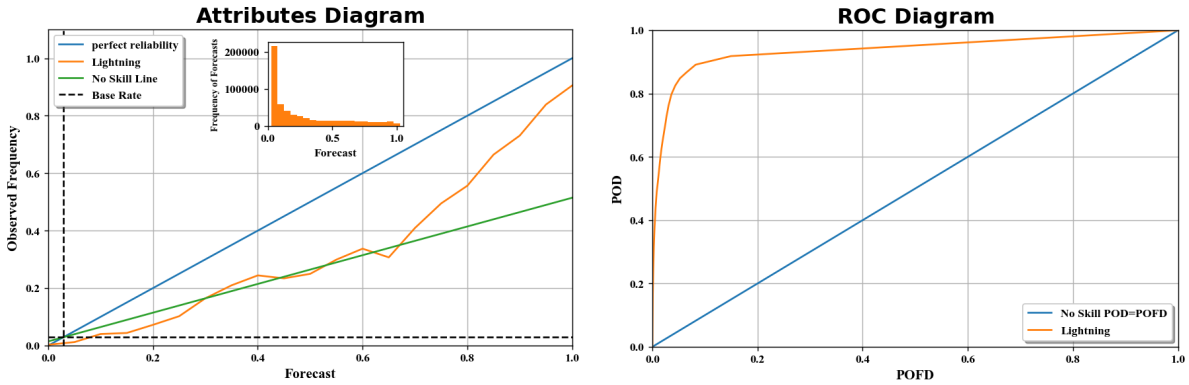
Using a single case such as this, also limits the sample size and produces the irregular nature of the plots (Figure 4.16). To reduce the impact of the sampling, probabilities will be binned in groups of 5 (i.e. 1-5, 6-10, 11-15, etc) for all the following plots. Forecast frequency histograms, excluding the zero bin, are shown to indicate the usage of the forecasts at the different bins. The zero bin was excluded because the zero forecast bin contains many more points than any of the other forecast bins since the majority of the area is covered by a null forecast.

The effect of changing the event definition for the Bennington, KS lightning plumes was shown in both the attributes (Figure 4.17) and ROC diagrams (Figure 4.18). Each subsequent plot in Figure 4.17 and Figure 4.18 shows the effect of increasing the radius of the event definition by 1 km from 0.5 km to 7.5 km.

Since the forecast was the same as the event definition changed the histograms in all the plots are identical. Outside of the the zero bin, the 1-5 and 91-95 bins both had the highest forecast frequency. As the distance away increases toward 7.5 km, the observed frequency of lightning flashes at the higher probabilities and lower probabilities got very close to the perfect reliability line. At a 7.5 km event definition, the line on the attributes diagrams lies on the reliability line for the lower probabilities. As the distance away was increased toward 7.5 km, the AUC decreased.

Based on Figures 4.17 and 4.18 the appropriate event definition was using a 7.5 km radius around the NLDN flashes. For the 7.5 km attributes diagram the resolution was 0.025 and reliability was 0.009. For the 7.5 km ROC diagram the AUC was 0.94. Using a 7.5 km event definition compared to using a 0.5 km event definition improved the reliability and resolution and the AUC was still above 0.70 for this case. The AUC changed less for lightning when using a 7.5 km compared to a 0.5 km event definition when compared to tornado. Additionally, further increasing the event definition up to 23.5 km was tested.

For the combined attributes diagram (Figure 4.19 (a)) the resolution was 0.011, reliability was 0.003, and uncertainty was 0.028.




(a) Attributes Diagram (b) ROC Diagram

Figure 4.19: The attributes diagram (a) and ROC diagram (b) for all lightning cases at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin. In the top middle of the attributes diagram there is a histogram of the frequency of the forecasts.

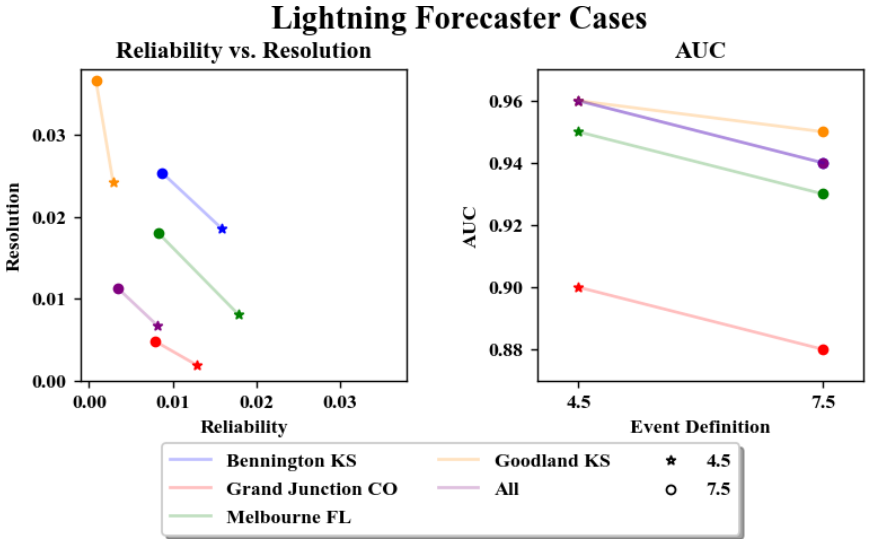


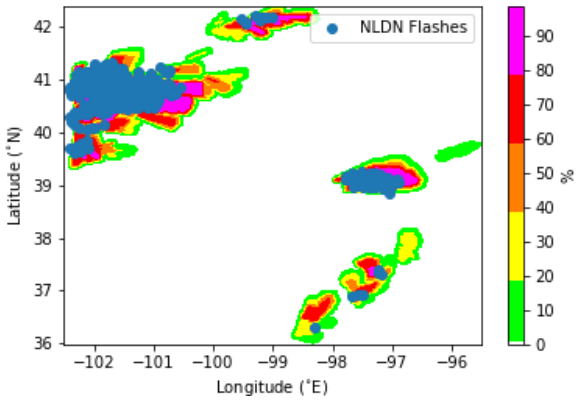

Figure 4.20: The forecaster reliability vs resolution terms (left) and the forecaster AUC values (right) for all lightning cases and the combined case for a 4.5 km and 7.5 km event definition.

For the combined ROC diagram (Figure 4.19 (b)) the AUC was 0.94. Overall, the lightning cases were fairly reliable especially at the higher probability values.

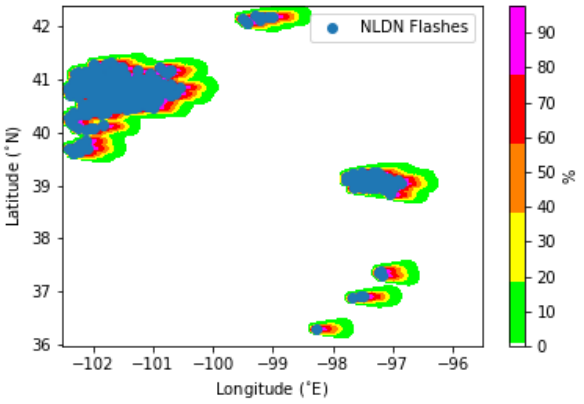
For all lightning cases and the combined case, going from a 4.5 km to a 7.5 km event definition improves the resolution and reliability values (Figure 4.20). For all of the cases and the combined case, except for the Goodland, KS and Melbourne, FL cases, there was equal improvement in the reliability and resolution values, when going from a 4.5 km to a 7.5 km event definition. For the Goodland, KS and

Melbourne FL case, there was a greater improvement in the resolution value than for the reliability value, when going from a 4.5 km to a 7.5 km event definition. For all cases and the combined cases, when going from a 4.5 km to a 7.5 km event definition the AUC value dropped. For all cases and the combined case, the AUC values with a 7.5 km event definition were still above 0.70.

### 4.2.1 Practically Perfect Plumes




(a) Forecaster Created Plume

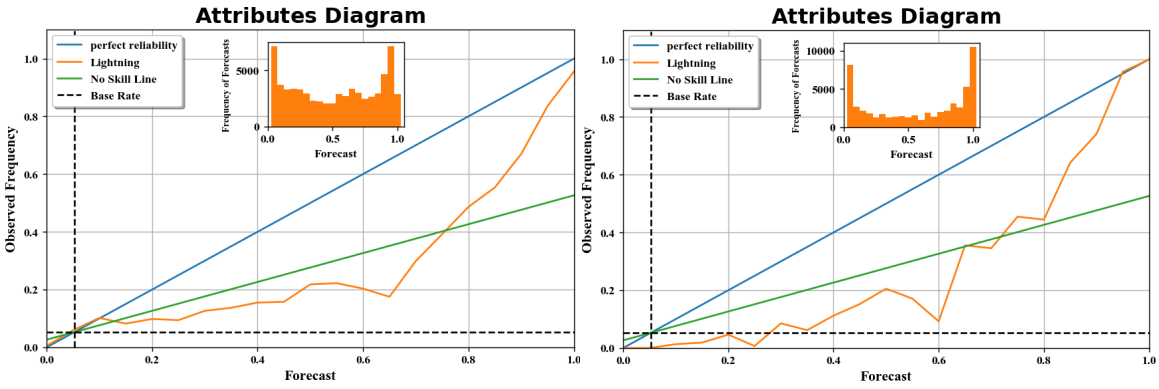



(b) Practically Perfect Plume

Figure 4.21: The NLDN flashes overlaid on top of the forecaster (a) and practically perfect (b) merged maximum plume at any time step over the duration of the Bennington, KS lightning case 25-26 May 2016.

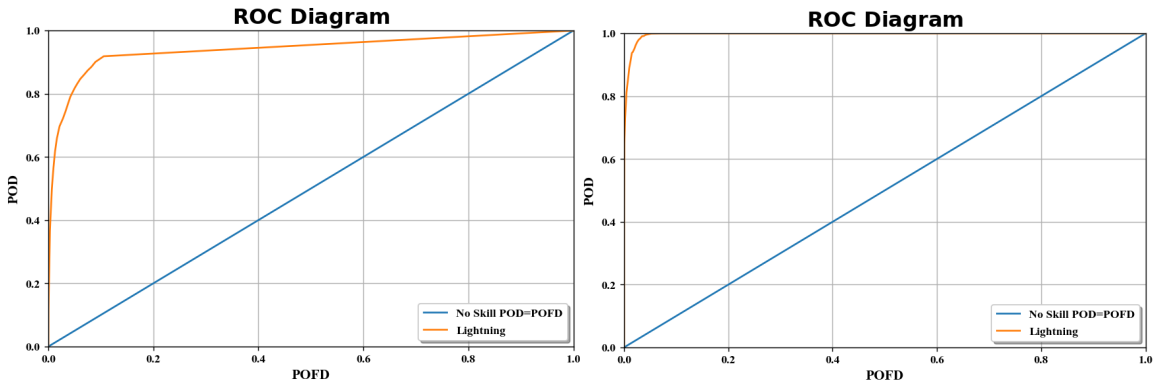
The concept of practically perfect forecasts was applied to the plumes to provide guidance as to what the "best" forecast will look like. The NLDN flashes overlaid on top of the practically perfect merged maximum plume at any time step over the duration of the Bennington, KS lightning case is shown in Figure 4.21 (b). Due to the forecaster forecasting the plume (Figure 3.12 (a)) well, there is not much different between the practically perfect and the forecaster created plume. However, the plumes in the southeastern corner of the domain are slightly different since there are only three NLDN flashes in that region.

For the Bennington, KS lightning case practically perfect attributes diagram (Figure 4.22 (b)), the resolution was 0.040 and the reliability was 0.005.




(a) Forecaster (b) Practically Perfect

Figure 4.22: The forecaster (a) and practically perfect (b) attributes diagram for the Bennington, KS lightning case 25-26 May 2016 at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin. In the top middle of the attributes diagrams there are histograms of the frequency of the forecasts.




(a) Forecaster (b) Practically Perfect

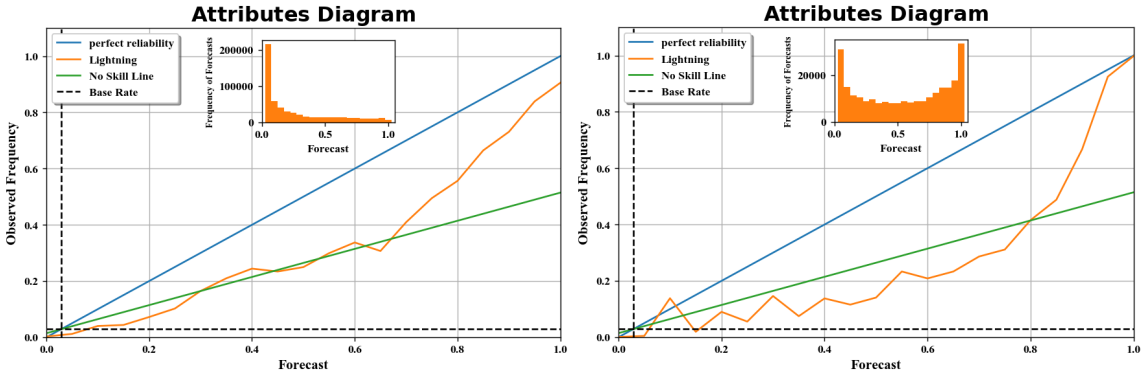
Figure 4.23: The forecaster (a) and practically perfect (b) ROC diagram for the Bennington, KS lightning case 25-26 May 2016 at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin.

The reliability and resolution values for the practically perfect attributes diagram indicate that the forecaster plume for this case could be better. Additionally, the attributes diagrams for the forecaster (Figure 4.22 (a)) and practically perfect plume are quite different from each other. At all of the probability values except the extremely high probability values the practically perfect attributes diagram is more reliable than the forecaster attributes

diagram. The histogram of the frequency of forecasts is quite different between the two attributes diagrams. For the practically perfect attributes diagram compared to the combined forecaster attributes diagram, there is a higher frequency of forecasts at the higher five bin forecast values and the 1-5 forecast bin.

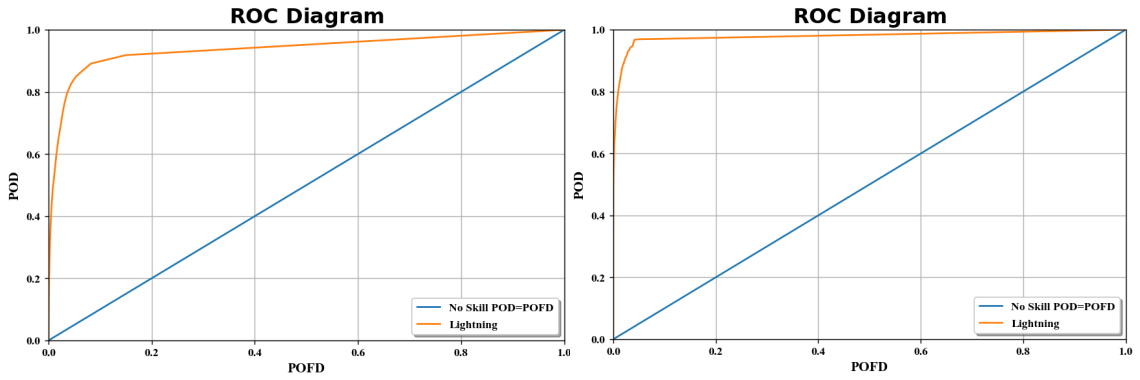
For the Bennington, KS lightning case practically perfect ROC diagram (Figure 4.23 (b)) the AUC was 1. The AUC values for the practically perfect ROC diagrams indicate that the forecaster plume for this case could be improved. Additionally, the ROC diagram for the forecaster (Figure 4.23 (a)) and practically perfect plumes are a little different.

For the combined practically perfect attributes diagram (Figure 4.24 (b)) the resolution was 0.019 and the reliability was 0.005. The resolution value for the practically perfect attributes diagram indicates that the forecaster plumes for all lightning cases could be improved. Additionally, the attributes diagrams for the forecaster (Figure 4.24 (a)) and practically perfect plumes are quite different from each other. At most of the probability values, the forecaster attributes diagram is more reliable than the practically perfect attributes diagram. At the extremely high probability values, is the only location where the practically perfect attributes diagram is more reliable than the forecaster attributes diagram. The histogram of the frequency of forecasts is quite different between the two attributes diagrams.




(a) Forecaster (b) Practically Perfect

Figure 4.24: The forecaster (a) and practically perfect (b) attributes diagram for all cases at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin. In the top middle of the attributes diagrams there are histograms of the frequency of the forecasts.




(a) Forecaster (b) Practically Perfect

Figure 4.25: The forecaster (a) and practically perfect (b) ROC diagram for all cases at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin.

For the practically perfect attributes diagram compared to the combined forecaster attributes diagram, there is a lower frequency of lightning flashes at all of the lower 5 bin forecast values excluding the zero bin.

For the combined practically perfect ROC diagram (Figure 4.25 (b)) the AUC was 0.98. The AUC values for the practically perfect ROC diagrams indicate that the forecaster plumes for all cases could be better. Additionally, the forecaster ROC diagram for the forecaster (Figure 4.25 (a)) and practically perfect plumes are very similar.

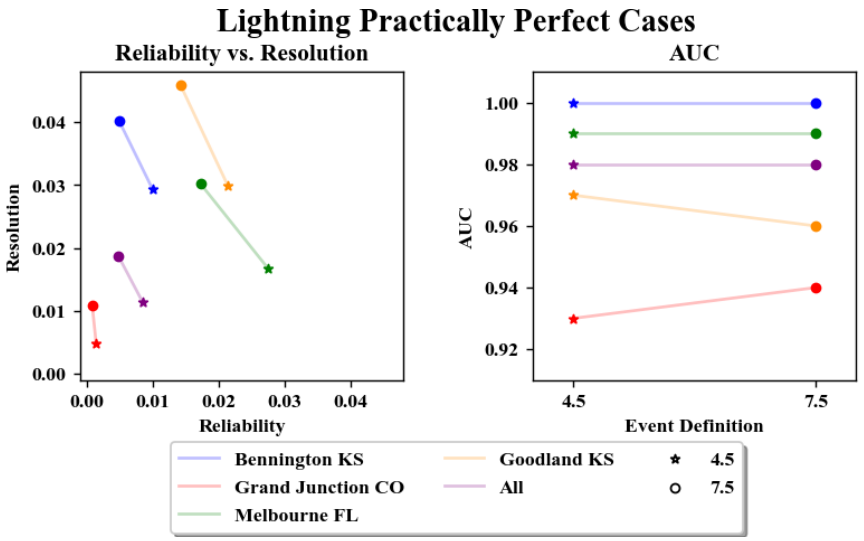



Figure 4.26: The practically perfect reliability vs resolution terms (left) and the practically perfect AUC values (right) for all lightning cases and the combined case for a 4.5 km and 7.5 km event definition.

For all practically perfect lightning cases and the practically perfect

combined lightning case, going from a 4.5 km to a 7.5 km event definition improves the resolution and reliability values (Figure 4.26). For all of the practically perfect lightning cases and the practically perfect combined lightning case there was a greater improvement in the resolution values than for the reliability values going from a 4.5 km to a 7.5 km event definition. For all practically perfect lightning cases except for Goodland, KS and the combined cases when going from a 4.5 km to a 7.5 km event definition the AUC values stayed the same or slightly increased.

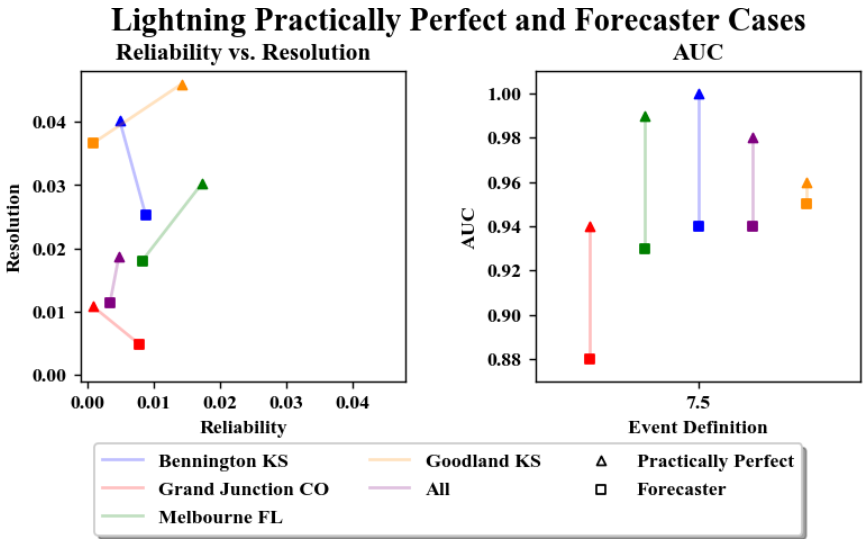



Figure 4.27: The practically perfect and forecaster reliability vs resolution terms (left) and the practically perfect and forecaster AUC values (right) for all lightning cases and the combined cases for a 7.5 km event definition.

For all forecaster cases and the forecaster combined cases the practically perfect plumes at a 7.5 km event definition improved the resolution values (Figure 4.27). For the forecaster Bennington, KS lightning and Grand Junction, CO case the practically perfect plumes at a 7.5 km event definition improved the reliability values. The practically perfect plumes improved the forecaster Bennington, KS lightning and Grand Junction, CO cases' resolution and reliability values the most. For all forecaster cases and the combined case the practically perfect plumes at a 7.5 km event definition improved the AUC value.

## 4.3 Discussion

For both tornado and lightning cases, a 7.5 km radius (neighborhood) around either the mesocyclone track or NLDN flashes defines an event. This is the neighborhood at which the tornado and lightning plumes create a useful forecast. This invokes a similar concept to the SPC convective outlooks where a 40 km radius defines a yes event. The results presented here show a way to perform gridded verification to verify tornado and lightning plumes. In this spring's HWT there was a mesocyclone tracking algorithm used with a 7.5 km radius to create automated guidance for the tornado plumes. At a 14.5 km event definition for the Bennington, KS tornado case all parts of the line on the attributes diagram were at or above the perfect reliability line, but there was a lot of underforecasting at the lower probabilities. At a 23.5 km event definition for the Bennington, KS lightning case all parts of the line on the attributes diagram were at or above the perfect reliability line but there was a lot of underforecasting at the lower and higher probabilities. Generally, underforecasting for tornado or lightning events is undesirable. Generally, for all cases going from a 4.5 to a 7.5 km event definition improved the resolution and reliability values and decreased the AUC. However, for the most part the AUC was still at or above 0.70.

The practically perfect attributes and ROC diagrams for both the tornado and lightning cases indicate the improvement that is possible in the forecaster created plumes. Overall, the practically perfect tornado attributes and ROC diagrams indicate there is more potential for improvement than the practically perfect lightning attributes and ROC diagrams. Based on Hitchens et al. (2013), it was thought that the practically perfect attributes diagram for lightning should be closer to the perfect reliability line and have better reliability values for all cases. However, for two of the lightning cases and the combined practically perfect attributes diagrams there was no improvement possible in terms of the reliability term of the BS. But, creating and verifying practically perfect plumes is a meaningful way to determine how good the forecaster plumes, attributes diagrams, and ROC diagrams could be.

## 4.4 Summary

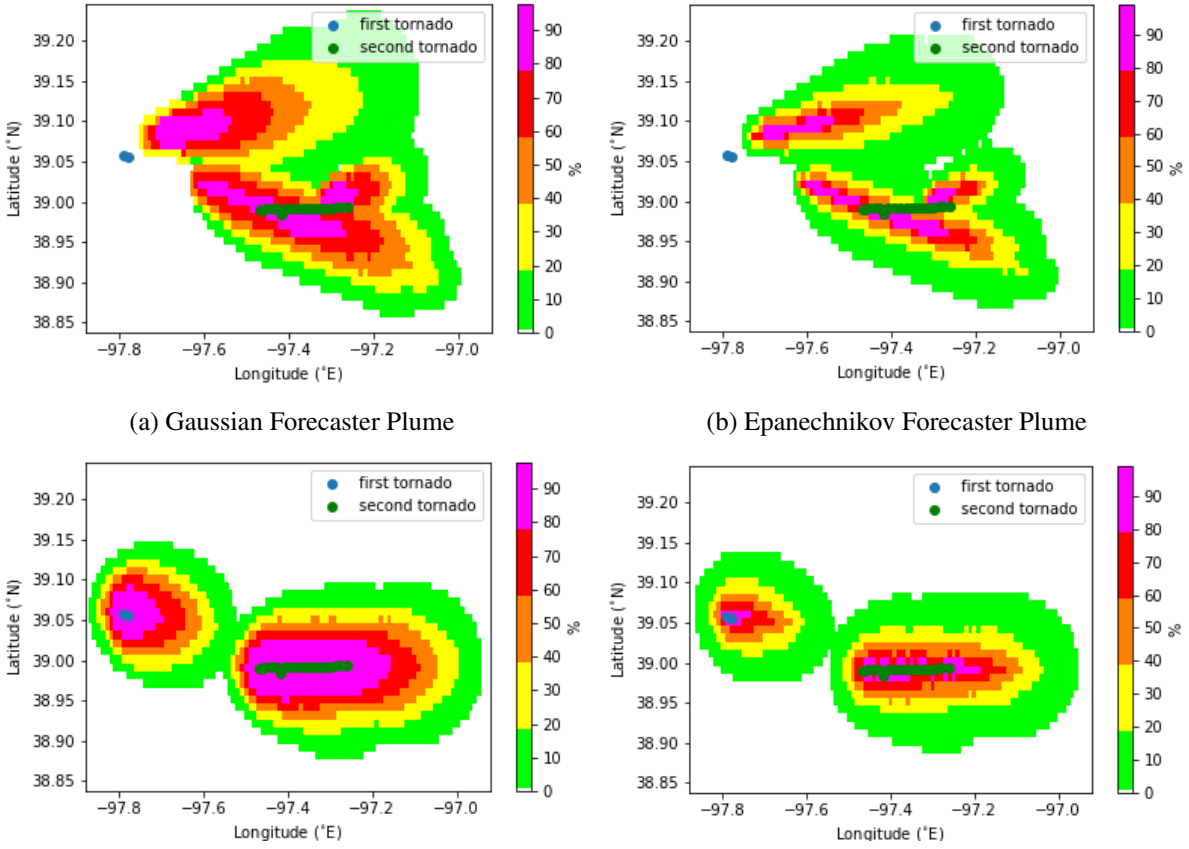
- Attributes diagrams and ROC diagrams were used to understand the consequences of varying event definitions on a gridded area.
- To address sample size, bins of 10 were used for tornado probabilities and bins of 5 were used for lightning probabilities.
- Cases for a given hazard varied widely in terms reliability values, resolution values, and AUC values.
- Resolution, reliability, and AUC values were better for lightning than tornado cases.
- For all cases going from a 4.5 km to a 7.5 km event definition improved the resolution and reliability values.
- For all cases going from a 4.5 km to a 7.5 km event definition decreased the AUC value, but it was still above 0.70.
- A 7.5 km radius around either the mesocyclone track or the NLDN flashes defines an event.
- A 7.5 km radius will be used for a mesocyclone tracking algorithm in the HWT.
- Practically perfect plumes indicate the possible improvement in the the tornado and lightning forecaster plumes.
- A 14.5 km event definition for the Bennington, KS tornado case was where all parts of the line on the attributes diagram were at or above the perfect reliability line.
- A 23.5 km event definition for the Bennington, KS lightning case was where all parts of the line on the attributes diagram were at or above the perfect reliability line.
- Practically perfect plumes allow for a calculation of the maximum possible reliability, resolution, and AUC values.

- Verifying practically perfect plumes determines how good the attributes and ROC diagrams could be.
- For a given event definition, the practically perfect lightning attributes diagrams indicated that there was not very much, if any, improvement possible in the lightning forecaster plumes.
- The practically perfect lightning ROC diagrams indicated that some improvement is possible in the lightning forecaster plumes.
- The practically perfect tornado attributes diagrams indicate that there was some improvement possible in the tornado forecaster plumes especially at the higher forecast values.
- The practically perfect tornado ROC diagrams indicate that there was more improvement possible in the tornado than the lightning forecaster plumes.

# Chapter 5

## Kernels

The next question this research is trying to address is whether the Gaussian or a different kernel (Epanechnikov, Quartic, or Triangular) should be applied to the tornado and lightning probabilities to create the plumes. The best kernel could be different between the tornado and lightning hazards. The ROC diagrams and AUC values did not appear to change appreciably between kernels.




(a) Gaussian Forecaster Plume

(b) Epanechnikov Forecaster Plume

(c) Gaussian Practically Perfect Plume

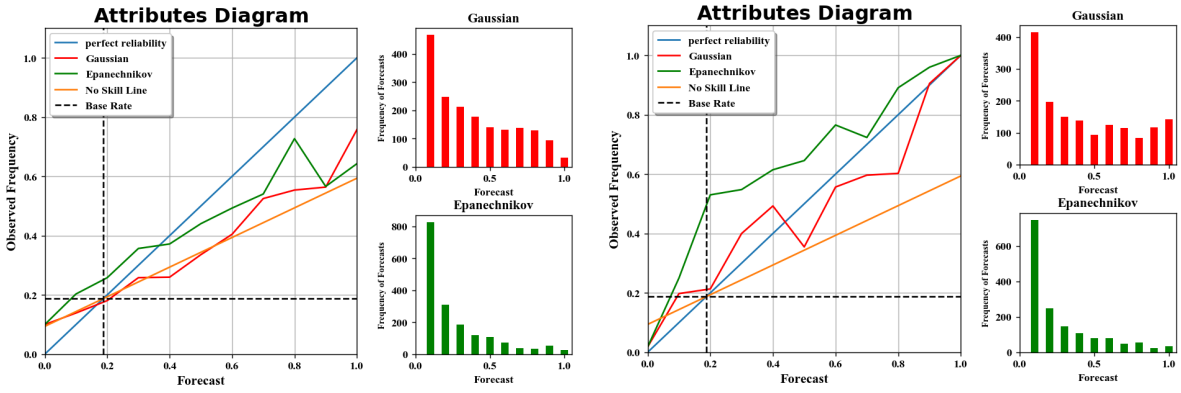
(d) Epanechnikov Practically Perfect Plume

Figure 5.1: The mesocyclone track overlaid on top of the Gaussian forecaster (a), Epanechnikov forecaster (b), practically perfect Gaussian (c), and practically perfect Epanechnikov (d) merged maximum forecasted probability at any time step over the duration of the Bennington, KS tornado case 25-26 May 2016.

## 5.1 Tornado

The mesocyclone track overlaid on top of forecaster and practically perfect merged maximum plume at any time step over the duration of the Bennington, KS tornado case for the Gaussian and Epanechnikov kernels are shown in Figure 5.1. The Epanechnikov kernel plume had somewhat fewer higher probability values and what appears to be significantly more of the lower probability values compared to the Gaussian kernel plume for both the forecaster and practically perfect plume. This was due to the Gaussian kernel going to zero probability at infinity while the Epanechnikov kernel goes to zero probability at 1.87 sigma where sigma was the standard deviation and the width of the Gaussian kernel. Sigma for all tornado and lightning cases was defined in the PHI Tool Code to be 500.

Scoring both the forecaster and practically perfect Epanechnikov plumes shows the impact of the decreased areal coverage of the higher probabilities. In Figure 5.2, the impact of the decreased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect plumes.

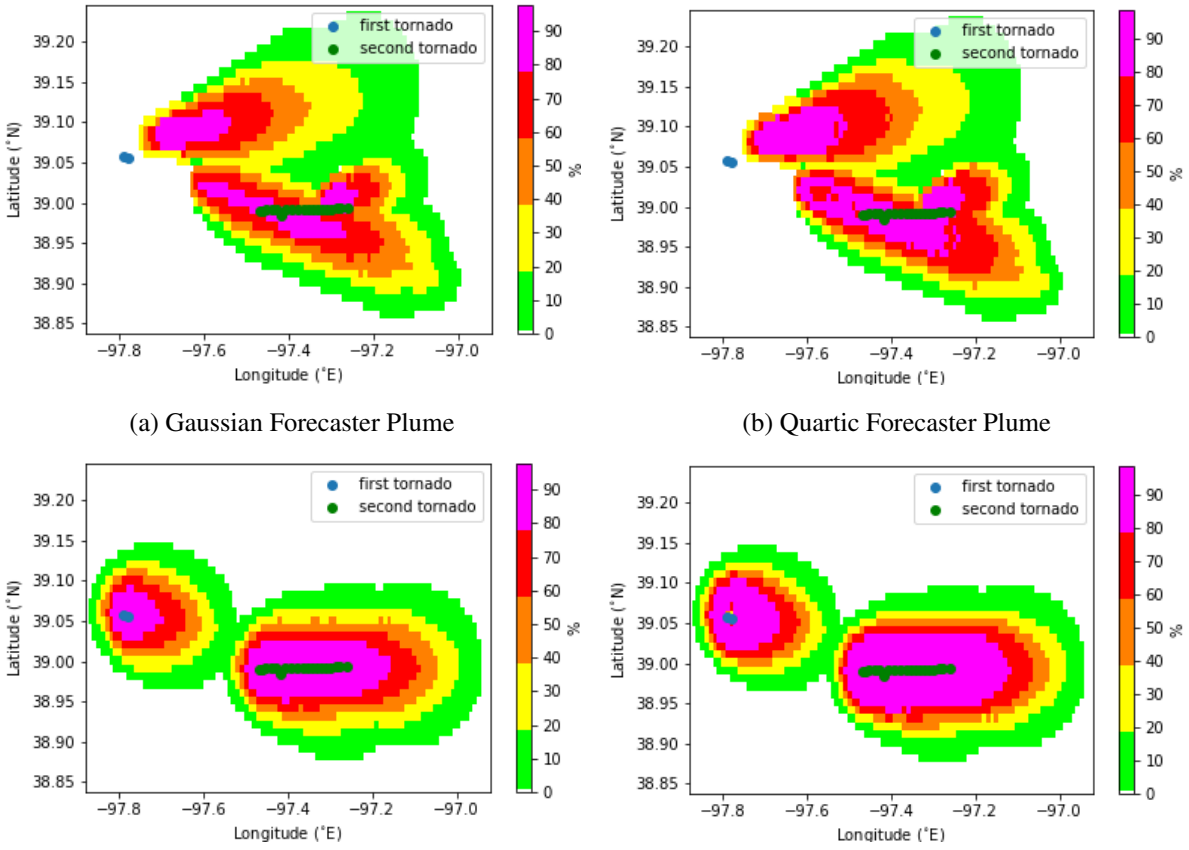



(a) Forecaster (b) Practically Perfect

Figure 5.2: The Gaussian and Epanechnikov forecaster (a) and practically perfect (b) attributes diagrams for the Bennington, KS tornado case 25-26 May 2016 at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Epanechnikov plume.

The resolution for the practically perfect Epanechnikov plumes decreased while the reliability increased compared to the Gaussian practically perfect plumes. The forecaster Epanechnikov plumes show a different behavior as the values of the reliability improved, but the values of resolution got worse.

The mesocyclone track overlaid on top of the forecaster and practically perfect merged maximum plume at any time step over the duration of the Bennington, KS tornado case for the Gaussian and Quartic kernels are shown in Figure 5.3.




(a) Gaussian Forecaster Plume

(b) Quartic Forecaster Plume

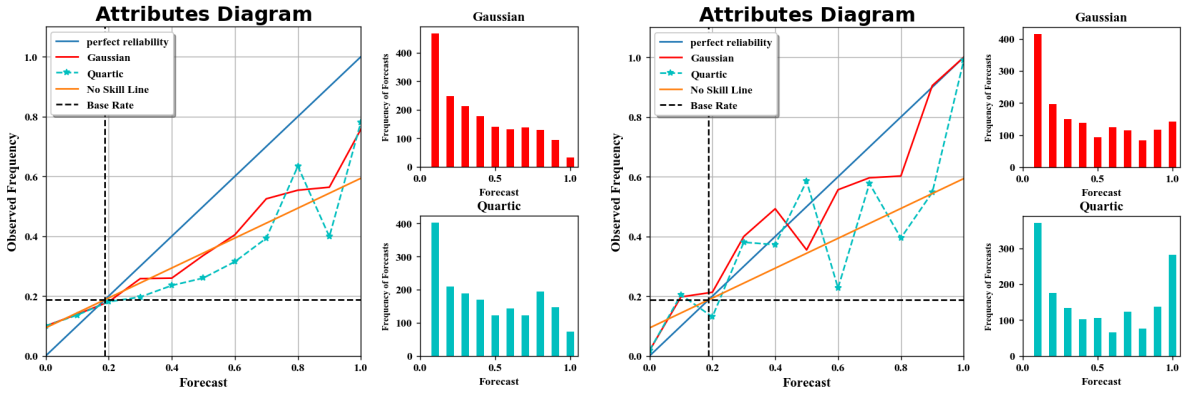
(c) Gaussian Practically Perfect Plume

(d) Quartic Practically Perfect Plume

Figure 5.3: The mesocyclone track overlaid on top of the Gaussian forecaster (a), Quartic forecaster (b), practically perfect Gaussian (c), and practically perfect Quartic (d) merged maximum forecasted probability at any time step over the duration of the Bennington, KS tornado case 25-26 May 2016.

The Quartic kernel plume had more higher probability values and about the same amount of the lower probability values compared to the Gaussian kernel plume for both the forecaster and practically perfect plumes. This was due to the Gaussian kernel going to zero probability at infinity while the Quartic kernel goes to zero probability at 1.87 sigma, similar to the Epanechnikov kernel, where sigma was the standard deviation and the width of the Gaussian kernel.

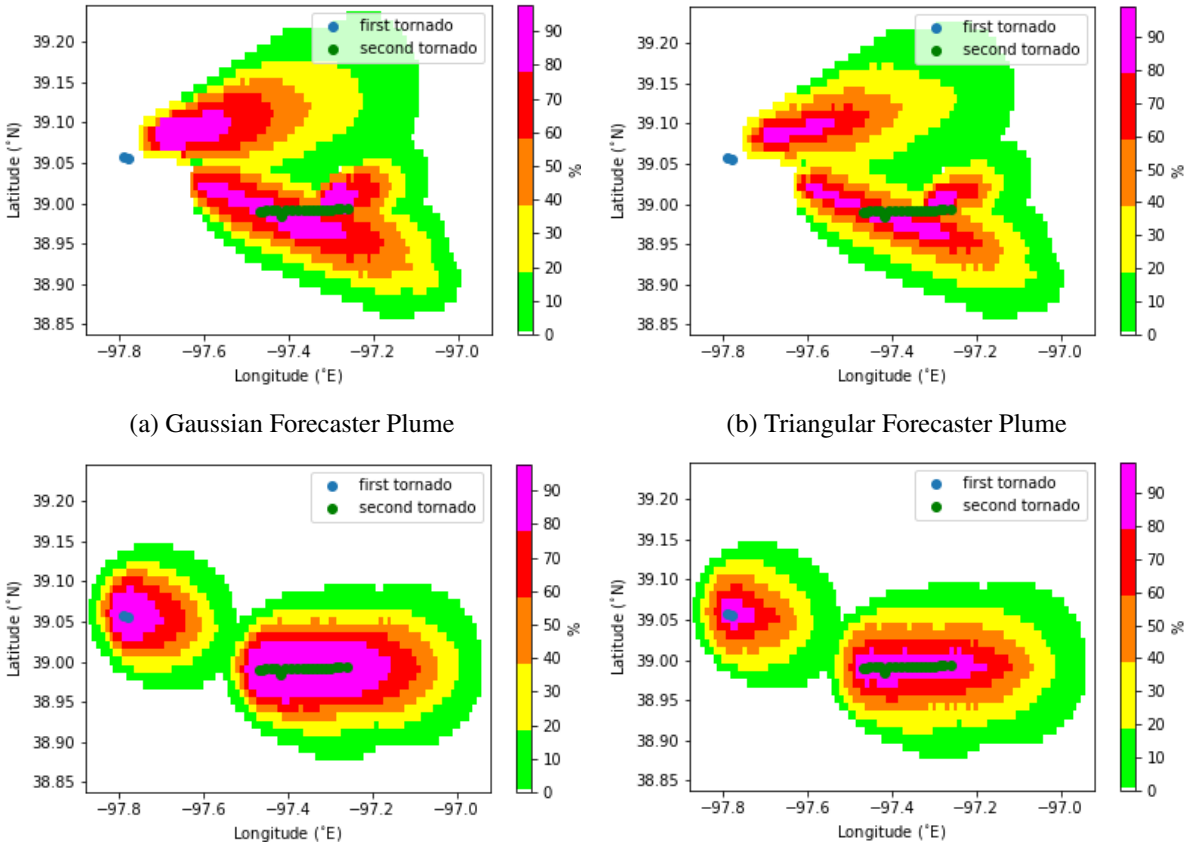
Scoring both the forecaster and practically perfect Quartic plumes shows the impact of the slightly increased areal coverage of the higher probabilities and the similarity to the Gaussian plumes. In Figure 5.4, the impact of the increased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect plumes. The resolution and reliability for the practically perfect Quartic plumes increased compared to the Gaussian practically perfect plumes. The forecaster Quartic plumes show a similar behavior as the values of the resolution improved, but the values of reliability got worse.




(a) Forecaster (b) Practically Perfect

Figure 5.4: The Gaussian and Quartic forecaster (a) and practically perfect (b) attributes diagrams for the Bennington, KS tornado case 25-26 May 2016 at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Quartic plume.

The mesocyclone track overlaid on top of the forecaster and practically perfect merged maximum plume at any time step over the duration of the Bennington, KS tornado case for the Gaussian and Triangular kernels are shown in Figure 5.5. Similar to the Epanechnikov kernel, the Triangular kernel plume had somewhat fewer higher probability values and more of the lower probability values compared to the Gaussian kernel plume for both the forecaster and practically perfect plume. This was due to the Gaussian kernel going to zero probability at infinity while the Triangular kernel goes to zero probability at one sigma where sigma was the standard deviation and the width of the Gaussian kernel.



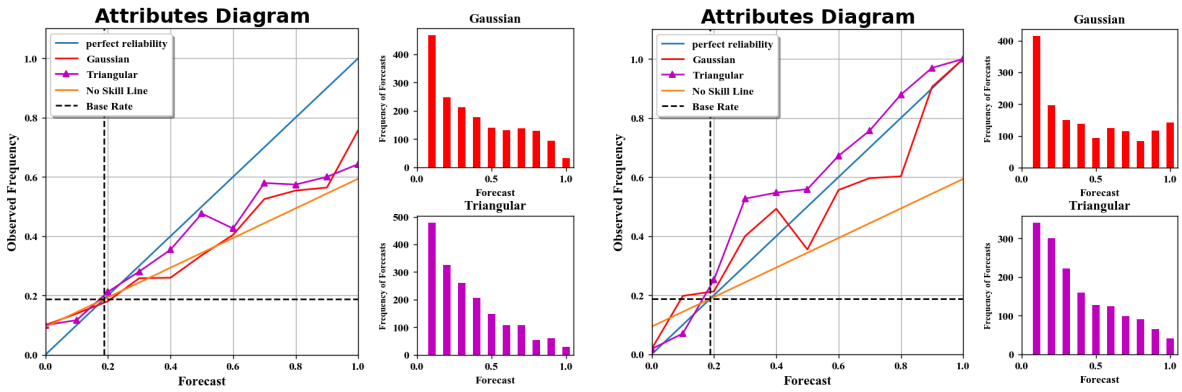

(a) Gaussian Forecaster Plume

(b) Triangular Forecaster Plume

(c) Gaussian Practically Perfect Plume

(d) Triangular Practically Perfect Plume

Figure 5.5: The mesocyclone track overlaid on top of the Gaussian forecaster (a), Triangular forecaster (b), practically perfect Gaussian (c), and practically perfect Triangular (d) merged maximum forecasted probability at any time step over the duration of the Bennington, KS tornado case 25-26 May 2016.




(a) Forecaster (b) Practically Perfect

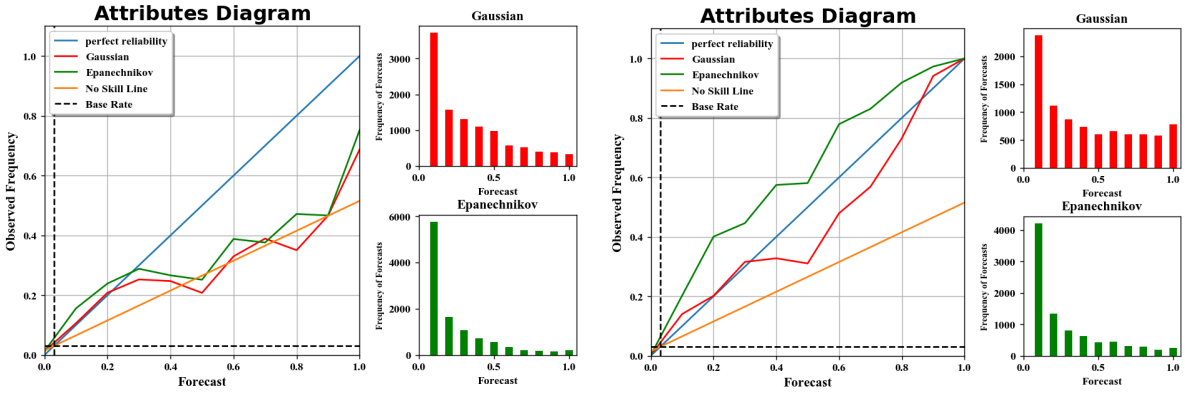
Figure 5.6: The Gaussian and Triangular forecaster (a) and practically perfect (b) attributes diagrams for the Bennington, KS tornado case 25-26 May 2016 at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Triangular plume.

The Triangular kernel had slightly more of the higher probabilities than the Epanechnikov kernel.

Scoring both the forecaster and practically perfect Triangular plumes shows the impact of the decreased areal coverage of the higher probabilities. In Figure 5.6, the impact of the decreased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect plumes. The resolution and reliability for the practically perfect Triangular plumes increased slightly compared to the Gaussian practically perfect plumes. The forecaster Triangular plumes show a different behavior as the values of the reliability improved, but the values of resolution stayed about the same. In Figure 5.6 the Triangular kernel line is slightly closer to the perfect reliability line than the Gaussian kernel line for the forecaster but not the practically perfect.

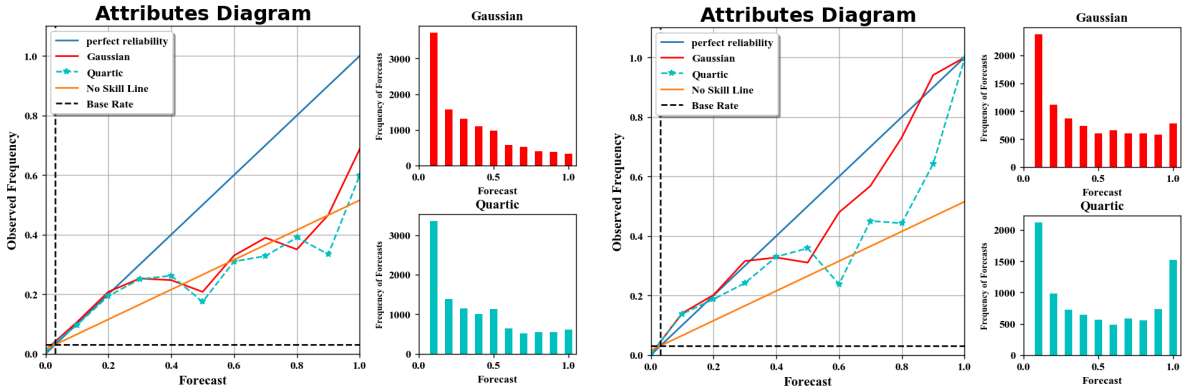
Scoring both the forecaster and practically perfect combined Meso Epanechnikov case shows the impact of the decreased areal coverage of the higher probabilities for all tornado Meso plumes. In Figure 5.7, the impact of the decreased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect combined

Meso case. The resolution for the practically perfect Epanechnikov combined Meso case decreased while the reliability increased compared to the practically perfect Gaussian d Meso case.




(a) Forecaster (b) Practically Perfect

Figure 5.7: The Gaussian and Epanechnikov (a) and practically perfect (b) Meso attributes diagrams for all cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of Forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Epanechnikov plume.



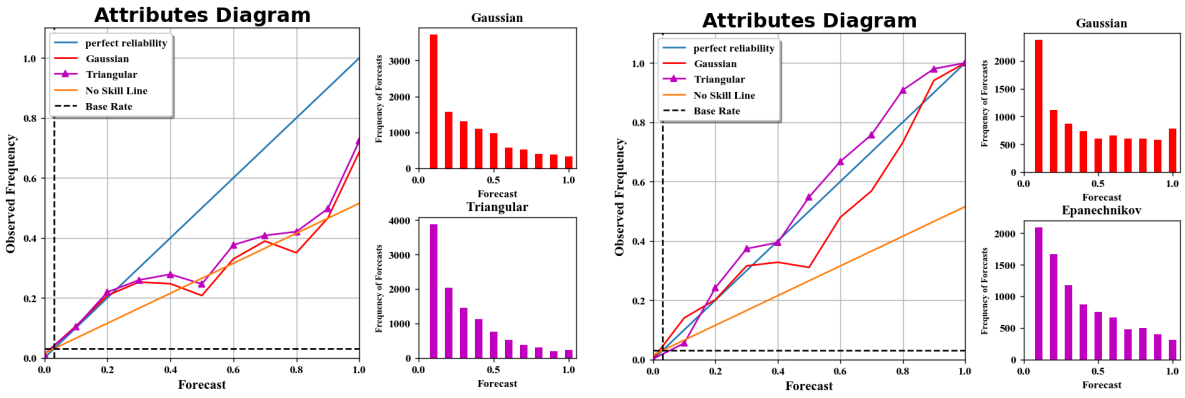

(a) Forecaster (b) Practically Perfect

Figure 5.8: The Gaussian and Quartic (a) and practically perfect (b) Meso attributes diagrams for all cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Quartic plume.

The forecaster combined Meso Epanechnikov case shows a different behavior as the values of the reliability improved, but the values of resolution got worse.

Scoring both the forecaster and practically perfect Quartic combined Meso case shows the impact of the slightly increased areal coverage of the higher probabilities and the similarity to the Gaussian combined Meso case for all tornado Meso plumes. In Figure 5.8, the impact of the increased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect combined Meso case. The resolution and reliability for the practically perfect Quartic combined Meso cases increased compared to the practically perfect Gaussian combined Meso case. The forecaster Quartic combined Meso case shows a similar behavior as the values of the resolution improved, but the values of reliability got worse.

Scoring both the forecaster and practically perfect Triangular combined Meso case shows the impact of the decreased areal coverage of the higher probabilities for all tornado Meso plumes. In Figure 5.9, the impact of the decreased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect combined Meso case.

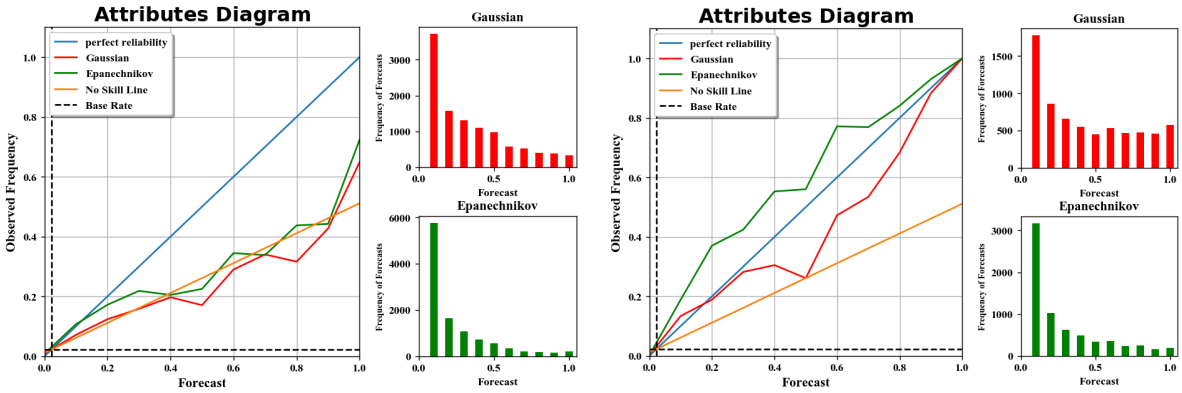



(a) Forecaster (b) Practically Perfect

Figure 5.9: The Gaussian and Triangular (a) and practically perfect (b) Meso attributes diagrams for all cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Triangular plume.

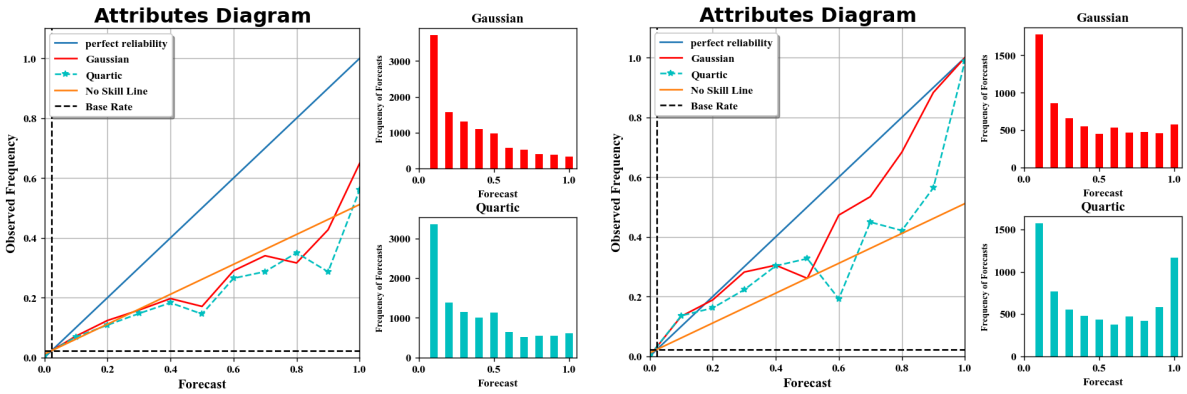
The resolution for the practically perfect Triangular combined Meso case increased while the reliability for the practically perfect Triangular combined Meso case decreased compared to the practically perfect Gaussian combined Meso case. The forecaster Triangular combined Meso case shows a different behavior as the values of the reliability improved, but the values of resolution got slightly worse.

Scoring both the forecaster and practically perfect combined Tor Official Epanechnikov case shows the impact of the decreased areal coverage of the higher probabilities for all tornado Tor Official plumes. In Figure 5.10, the impact of the decreased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect combined Tor Official case. The resolution for the practically perfect Epanechnikov combined Tor Official case decreased while the reliability increased compared to the practically perfect Gaussian combined Tor Official case. The forecaster combined Tor Official Epanechnikov case shows a different behavior as the values of the reliability improved, but the values of resolution got worse.




(a) Forecaster (b) Practically Perfect

Figure 5.10: The Gaussian and Epanechnikov (a) and practically perfect (b) Tor Official attributes diagrams for all cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Epanechnikov plume.

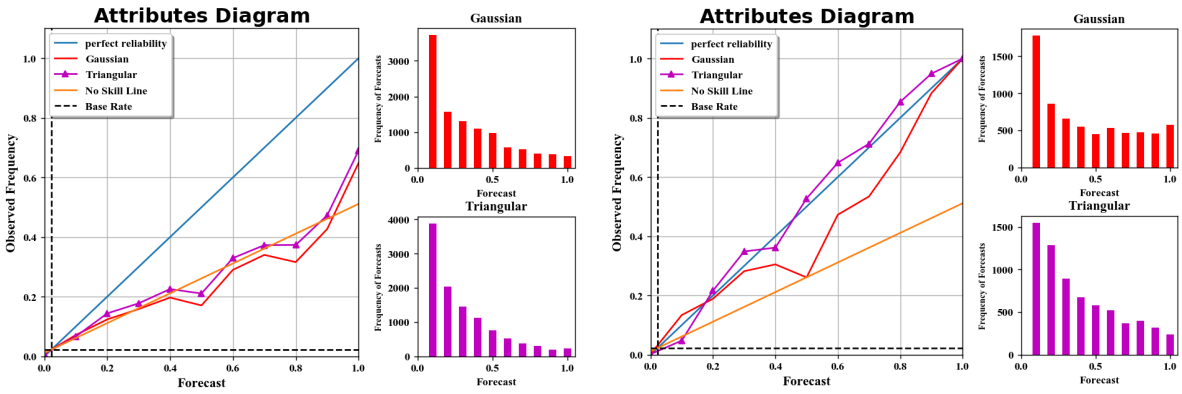



(a) Forecaster (b) Practically Perfect

Figure 5.11: The Gaussian and Quartic (a) and practically perfect (b) Tor Official attributes diagrams for all cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Quartic plume.

Scoring both the forecaster and practically perfect Quartic combined Tor Official case shows the impact of the slightly increased areal coverage of the higher probabilities and the similarity to the Gaussian combined Tor Official case for all tornado Tor Official plumes. In Figure 5.11, the impact of the increased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect combined Tor Official case. The resolution and reliability for the practically perfect Quartic combined Tor Official cases increased compared to the practically perfect Gaussian combined Tor Official case. The forecaster Quartic combined Tor Official case shows a similar behavior as the values of the resolution improved, but the values of reliability got worse.

Scoring both the forecaster and practically perfect Triangular combined Tor Official case shows the impact of the decreased areal coverage of the higher probabilities for all tornado Tor Official plumes. In Figure 5.12, the impact of the decreased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect combined Tor Official case.




(a) Forecaster (b) Practically Perfect

Figure 5.12: The Gaussian and Triangular (a) and practically perfect (b) Tor Official attributes diagrams for all cases at a 7.5 km radial distance away from the mesocyclone track using a 10 probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Triangular plume.

The resolution for the practically perfect Triangular combined Tor Official case increased while the reliability for the practically perfect Triangular combined Tor Official case decreased compared to the practically perfect Gaussian combined Tor Official case. The forecaster Triangular combined Tor Official case shows a different behavior as the values of the reliability improved, but the values of resolution got slightly worse.

For all cases and the combined cases except for the Columbia, SC Meso the Quartic kernel had the worst reliability out of all the kernels (Figure 5.13). For all cases and the combined cases except for the Dodge City, KS and Tor Official combined cases the Quartic kernel had a higher resolution than the Gaussian kernel. For most of those cases the Quartic kernel had the highest resolution out of all kernels. For all cases and the combined cases the Epanechnikov kernel has the worst resolution out of all the kernels. For all cases and the combined cases except the Columbia, SC Meso case the Epanechnikov kernel had a lower reliability value than the Gaussian kernel. For all cases and the combined cases except the Columbia, SC Meso and the combined Meso cases the Triangular kernel had a lower reliability value than the Gaussian kernel. For all cases and the combined cases except the Denver, CO case the Triangular kernel had a lower resolution value than the Gaussian

kernel. For all cases and the combined cases except for the Columbia, SC Meso case the Triangular and Epanechnikov kernels offered some improvement over the Gaussian kernel.

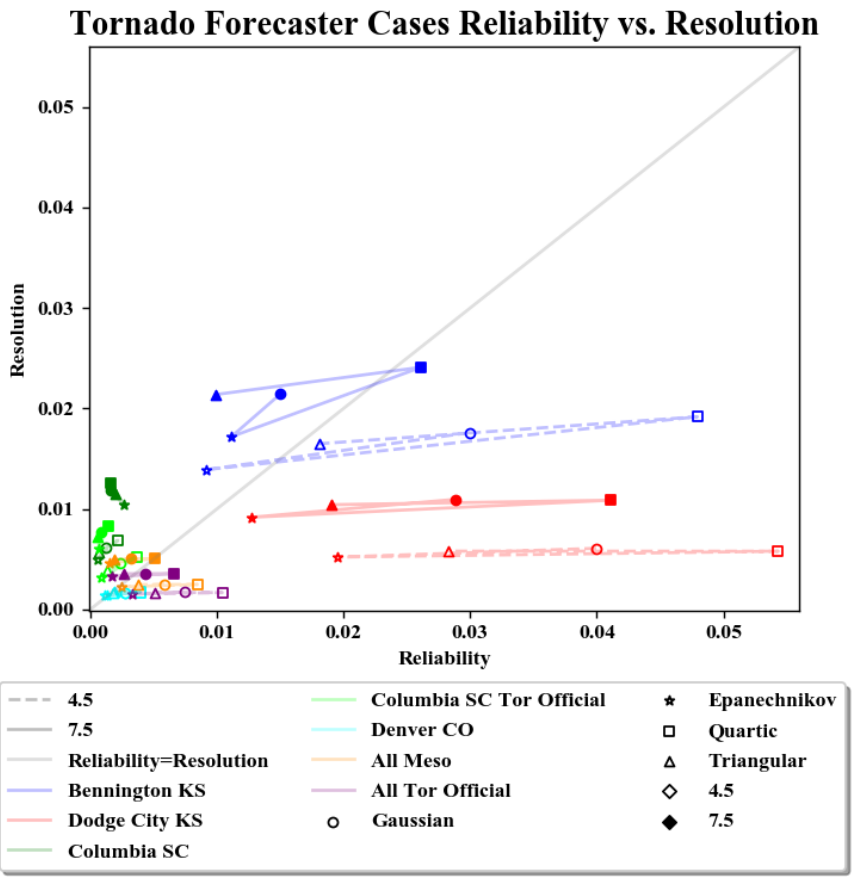
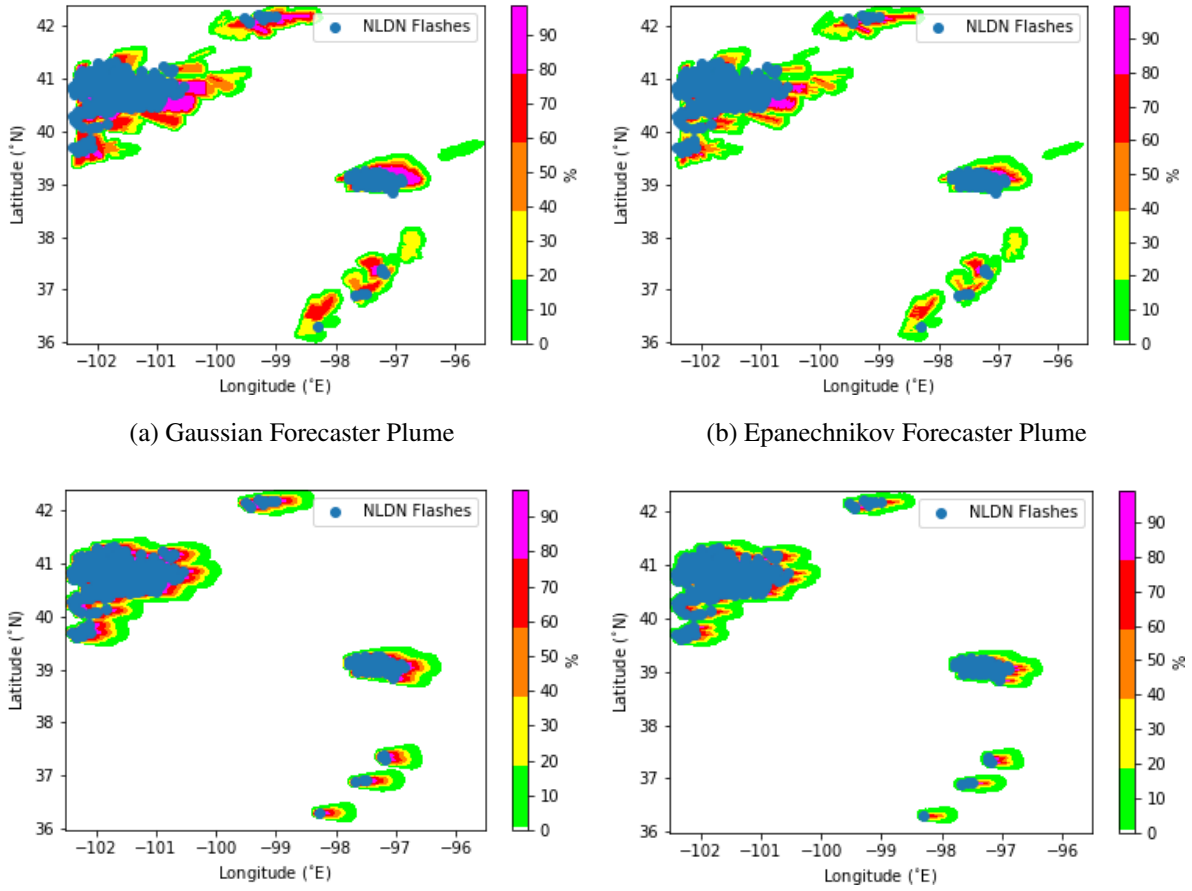



Figure 5.13: The forecaster reliability vs resolution terms for all tornado cases and the combined cases for a 7.5 km and 4.5 km event definition and all kernels.

For all cases and the combined cases using a 7.5 km compared to a 4.5 km event definition for all kernels improved the resolution values (Figure 5.13). For the Bennington, KS tornado, Denver, CO, and Columbia, SC Meso cases using a 4.5 km compared to a 7.5 km event definition for the Epanechnikov kernel improved the reliability values. For the Denver, CO and Columbia, SC Meso cases using a 4.5 km compared to a 7.5 km event definition for the Gaussian and Triangular kernels improved the reliability values. For the Denver, CO case using a 4.5 km compared to a 7.5 km event definition for the Quartic kernel improved the reliability value.

## 5.2 Lightning

The NLDN flashes overlaid on top of the forecaster and practically perfect merged maximum plume at any time step over the duration of the Bennington, KS lightning case for the Gaussian and Epanechnikov kernels are shown in Figure 5.14. The Epanechnikov kernel plume had somewhat less higher probability values and more of the lower probability values compared to the Gaussian kernel plume for both the forecaster and practically perfect plume. This was due to the reasons mentioned in section 5.1.




(a) Gaussian Forecaster Plume

(b) Epanechnikov Forecaster Plume

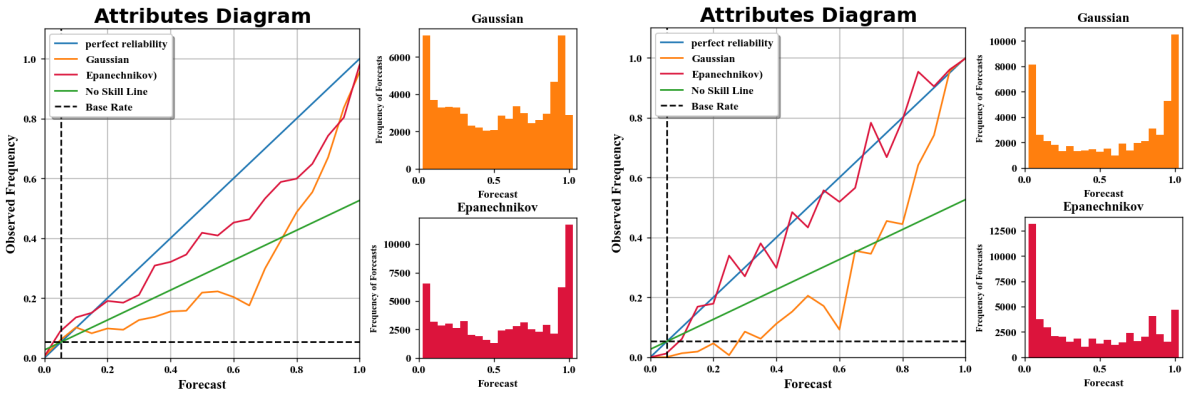
(c) Gaussian Practically Perfect Plume

(d) Epanechnikov Practically Perfect Plume

Figure 5.14: The NLDN flashes overlaid on top of the Gaussian forecaster (a), Epanechnikov forecaster (b), practically perfect Gaussian (c), and practically perfect Epanechnikov (d) merged maximum forecasted probability at any time step over the duration of the Bennington, KS lightning case 25-26 May 2016.

Scoring both the forecaster and practically perfect Epanechnikov plumes shows the impact of the decreased areal coverage of the higher probabilities. In Figure 5.15, the impact of the decreased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect plumes. The resolution and reliability for the practically perfect Epanechnikov plumes decreased compared to the Gaussian practically perfect plumes. The forecaster Epanechnikov plumes show a similar behavior as the values of the reliability improved, but the values of resolution got worse. At most of the probability values the practically perfect attributes diagram was more reliable than the forecaster attributes diagram for the Epanechnikov but not the Gaussian kernel.

The NLDN flashes overlaid on top of the forecaster and practically perfect merged maximum plume at any time step over the duration of the Bennington, KS lightning case for the Gaussian and Quartic kernels are shown in Figure 5.16.

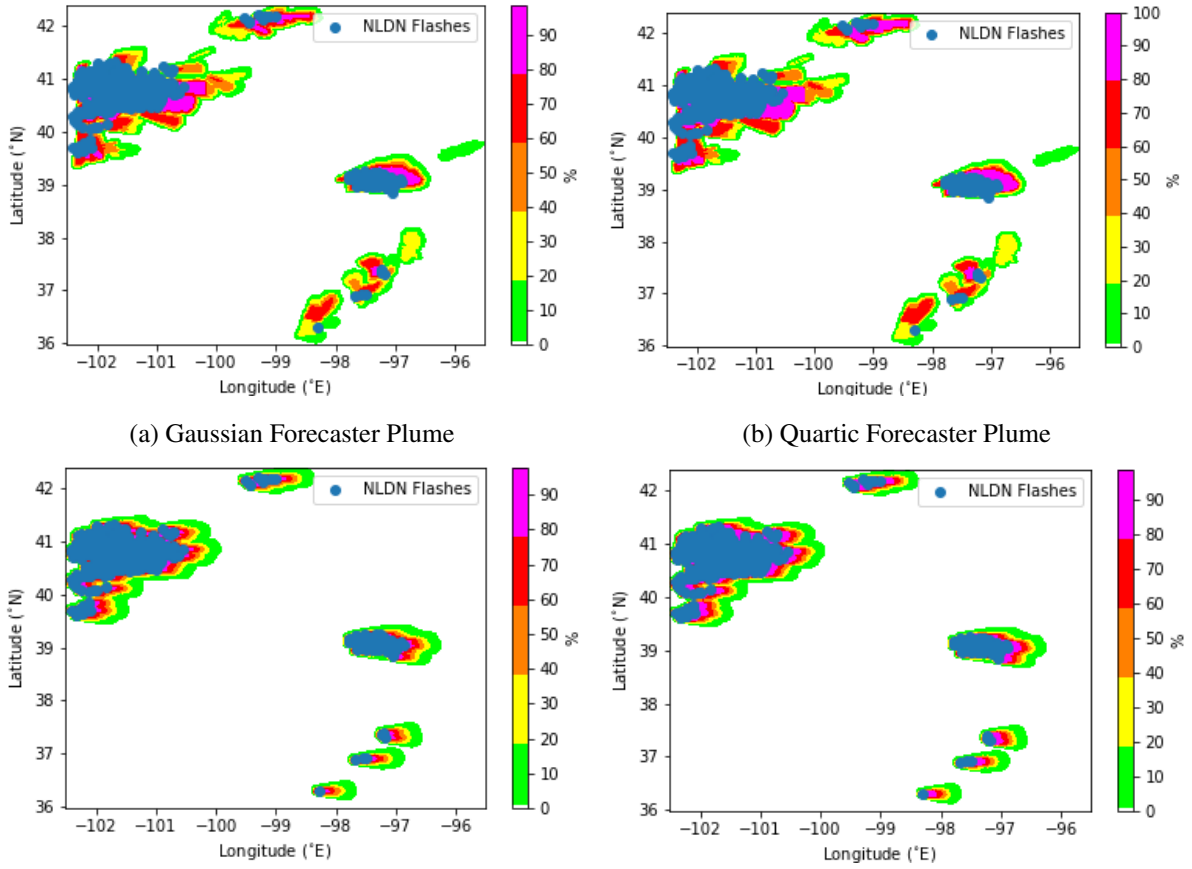



(a) Forecaster (b) Practically Perfect

Figure 5.15: The Gaussian and Epanechnikov forecaster (a) and practically perfect (b) attributes diagrams for the Bennington, KS lightning case 25-26 May 2016 at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Epanechnikov plume.

The Quartic kernel plume had more higher probability values and about the same amount of the lower probability values compared to the Gaussian kernel plume for both the forecaster and practically perfect plume. This was due to the reasons mentioned in section 5.1.

Scoring both the forecaster and practically perfect Quartic plumes shows the impact of the slightly increased areal coverage of the higher probabilities and the similarity to the Gaussian plumes. In Figure 5.17, the impact of the increased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect plumes.



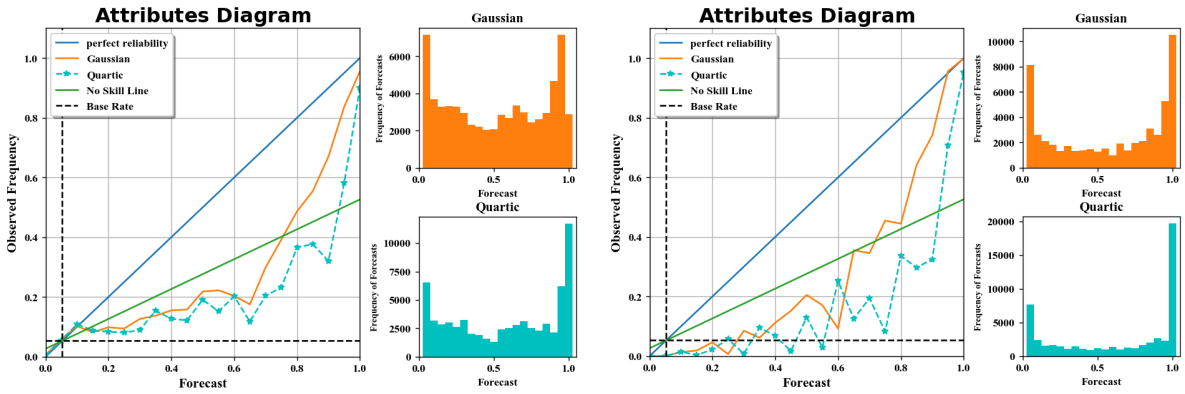

(a) Gaussian Forecaster Plume

(b) Quartic Forecaster Plume

(c) Gaussian Practically Perfect Plume

(d) Quartic Practically Perfect Plume

Figure 5.16: The NLDN flashes overlaid on top of the Gaussian forecaster (a), Quartic forecaster (b), practically perfect Gaussian (c), and practically perfect Quartic (d) merged maximum forecasted probability at any time step over the duration of the Bennington, KS lightning case 25-26 May 2016.

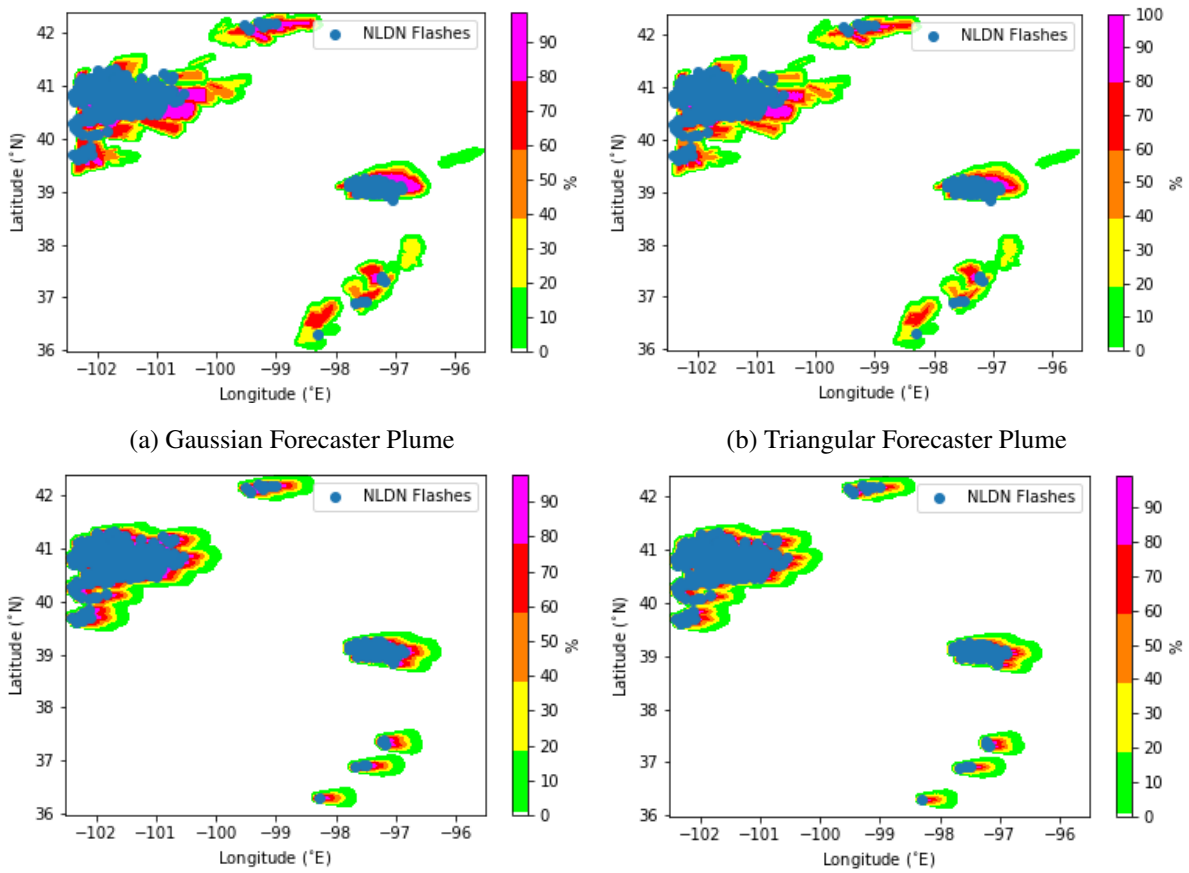



(a) Forecaster (b) Practically Perfect

Figure 5.17: The Gaussian and Quartic forecaster (a) and practically perfect (b) attributes diagrams for the Bennington, KS lightning case 25-26 May 2016 at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Quartic plume.

The resolution and reliability for the practically perfect Quartic plumes increased compared to the Gaussian practically perfect plumes. The forecaster Quartic plumes shows a similar behavior as the values of the resolution improved, but the values of reliability got worse. At most of the probability values the forecaster attributes diagram was more reliable than the practically perfect attributes diagram for both kernels.

The NLDN flashes overlaid on top of the forecaster and practically perfect merged maximum plume at any time step over the duration of the Bennington, KS lightning case for the Gaussian and Triangular kernels are shown in Figure 5.18. The Triangular, similarly to the Epanechnikov, kernel plume had somewhat less higher probability values and considerably more of the lower probability values compared to the Gaussian kernel plume for both the forecaster and practically perfect plumes. This was due to the same reason as mentioned in section 5.1.




(a) Gaussian Forecaster Plume

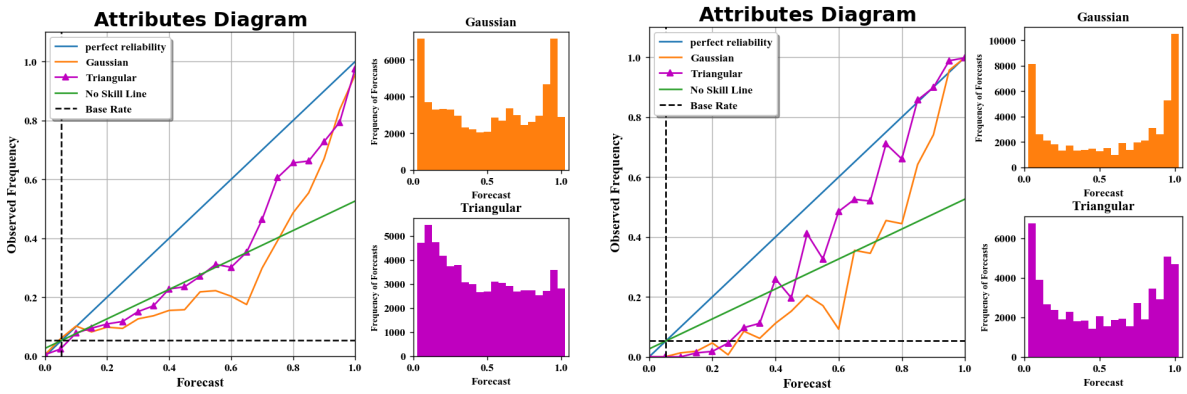
(b) Triangular Forecaster Plume

(c) Gaussian Practically Perfect Plume

(d) Triangular Practically Perfect Plume

Figure 5.18: The NLDN flashes overlaid on top of the Gaussian forecaster (a), Triangular forecaster (b), practically perfect Gaussian (c), and practically perfect Triangular (d) merged maximum forecasted probability at any time step over the duration of the Bennington, KS lightning case 25-26 May 2016.

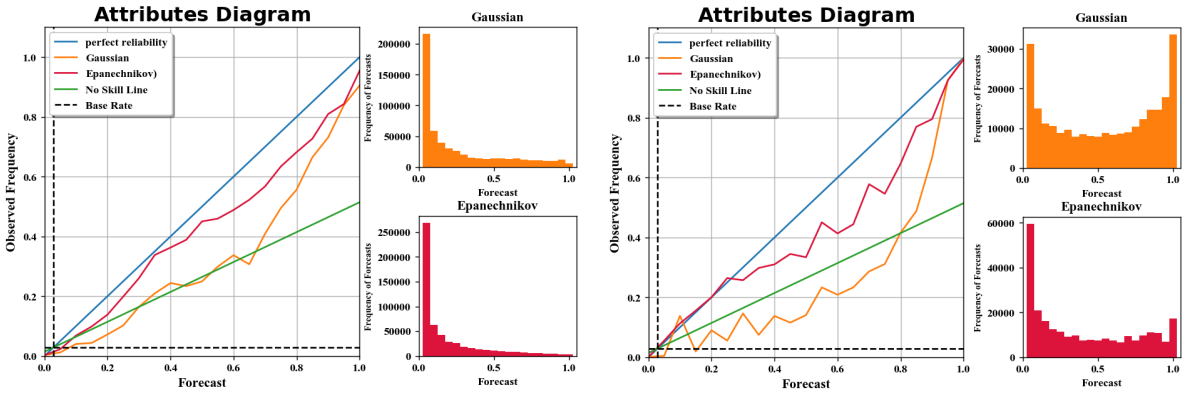
Scoring both the forecaster and practically perfect Triangular plumes shows the impact of the decreased areal coverage of the higher probabilities. In Figure 5.19, the impact of the decreased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect plumes. The resolution and reliability for the practically perfect Triangular plumes decreased compared to the Gaussian practically perfect plumes. The forecaster Triangular plumes show a similar behavior as the values of the reliability improved, but the values of resolution got worse. At most of the probability values the practically perfect attributes diagram was more reliable than the forecaster attributes diagram for both kernels.




(a) Forecaster (b) Practically Perfect

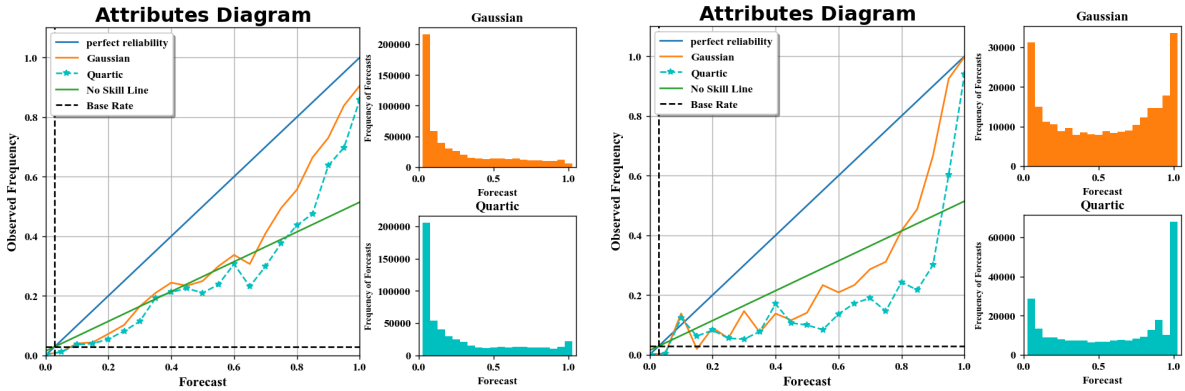
Figure 5.19: The Gaussian and Triangular forecaster (a) and practically perfect (b) attributes diagrams for the Bennington, KS lightning case 25-26 May 2016 at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Triangular plume.

Scoring both the forecaster and practically perfect Epanechnikov combined case shows the impact of the decreased areal coverage of the higher probabilities for all lightning cases. In Figure 5.20, the impact of the decreased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect combined case. The resolution and reliability for the practically perfect Epanechnikov combined case decreased compared to the practically perfect Gaussian combined case. The forecaster Epanechnikov combined case shows a similar behavior as the values of the reliability improved, but the values of resolution got worse. At most of the probability values the forecaster attributes diagram was more reliable than the practically perfect attributes diagram for the Epanechnikov but not the Gaussian kernel. At the extremely high and low probability values the practically perfect attributes diagram was more reliable than the forecaster attributes diagram for the Gaussian and Epanechnikov kernel.




(a) Forecaster (b) Practically perfect

Figure 5.20: The Gaussian and Epanechnikov forecaster (a) and practically perfect (b) attributes diagrams for all lightning cases at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Epanechnikov plume.



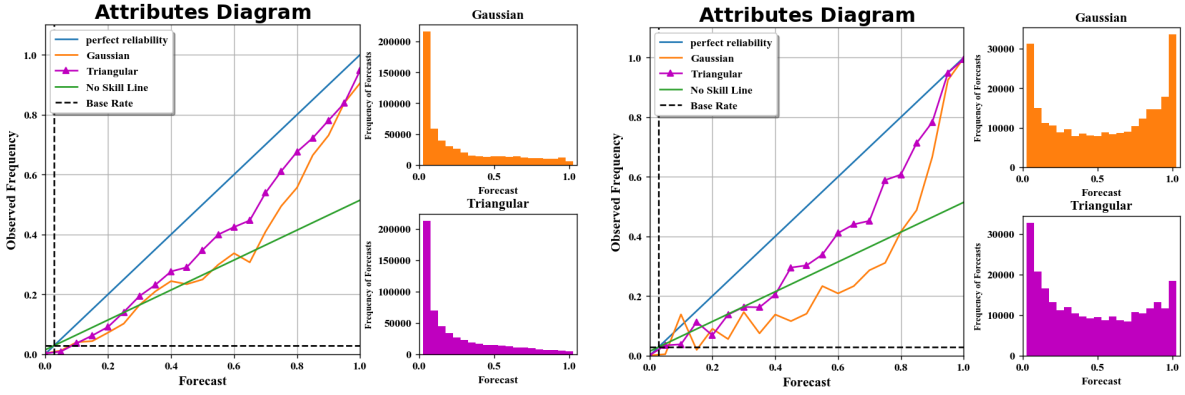

(a) Forecaster (b) Practically Perfect

Figure 5.21: The Gaussian and Quartic forecaster (a) and practically perfect (b) attributes diagrams for all lightning cases at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Quartic plume.

Scoring both the forecaster and practically perfect Quartic combined case shows the impact of the slightly increased areal coverage of the higher probabilities and the similarity to the Gaussian combined case for all lightning cases.In Figure 5.21, the impact of

the increased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect combined case. The resolution and reliability for the practically perfect Quartic combined case increased compared to the practically perfect Gaussian combined case. The forecaster Quartic combined case shows a similar behavior as the values of the resolution improved, but the values of reliability got worse. At most of the probability values the forecaster attributes diagram was more reliable than the practically perfect attributes diagram for both kernels.

Scoring both the forecaster and practically perfect Triangular combined case shows the impact of the decreased areal coverage of the higher probabilities for all lightning cases. In Figure 5.22, the impact of the decreased areal coverage is depicted in the histogram bins of probability for both the forecaster and practically perfect combined case. The resolution and reliability for the practically perfect Triangular combined case decreased compared to the practically perfect Gaussian combined case. The forecaster Triangular combined case show a similar behavior as the values of the reliability improved, but the values of resolution got worse.




(a) Forecaster (b) Practically Perfect

Figure 5.22: The Gaussian and Triangular forecaster (a) and practically perfect (b) attributes diagrams for all lightning cases at a 7.5 km radial distance away from the NLDN flashes using a five probability (forecast) bin. On the top right beside the attributes diagrams is the frequency of forecasts for the Gaussian plume. On the bottom right beside the attributes diagrams is the frequency of forecasts for the Triangular plume.

At the extremely high probability values was the only location where the practically perfect attributes diagram was more reliable than the forecaster attributes diagram for the Gaussian and Triangular kernels.

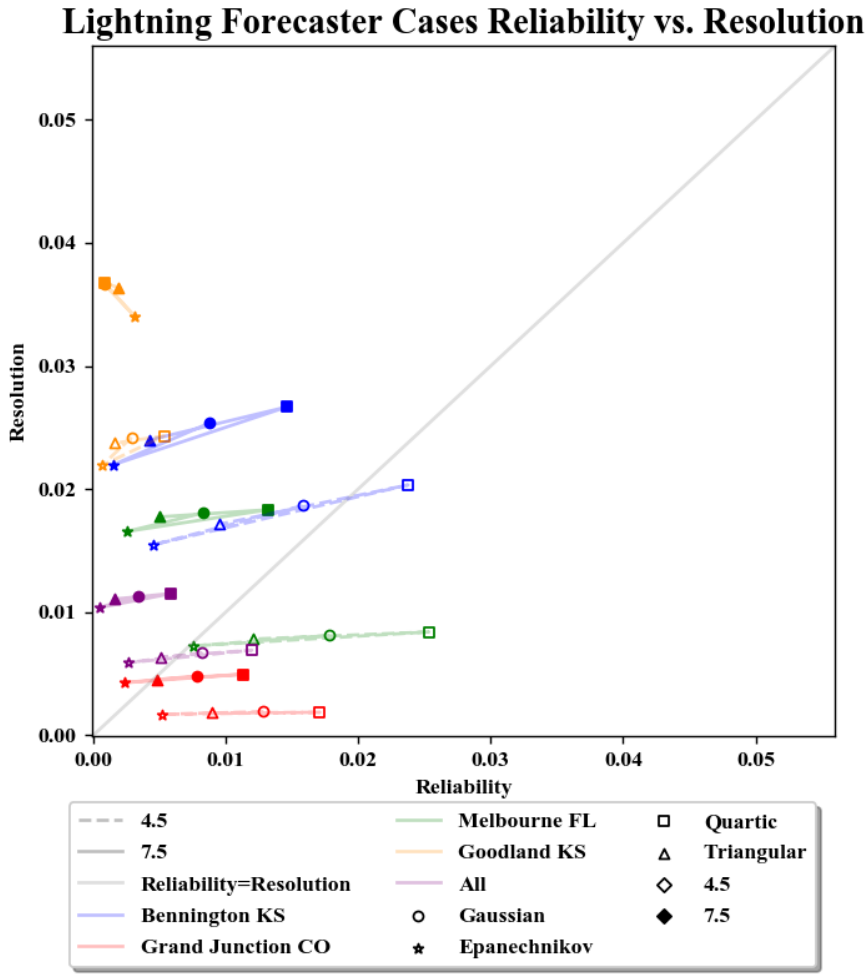



Figure 5.23: The forecaster reliability vs resolution terms for all lightning cases and the combined case for a 7.5 km and 4.5 km event definition and all kernels.

For all cases and the combined case except for the Goodland, KS case the Quartic kernel had the highest reliability out of all the kernels (Figure 5.23). For all cases and the combined case the Quartic kernel had the highest resolution out of all the kernels. For all cases and the combined case the Epanechnikov kernel had the lowest resolution out of all the kernels. For all cases and the combined case using a 7.5 km event definition for all kernels improved the resolution values (Figure 5.23). For all cases and the combined case except for the Goodland, KS case using a 7.5 km event definition for all kernels improved the reliability values. For all cases and the combined case except for the Goodland,

KS case the Epanechnikov had the lowest reliability out of all the kernels. For all cases and the combined case except for the Goodland, KS case the Triangular kernel had a lower reliability than the Gaussian kernel. For all cases the Triangular kernel had a lower resolution than the Gaussian kernel.

For the Goodland, KS case using the Triangular and Epanechnikov kernel with a 4.5 km event definition improved the reliability values. For all cases and the combined case using a 4.5 km event definition the individual kernels had the same relative improvement as using a 7.5 km event definition excluding the Goodland, KS case.

## 5.3 Discussion

Generally, for both the tornado and lightning cases using either the Epanechnikov or the Triangular kernel provided some improvement in the reliability term and no improvement in the resolution term. For the tornado cases there were less improvement and more exceptions than for lightning. For the tornado and lightning cases changing the event definition from 4.5 km to 7.5 km improves the resolution terms for all four kernels. Generally, for lightning, changing the event definition from 4.5 km to 7.5 km improves the reliability terms for all four kernels. The results presented in this section are very limited which makes generalization challenging. The questions that arise from these results are is it worth it to use a different kernel for tornado than for lightning and is the improvement in the Epanechnikov and/or Triangular kernels for lightning enough to justify making a change. For lightning the improvement in the reliability for the Epanechnikov compared to Gaussian kernel counteracts the worsening of the resolution for the Epanechnikov compared to the Gaussian kernel. The smaller area of higher probabilities for the Epanechnikov and/or Triangular should out-perform the Gaussian distribution on a 'perfect' forecast. But if there is some error on the track location relative to the plume, the penalty will be greater as a result.

## 5.4 Summary

- Additional kernels (Epanechnikov, Triangular, Quartic) were examined to see if there was added value in different methods of probability interpolation beyond the Gaussian method first tested in the HWT.
- Both the Epanechnikov and the Triangular kernels had fewer higher probability values and increased coverage of lower probability values than the Gaussian kernel.
- The Epanechnikov and Triangular kernel provided improvement in the reliability term for both the tornado and lighting cases.
- The Epanechnikov kernel provided the most improvement in the reliability term for both the tornado and lightning cases.
- The Quartic kernel had the best resolution term, but the worst reliability term out of all four kernels for both the tornado and lightning cases for a 4.5 km and a 7.5 km event definition.
- Changing from a 4.5 km to a 7.5 km event definition improved the resolution terms for all four kernels for both the tornado and lightning cases.
- There was one lightning and three tornado cases that had improved reliability when using a 4.5 km event definition instead of using a 7.5 km event definition for the Epanechnikov kernel.
- Generally, the ROC diagrams and AUC values did not change when using the four different kernels.

# Chapter 6

# Summary and Conclusion

The current watch/warning paradigm poses a number of challenges outlined in NWS assessments following large events such as the 2011 Joplin, MO tornado (National Weather Service 2011b). FACETs was created in part to address these problems and allow for better communication of risks from the NWS to end-users and the public (Rothfusz et al. 2018). FACETs is a proposed framework to modify the current deterministic watch and warning system to a continuum of information with the use of PHI. During 2016 and 2017 experiments in the HWT, NWS forecasters created PHI plumes to produce a gridded probabilistic risk of severe and hazardous weather including lightning, wind, hail, and tornadoes (Karstens et al. 2018). The PHI Tool currently uses a Gaussian kernel applied to the probabilities to create the gridded probability plumes. The goals of this study were to (1) provide an event definition for each hazard, tornado and lightning, (2) create a method to verify the tornado and lightning plumes that were created by forecasters and the automated guidance (for lightning only) in the NOAA HWT, and (3) determine if the Gaussian kernel or a different kernel should be used to be applied to the forecaster and automated (for the lightning hazard only) probabilities to create the plumes for both the tornado and lightning hazards.

## 6.1 Event Definition and Practically Perfect Plumes

Limiting the event definition to an extremely precise location, such as 0.5 km away from a specific point of a hazard provided little meaning in terms of verification results. The forecast versus observed frequency shown in the attributes diagram in Chapter 4 and in terms of both reliability and resolution did not adequately represent what was a subjectively good forecast for both the tornado and lightning hazards. Therefore, using the concept of neighborhood verification from Ebert (2008) and Ebert (2009) a 4.5 km and 7.5 km radius away

from the mesocyclone track or NLDN flashes for the tornado or lightning cases was used. Generally, for both the tornado and lightning, and the combined tornado and combined lightning forecaster cases using a 7.5 km event definition improved the appearance of the attributes diagram and the reliability and resolution terms for the forecaster plumes. However, a 7.5 km event definition reduced the AUC calculated within the ROC diagrams when compared to a 4.5 km definition for both tornado and lightning hazards. Yet the AUC remained greater than 0.7 for both hazards at a 7.5 km event definition which indicates that a yes forecast can be discriminated. This is the primary use of the AUC, so even though it will decrease, the improvement in the resolution (increase) and reliability (decrease) by changing the event definition is justified. Therefore, based on these limited results, when verifying the tornado and lightning forecaster plumes an event definition or neighborhood of 7.5 km should be used.

The concept of practically perfect forecasts from Hitchens et al. (2013) was used to create practically perfect plumes to determine what the best possible forecast could be knowing the results and limiting the constraints to that of what the forecaster had. This provides a way to fairly judge the forecaster skill and to evaluate different event definitions. Practically perfect plumes were created for the tornado and lightning cases using a 4.5 km and a 7.5 km radius around the mesocyclone or NLDN flashes. Generally, for both the tornado and lightning and the combined tornado and combined lightning practically perfect cases using the Gaussian kernel and a 7.5 km compared to a 4.5 km event definition improved the appearance of the attributes diagram and the reliability and resolution terms. However, for both the tornado and lightning and the combined tornado and combined lightning practically perfect cases using a 7.5 km compared to a 4.5 km event definition, the ROC diagrams had a similar appearance and the AUC stayed about the same. Generally, for a 7.5 km event definition using the Gaussian kernel for both the tornado and lightning and the combined tornado and lightning practically perfect compared to the forecaster cases improved the appearance of the attributes diagrams and ROC diagrams and the reliability, resolution, and

AUC terms. Based on these limited results, to determine the optimal tornado and lightning attributes diagrams, ROC diagrams, resolution terms, reliability terms, and AUC values could be using a 7.5 km event definition practically perfect plumes, attributes diagrams, and ROC diagrams, need to be created.

## 6.2 Kernels

For the four tornado and lightning cases the Gaussian kernel in addition to the Epanechnikov, Quartic, and Triangular kernels were tested for the smoothing of the probabilities of the plumes. When using the Epanechnikov and/or Trinagular kernel there is the implication that the forecaster is more confidant in their forecast than when using the Gaussian kernel. Additionally, using the Epanechnikov kernel could indicate to the general public that the forecaster is more confident in their forecast. Also for lightning when using either the Epanechnikov or Triangular kernel some of the lower probability values are removed a distance away from the storm. This could cause isolated flashes that occur a distance away from the storm to be missed. This could cause problems for the forecaster as well as the general public.

For both the tornado and lightning and the combined tornado and combined lightning forecaster cases, the ROC diagrams and AUC did not appear to notably change between the four kernels. However, for both the tornado and lightning and the combined tornado and combined lightning forecaster cases the AUC value for the Epanechnikov kernel was the lowest. Generally, for both the tornado and lightning and the combined tornado and combined lightning forecaster cases the Quartic kernel had the highest resolution and reliability terms out of all four kernels and the Epanechnikov kernel had the lowest resolution and reliability terms out of all four kernels. Generally, for both the tornado and lightning and the combined tornado and combined lightning forecaster cases when using a 7.5 km compared to a 4.5 km event definition the reliability values were lower and the resolution values were higher. For the lightning combined case the Epanechnikov kernel reliability term is

approximately an order of magnitude lower and the resolution term is slightly lower when compared to the Gaussian kernel. For the tornado combined cases, the resolution for all four kernels are approximately the same and the differences between the reliability values across all four kernels were minor. Therefore, based on these limited results for lightning cases using a 7.5 km event definition the Epanechnikov kernel and for tornado cases using a 7.5 km event definition the Gaussian kernel should be applied to the probabilities to create the plumes.

## 6.3 Future Work

One of the limitations in this work is a limited dataset. So, using additional cases across multiple years and forecasters would be an avenue for future work. Additionally, different kernels could be used to see if any of them offer more improvement. Additionally, new kernels could be developed and used to see if those kernels offer more improvement over the Gaussian than the Epanechnikov, Quartic, and Triangular kernels. The ideal kernel would be a kernel that is able able to improve reliability with the least cost of worsening the reliability when compared to the Gaussian kernel. More work could be done with the practically perfect plumes to see if they could be created more similarly to how the forecasters used the PHI Tool to create the plumes in the HWT. This could be done to see how that would change the practically perfect plumes, attributes diagrams, and ROC diagrams. Additionally, work could be done to verify the plumes temporally and spatially instead of just spatially.

## 6.4 Summary

- A goal of this study was to provide an event definition for each hazard, tornado and lightning.

- A goal of this study was to create a method to verify the areal coverage of the tornado and lightning probabilities.
- A goal of this study also aimed to determine if the Gaussian kernel used by default in the HWT was most appropriate kernel to create the spatial coverage of probabilities from the peak values defined by the forecaster.
- Multiple kernel smoothers, Epanechnikov, Quartic, and Triangular, were tested in addition to the Gaussian kernel to to see if the areal coverage of probabilities provided by each of these produced better results.
- Multiple event definitions from 0.5 to 7.5 km were used to determine the optimal distance for a reference class for the forecast probabilities of each hazard (tornado and lightning).
- A neighborhood of 4.5 and 7.5 km from the mesocyclone or NLDN CG location provided a comparison for the trade-off between resolution, reliability and AUC for both tornado and lightning hazards, respectively.
- It was found for lightning cases the Epanechnikov kernel and for tornado cases the Gaussian kernel should be applied to the probabilities to create the plumes.
- Future work would be establishment of a larger dataset and including additional cases from multiple forecasters to determine if the results from this small sample size can be generalized.

## Bibliography


Ahijevych, D., E. Gilleland, B. G. Brown, and E. E. Ebert, 2009: Application of Spatial Verification Methods to Idealized and NWP-Gridded Precipitation Forecasts. *Weather and Forecasting*, **24 (6)**, 1485–1497, doi:10.1175/2009WAF2222298.1, URL https://journals.ametsoc.org/doi/abs/10.1175/2009WAF2222298.1, publisher: American Meteorological Society.

Bates, F. C., 1962: Severe Local Storm Forecasts and Warnings and the General Public. *Bulletin of the American Meteorological Society*, **43 (7)**, 288–291, doi:10.1175/1520-0477-43.7.288, URL http://journals.ametsoc.org/doi/10.1175/1520-0477-43.7.288.

Berg, R., 2017: Hurricane Hermine. Tech. rep., National Hurricane Center Tropical Cyclone Report, 63 pp. URL https://www.nhc.noaa.gov/data/tcr/AL092016_Hermine.pdf.

Brooks, H. E., C. A. Doswell, and M. P. Kay, 2003: Climatological Estimates of Local Daily Tornado Probability for the United States. *Weather and Forecasting*, **18**, 15.

Calhoun, K. M., T. Meyer, and D. Kingfield, 2017: Storm-based Cloud-to-Ground Lightning Probabilities and Warnings. *AGU Fall Meeting*, AGU, New Orleans, LA, AE11A–08, URL https://agu.confex.com/agu/fm17/meetingapp.cgi/Paper/225133.

Cartier, K. M. S., 2017: New Model Predicts Lightning Strikes; Alert System to Follow. *Eos, 98*, URL https://doi.org/10.1029/2017EO088591, published on 11 December 2017.

Cintineo, J. L., and Coauthors, 2018: The NOAA/CIMSS ProbSevere Model: Incorporation of Total Lightning and Validation. *Weather and Forecasting*, **33 (1)**, 331–345, doi:10.1175/WAF-D-17-0099.1, URL https://journals.ametsoc.org/waf/article/33/1/331/40990/The-NOAACIMSS-ProbSevere-Model-Incorporation-of.

Coleman, T. A., K. R. Knupp, J. Spann, J. B. Elliott, and B. E. Peters, 2011: The History (and Future) of Tornado Warning Dissemination in the United States. *Bulletin of the American Meteorological Society*, **92 (5)**, 567–582, doi:10.1175/2010BAMS3062.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2010BAMS3062.1.

Cummins, K. L., and M. J. Murphy, 2009: An Overview of Lightning Locating Systems: History, Techniques, and Data Uses, With an In-Depth Look at the U.S. NLDN. *IEEE Transactions on Electromagnetic Compatibility*, **51 (3)**, 499–518, doi:10.1109/TEMC.2009.2023450, conference Name: IEEE Transactions on Electromagnetic Compatibility.

Doswell, C. A., A. R. Moller, and H. E. Brooks, 1999: Storm Spotting and Public Awareness since the First Tornado Forecasts of 1948. *Weather and Forecasting*, **14**, 544–557, doi:10.1175/1520-0434(1999)014⟨0544:SSAPAS⟩2.0.CO;2, URL https://journals.ametsoc.org/doi/pdf/10.1175/1520-0434%281999%29014%3C0544%3ASSAPAS%3E2.0.CO%3B2.

Ebert, E. E., 2008: Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Met. Apps*, **15 (1)**, 51–64, doi:10.1002/met.25, URL http://doi.wiley.com/10.1002/met.25.

Ebert, E. E., 2009: Neighborhood Verification: A Strategy for Rewarding Close Forecasts. *Wea. Forecasting*, **24 (6)**, 1498–1510, doi:10.1175/2009WAF2222251.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2009WAF2222251.1.

Ferree, J. T., D. Freeman, and E. Jacks, 2007: Storm-based Warnings Changes to NWS Warnings for the Digital Age. *35th Conf. on Broadcast Meteorology*, Amer. Meteor. Soc., San Antonio, TX, 5.6, URL https://ams.confex.com/ams/87ANNUAL/techprogram/paper_120818.htm.

Flora, M. L., P. S. Skinner, C. K. Potvin, A. E. Reinhart, T. A. Jones, N. Yussouf, and K. H. Knopfmeier, 2019: Object-Based Verification of Short-Term, Storm-Scale Probabilistic Mesocyclone Guidance from an Experimental Warn-on-Forecast System. *Weather and Forecasting*, **34 (6)**, 1721–1739, doi:10.1175/WAF-D-19-0094.1, URL https://journals.ametsoc.org/waf/article/34/6/1721/344744/ObjectBased-Verification-of-ShortTerm-StormScale.

Gallo, B. T., and Coauthors, 2017: Breaking New Ground in Severe Weather Prediction: The 2015 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. *Weather and Forecasting*, **32 (4)**, 1541–1568, doi:10.1175/WAF-D-16-0178.1, URL https://journals.ametsoc.org/doi/abs/10.1175/WAF-D-16-0178.1.

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of Spatial Forecast Verification Methods. *Wea. Forecasting*, **24 (5)**, 1416–1430, doi:10.1175/2009WAF2222269.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2009WAF2222269.1.

Harrison, D., 2018: Correcting, Improving, and verifying Automated Guidance in a New Warning Paradigm. M.S. thesis, School of Meteorology, University of Oklahoma, 109 pp., Norman, OK.

Harrison, D. R., and C. D. Karstens, 2017: A Climatology of Operational Storm-Based Warnings: A Geospatial Analysis. *Weather and Forecasting*, **32 (1)**, 47–60, doi:10.1175/WAF-D-15-0146.1, URL http://journals.ametsoc.org/doi/10.1175/WAF-D-15-0146.1.

Hitchens, N. M., H. E. Brooks, and M. P. Kay, 2013: Objective Limits on Forecasting Skill of Rare Events. *Wea. Forecasting*, **28**, 525–534, doi:10.1175/WAF-D-12-00113.1, URL http://journals.ametsoc.org/doi/abs/10.1175/WAF-D-12-00113.1.

Karstens, C. D., and Coauthors, 2015: Evaluation of a Probabilistic Forecasting Methodology for Severe Convective Weather in the 2014 Hazardous Weather Testbed. *Weather and Forecasting*, **30 (6)**, 1551–1570, doi:10.1175/WAF-D-14-00163.1, URL http://journals.ametsoc.org/doi/10.1175/WAF-D-14-00163.1.

Karstens, C. D., and Coauthors, 2018: Development of a Human-Machine Mix for Forecasting Severe Convective Events. *Weather and Forecasting*, **33 (3)**, 715–737, doi:10.1175/WAF-D-17-0188.1, URL https://journals.ametsoc.org/waf/article/33/3/715/40042/Development-of-a-HumanMachine-Mix-for-Forecasting.

Krocak, M., 2017: Establishing a Baseline: What We Know about Tornado Warning Reception, Comprehension and Response. *7th Research to Operations (7R2O) - Poster Session*, Amer. Meteor. Soc., Seattle, WA, URL https://ams.confex.com/ams/97Annual/webprogram/Paper307718.html.

Lemon, L. R., and M. Umscheid, 2008: The Greensburg, KS tornadic storm: a storm of extremes. *24th Conference on Severe Local Storms*, Amer. Meteor. Soc., Savannah, GA, 2.4, URL https://ams.confex.com/ams/24SLS/webprogram/Paper141811.html.

Lu, C., H. Yuan, B. E. Schwartz, and S. G. Benjamin, 2007: Short-Range Numerical Weather Prediction Using Time-Lagged Ensembles. *Weather and Forecasting*, **22 (3)**, 580–595, doi:10.1175/WAF999.1, URL https://journals.ametsoc.org/waf/article/22/3/580/38927/ShortRange-Numerical-Weather-Prediction-Using.

Mason, I., 1982: A Model for Assesment of Weather Forecasts. *Austrailian Meteorological Magazine*, **30 (4)**, 291–303.

National Weather Service, 2011a: The Historic Tornadoes of April 2011. NOAA National Disaster Survey Rep., URL https://repository.library.noaa.gov/view/noaa/6977, 76 pp.

National Weather Service, 2011b: Joplin, Missouri, Tornado – May 22, 2011. NWS Central Region Service Assessment, URL https://repository.library.noaa.gov/view/noaa/6576, 35 pp.

National Weather Service, nda: May 24th, 2017 Tornado Event. NOAA, Accessed 23 March 2020, URL https://www.weather.gov/cae/may_24_2017_event_review.html.

National Weather Service, ndb: Severe Weather Hits Western/Central KS on 24 May 2016. NOAA, Accessed 23 March 2020, URL https://www.weather.gov/ddc/24may2016SevereWx.

National Weather Service, ndc: Wednesday, May 25th, 2016: Long Track Tornado Hits North Central Kansas. NOAA, Accessed 23 March 2020, URL https://www.weather.gov/top/LongTrackTornadoHitsNorthCentralKS.

Roberts, N. M., and H. W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Mon. Wea. Rev.*, **136 (1)**, 78–97, doi:10.1175/2007MWR2123.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2007MWR2123.1.

Rothfusz, L. P., R. Schneider, D. Novak, K. Klockow-McClain, A. E. Gerard, C. Karstens, G. J. Stumpf, and T. M. Smith, 2018: FACETs: A Proposed Next-Generation Paradigm

for High-Impact Weather Forecasting. *Bulletin of the American Meteorological Society*, **99 (10)**, 2025–2043, doi:10.1175/BAMS-D-16-0100.1, URL http://journals.ametsoc.org/doi/10.1175/BAMS-D-16-0100.1.

Stumpf, G. J., C. D. Karstens, and L. P. Rothfusz, 2015: Probabilistic Hazard Information (PHI): Highlighting the benefits via new verification techniques for FACETS. *Third Conf. on Weather Warnings and Communication*, Amer. Meteor. Soc., Raleigh, NC, 5.7, URL https://ams.confex.com/ams/43BC3WxWarn/webprogram/Paper272745.html.

Uccellini, L. W., and J. E. Ten Hoeve, 2019: Evolving the National Weather Service to Build a Weather-Ready Nation: Connecting Observations, Forecasts, and Warnings to Decision-Makers through Impact-Based Decision Support Services. *Bulletin of the American Meteorological Society*, **100 (10)**, 1923–1942, doi:10.1175/BAMS-D-18-0159.1, URL https://journals.ametsoc.org/bams/article/100/10/1923/344812/Evolving-the-National-Weather-Service-to-Build-a.