CASE STUDIES OF CLEAR-AIR TURBULENCE: EVALUATION AND VERIFICATION OF NEW FORECASTING TECHNIQUES

by

EMILY NICHOLE WILSON

(Under the Direction of John A. Knox)

ABSTRACT

Clear-air turbulence (CAT) is an important and unsolved problem within the aviation industry and the atmospheric sciences. This thesis examined CAT using three case studies of turbulence outbreaks in cyclonic upper-level flow in December 2010, January 2011, and September 2011. High-resolution model output was utilized to create turbulence forecasts using six turbulence indices (Ellrod-Knox, Ellrod-Knapp, Lighthill-Ford, Richardson number, frontogenesis, and vertical wind shear), and these forecasts were compared to actual eddy dissipation rate (EDR) turbulence reports. Verification statistics and ROC (relative operating characteristic) curves were produced to determine which forecast metric had the most skill during these outbreaks. In terms of the area under the ROC curve, the Ellrod-Knox method performed the best in the December and January cases, for all turbulence intensities and moderate-or-greater turbulence. For the September case, the Lighthill-Ford method performed the best for all turbulence intensities, while the Ellrod-Knapp index was best for moderate-or-greater turbulence.

INDEX WORDS: clear-air turbulence, aviation, forecasting

CASE STUDIES OF CLEAR-AIR TURBULENCE: EVALUATION AND VERIFICATION OF NEW FORECASTING TECHNIQUES

by

EMILY NICHOLE WILSON

B.S., The University of Georgia, 2010

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2012

CASE STUDIES OF CLEAR-AIR TURBULENCE: EVALUATION AND VERIFICATION OF NEW FORECASTING TECHNIQUES

by

Emily Nichole Wilson

Major Professor: John A. Knox

Committee: Thomas L. Mote
J. Marshall Shepherd

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2012

DEDICATION

I would like to dedicate this thesis to my family. I decided at the age of seven that I wanted to become a meteorologist, and from then on, my family always supported me and encouraged me to achieve my dream. They always believed in me, even when I didn't believe in myself, and without all of their love and encouragement throughout the years, I would not be where I am today.

ACKNOWLEDGMENTS

First, I would like to thank Dr. John Knox (UGA) for all of his input and assistance with writing this thesis, for serving as my graduate advisor, and for all of his encouragement and support during both my undergraduate and graduate careers. I would also like to thank Dr. Thomas Mote (UGA) and Dr. Marshall Shepherd (UGA) for serving on my committee, as well as for their support and inspiration during my undergraduate and graduate careers. Thank you to Dr. Robert Sharman (NCAR) for going above and beyond the call of duty in supplying and helping me with the data for this thesis, as well as for his hospitality during my visit to NCAR. Thank you to Victor Gensini for all of his help with technology issues, as well as for his GIS expertise. I would also like to say thank you to Dr. Steve Silberberg (AWC) for providing a list of potential CAT outbreaks for this thesis. To my friends at the UGA Climate Research Laboratory, thank you for all of the good times and laughter, and for making each day of my graduate school experience more enjoyable than the last. Finally, I would like to thank anyone who has helped, supported, and/or encouraged me at any point throughout my life. I would not be here without all of you!

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Clear-air turbulence (CAT) has long been a serious concern for the aviation industry, as well as an unsolved problem in aviation meteorology. Simply put, CAT can be described as "bumpiness in flight" (Ellrod et al. 2003). Several factors impact the "bumpiness" that is felt by the aircraft: the design of the aircraft, the weight of the aircraft, the speed of the aircraft, the pilot input, and, most importantly for the purposes of this thesis, the state of the atmosphere (Ellrod et al. 2003).

According to the *Encyclopedia of Atmospheric Sciences*, CAT is defined as aircraft turbulence that occurs at 500 hPa (about 5.6 km) or higher, either in cloud-free conditions or within stratiform clouds (Ellrod et al. 2003). Reports of CAT date back to the 1940s (Baughman 1946, Ellrod et al. 2003). CAT was first encountered in World War II by high-flying aircraft, because these planes could easily reach the altitudes at which this type of turbulence most commonly occurs. By the 1950s, CAT was a recognized aviation hazard (Chambers 1955).

Even today, despite the advances in aviation technology, CAT is still a serious problem, as aircraft encounters with CAT account for a significant percentage of weather-related commercial aircraft incidents (Sharman et al. 2006). Aircraft encounters with turbulence continue to be a serious source of occupant injuries. According to Sharman et al. (2006), one major airline reported almost 400 injury-causing turbulence encounters over a period of three years. One air carrier estimated that it pays out tens of millions of dollars per year for customer

injury claims, and that it loses about 7,000 days in employee injury-related disabilities (Sharman et al. 2006). Repeated turbulence encounters over the lifetime of the aircraft may lead to metal fatigue and, in extremely rare cases, structural failure of the aircraft (Ellrod et al. 2003). This can cause aircraft to be taken out of service, also creating a financial loss for the airlines.

Table 1.1 displays CAT intensities and their effect on passengers, as well as the effect on the aircraft itself. It is clear that CAT is still a significant threat to aviation, and that finding better methods of CAT prediction would prove to be extremely beneficial to airlines and their passengers, while also helping to answer scientific questions about CAT.

Table 1.1: CAT intensities and their effects on aircraft and passengers (Lane et al. 2012).

Turbulence Category	Aircraft Response	Aircraft Vertical Acceleration Magnitude (g)	Passenger Experience	Approximate $\varepsilon^{1/3}$ ($m^{2/3}s^{-1}$) for B737, B757 aircraft
Light	Momentarily causes slight, erratic changes in altitude and/or attitude (pitch, roll, yaw).	0.2 – 0.5	A slight strain against seat belts. Unsecured objects may be displaced slightly. Food service may be conducted with little difficulty walking.	0.1 – 0.3
Moderate	Similar to 'Light Turbulence' but greater intensity. Changes in altitude, attitude, and/or airspeed occur. The aircraft remains in control at all times.	0.5 – 1.0	Definite strain against seat belts. Unsecured objects are dislodged. Food service and walking are difficult.	0.3 – 0.5
Severe	Large, abrupt changes in altitude, attitude, and/or airspeed. Aircraft may be momentarily out of control.	1.0 – 2.0	Occupants are forced violently against seat belts. Unsecured objects are tossed about. Food service and walking are impossible.	0.5 – 0.7
Extreme	The aircraft is violently tossed about and is practically impossible to control. It may cause structural damage.	> 2.0	Truly frightening.	> 0.7

CHAPTER 2

LITERATURE REVIEW

Forecasting outbreaks of CAT has proven challenging over the past several decades, as none of the current numerical weather prediction (NWP) models can resolve features at the scale at which CAT occurs. The resolution of current NWP models is about two orders of magnitude too coarse to resolve aircraft-scale turbulence, which is, at most, only a few hundred meters (Sharman et al. 2006). Assuming the large-scale forecasts are sufficiently accurate, the turbulence forecasting problem is then one of identifying large-scale features that are conducive to the formation of aircraft-scale eddies (Sharman et al. 2006). The large-scale feature that is most commonly associated with CAT is frontogenesis (Baughman 1946, Dutton and Panofsky 1970, Ellrod and Knapp 1992).

The notion that CAT occurs in regions of frontogenesis dates back to World War II. In his paper, Baughman (1946) used a case study to suggest that severe turbulence was occurring over the North Atlantic Ocean in regions where there was a sharp air mass contrast and strong wind shear. Baughman (1946) concluded that that turbulence can occur where a frontal surface is accompanied by strong wind shear. Decades later, Dutton and Panofsky (1970) echoed the importance of frontogenesis by finding that CAT had the highest probability of occurring in regions of strong vertical wind shear and strong horizontal temperature gradients.

Even today, the most commonly used CAT index, the turbulence index (TI), is based on the notion that CAT occurs in regions of frontogenesis. Ellrod and Knapp (1992) created the TI

based on vertical wind shear and deformation. The TI is derived from Petterssen's frontogenetic intensity equation, which relates frontogenesis to increased vertical wind shear (Ellrod and Knapp 1992). An increase in vertical wind shear should lead to an increased probability of turbulence occurring. Two case studies were examined to illustrate TI's ability to detect CAT, and it was determined that the TI correctly forecast between two-thirds and three-fourths of the CAT events (Ellrod and Knapp 1992). This method became the most widely used index for forecasting CAT because it is simple to calculate; all of the terms in the equation are kinematic, rather than dynamic.

While most of the early CAT indices relate turbulence with regions of frontogenesis, Knox (1997) suggested that CAT can occur in strongly anticyclonic regions as well. It is possible that strongly anticyclonic regions can generate CAT through means that are not accounted for in conventional CAT theory, including geostrophic adjustment (Knox 1997). Knox (1997) compared several different CAT forecasting methods, and found that they may be correctly predicting CAT for the wrong reasons, or incorrectly predicting smooth conditions in regions of turbulence caused by processes other than frontogenesis.

Ellrod and Knox (2010) then attempted to improve the TI, based on the insight in Knox (1997), by adding a divergence trend term to create the divergence-modified turbulence index (DTI). The proposed change to the TI is to add a proxy term for the divergence tendency to account for CAT in situations of rapidly changing divergence associated with anticyclonic flow (both shear and curvature), and in cyclonic regions not in gradient balance (Ellrod and Knox 2010). Since most turbulence indices use frontogenesis as a proxy for turbulence, the addition of the divergence trend term allows for non-frontogenetical situations to be taken into account in forecasting turbulence. Two case studies of CAT outbreaks were provided as supporting

evidence that the addition of the divergence trend into the TI makes it a better turbulence diagnostic. For these two cases, the DTI showed an improvement over the TI (Ellrod and Knox 2010).

Sharman et al. (2006) took a different approach to CAT forecasting by creating a product called Graphical Turbulence Guidance (GTG). The essence of the GTG forecasting method is to integrate a combination of many separate turbulence diagnostics, with each diagnostic weighted to get the best agreement with available observations (Sharman et al. 2006). The GTG is composed of ten different turbulence diagnostics at upper levels and nine different diagnostics at mid-levels, which can be found in Table 2.1 (Sharman et al. 2006). To test the effectiveness of the GTG, CAT outbreaks were predicted using the GTG and using each of its components. Each of these forecasts was then compared to PIREPs for verification. It was determined that the GTG combination performed better than any single turbulence diagnostic (Sharman et al. 2006). However, out of all of the different diagnostics used in GTG, the single best diagnostic appears to be the frontogenesis function (Sharman et al. 2006).

Knox et al. (2008) had the goal of trying to place CAT forecasting in a more theoretical framework, using the Lighthill-Ford theory of spontaneous imbalance (Ford 1994) to predict episodes of CAT. The right-hand side of the Lighthill-Ford equation, which can be seen in Equation 2.1, contains three terms that can be interpreted as forcing gravity waves (Knox et al. 2008).

$$R = \underbrace{\frac{\partial}{\partial t}\left(\nabla \cdot \mathbf{G}\right)}_{\text{Term 1}} + \underbrace{f\mathbf{k} \cdot \nabla \times \mathbf{G}}_{\text{Term 2}} + \underbrace{\frac{g}{2}\frac{\partial}{\partial t}\nabla^2\left(h - h_0\right)^2}_{\text{Term 3}} \tag{2.1}$$

$$\mathrm{G} = \boldsymbol{u}\boldsymbol{\nabla} \cdot (h\boldsymbol{u}) + (h\boldsymbol{u} \cdot \boldsymbol{\nabla})\boldsymbol{u} \tag{2.2}$$

In Equation 2.1, f is the Coriolis force, h is the layer depth, h_0 is the layer depth away from the region containing vertical motion, and G can be calculated using Equation 2.2 (Knox et al. 2008). In Equation 2.1, the first term is related to the horizontal divergence, the second term is a combination of the horizontal divergence and the relative vorticity, and the third term is related to the geopotential height (Knox et al. 2008). After performing a scale analysis appropriate for mid-latitude, synoptic scale flows, it was determined that the leading-order forcing term in the Lighthill-Ford equation is the advection of relative vorticity (Knox et al. 2008). This means that areas with large relative vorticity advections should lead to spontaneous gravity wave generation. The scaled version of the Lighthill-Ford equation can be seen in Equation 2.3, where D is the horizontal divergence, ζ is the relative vorticity, and J is the Jacobian (Knox et al. 2008).

$$R = f\mathbf{V}\cdot\nabla\zeta + 2Df\zeta - f\mathbf{k}\cdot\mathbf{V}\times\nabla D - 2\frac{\partial}{\partial t}J(u,v) \tag{2.3}$$

Their theory was supported by a case study of a CAT outbreak over the Upper Mississippi and Ohio Valleys, in which PIREPs were overlaid on a map of the turbulent kinetic energy (TKE) dissipation rates. The turbulence reports coincided with large TKE dissipation rates, while null reports of turbulence appeared in areas with low TKE dissipation rates (Knox et al. 2008). A seasonal examination of TKE dissipation and turbulence revealed that maximum TKE dissipation rates appeared to be correlated to PIREPs of experienced turbulence (Knox et al. 2008). The seasonal examination showed that the Lighthill-Ford approach performed better than Sharman et al.'s (2006) GTG method (Knox et al. 2008).

Plougonven et al. (2009) took issue with Knox et al.'s (2008) application of the Lighthill-Ford theory to gravity waves. They disagreed with the claim that Lighthill-Ford theory could be used to predict the location and intensity of spontaneous gravity waves in mid-latitude flows (Plougonven et al. 2009). It was argued that the spatial scales between large-scale gravity waves

and the small-scale motions that generate the waves were ignored, and that forcing terms in the equation contribute to balanced motions, such as frontogenesis, not just gravity wave generation (Plougonven et al. 2009). In reply to Plougonven et al. (2009), Knox et al. (2009) explained that they were testing the applicational limits of the Lighthill-Ford theory, not its theoretical limits. They refuted Plougonven et al.'s (2009) claim that frontogenesis explained the results of Knox et al. (2008) by showing a case in which Lighthill-Ford forcing was mostly associated with frontolytic situations.

Most recently, McCann et al. (2012) improved on Knox et al.'s (2008) Lighthill-Ford method of predicting turbulence. This led to the creation of the gravity wave modified TKE equation, also known as the ULTURB method (McCann et al. 2012). The ULTURB method outperformed the Knox et al. (2008) application of Lighthill-Ford and Sharman et al.'s (2006) GTG in the prediction of CAT for all turbulence intensities (McCann et al. 2012).

Table 2.1: CAT diagnostics in GTG2 (modified from Sharman et al. 2006).

GTG2 Upper Level Diagnostics	GTG2 Mid-Level Diagnostics
Colson-Panofsky Index (Colson and Panofsky 1965)	Turbulence Index 1 (Ellrod and Knapp 1992)
Richardson Number (Endlich 1964; Kronebach 1964)	Wind Speed X Horizontal Deformation (Reap 1996)
Diagnostic Turbulent Kinetic Energy Formulation (Marroquin 1998)	Absolute Value “Inertial Advection-Centrifugal Wind” (McCann 2001)
Frontogenesis Function in Isentropic Coordinates (Bluestein 1993)	Horizontal Temperature Gradient (Buldovskii et al. 1976)
Unbalanced Flow Diagnostic (Knox 1997)	Wind Speed (Endlich 1964)
Horizontal Temperature Gradient (Buldovskii et al. 1976)	North Carolina State University Index (Kaplan et al. 2004)
Turbulence Index 1 (Ellrod and Knapp 1992)	Structure Function Derived Eddy Dissipation Rate (Lindborg 1999)
North Carolina State University Index (Kaplan et al. 2004)	Structure Function Derived Sigma Vertical Velocity (Frehlich and Sharman 2004)
Structure Function Derived Eddy Dissipation Rate (Lindborg 1999)	Frontogenesis Function in Pressure Coordinates (Bluestein 1993)
Structure Function Derived Sigma Vertical Velocity (Frehlich and Sharman 2004)	

CHAPTER 3

RESEARCH QUESTIONS

This thesis examined case studies of CAT outbreaks, and multiple turbulence forecasts were generated for each outbreak using several different methods of CAT prediction. Each forecast was verified by comparing the model output to actual turbulence observations. The synoptic conditions under which these CAT outbreaks occurred were also examined, to see if there were any similarities between the three cases. This thesis attempts to answer several research questions, namely:

1. Is the Lighthill-Ford method of forecasting CAT (Knox et al. 2008) effective at predicting episodes of CAT in three case studies?
2. Is the Lighthill-Ford method complementary to other turbulence indices, such as the improved Ellrod method (Ellrod and Knox 2010)? How do the different CAT forecasts from different indices compare? Do they forecast CAT in similar areas?
3. What are the synoptic patterns in the regions in which CAT outbreaks occur?

The ultimate goal of this thesis is to make an original contribution to the quest for a better CAT diagnostic.

CHAPTER 4

DATA

The oldest and most common source of turbulence data is the pilot report (PIREP). PIREPs contain information about the turbulence intensity, the turbulence level, and the location of the turbulence. However, there are some well-known issues with PIREPs of turbulence. According to Schwartz (1996), the PIREP system was not designed with research in mind, so these reports often have serious shortcomings. PIREPs are sporadic in space and time, provide only a subjective measure of turbulence, and include few negative reports of turbulence (Takacs et al. 2005). Turbulence location is a significant potential error in PIREPs, because the reported location of turbulence may not actually be the location at which turbulence was experienced, due to delays in reporting and the high speed of the aircraft. Sharman et al. (2006) found that PIREPs may have reporting errors up to 50 kilometers horizontally, 70 meters vertically, and 200 seconds away from the source of the turbulence. In extreme turbulence events, pilots may submit the report at a location far from the actual turbulence event, or they may forget to submit a PIREP altogether.

Recently, eddy dissipation rate (EDR) data has become an alternative method of recording aircraft turbulence information. EDR is an in situ method of measuring the turbulence experienced by the aircraft, and it removes the subjective nature of PIREPs (Takacs et al. 2005). The EDR algorithm relates the eddy dissipation rate to aircraft vertical acceleration (Takacs et al. 2005). EDR values and their associated turbulence intensity can be seen in Table 1.1. This new

method also provides researchers with much more data on turbulence. EDR observations are transmitted approximately every minute, and both the peak and median values are recorded (Takacs et al. 2005). This is a vast improvement over PIREPs, which pilots often do not report during the peak of a turbulent episode.

The EDR data that was used in this thesis contained information on the date, time, location, altitude, and turbulence intensity for each report of turbulence or null turbulence. For the December 2010 case, there were initially 6,625 EDR observations. For the January 2011 outbreak, there were initially 4,400 EDR observations. For the September 2011 case, there were initially 951 EDR observations. The EDR data was provided by Dr. Robert Sharman at NCAR.

To create and analyze the different CAT forecasts, high resolution model output from the Weather Research and Forecasting Rapid Refresh model (WRF-RR) was used. The WRF-RR has recently replaced the Rapid Update Cycle (RUC) model. The WRF-RR has a horizontal resolution of 13 kilometers, and can be updated every hour to make aviation forecasts (Benjamin and Sahm 2011). Dr. Robert Sharman, at NCAR, supplied this model output.

To remove turbulence reports due to convection, it was necessary to compare archived radar data to EDR observations of turbulence. This data came from the Iowa State Mesonet, because the radar data could be downloaded as a georeferenced raster layer, which was easily imported into ArcMap for analysis. This radar data can be found at http://mesonet.agron.iastate.edu/current/mcview.phtml. To analyze the synoptic conditions at the time of the turbulence outbreaks, archived weather maps were needed. This weather data came from the Storm Prediction Center's Mesoanalysis Archive. This archive provides hourly analysis products for a variety of meteorological variables. The Mesoanalysis Archive can be found at http://www.spc.noaa.gov/exper/ma_archive/.

CHAPTER 5

METHODOLOGY

5.1 Case Study Periods

For this thesis, three outbreaks of clear air turbulence were examined. These case studies were chosen from a list of CAT outbreaks provided by Dr. Steve Silberberg at the NOAA/NWS/NCEP/Aviation Weather Center in Kansas City, MO. The first outbreak occurred from 1800 UTC on 26 December 2010 to 1800 UTC on 28 December 2010. The second outbreak was somewhat shorter, beginning at 1800 UTC on 06 January 2011 and ending at 1800 UTC on 07 January 2011. The December and January cases were selected due to their lack of convection throughout the duration of the outbreaks. The last turbulence outbreak that was examined started at 1200 UTC on 22 September 2011 and ended six hours later at 1800 UTC on 22 September 2011. This case was chosen because it was the most recent CAT outbreak at the time that case studies were being determined, as well as for its relative lack of convection when compared to other potential cases. For each of these outbreaks, turbulence forecasts were produced at six-hour intervals (0000 UTC, 0600 UTC, 1200 UTC, and 1800 UTC) for the entirety of the events. Turbulence forecasts were produced every 1,000 feet (0.3 km), starting at an altitude of 19,000 feet (5.8 km) and ending at 45,000 feet (13.7 km).

5.2 Turbulence Indices

For each case study, several different turbulence forecasts were generated. Forecasts were created using six different CAT forecasting indices: 1) the Ellrod-Knapp (1992) method, 2) the Ellrod-Knox (2010) method, 3) the Richardson number, 4) vertical wind shear, 5) frontogenesis, and 6) the Sharman (2011, personal communication) method of calculating Lighthill-Ford. The calculation of these methods is described below.

The Ellrod-Knapp (1992) method is the most widely used CAT forecasting index. It is simple to calculate, as it considers turbulence to be the product of deformation (DEF) and vertical wind shear (VWS) (Ellrod and Knapp 1992). The deformation term contains components of shearing deformation (DSH) and stretching deformation (DST) (Ellrod and Knapp 1992). The final calculation of the Ellrod-Knapp index can be seen in Equation 5.1, while all of its component equations can be found in Equations 5.2 through 5.5 (Ellrod and Knapp 1992).

$$\text{TI} = \text{VWS} \times \text{DEF} \tag{5.1}$$

$$\text{VWS} = \frac{\left(\Delta u^2 + \Delta v^2\right)^{1/2}}{\Delta z} \tag{5.2}$$

$$\text{DEF} = \left(DST^2 + DSH^2\right)^{1/2} \tag{5.3}$$

$$\text{DST} = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \tag{5.4}$$

$$\text{DSH} = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \tag{5.5}$$

Augmenting the Ellrod-Knapp (1992) method, the Ellrod-Knox (2010) method adds a divergence trend term to the product of deformation and vertical wind shear. This can be seen in Equation 5.6, where DTI is the Ellrod-Knox method, TI is Equation 5.1, and DVT is a divergence trend that can be calculated using Equation 5.7, where h1 and h2 are two different

forecast times (Ellrod and Knox 2010). This method has been shown (Ellrod and Knox 2010) to provide improved forecast accuracy in limited tests.

$$\text{DTI} = \text{TI} + \text{DVT} \tag{5.6}$$

$$\text{DVT} = \text{C}\left[\left(\frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta y}\right)_{h2} - \left(\frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta y}\right)_{h1}\right] \tag{5.7}$$

Regions in which there are low values of the Richardson number have long been associated with CAT (Endlich 1964). The formula for the Richardson number can be seen in Equation 5.8, while the components of the Richardson number formulation are displayed in Equation 5.9 and Equation 5.10 (Endlich 1964, Sharman et al. 2006). The Richardson number forecast that was computed by the GTG diagnostic was actually calculated as one divided by the Richardson number, because it provided a climatological improvement for this index (Sharman, 2012, personal communication). In these equations, θ is the potential temperature, g is the acceleration due to gravity, z is the vertical direction, and $\vec{v}$ is the horizontal wind vector.

$$\text{Ri} = \frac{N^2}{S_v^2} \tag{5.8}$$

$$N^2 = \frac{g}{\theta}\frac{\partial \theta}{\partial z} \tag{5.9}$$

$$S_v = \left|\frac{\partial \vec{v}}{\partial z}\right| \tag{5.10}$$

CAT has also long been associated with regions of the atmosphere undergoing frontogenesis, due to the sharp air mass contrasts and strong wind shear (Baughman 1946, Dutton and Panofsky 1970, Ellrod and Knapp 1992). Frontogenesis was calculated in isentropic coordinates using the Bluestein frontogenesis function (Bluestein 1993, Sharman et al. 2006). The formula for the Bluestein frontogenesis function can be seen in Equation 5.11, where θ is the potential temperature (Bluestein 1993, Sharman et al. 2006).

$$F = \frac{D}{Dt}|\nabla\theta| \tag{5.11}$$

The McCann et al. (2012) ULTURB method is a gravity wave modified turbulent kinetic energy (TKE) equation, which is based on the Lighthill-Ford theory. An alternative to the ULTURB method, the Sharman (2011, personal communication) method also has its roots in the Lighthill-Ford theory. Instead of converting Lighthill-Ford forcing terms to TKE dissipation, this approach proxies the effect of inertia-gravity waves on the environment by dividing Knox et al.'s (2008) $|R|^{1/2}$ by the Richardson number, and can be seen in Equation 5.12. This achieves an effect similar to ULTURB in that the impact of inertia-gravity waves is greatest in regions of already low Richardson number, but lacks the gravity wave physics included in the ULTURB method.

$$I = \frac{|R|^{1/2}}{MAX(Ri,10^{-2})} \tag{5.12}$$

5.3 Observational Data

Data from the WRF-RR model was used to create the turbulence forecasts for this thesis. This model output was fed through Dr. Robert Sharman's GTG program, which created individual turbulence forecasts for the six indices of interest. Originally, this created forecast graphics like the ones shown in Figure 5.1. However, these forecast graphics could not easily be put into ArcMap for analysis, so the model forecast data was then output into a two-dimensional array in ASCII format. The model output for each turbulence index was normalized on a 0 to 1 scale, to make the model data similar to reported EDR values (Sharman 2012, personal communication). This normalization was done by comparing observed EDR turbulence values to the actual values of the different turbulence indices over a climatological period of several years, and then calculating the median index value that corresponds to each turbulence intensity

category (Sharman et al. 2006). The normalized turbulence intensity thresholds were then chosen by using the breakpoints determined by a piecewise linear function (Sharman et al. 2006). This process of normalizing the model data should not affect the interpretation of the results, as compared to other studies that did not use normalized data, because this is a linear remapping process (Sharman 2012, personal communication). This normalization process allows the turbulence forecasts to be easily compared to EDR observations for forecast verification. Table 5.1 shows the range of values for each turbulence intensity using this normalized scale.

This turbulence forecast data could then be imported into ArcMap by the coordinates of each model grid point using the Add XY Data option. This process allowed the model data to be imported into ArcMap as a feature layer. This feature layer was then turned into a raster layer using the Feature to Raster conversion tool. This process created separate raster layers and separate attribute tables for each turbulence index. This procedure created a total of 2,496 turbulence forecast maps for the three case studies (16 total forecast times, 6 forecast indices at each time, and 26 flight levels for each index).

Next, the EDR turbulence report data was imported into ArcMap by the latitude and longitude coordinates of each turbulence report using the Add XY Data option. This data was then saved as a shapefile, which created an attribute table containing the date, time, location, altitude, and severity of each piece of EDR data. This process created a total of 16 EDR shapefiles, one for the time of each turbulence forecast.

Once each forecast raster layer and EDR shapefile was created, the EDR shapefile was overlaid with the forecast layer. Then, the Extract Values to Points tool was used to extract the forecasted turbulence value at the location of each turbulence report. To complete this process for all of the data, this step needed to be repeated 2,496 times, one time for each of the

turbulence forecast maps. ArcMap's model builder feature was employed to better process the turbulence forecast maps, and Figure 5.2 shows the model that was created and used for this thesis. This model was designed to automatically extract the forecasted turbulence value at each EDR point for each turbulence forecast map. This process created a new shapefile, with an attribute table containing the information about the turbulence report, as well as the forecasted turbulence value at that specific location.

Once the above process had been completed for all of the forecast maps, the data was ready to be exported for analysis and forecast verification. In order to do this, the turbulence reports had to be queried by altitude, because the turbulence model extracted forecast values for every EDR point that was overlaid with the forecast map, regardless of the altitude of the EDR observation. This step involved using the Select by Attributes feature in ArcMap. The turbulence data was queried by altitude range (for example, all reports between 20,000 feet (6.10 km) and 21,000 feet (6.40 km)) to make sure that all turbulence reports within a flight level were accounted for. This data was then exported as a database table, which could be put into Excel for analysis and forecast verification.

It was important to make sure the turbulence reports were related solely to CAT, so the data needed to be quality controlled in two ways. First, in order to exclude any possible turbulence related to mountain waves, which these forecasting methods do not attempt to address, any turbulence data west of a north-south line through Denver (104^{o}W) were ignored. This step involved sorting the turbulence report data by longitude value and deleting any report with a longitude larger than 104^{o}. As shown in Wolff and Sharman (2008), this should eliminate nearly all mountain-related turbulence reports.

Second, in order to avoid any turbulence related to convection, turbulence reports were compared with archived radar data. Any turbulence report that was within 50 miles (80.5 km) of high radar reflectivity (50 dBZ or greater) was ignored, as in Ellrod and Knox (2010). To complete this step, archived radar data was downloaded from the Iowa State Mesonet in an already georeferenced raster format, so the images could be easily put into ArcMap. Once the radar images were in ArcMap, all reflectivities of less than 50 dBZ were blacked out, leaving only areas with high radar reflectivity on the map. The EDR shapefiles were overlaid with the radar raster layer, and the Measure tool was used to measure the distance to any turbulence reports within the vicinity of convection. Any reports of turbulence that were within the 50-mile (80.5 km) buffer around an area of strong convection were deleted from the dataset. While this may not filter out all of the turbulence reports that are related to convection (Lane et al. 2012), it should filter out a large percentage of non-CAT events.

5.4 Forecast Verification

Once all of the CAT cases were analyzed, the CAT forecasts were verified and skill statistics were computed in order to determine which CAT index did the best job at forecasting these outbreaks. According to Jolliffe and Stephenson (2003), for a yes/no type forecast, such as turbulence, there are two ways for a forecast to be correct, as well as two ways for the forecast to be incorrect. Table 5.2 illustrates these four possible forecast outcomes, and these possible forecast outcomes were used in calculating several forecast skill statistics. For the calculation of all statistics, term “a” refers to the hits, term “b” refers to the false alarms, term “c” refers to the misses, and term “d” refers to the correct rejections. These associations are used in the equations throughout this section, and they can be seen in Table 5.2.

Two of the most important skill statistics are the probability of detection yes (POD_y), often referred to as the hit rate, and the probability of detection no (POD_n). The hit rate is defined as the proportion of occurrences that were correctly forecast, and it can be calculated by dividing the number of hits by the sum of the hits and misses, which can be seen in Equation 5.13 (Jolliffe and Stephenson 2003). The POD_n can be defined as the proportion of correct rejections, or null occurrences of turbulence, that were correctly forecast. The POD_n can be calculated by dividing the number of correct rejections by the sum of the correct rejections and false alarms, which can be seen in Equation 5.14 (Jolliffe and Stephenson 2003).

$$POD_y = \frac{a}{a+c} \tag{5.13}$$

$$POD_n = \frac{d}{b+d} \tag{5.14}$$

$$POFD = 1 - POD_n \tag{5.15}$$

$$FAR = \frac{b}{a+b} \tag{5.16}$$

The false alarm rate, sometimes called the probability of false detection (POFD), is defined as the proportion of non-occurrences that were incorrectly forecast, and it can be calculated by dividing the number of false alarms by the sum of the false alarms and correct rejections, or by subtracting the POD_n value from 1 (Jolliffe and Stephenson 2003). The formula for POFD can be seen in Equation 5.15. Another important skill statistic is the false alarm ratio (FAR). The FAR is the probability of a false alarm occurring for a forecasted event, and it can be calculated by dividing the false alarms by the sum of the hits and the false alarms (Jolliffe and Stephenson 2003). This formula is shown in Equation 5.16.

Some of the more advanced skill statistics include the Heidke Skill Score (HSS), the True Skill Statistic (TSS), and the Critical Success Index (CSI). The HSS accounts for the proportion of forecasts that were correct after eliminating the forecasts that were correct due to chance

(Woodcock 1976, Jolliffe and Stephenson 2003). The formula for HSS can be seen in Equation 5.17, where PC is the proportion of correct forecasts, and E is the proportion of forecasts that would have been correct if forecasts and observations were independent (Jolliffe and Stephenson 2003). A perfect forecast should have an HSS of 1, while a forecast with no skill would have a value of -1 (Jolliffe and Stephenson 2003). The TSS, which can be seen in Equation 5.20, is useful for examining how well the index of interest performed at separating the "yes" events, where turbulence was observed, and the "no" events, where turbulence was not observed (Woodcock 1976, Jolliffe and Stephenson 2003). Similar to the HSS, a perfect forecast would have a TSS value of 1, while a forecast with no skill would have a TSS of -1 (Jolliffe and Stephenson 2003). CSI is defined as the probability of having a forecast hit given that the event of interest was either forecast, observed, or both (Jolliffe and Stephenson 2003). The formula for CSI is given by Equation 5.21 (Jolliffe and Stephenson 2003). A perfect forecast would have a CSI of 1, and a forecast with no skill would have a CSI of 0 (Jolliffe and Stephenson 2003).

$$\text{HSS} = \frac{PC-E}{1-E} \tag{5.17}$$

$$\text{PC} = \frac{a+d}{n} \tag{5.18}$$

$$\text{E} = \left(\frac{a+c}{n}\right)\left(\frac{a+b}{n}\right) + \left(\frac{b+d}{n}\right)\left(\frac{c+d}{n}\right) \tag{5.19}$$

$$\text{TSS} = \frac{ad-bc}{(a+c)(b+d)} \tag{5.20}$$

$$\text{CSI} = \frac{a}{a+b+c} \tag{5.21}$$

All of the forecast skill statistics were calculated using formulas in an Excel spreadsheet. In Excel, each turbulence report was categorized as a hit, miss, false alarm, or correct rejection. A hit occurred when both the EDR observation and the forecasted turbulence value were equal to or above the threshold of interest. A miss occurred when the EDR observation was at or above the threshold of interest, but the forecasted turbulence value was not. A false alarm occurred

when the forecasted turbulence value was greater than or equal to the threshold of interest, but the observed EDR value was not. A correct rejection occurred when both the EDR observation and the forecasted turbulence value were less than the threshold of interest. The total numbers of each of these outcomes were summed, and these totals were used in the formulas to calculate the statistics. Each skill statistic was calculated for all turbulence events, as well as for the moderate-or-greater turbulence events.

Once forecast skill statistics had been calculated, they were used to produce Relative Operating Characteristic (ROC) curves. The ROC curve illustrates the skill of the forecast, with the hit rate, or probability of detection yes (POD_y), plotted along the Y-axis, and the false alarm rate, or one minus the probability of detection no ($1 - POD_n$), plotted along the X-axis (Jolliffe and Stephenson 2003). The diagonal line in the center of the ROC diagram denotes zero forecast skill, and the further that the curve extends toward the top-left side of the diagram, the more forecast skill the metric has (Jolliffe and Stephenson 2003). These ROC curves were plotted in Excel, using several different EDR thresholds. The EDR threshold values of 0.10, 0.15, 0.25, 0.30, 0.35, and 0.40 were used to plot the ROC curves for all of the turbulence events. For the moderate-or-greater level, EDR thresholds of 0.31, 0.35, and 0.40 were used to produce ROC curves for the turbulence indices.

Perhaps the best metric for examining the skill of a forecast is the area under the curve (AUC). The AUC is a method by which the forecast skill can be estimated from the data used to produce a ROC curve. The AUC has a maximum value of 1, while an AUC value larger than 0.500 shows that the index of interest has some forecast skill (Jolliffe and Stephenson 2003). For this thesis, the AUC was calculated in Excel using Equation 5.22, which employs the trapezoid rule for finding the area under a curve (Massachusetts Water Resources Authority 2005).

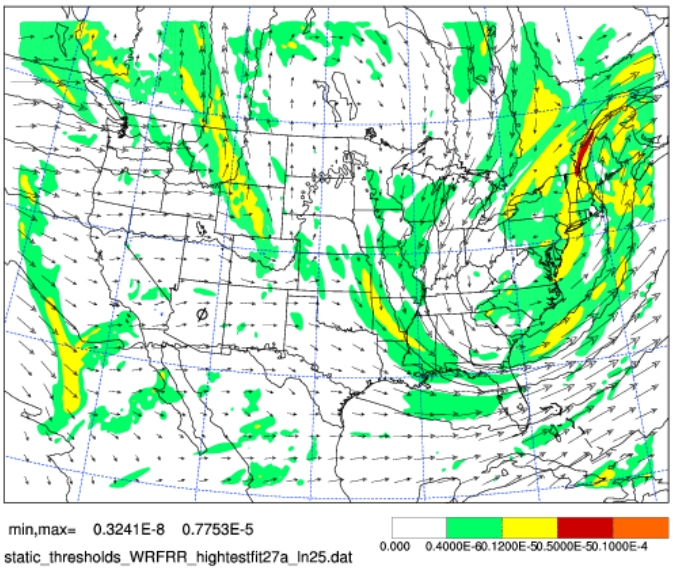
$$\text{AUC} = \left(\frac{A1+A2}{2}\right)(B1 - B2) \quad (5.22)$$

In this equation, A1 and A2 correspond to the POD_y values of two cells in Excel, while B1 and B2 correspond to the $1 - POD_n$ values of the same two cells. Based on the number of points on the ROC curves, this formula created seven trapezoids under each curve for all of the turbulence events. For the moderate-or-greater events, this formula created four trapezoids, meaning that the moderate-or-greater AUC calculations may have slightly more error than the AUC values for all turbulence events.

5.5 Meteorological Analysis

The final step in this thesis was to examine the synoptic conditions around the time of the CAT outbreaks. Archived weather data from the Storm Prediction Center's Mesoanalysis Archive was compared to each of the CAT outbreaks to see if there were any similarities in synoptic forcing between the cases. Specifically, variables that have long been associated with CAT were examined. Such variables included the location of troughs, the location and speed of jet streams and jet streaks, and areas of frontogenesis. To determine the location of troughs and jet streams in relation to turbulence, 300-hPa height, divergence, and wind maps were compared to the turbulence forecast maps. To determine the location of areas undergoing frontogenesis, 700 to 500-hPa mean Petterssen frontogenesis maps were compared to regions of forecasted turbulence. Observed turbulence events (forecast hits and forecast misses) were also compared to these synoptic features in order to examine the conditions under which turbulence actually occurred.

(a)



(b)

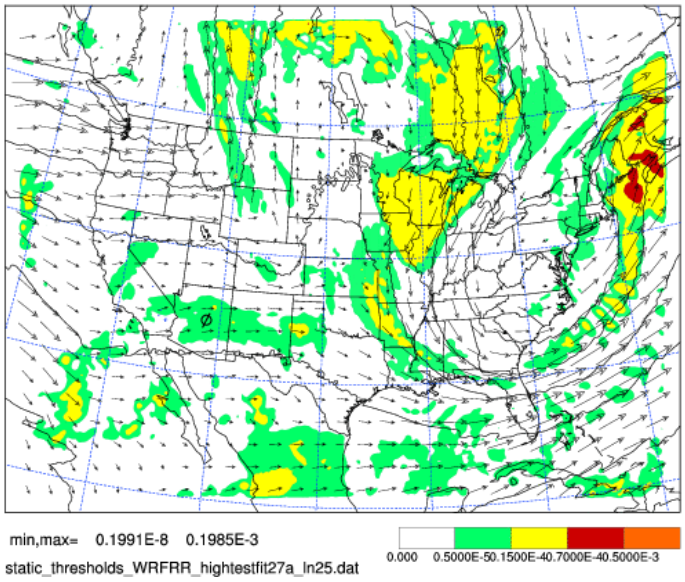


Figure 5.1: WRF-RR turbulence forecasts using (a) the Ellrod-Knapp index and (b) the Lighthill-Ford method for 0000 UTC on 27 December 2010 at 30,000 feet (9.1 km).

Table 5.1: Turbulence intensities and their normalized forecast index values (Sharman 2012, personal communication).

Turbulence Intensity	Normalized Turbulence Index Values
Null	0.00 to 0.14
Light	0.15 to 0.30
Moderate	0.31 to 0.53
Severe	0.54 to 0.79
Extreme	0.80 to 1.00

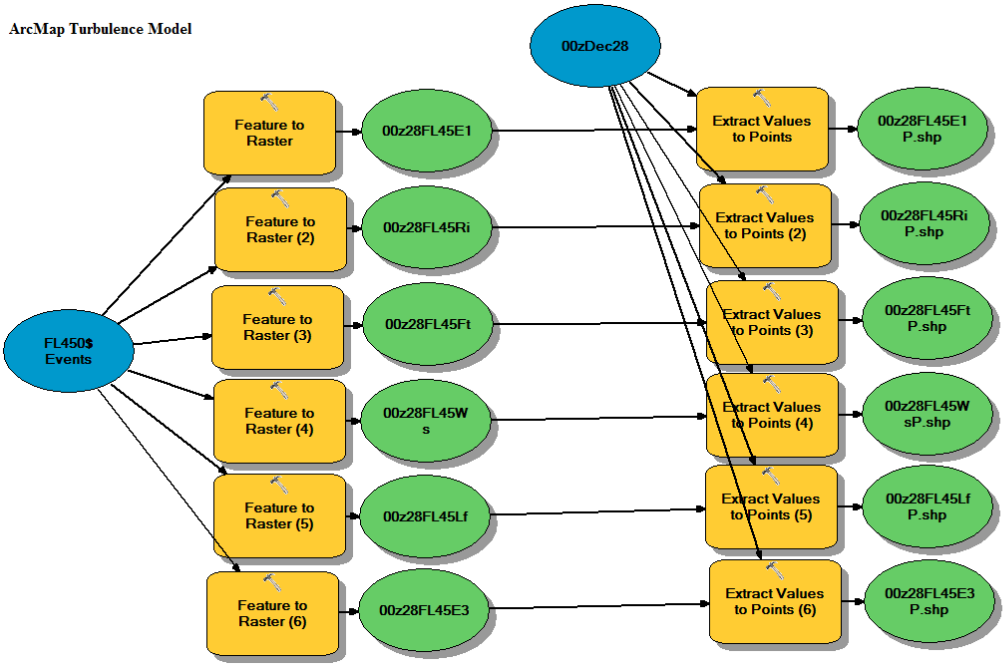


Figure 5.2: Model created in ArcMap for finding the forecasted turbulence value at each EDR report location.

Table 5.2: Sample CAT Forecast Verification (modified from Jolliffe and Stephenson 2003).

	Observed CAT	Observed No CAT
Forecast CAT	(a) Hit	(b) False Alarm
Forecast No CAT	(c) Miss	(d) Correct Rejection

CHAPTER 6

RESULTS

6.1 December 2010 Case

a. Overview

This CAT outbreak occurred from 1800 UTC on 26 December 2010 until 1800 UTC on 28 December 2010. During this turbulence outbreak, there was very little convection across the contiguous United States. The only deep convection during this period occurred in southwestern Minnesota and southwestern Iowa at 1800 UTC on 28 December, and this resulted in only eleven EDR reports being removed from the dataset. After removing all reports west of 104^{o}W, as well as all reports within 50 miles (80.5 kilometers) of convection, this trimmed the dataset down to 3,741 EDR observations. Figure 6.1 spatially illustrates the turbulence reports by intensity. The majority of the moderate turbulence reports were clustered in the Mid-Atlantic region, with a few in the Southeastern U.S., as well as some scattered over the Central Plains. Table 6.1 shows the numbers of turbulence reports broken down into their respective intensity categories.

b. Synoptic Setup

As can be seen in Figure 6.2a, beginning at 1800 UTC on 28 December, there was a deep 300-hPa trough over the eastern third of the United States, with a sharp ridge building over the Central Plains region. The jet core had wind speeds in excess of 140 knots (72.0 m s^{-1}) at this time. The region with wind speeds of at least 100 knots (51.4 m s^{-1}) looped from Wisconsin,

south to Florida, and back north to Maine. Between 700 hPa and 500 hPa, there were three pockets of moderate frontogenesis occurring across the United States: one over the Mid-Atlantic and New England area, one over the Gulf of Mexico, and one over the Midwest.

The most notable synoptic features on the 0000 UTC 27 December maps were a 300-hPa closed low that had formed over the Ohio River Valley, and the region of frontogenesis over New England that had strengthened significantly. The area of frontogenesis over New England at this time was the strongest out of all of three case studies. These features can be seen in Figure 6.2b.

As illustrated by Figure 6.2c, at 0600 UTC on 27 December, the closed upper level low continued to move towards the Atlantic Ocean, while the ridge progressed eastward toward the Great Plains. During this time, there was an impressive 160-knot (82.3 m s^{-1}) jet core located over the Atlantic Ocean, near the Carolina coastline. The strong region of frontogenesis over New England had weakened slightly during this time as well.

By 1200 UTC on 27 December, whose synoptic conditions are shown in Figure 6.2d, the upper level low was centered over the New England coast. By this time, the strongest region of frontogenesis had moved over the extreme northeast corner of Maine. There was also a moderate region of frontogenesis located over southern Florida, while weaker regions of frontogenesis developed over Kansas and the Virginia coast, as well as a larger band extending from Missouri to Mississippi.

At 1800 UTC on 27 December, an upper level low was located over Maine, and the core of the jet stream had moved out over the Atlantic Ocean. A region of moderate frontogenesis had moved from Canada to the Dakotas, while other regions of moderate frontogenesis developed over Florida. The synoptic conditions at this time can be seen in Figure 6.2e.

Upon examination of Figure 6.2f, it can be seen that at 0000 UTC on 28 December, the ridge was centered over the Upper Midwest, and the deep trough was lifting out of the Mid-Atlantic region. Meanwhile, another trough was forming over the desert Southwest. Between 700 hPa and 500 hPa, an area of moderate frontogenesis remained focused over the Dakotas. The moderate region of frontogenesis that was previously located over Florida had moved out over the Atlantic Ocean, leaving behind weak regions of frontogenesis in Georgia and Florida. Lastly, a small region of frontogenesis formed along the Kansas/Colorado border.

As illustrated by Figure 6.2g, at 0600 UTC on 28 December, the deep trough was lifting out of New England, however, a small jet maximum with winds of 120 knots (61.7 m s^{-1}) developed over western New York. The region of moderate frontogenesis that had been located over the Dakotas had moved eastward toward Minnesota, while smaller areas of weak frontogenesis developed over North Carolina and New England. The frontogenetical region along the Kansas/Colorado border increased slightly in size.

By 1200 UTC on 28 December, the first jet maximum had moved away from the New England coast, while a secondary jet stream maximum formed over Vermont. By 1800 UTC on 28 December, this jet maximum had moved out over the Atlantic Ocean. At 1200 UTC on 28 December, the region of moderate frontogenesis over Minnesota strengthened slightly, and began moving eastward into Wisconsin. The region of frontogenesis that was located over New England also increased slightly in strength, with the strongest frontogenesis occurring back into Canada. By 1800 UTC on 28 December, there was not much in the way of frontogenesis in the United States. These last two sets of synoptic conditions can be seen in Figures 6.2h and 6.2i.

c. Forecast Verification

Upon examination of turbulence forecast maps, it is clear that the areas of forecasted turbulence follow the upper-level features in the atmosphere. Figures 6.3a through 6.3f display turbulence forecasts for 0000 UTC on 27 December at 30,000 feet (9.1 km). In these figures, the deep trough over the eastern United States is clearly visible, and most of the indices predicted increased turbulence in the same region where the core of the jet stream was located at this time. The forecast maps even appear to catch some of the regions of frontogenesis; however, the frontogenesis forecast metric missed several areas of diagnosed frontogenesis. For example, in Figure 6.2b, there is a large area of frontogenesis stretching from the Gulf of Mexico into Florida, however, this area of frontogenesis is not present in Figure 6.3f, the frontogenesis forecast metric. However, this discrepancy may be due to the fact that the GTG calculated frontogenesis using the Bluestein equation, which calculated frontogenesis using a three-dimensional surface, while the SPC weather maps show Petterssen frontogenesis, which calculates frontogenesis using a two-dimensional surface (Ray 1986, Sharman et al. 2006).

To further relate turbulence to synoptic weather patterns, the locations of observed turbulence events (forecast hits and forecast misses) were plotted and compared to the large-scale weather patterns. Figures 6.4a through 6.4f illustrate the locations of observed turbulence events at 0000 UTC on 27 December 2010. Upon comparing these figures to the synoptic maps in Figure 6.2b, it is evident that turbulence occurred within the trough that covered the eastern third of the United States during this time period. Almost all of the turbulence events were located within the trough, and a large majority of these turbulence events coincided with the position of the jet stream. The large region of diagnosed frontogenesis appeared to solely account for few reports of turbulence during this time period. Interestingly, no turbulence was

reported over the Gulf of Mexico/southwestern Florida or over New England, the two regions that had the strongest diagnosed frontogenesis.

Also of importance is a cluster of turbulence located over South Dakota and Minnesota. This turbulence appeared to be associated with a ridge that was moving into this region. It is interesting to note that the Ellrod-Knox index, which includes a divergence trend term to forecast turbulence in anticyclonic situations, appeared to catch most of these turbulence events. The Lighthill-Ford method, which is most appropriate for troughs (Paul Williams 2012, personal communication), missed forecasting all of the turbulence in this region. For the most part, all six forecast indices showed turbulence occurring in similar locations along the synoptic scale atmospheric features during this case study. All six indices correctly forecasted some of the turbulence associated with the trough. The Ellrod-Knox index had the best turbulence forecast in the trough, while frontogenesis had the fewest number of forecast hits within the trough. Only the Ellrod-Knox and Ellrod-Knapp indices correctly predicted the turbulence associated with the ridge that was moving into the central United States.

i. All Turbulence

During this turbulence outbreak, in terms of the hit rate, the Ellrod-Knox index performed the best at predicting incidences of turbulence. It had the highest number of correct hits, 288, giving it a hit rate of 0.692. The Ellrod-Knapp index was a close second, with a hit rate of 0.625. However, if Figure 6.3b is examined, it is evident that the Ellrod-Knox forecast map had the most area shaded for potential turbulence, so it makes sense that this index had the highest false alarm rate, with a POFD value of 0.391. In terms of the hit rate, frontogenesis performed the worst at predicting turbulence during this outbreak, followed by the Lighthill-Ford

method. However, when predicting null turbulence events, Frontogenesis performed the best, while the Lighthill-Ford method came in at a close second, with POD_n results of 0.968 and 0.904 respectively. The indices with the highest hit rates, the Ellrod-Knox index and the Ellrod-Knapp index, had the lowest POD_n values.

Frontogenesis and the Lighthill-Ford method had the two highest values for the Heidke Skill Score, with HSS values of 0.252 and 0.241. This means that, according to this metric, frontogenesis and the Lighthill-Ford method had the most forecast skill. According to the HSS, the Ellrod-Knox index and the Ellrod-Knapp index both had the least forecast skill during this outbreak, with scores of 0.146 and 0.179, meaning that the forecast skill of these two indices is due mostly to chance.

Vertical wind shear and the Ellrod-Knapp index had the two highest scores for the True Skill Statistic, with values of 0.320 and 0.318 respectively. According to this metric, while forecasting this turbulence outbreak, these indices were the best at being able to discriminate between observed "yes" and "no" turbulence events. Conversely, forecasting turbulence using frontogenesis led to the least ability to discriminate between these types of events, with the smallest TSS value of 0.193.

The Richardson number had the highest Critical Success Index, with a value of 0.202. This means that the Richardson number had the highest probability of having a forecast hit for a forecasted turbulence event. The Lighthill-Ford method was not far behind, with a CSI of 0.201. The Ellrod-Knox index and frontogenesis tied for the worst CSI score, with both indices having values of 0.183. For these two indices, the forecasted turbulence and the observed turbulence events did not correspond well.

Tables 6.2a through 6.2f show the number of hits, misses, false alarms, and correct rejections for each turbulence index. Table 6.3 shows the forecast verification statistics for each of the turbulence indices.

Upon examination of the ROC curve for this event, it is clear that all six indices had some skill in forecasting this event, as all six curves are above the no-skill line. Similarly, the AUC values for each index show that all six indices showed some forecast skill during this event, as none of the indices had AUC values less than 0.500. According to the AUC values, the Ellrod-Knapp index had the most skill at forecasting turbulence during this outbreak; it had an AUC value of 0.746. Frontogenesis performed the worst, with an AUC value of 0.609. Figure 6.5 displays the ROC curves for the six turbulence indices during the December turbulence outbreak, while Table 6.4 shows the AUC values for each turbulence index.

ii. Moderate-or-Greater Turbulence

For the moderate-or-greater turbulence events during this outbreak, results were similar to those from all turbulence events. In terms of the hit rate, the Ellrod-Knox index proved to be the best forecast metric during this turbulence outbreak. For these events, vertical wind shear had the highest false alarm rate, with the Ellrod-Knox index coming in second. For moderate-or-greater reports during this outbreak, the Lighthill-Ford method had the worst hit rate out of the six indices. When predicting situations where turbulence was not expected, the Lighthill-Ford method performed the best, with a POD_n value of 0.994. Vertical wind shear performed the worst at predicting null instances of turbulence, with a POD_n value of 0.884.

For the moderate-or-greater turbulence events, the Richardson number and frontogenesis had the two highest values for the Heidke Skill Score, with HSS values of 0.299 and 0.264.

These values mean that these two indices showed the most forecast skill for predicting moderate-or-greater turbulence during this event. According to this metric, vertical wind shear had the least forecast skill at predicting moderate-or-greater turbulence during this outbreak, with an HSS value of 0.104.

The Ellrod-Knox index and the Ellrod-Knapp index had the two highest values for the True Skill Statistic, with values of 0.466 and 0.388 respectively. This means that while forecasting moderate-or-greater turbulence during this outbreak, these indices were the best at distinguishing between observed and non-observed turbulence events. According to this metric, the Lighthill-Ford method was the worst at distinguishing between these two types of events while forecasting moderate-or-greater turbulence during this case.

The Richardson number had the highest Critical Success Index, with a value of 0.188. This means that the Richardson number had the highest probability of having a forecast hit, depending on whether a moderate-or-greater turbulence event was forecasted, observed, or both. Vertical wind shear had the worst CSI score for moderate-or-greater turbulence, with a value of 0.072. This means that vertical wind shear did not correctly predict a large fraction of the moderate-or-greater turbulence events.

Tables 6.5a through 6.5f show the number of moderate-or-greater turbulence event hits, misses, false alarms, and correct rejections for each turbulence index. Table 6.6 shows the forecast verification statistics for the moderate-or-greater turbulence events.

Upon examination of the ROC curve for the moderate-or-greater events, it is clear that all six indices had some skill in forecasting this event, as all six curves are above the no-skill line. The AUC values for each index also show that all six indices showed some forecast skill during this event, as none of the indices had AUC values less than 0.500. According to the AUC values,

the Ellrod-Knox index had the most skill at forecasting moderate-or-greater turbulence during this outbreak; it had an AUC value of 0.745. The Lighthill-Ford method performed the worst at the moderate-or-greater level, with an AUC value of 0.578. Figure 6.6 displays the moderate-or-greater ROC curves for the six turbulence indices during the December turbulence outbreak, while Table 6.7 shows the AUC values for each turbulence index.

6.2 January 2011 Case

a. Overview

This CAT outbreak occurred from 1800 UTC on 06 January 2011 until 1800 UTC on 07 January 2011. Similar to the previous case, there was very little convection across the United States during this time period. There was some convection in southern Florida at 0000 UTC on 07 January. There was also some convection in southern Texas and western New York between at 0600 UTC on 07 January. This resulted in only five EDR reports being removed from the dataset due to possible convection contamination. After removing all reports west of 104^{o}W, and all reports within 50 miles (80.5 km) of convection, this trimmed the dataset down to 2,499 reports. Figure 6.7 spatially illustrates the turbulence reports by intensity level, while Table 6.8 shows the number of turbulence reports broken down into the turbulence intensity categories. The majority of moderate turbulence observations were centered over the Great Plains and the Gulf Coast, with a few scattered reports over the Ohio River Valley.

b. Synoptic Setup

Upon examination of Figure 6.8a, it is clear that at 1800 UTC on 06 January, there was a large 300-hPa trough over the eastern half of the United States. What is particularly interesting

about this synoptic setup is that the polar and subtropical jets appeared to be merging over the Gulf Coast region. During this time period, there were two jet streaks with windspeeds of 120 knots (61.7 m s^{-1}), one located over the Gulf Coast, and the other located over the Dakotas. Additionally, there were four regions of frontogenesis across the United States. The strongest of the four was moving into North Dakota, while the other three were located over the Gulf Coast, over the Carolina coast, and over New England.

At 0000 UTC on 07 January, the trough had progressed further to the east, and a ridge was beginning to build into the Southern Plains. The two jet streams remained merged, with two 120-knot (61.7 m s^{-1}) jet core maxima. There was not a lot of frontogenesis at this time, only two weak areas over the Gulf of Mexico and over Wisconsin. These features can be seen in Figure 6.8b.

By 0600 UTC, whose synoptic conditions are illustrated in Figure 6.8c, the large trough was still located over the eastern third of the United States, and the ridge continued to build in over the Great Plains. A large 120-knot (61.7 m s^{-1}) jet core stretched from the Dakotas to Kentucky, with a 140-knot (72.0 m s^{-1}) jet streak developing over Iowa. A region of moderate frontogenesis that had been located over Canada began making its way into Montana and North Dakota.

At 1200 UTC on 07 January, the 140-knot (72.0 m s^{-1}) jet core had moved over Kentucky and Tennessee, while the region of 120-knot (61.7 m s^{-1}) wind speeds extended from the Dakotas, all the way to the Georgia/South Carolina coast. By the end of this outbreak, at 1800 UTC on 07 January, the two jet streams remained converged, with a 140-knot (72.0 m s^{-1}) jet core extending from Missouri to North Carolina. At 1200 UTC, the region of frontogenesis over the Dakotas had weakened, while another large region of weak frontogenesis developed over

Nebraska, Kansas, and Oklahoma. By 1800 UTC, the regions of weak frontogenesis over the Dakotas and the Central Plains had merged into one large region. Two other regions of weak frontogenesis developed over Texas and the Gulf Coast states. Figures 6.8d and 6.8e display the synoptic conditions at 1200 UTC and 1800 UTC.

c. Forecast Verification

Figures 6.9a through 6.9f show the forecasted turbulence intensities for the six turbulence indices at 0600 UTC on 07 January 2011 at an altitude of 30,000 feet (9.1 km). Similar to the December case, these turbulence forecasts appear to predict an increased chance of turbulence in the vicinity of the upper-level features, especially along the subtropical jet stream in this case. Three of the indices also appear to catch the region of moderate frontogenesis moving into the Dakotas, as increased turbulence was predicted in this region. Once again, the frontogenesis metric did not forecast one of the large areas of diagnosed frontogenesis.

To further relate turbulence to synoptic weather patterns, the locations of observed turbulence events (forecast hits and forecast misses) were plotted and compared to the large-scale weather patterns. Figures 6.10a through 6.10f illustrate the locations of observed turbulence events at 0600 UTC on 07 January 2011. Upon comparing these figures to the synoptic maps in Figure 6.8b, it is evident that turbulence occurred within the trough that covered the eastern half of the United States during this time period. Almost all of the turbulence events are located within the trough, and a large majority of these turbulence events coincided with the position of either of the jet streams, with more turbulence being reported along the southern jet stream. Interestingly, the Lighthill-Ford method was the only turbulence index to correctly forecast the turbulence that occurred over Paducah, KY. This turbulence occurred in the middle of a deep trough, which is where the Lighthill-Ford method should

perform the best. For this case, frontogenesis did not appear to be a major cause of turbulence events. For example, at 0600 UTC on 07 January 2011, there was only one turbulence report located along the Texas/Oklahoma border that could have been associated with a small region of diagnosed frontogenesis located over the Texas panhandle. During this case study, all six forecast indices showed turbulence occurring in similar locations in the troughs and along the jet streams.

i. All Turbulence

Also similar to the December case, the Ellrod-Knox index performed well in terms of the hit rate during this turbulence outbreak. This index had the highest hit rate, with a value of 0.734. The Ellrod-Knox index also had the highest false alarm rate, with a value of 0.361, following closely behind was vertical wind shear, with a false alarm rate of 0.349. Frontogenesis and the Lighthill-Ford method had the two lowest hit rates of 0.342 and 0.476 respectively. In terms of predicting reports of null turbulence, frontogenesis and the Lighthill-Ford method performed the best, with POD_n values of 0.930 and 0.783. The Ellrod-Knox index had the lowest POD_n score, followed closely by vertical wind shear.

The indices with the highest values for the Heidke Skill Score were frontogenesis and the Ellrod-Knapp index, with HSS values of 0.305 and 0.259. This means that these two indices had the most forecast skill during this outbreak. According to this index, vertical wind shear and the Lighthill-Ford method were correctly predicting turbulence mostly by chance with little actual forecast skill, because they had the lowest Heidke Skill Score values of 0.186 and 0.202.

The Ellrod-Knapp and Ellrod-Knox indices had the two highest values for the True Skill Statistic, with TSS scores of 0.392 and 0.373 respectively. These scores mean that during this

outbreak, these two indices were the best at differentiating between observed and non-observed turbulence events. During this event, the forecast produced by the Lighthill-Ford method performed the poorest at differentiating between these two types of events, with a True Skill Score of 0.259.

The Ellrod-Knapp index had the highest Critical Success Index, with a value of 0.262. This means that the forecast produced by the Ellrod-Knapp index had the largest proportion of correctly predicted turbulence events. The Lighthill-Ford method performed the worst, with a CSI score of 0.215. The Lighthill-Ford method had the fewest correctly forecasted turbulence events.

Tables 6.9a through 6.9f show the number of hits, misses, false alarms, and correct rejections for each turbulence index. Table 6.10 shows the forecast verification statistics for each of the turbulence indices.

Upon examination of the ROC curve for this event, it is clear that all six indices had some skill in forecasting this event, as all six curves are above the no-skill line. Once again, the AUC values for each index also show that all six indices showed some forecast skill during this event, as none of the indices had AUC values less than 0.500. According to the AUC values, the Ellrod-Knox index had the most skill at forecasting turbulence during this outbreak; it had an AUC value of 0.758. The Lighthill-Ford method performed the worst, with an AUC value of 0.674. Figure 6.11 displays the ROC curves for the six turbulence indices during the January turbulence outbreak, while Table 6.11 shows the AUC values for each turbulence index.

ii. Moderate-or-Greater Turbulence

For the moderate-or-greater turbulence events, the Ellrod-Knox index once again proved to have the best hit rate during this event, with a value of 0.486. The Lighthill-Ford method and frontogenesis both performed poorly during this event, with hit rates of 0.156 and 0.294 respectively. In terms of predicting areas of null turbulence, the Lighthill-Ford method performed the best. This method had a very large POD_n value of 0.976. While vertical wind shear was second best at predicting reports of turbulence, it had the lowest POD_n value of 0.833.

For the moderate-or-greater turbulence events, frontogenesis and the Richardson number had the two highest values for the Heidke Skill Score, with scores of 0.269 and 0.214. These values mean that while forecasting for moderate-or-greater intensity turbulence, these two indices had the most forecast skill. According to this metric, vertical wind shear had the least forecast skill during this outbreak, with a HSS of 0.100.

For moderate-or-greater turbulence events during this outbreak, the Ellrod-Knox index had the largest True Skill Statistic, with a value of 0.343. This means that while forecasting moderate-or-greater turbulence during this outbreak, the Ellrod-Knox index was the best at discriminating between "yes" and "no" turbulence events. The Lighthill-Ford method was the worst at discriminating between these events, with the smallest TSS of 0.132.

While forecasting moderate-or-greater turbulence, frontogenesis had the highest Critical Success Index, with a value of 0.177. This means that frontogenesis had the highest probability of having a forecast hit, depending on whether a moderate-or-greater turbulence event was forecasted, observed, or both. Vertical wind shear had the lowest CSI at the moderate-or-greater turbulence level, with a value of 0.089.

Tables 6.12a through 6.12f show the number of moderate-or-greater event hits, misses, false alarms, and correct rejections for each turbulence index. Table 6.13 displays the forecast verification statistics for all moderate-or-greater turbulence reports.

Upon examination of the ROC curve for moderate-or-greater turbulence during this event, it is clear that all six indices had some skill, as all six curves are above the no-skill line. The AUC values for each index show that all six indices showed some forecast skill during this event, as none of the indices had AUC values less than 0.500. According to the AUC values for moderate-or-greater turbulence events, the Ellrod-Knox index had the most forecast skill during this outbreak; it had an AUC value of 0.679. The Lighthill-Ford method performed the worst, with an AUC value of 0.566. Figure 6.12 displays the moderate-or-greater ROC curves for the six turbulence indices during the January turbulence outbreak, while Table 6.14 shows the AUC values for each turbulence index.

6.3 September 2011 Case

a. Overview

This CAT outbreak occurred from 1200 UTC on 22 September 2011 until 1800 UTC on 22 September 2011. Figure 6.13 spatially illustrates the turbulence reports by intensity level, while Table 6.15 shows the number of turbulence reports within each intensity category. The majority of moderate reports of turbulence were centered over the Upper Midwest and the Mid-Atlantic region.

Unlike the previous two cases, there was some convection across portions of the United States during this turbulence outbreak. Figures 6.14a and 6.14b show the turbulence reports overlaid with convection that is 50 dBZ or greater. One hundred and twenty EDR reports had to

be removed from the dataset because they were located too close to deep convection. After removing all reports west of 104°W, and all reports within 50 miles (80.5 km) of convection, this trimmed the dataset to 625 reports.

b. Synoptic Setup

During this turbulence outbreak, there was a large trough located over the central United States. At 1200 UTC, a swath of 100-knot (51.4 m s^{-1}) winds stretched from Missouri north to Michigan, and by 1800 UTC, a small 120-knot (61.7 m s^{-1}) jet streak had developed over Wisconsin and Michigan. At 1200 UTC, there were two regions of weak frontogenesis occurring in the 700 to 500 hPa layer, one through the South-Central Plains, and the other over the Dakotas. By 1800 UTC, these frontogenesis regions had weakened and broken into four smaller regions: two were over the Dakotas, one was over Michigan, and the last stretched from Oklahoma to Indiana. Figure 6.15a illustrates the synoptic conditions at 1200 UTC, while Figure 6.15b shows the large scale weather features at 1800 UTC.

c. Forecast Verification

Just like the December and January cases, the turbulence indices appear to predict an increased chance of turbulence along certain upper-level features, for example, along troughs and in regions of frontogenesis. On this date, the Richardson Number was the only index forecasting a small area with an increased chance for severe turbulence. Figures 6.16a through 6.16f illustrate turbulence forecasts for 1800 UTC on 22 September 2011 for an altitude of 30,000 feet (9.1 km).

To further relate turbulence to synoptic weather patterns, the locations of observed turbulence events (forecast hits and forecast misses) were plotted and compared to the large-scale weather patterns. Figures 6.17a through 6.17f illustrate the locations of observed turbulence events at 1800 UTC on 22 September 2011. Upon comparing these figures to the synoptic maps in Figure 6.15b, it is evident that turbulence occurred within the trough that covered the eastern two-thirds of the United States during this time period. Almost all of the turbulence events are located within the trough, and a large majority of these turbulence events, especially those over the northern and Central Plains and the Upper Midwest, coincided with the fast wind speeds associated with the jet stream. There was also an area of high wind speeds stretching from West Virginia to New England that was associated with many turbulence observations. In this case study, there were several areas of diagnosed frontogenesis that were associated with reports of turbulence, however, all of these areas were in regions that had high wind speeds as well. For example, at 1800 UTC on 22 September 2011, there appeared to be a nice correlation between observed turbulence, the core of the jet stream, and diagnosed frontogenesis over an area that stretched from Illinois to Michigan. Interestingly, the Lightill-Ford method had more forecast hits along the eastern coast of the United States, a region that was in the outer edge of the trough. These hits may have been associated with scattered convection that was occurring along the east coast during this time. During this case study, all six forecast indices showed turbulence occurring in fairly similar locations within the trough and along the jet stream.

i. All Turbulence

Again, similar to both the December and January cases, the Ellrod-Knox index did well at forecasting turbulence events during this outbreak. It had the highest hit rate of the six indices, with a value of 0.738. The Ellrod-Knapp index also had a high hit rate of 0.672. Using frontogenesis as a proxy for turbulence did not work as well during this outbreak, as this method had the lowest hit rate, with a value of 0.328. In terms of forecasting areas of null turbulence, frontogenesis and the Lighthill-Ford method both did well. The Richardson number had a POD_n of 0.830, while the Lighthill-Ford method had a POD_n value of 0.710. The Ellrod-Knox index had the lowest POD_n, with a score of 0.430.

The indices with the highest values for the Heidke Skill Score were the Lighthill-Ford method and the Richardson number, with scores of 0.272 and 0.223. This means that these two indices had the most forecast skill during this turbulence outbreak, and that most of their correct forecasts were not due to chance. According to this metric, vertical wind shear and the Ellrod-Knox index had the least forecast skill during this turbulence outbreak, because they had the two lowest Heidke Skill Scores, with values of 0.077 and 0.126 respectively.

The Lighthill-Ford method and the Richardson number had the two highest values for the True Skill Statistic, meaning that their forecasts were the best at distinguishing between "yes" and "no" turbulence events. The TSS scores of the Lighthill-Ford method and the Richardson number were 0.295 and 0.257. During this event, the forecast produced by vertical wind shear performed the worst at distinguishing between these types of events, with a True Skill Score of 0.093.

The Lighthill-Ford method had the highest Critical Success Index, with a value of 0.344. This means that the Lighthill-Ford method had the highest fraction of turbulence events that were

correctly predicted. Frontogenesis had the worst CSI score of 0.233, meaning that this index had the smallest fraction of correctly predicted turbulence events.

Tables 6.16a through 6.16f show the number of hits, misses, false alarms, and correct rejections for each turbulence index. Table 6.17 shows the forecast verification statistics for each of the turbulence indices.

Upon examination of the ROC curve for this event, it is clear that all six indices had some skill in forecasting this event, as all six curves are above the no-skill line. Once again, the AUC values for each index also show that all six indices showed some forecast skill during this event, as none of the indices had AUC values less than 0.500. According to the AUC values, the Lighthill-Ford method had the most skill at forecasting turbulence during this outbreak; it had an AUC value of 0.687. Vertical wind shear performed the worst, with an AUC value of 0.590. Figure 6.18 displays the ROC curves for the six turbulence indices during the September turbulence outbreak, while Table 6.18 shows the AUC values for each turbulence index.

ii. Moderate-or-Greater Turbulence

For the moderate-or-greater events, once again, the Ellrod-Knox index had the highest hit rate out of the six indices, with a value of 0.449. At the moderate-or-greater scale, the Lighthill-Ford method and frontogenesis performed the worst at predicting incidences of turbulence. They had the lowest hit rates of all methods, with a value of 0.184. In terms of predicting areas of null turbulence at the moderate-or-greater scale, frontogenesis and the Lighthill-Ford method performed almost equally well. Frontogenesis had a POD_n value of 0.938, with the Lighthill-Ford method following closely behind with a POD_n of 0.931. While the Ellrod-Knox index did well at forecasting areas where turbulence was expected, it had the lowest POD_n of 0.727.

For the moderate-or-greater turbulence events, the Richardson number and frontogenesis had the two highest values for the Heidke Skill Score, with values of 0.152 and 0.126. These values mean that these indices had the most skill at forecasting moderate-or-greater turbulence, and that their correct forecasts were not due to chance. According to this metric, vertical wind shear had the least skill at forecasting moderate-or-greater turbulence events during this outbreak, with a HSS value of 0.065.

For moderate-or-greater turbulence events during this outbreak, the Richardson number had the largest True Skill Statistic, with a value of 0.205. This means that while forecasting moderate-or-greater turbulence during this outbreak, this index was the best at differentiating between observed and non-observed turbulence events. The Lighthill-Ford method had the lowest True Skill Statistic, with a value of 0.114. This means that the Lighthill-Ford method was the worst at distinguishing between observed and non-observed events while forecasting turbulence of at least moderate intensity during this outbreak.

The Richardson number had the highest Critical Success Index, with a value of 0.134. This means that the Richardson number had the highest probability of forecasting a moderate-or-greater turbulence hit, depending on whether a turbulence event was forecasted, observed, or both. Vertical wind shear had the smallest CSI, with a score of 0.095.

Tables 6.19a through 6.19f show the number of moderate-or-greater event hits, misses, false alarms, and correct rejections for each turbulence index. Table 6.20 shows the forecast verification statistics for the moderate-or-greater turbulence reports for this case.

Upon examination of the ROC curve for this event, it is clear that all six indices had some skill in forecasting moderate-or-greater turbulence during this outbreak, as all six curves are above the no-skill line. The AUC values for each index show that all six indices showed some

skill when forecasting moderate-or-greater turbulence, as none of the indices had AUC values less than 0.500. According to the AUC values, the Ellrod-Knox index had the most forecast skill at predicting moderate-or-greater turbulence during this outbreak; it had an AUC value of 0.708. The Lighthill-Ford method performed the worst, with an AUC value of 0.558. Figure 6.19 displays the ROC curves for the six turbulence indices during the September turbulence outbreak, while Table 6.21 shows the AUC values for each turbulence index.

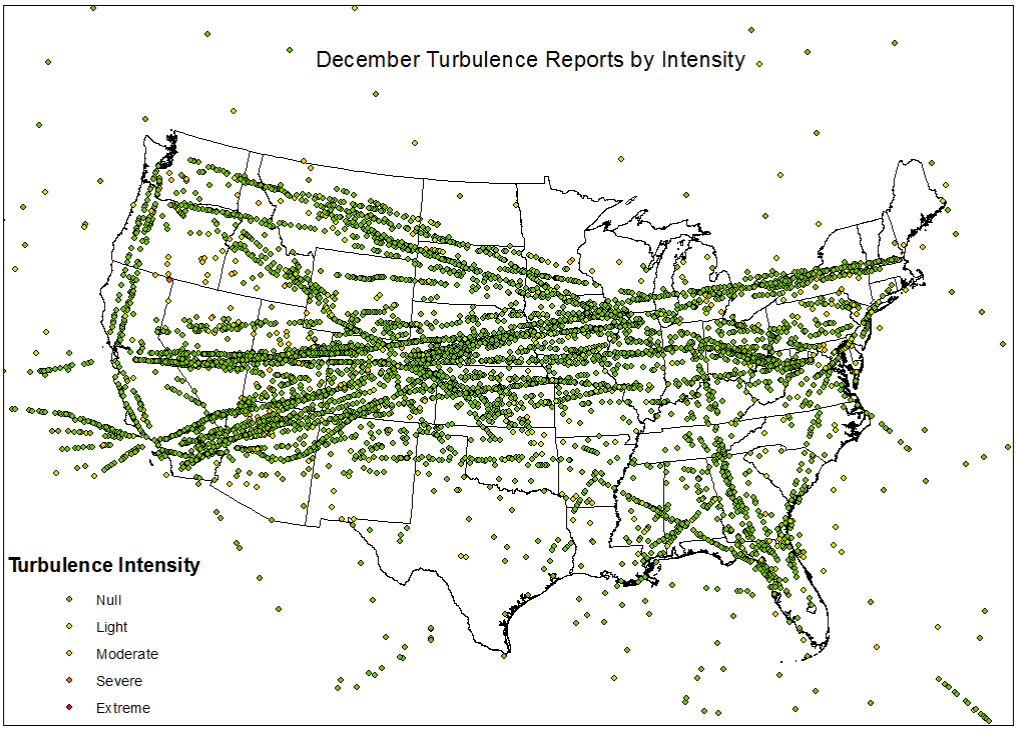
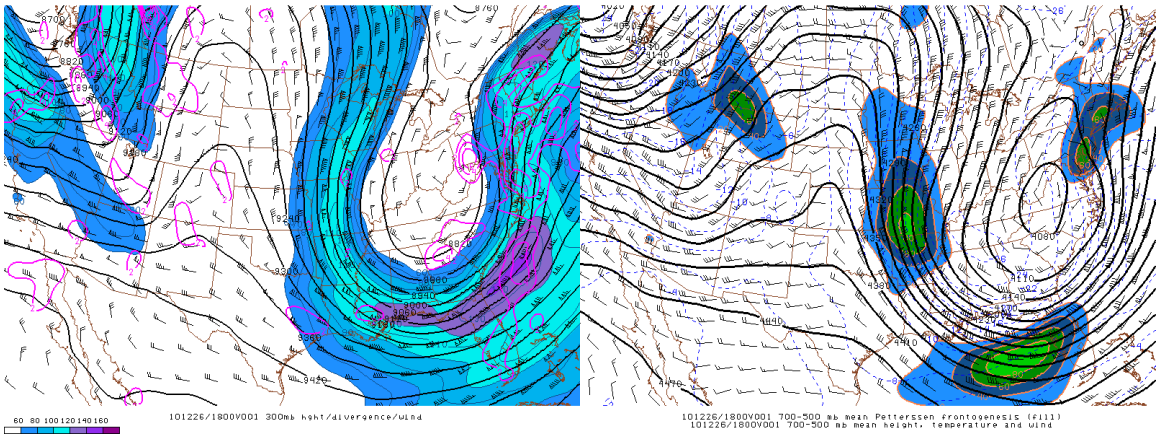


Figure 6.1: EDR data from the December 2010 turbulence outbreak, sorted by turbulence intensity.

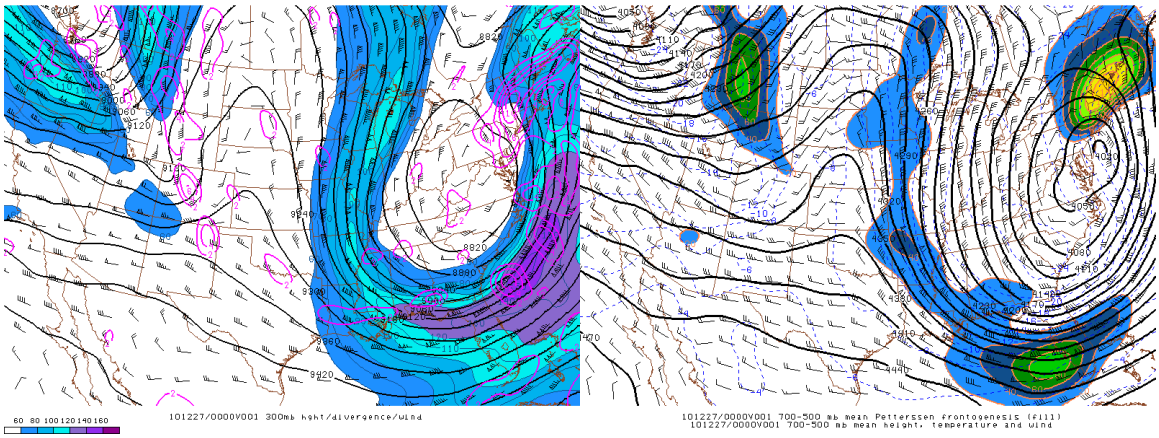
Table 6.1: Turbulence reports by intensity for the December 2010 outbreak.

	Null	Light	Moderate	Severe
All Turbulence Reports	3280	387	74	0

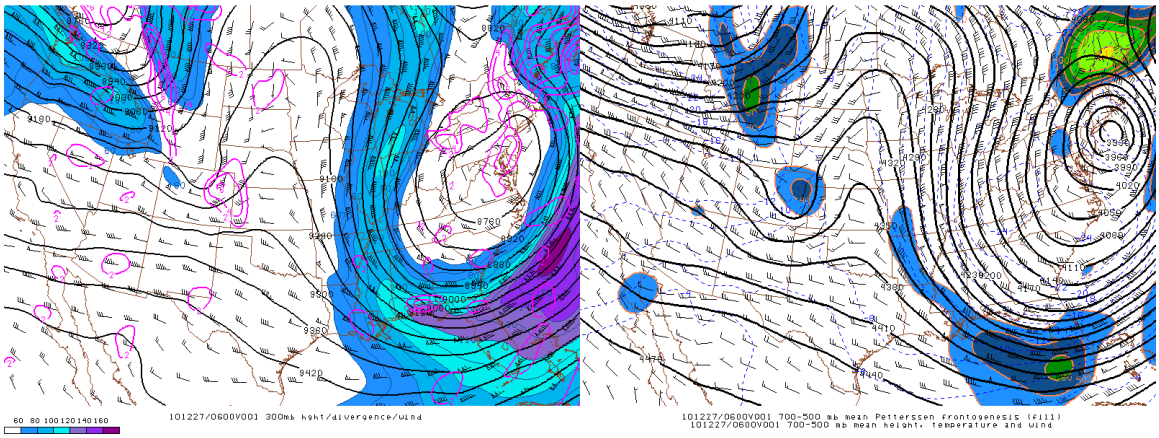
(a)


101226/1800V001 300mb hght/divergence/wind
60 80 100 120 140 160
101226/1800V001 700-500 mb mean Petterssen frontogenesis (fill)
101226/1800V001 700-500 mb mean height, temperature and wind

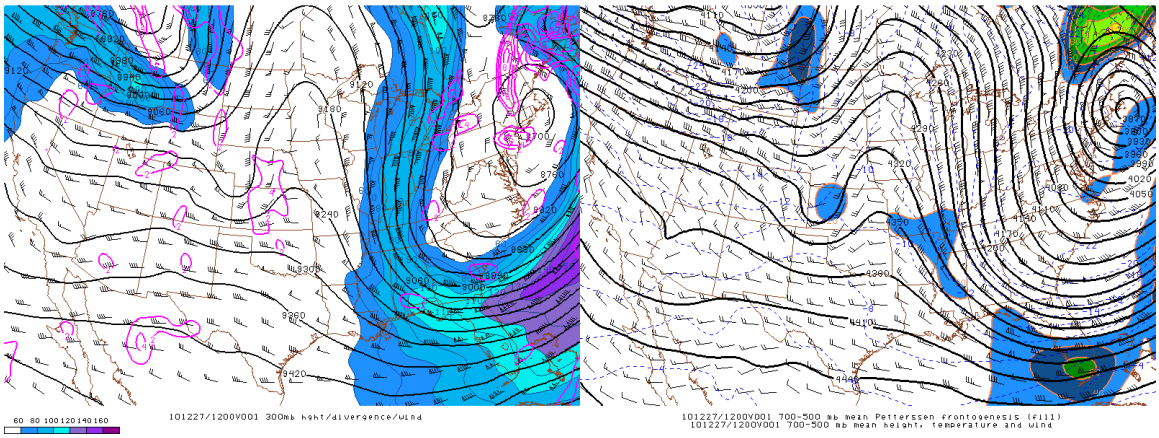
(b)


101227/0000V001 300mb hght/divergence/wind
60 80 100 120 140 160
101227/0000V001 700-500 mb mean Petterssen frontogenesis (fill)
101227/0000V001 700-500 mb mean height, temperature and wind

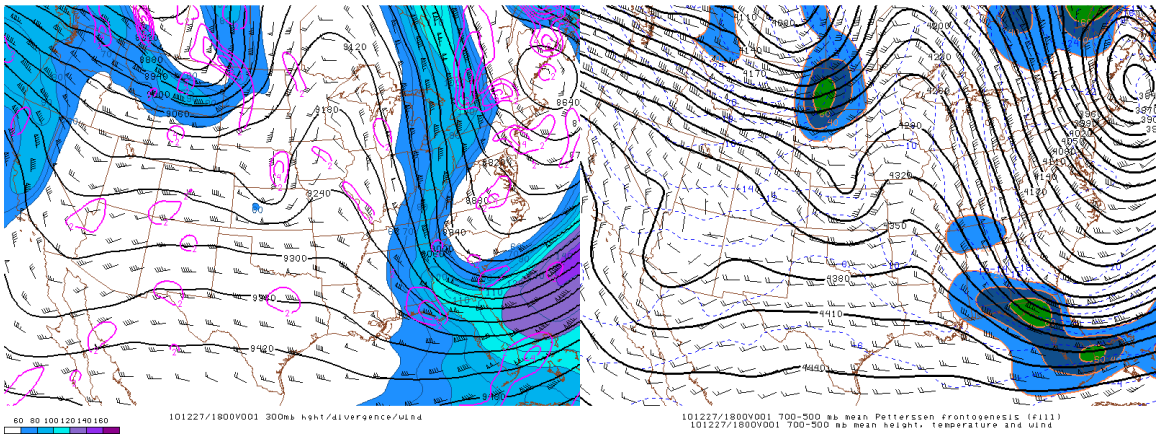
(c)


101227/0600V001 300mb hght/divergence/wind
60 80 100 120 140 160
101227/0600V001 700-500 mb mean Petterssen frontogenesis (fill)
101227/0600V001 700-500 mb mean height, temperature and wind

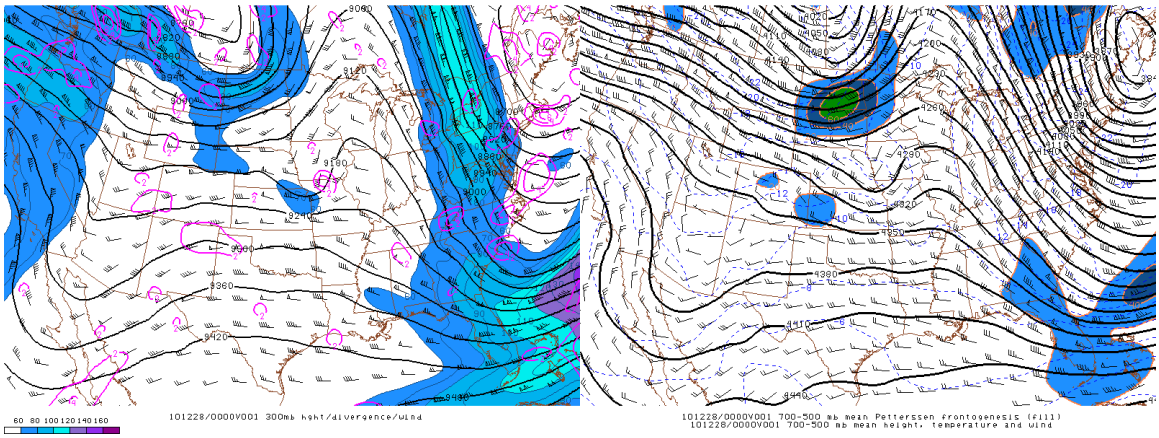
(d)


101227/1200V001 300mb hght/divergence/wind
60 80 100 120 140 160
101227/1200V001 700-500 mb mean Petterssen frontogenesis (fill)
101227/1200V001 700-500 mb mean height, temperature and wind

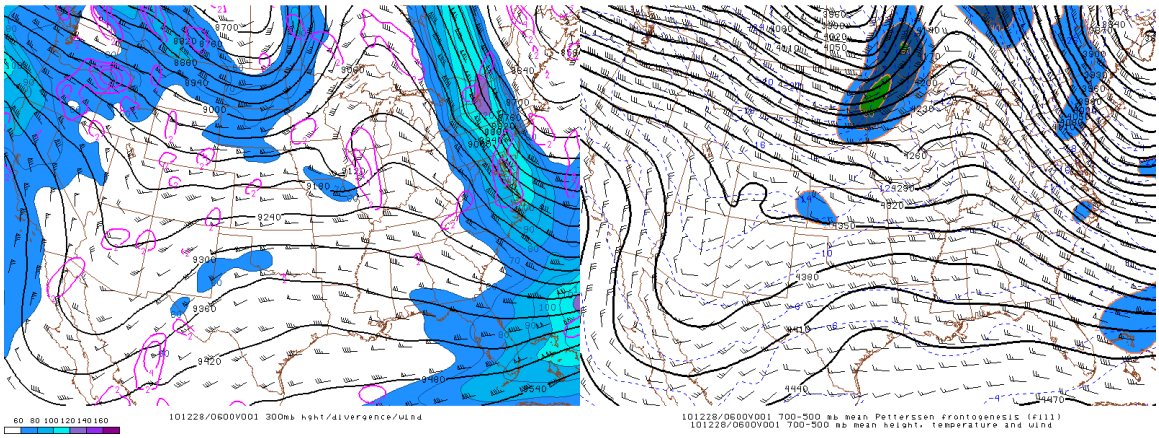
(e)


101227/1800V001 300mb hght/divergence/wind
60 80 100 120 140 160
101227/1800V001 700-500 mb mean Petterssen frontogenesis (fill)
101227/1800V001 700-500 mb mean height, temperature and wind

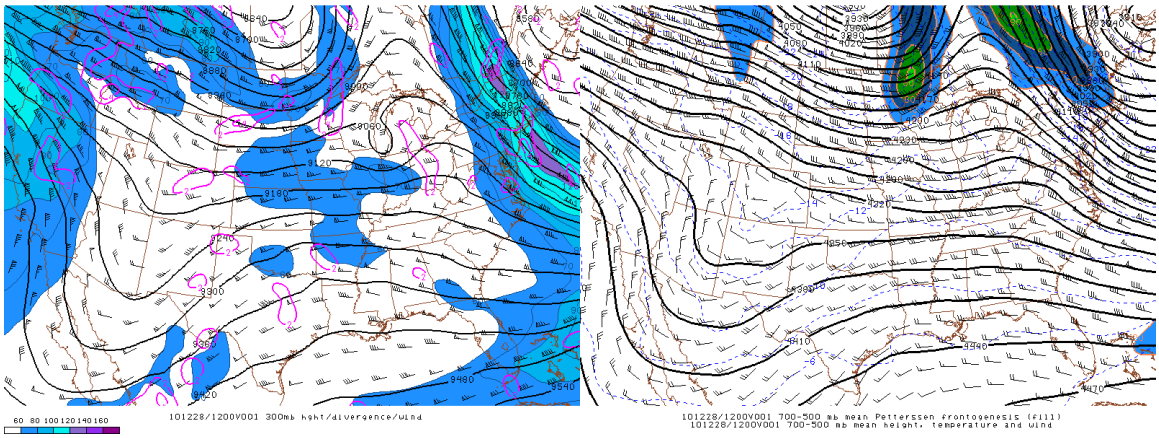
(f)


101228/0000V001 300mb hght/divergence/wind
60 80 100 120 140 160
101228/0000V001 700-500 mb mean Petterssen frontogenesis (fill)
101228/0000V001 700-500 mb mean height, temperature and wind

(g)


101228/0600V001 300mb hght/divergence/wind
60 80 100 120 140 160
101228/0600V001 700-500 mb mean Petterssen frontogenesis (fill)
101228/0600V001 700-500 mb mean height, temperature and wind

(h)


101228/1200V001 300mb hght/divergence/wind
60 80 100 120 140 160
101228/1200V001 700-500 mb mean Petterssen frontogenesis (fill)
101228/1200V001 700-500 mb mean height, temperature and wind

(i)

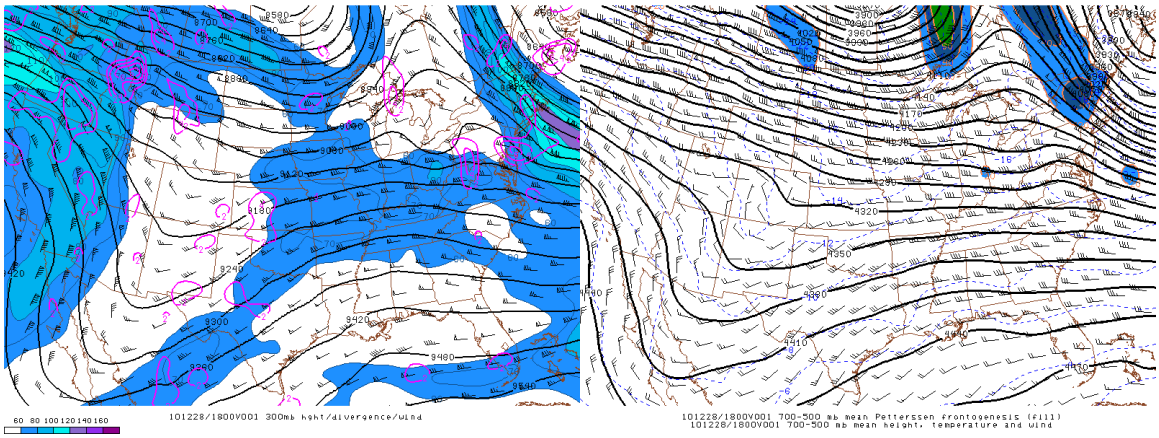
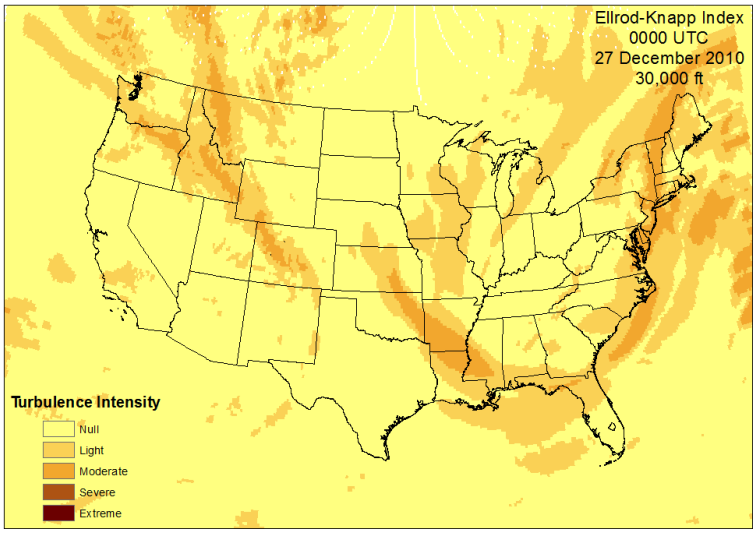
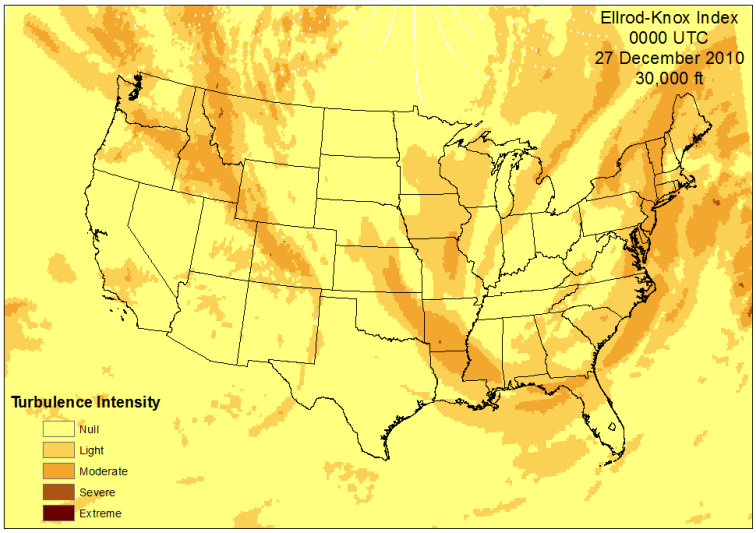

101228/1800V001 300mb hght/divergence/wind
60 80 100 120 140 160
101228/1800V001 700-500 mb mean Petterssen frontogenesis (fill)
101228/1800V001 700-500 mb mean height, temperature and wind

Figure 6.2: Synoptic maps for the December 2010 outbreak at (a) 1800 UTC on 26 December 2010, (b) 0000 UTC on 27 December 2010, (c) 0600 UTC on 27 December 2010, (d) 1200 UTC on 27 December 2010, (e) 1800 UTC on 27 December 2010, (f) 0000 UTC on 28 December 2010, (g) 0600 UTC on 28 December 2010, (h) 1200 UTC on 28 December 2010, and (i) 1800 UTC on 28 December 2010. In the left column, 300 hPa heights, windspeeds, and divergence are shown. Heights are shown by the solid black contours, divergence is shown by the purple contours, and wind speeds are illustrated by the shaded area. In the right column, 700 hPa to 500 hPa Petterssen frontogenesis, heights, and temperatures are illustrated. Heights are shown by the solid black contours, temperatures are illustrated using the dashed blue contours, and mean frontogenesis is shown in the shaded area. (Source: Storm Prediction Center)

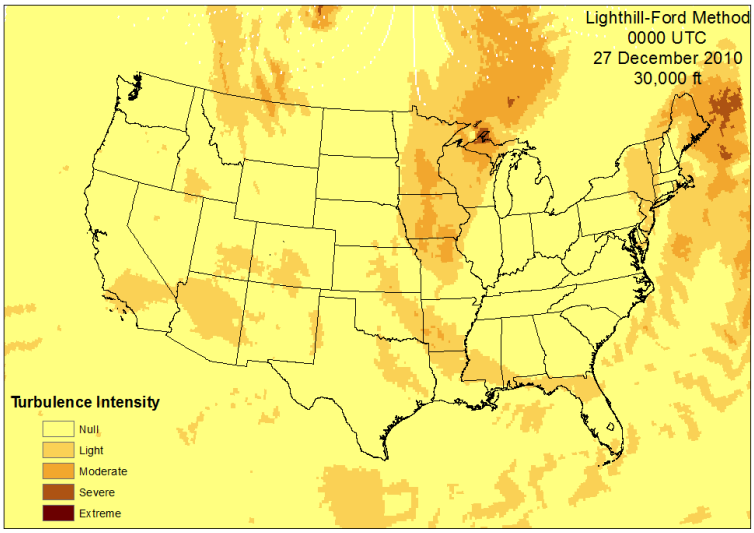
(a)


Ellrod-Knapp Index
0000 UTC
27 December 2010
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

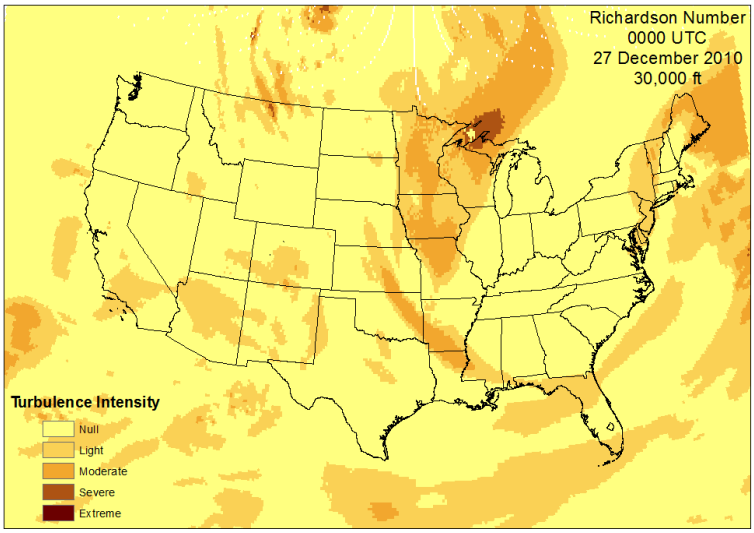
(b)


Ellrod-Knox Index
0000 UTC
27 December 2010
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

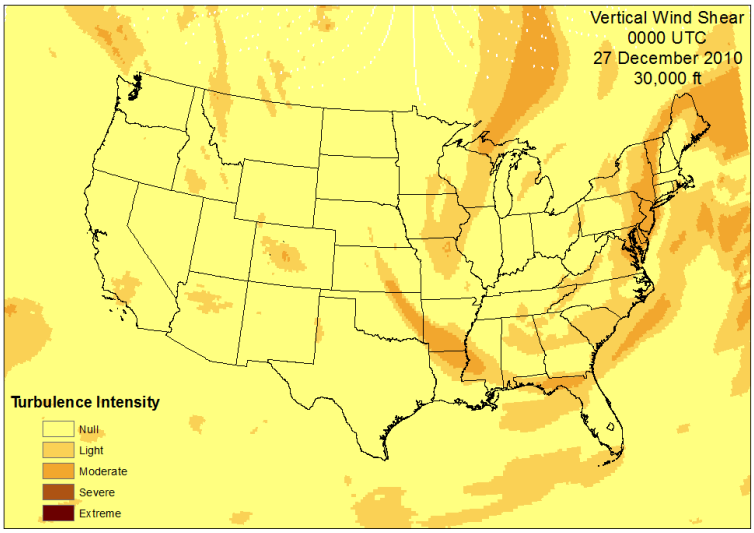
(c)


Lighthill-Ford Method
0000 UTC
27 December 2010
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

(d)


Richardson Number
0000 UTC
27 December 2010
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

(e)


Vertical Wind Shear
0000 UTC
27 December 2010
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

(f)

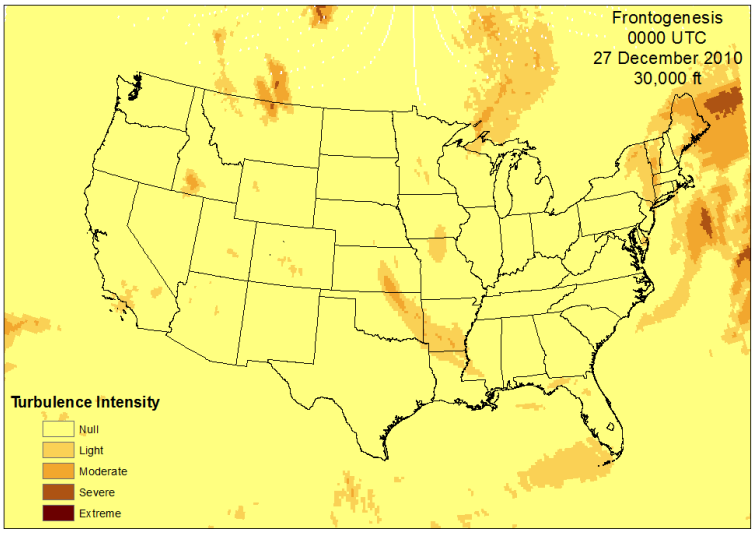
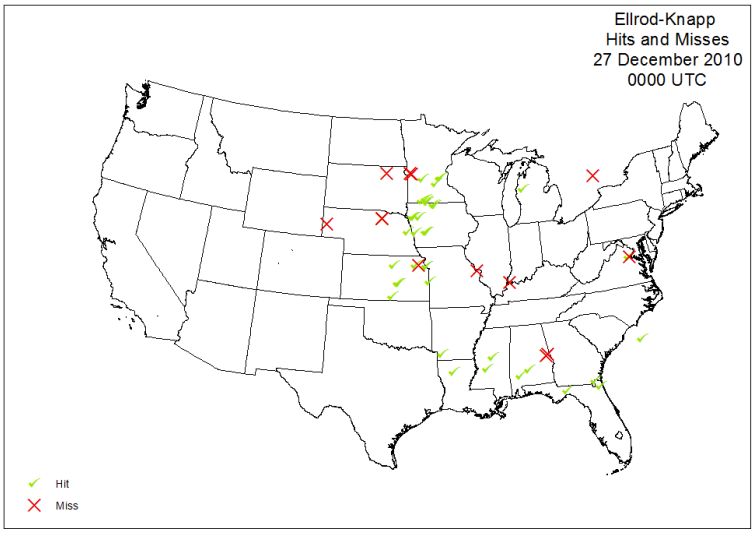
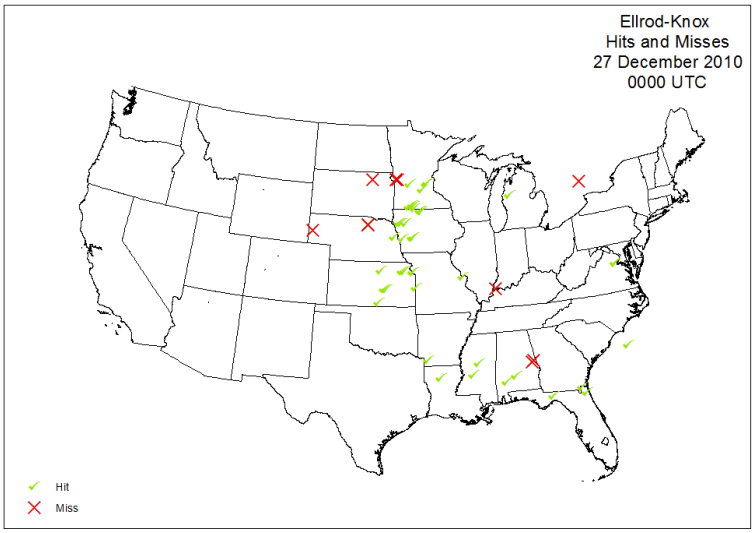

Frontogenesis
0000 UTC
27 December 2010
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

Figure 6.3: Turbulence forecasts using (a) the Ellrod-Knapp index, (b) the Ellrod-Knox index, (c) the Lighthill-Ford method, (d) the Richardson number, (e) vertical wind shear, and (f) frontogenesis for 0000 UTC on 27 December 2010 for the altitude of 30,000 feet (9.1 km).

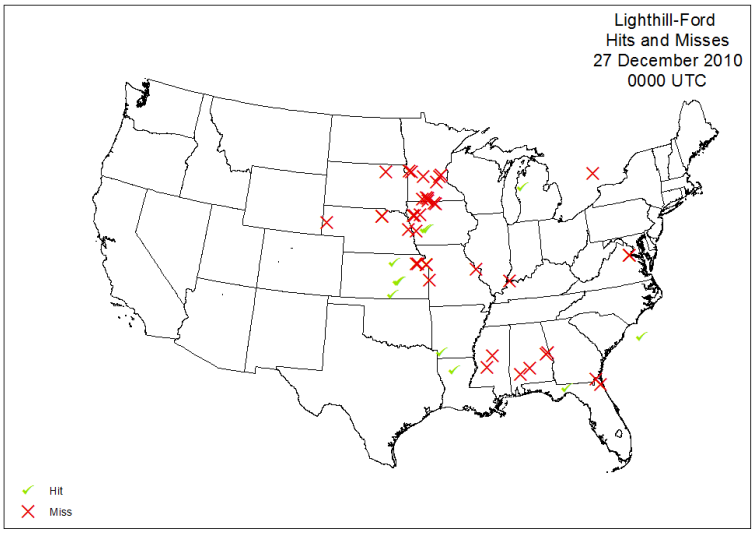
(a)


Ellrod-Knapp
Hits and Misses
27 December 2010
0000 UTC
Hit
Miss

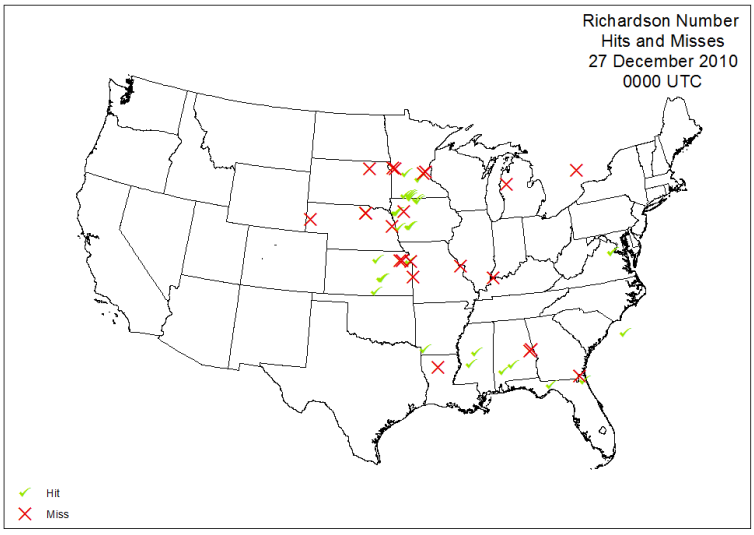
(b)


Ellrod-Knox
Hits and Misses
27 December 2010
0000 UTC
Hit
Miss

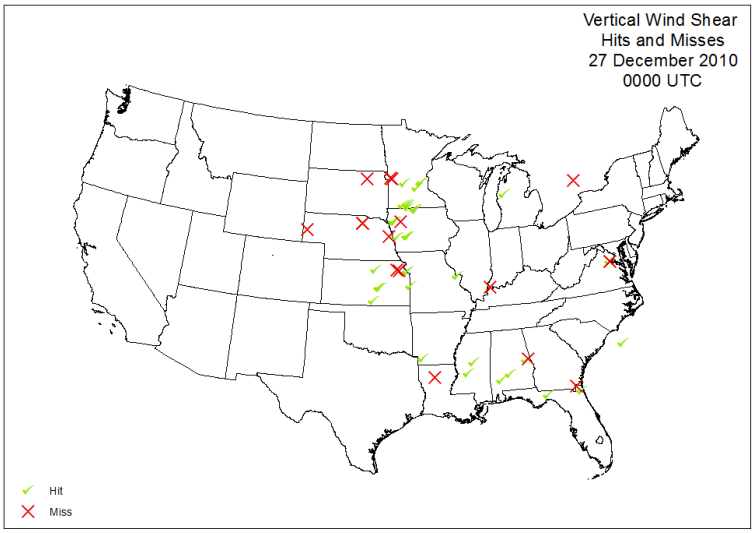
(c)


Lighthill-Ford
Hits and Misses
27 December 2010
0000 UTC
Hit
Miss

(d)


Richardson Number
Hits and Misses
27 December 2010
0000 UTC
Hit
Miss

(e)


Vertical Wind Shear
Hits and Misses
27 December 2010
0000 UTC
Hit
Miss

(f)

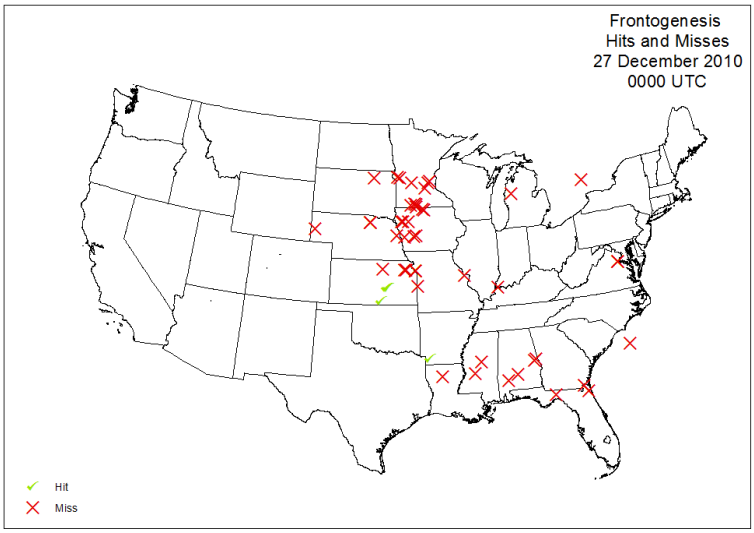

Frontogenesis
Hits and Misses
27 December 2010
0000 UTC
Hit
Miss

Figure 6.4: Turbulence forecast hits and misses using (a) the Ellrod-Knapp index, (b) the Ellrod-Knox index, (c) the Lighthill-Ford method, (d) the Richardson number, (e) vertical wind shear, and (f) frontogenesis for 0000 UTC on 27 December 2010.

Table 6.2: Forecast verifications for all turbulence intensities from the December 2010 turbulence outbreak.

(a)

Ellrod-Knapp	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 288	*False Alarms* 1007
Forecasted No Turbulence	*Misses* 173	*Correct Rejections* 2273

(b)

Ellrod-Knox	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 319	*False Alarms* 1284
Forecasted No Turbulence	*Misses* 142	*Correct Rejections* 1996

(c)

Lighthill-Ford	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 156	*False Alarms* 314
Forecasted No Turbulence	*Misses* 305	*Correct Rejections* 2966

(d)

Richardson Number	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 194	*False Alarms* 498
Forecasted No Turbulence	*Misses* 267	*Correct Rejections* 2782

(e)

Vertical Wind Shear	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 287	*False Alarms* 994
Forecasted No Turbulence	*Misses* 174	*Correct Rejections* 2286

(f)

Frontogenesis	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 104	*False Alarms* 106
Forecasted No Turbulence	*Misses* 357	*Correct Rejections* 3174

Table 6.3: Forecast verification statistics for the December 2010 outbreak.

Turbulence Index	POD_y	POD_n	POFD	FAR	HSS	TSS	CSI
Ellrod-Knapp	0.625	0.693	0.307	0.778	0.179	0.318	0.196
Ellrod-Knox	0.692	0.609	0.391	0.801	0.146	0.301	0.183
Lighthill-Ford	0.338	0.904	0.096	0.668	0.241	0.243	0.201
Richardson Number	0.421	0.848	0.152	0.720	0.221	0.269	0.202
Vertical Wind Shear	0.623	0.697	0.303	0.776	0.181	0.320	0.197
Frontogenesis	0.226	0.968	0.032	0.505	0.252	0.193	0.183

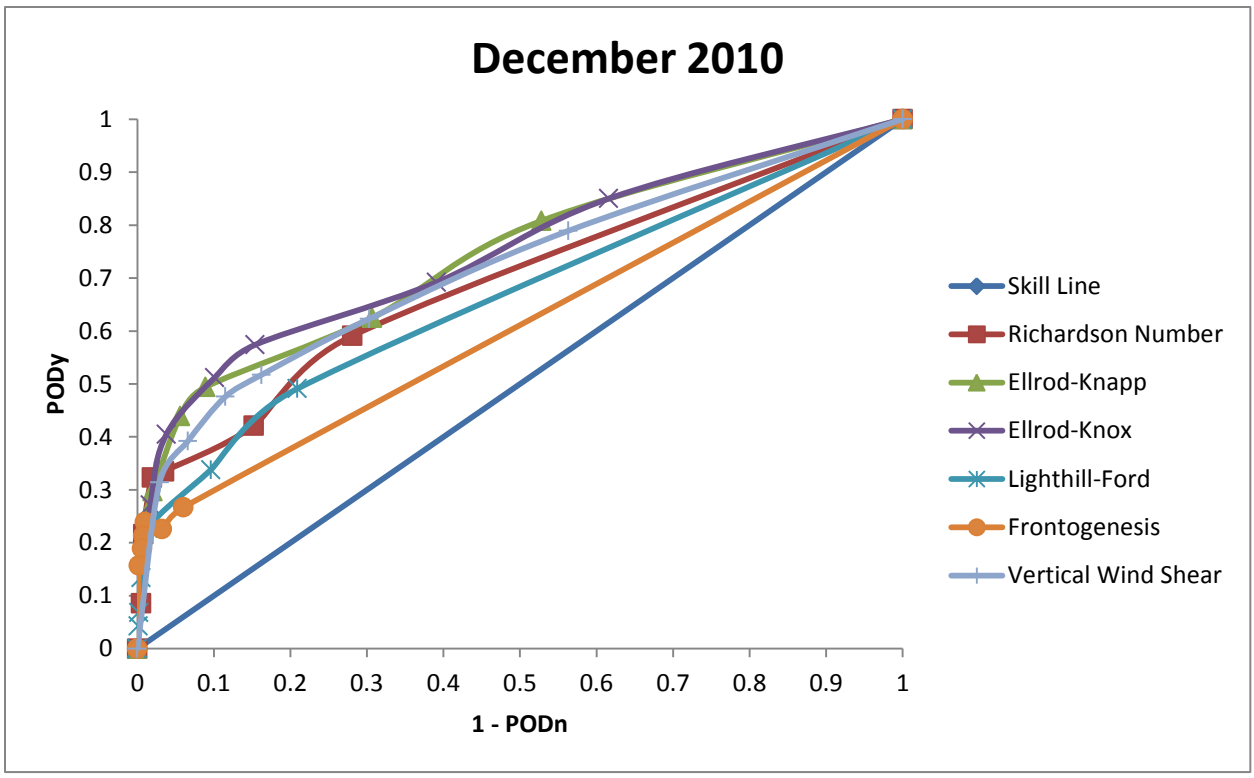


Figure 6.5: ROC curve for the turbulence indices during the December 2010 outbreak.

Table 6.4: AUC values for the six turbulence indices during the December 2010 outbreak.

Turbulence Index	AUC
Ellrod-Knox	0.746
Ellrod-Knapp	0.740
Vertical Wind Shear	0.717
Richardson Number	0.690
Lighthill-Ford	0.662
Frontogenesis	0.609

Table 6.5: Forecast verifications for moderate-or-greater (MOG) turbulence from the December 2010 turbulence outbreak.

(a)

Ellrod-Knapp	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 33	*False Alarms* 212
Forecasted No MOG Turbulence	*Misses* 41	*Correct Rejections* 3455

(b)

Ellrod-Knox	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 42	*False Alarms* 374
Forecasted No MOG Turbulence	*Misses* 32	*Correct Rejections* 3293

(c)

Lighthill-Ford	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 12	*False Alarms* 21
Forecasted No MOG Turbulence	*Misses* 62	*Correct Rejections* 3646

(d)

Richardson Number	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 30	*False Alarms* 86
Forecasted No MOG Turbulence	*Misses* 44	*Correct Rejections* 3581

(e)

Vertical Wind Shear	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 36	*False Alarms* 426
Forecasted No MOG Turbulence	*Misses* 38	*Correct Rejections* 3241

(f)

Frontogenesis	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 19	*False Alarms* 44
Forecasted No MOG Turbulence	*Misses* 55	*Correct Rejections* 3623

Table 6.6: Moderate-or-greater turbulence forecast verification statistics for the December 2010 outbreak.

Turbulence Index	POD_y	POD_n	POFD	FAR	HSS	TSS	CSI
Ellrod-Knapp	0.446	0.942	0.058	0.865	0.182	0.388	0.115
Ellrod-Knox	0.568	0.898	0.102	0.899	0.143	0.466	0.094
Lighthill-Ford	0.162	0.994	0.006	0.636	0.215	0.156	0.126
Richardson Number	0.405	0.977	0.023	0.741	0.299	0.382	0.188
Vertical Wind Shear	0.486	0.884	0.116	0.922	0.104	0.370	0.072
Frontogenesis	0.257	0.988	0.012	0.698	0.264	0.245	0.161

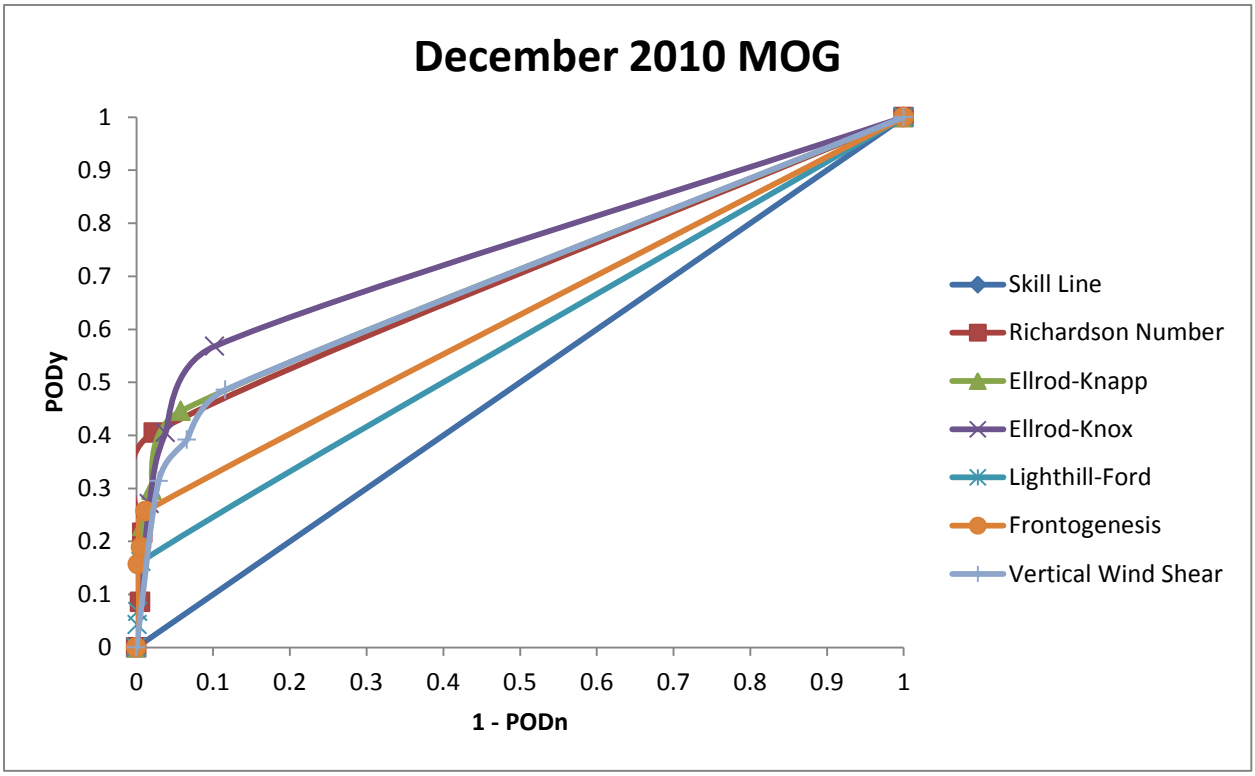


Figure 6.6: ROC curve for the turbulence indices at the moderate-or-greater scale during the December 2010 outbreak.

Table 6.7: AUC values for the six turbulence indices at the moderate-or-greater scale during the December 2010 outbreak.

Turbulence Index	AUC
Ellrod-Knox	0.745
Ellrod-Knapp	0.699
Vertical Wind Shear	0.696
Richardson Number	0.692
Frontogenesis	0.623
Lighthill-Ford	0.578

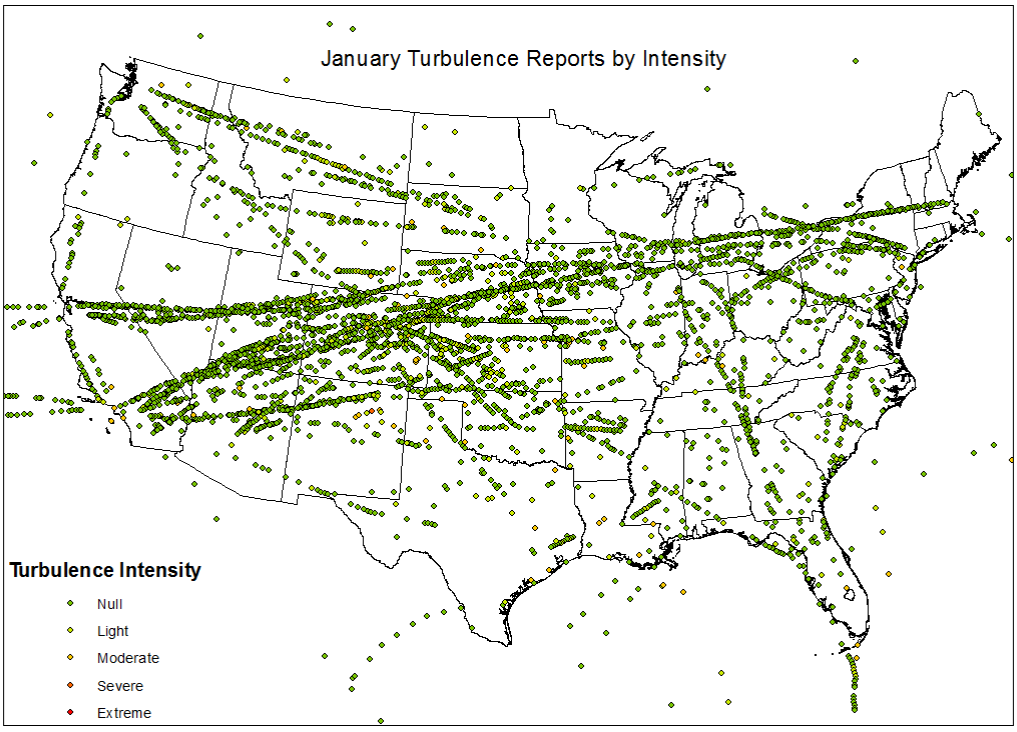
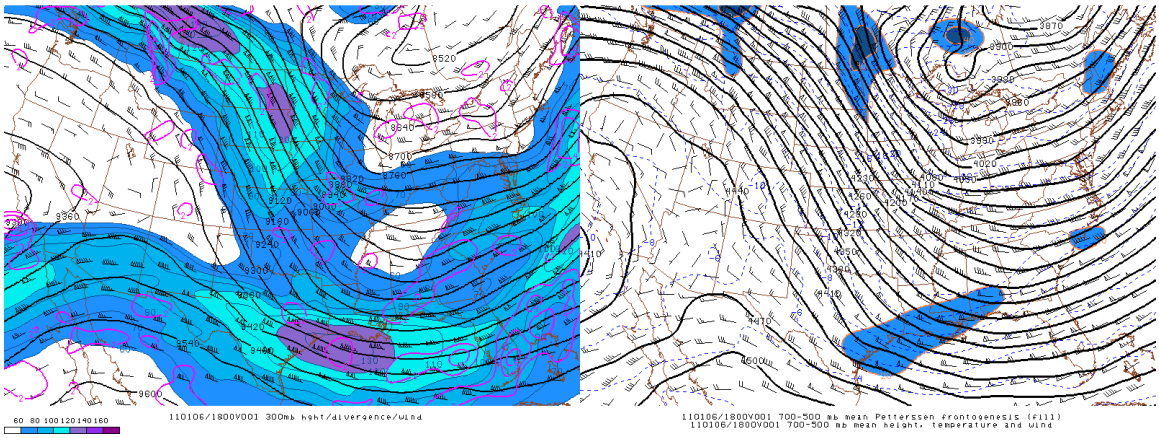


Figure 6.7: EDR data from the January 2011 turbulence outbreak, sorted by turbulence intensity.

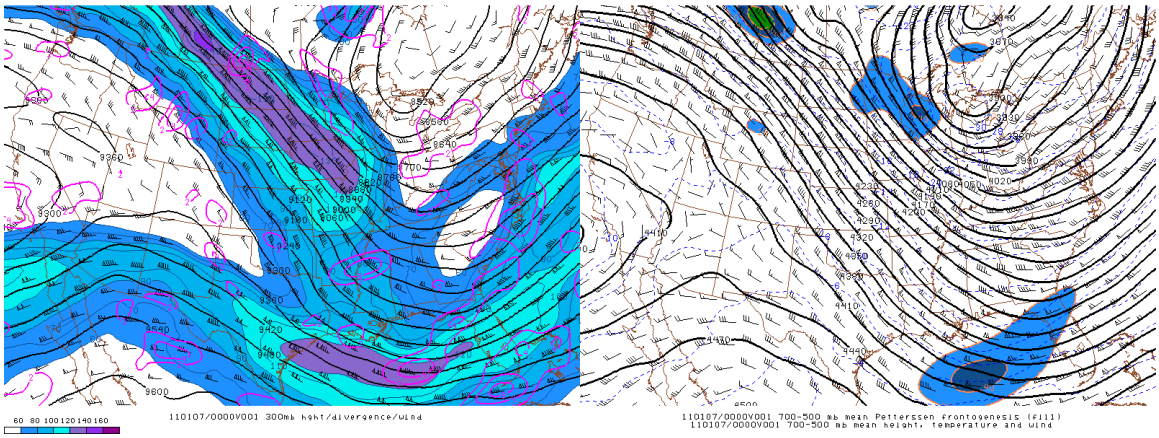
Table 6.8: Turbulence reports by intensity for the January 2011 outbreak.

	Null	Light	Moderate	Severe
All Turbulence Reports	2119	271	109	0

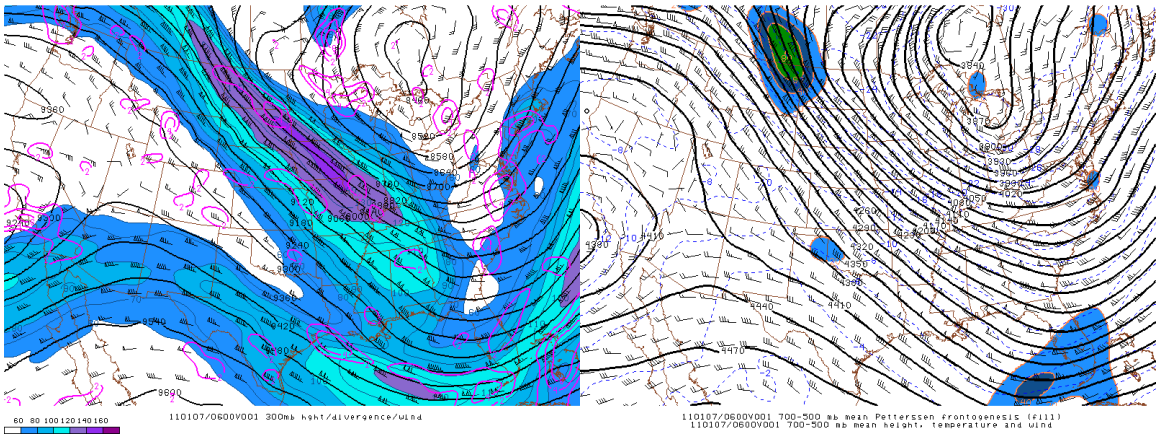
(a)


110106/1800V001 300mb hght/divergence/wind
60 80 100 120 140 160
110106/1800V001 700-500 mb mean Petterssen frontogenesis (fill)
110106/1800V001 700-500 mb mean height, temperature and wind

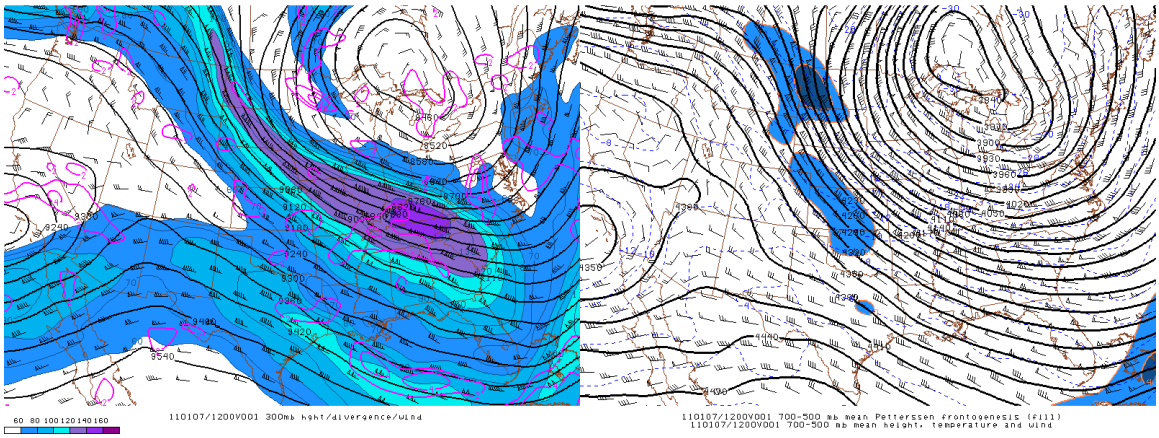
(b)


110107/0000V001 300mb hght/divergence/wind
60 80 100 120 140 160
110107/0000V001 700-500 mb mean Petterssen frontogenesis (fill)
110107/0000V001 700-500 mb mean height, temperature and wind

(c)


110107/0600V001 300mb hght/divergence/wind
60 80 100 120 140 160
110107/0600V001 700-500 mb mean Petterssen frontogenesis (fill)
110107/0600V001 700-500 mb mean height, temperature and wind

(d)



(e)

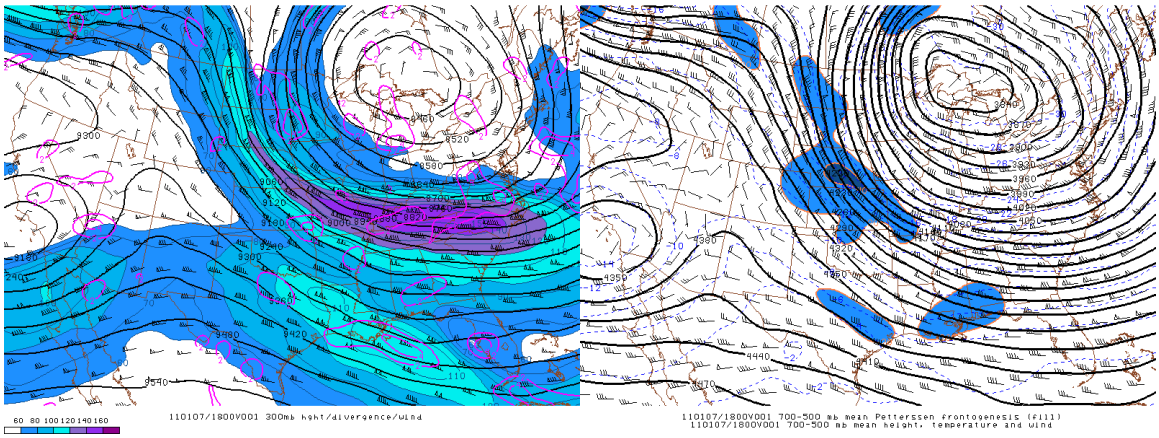
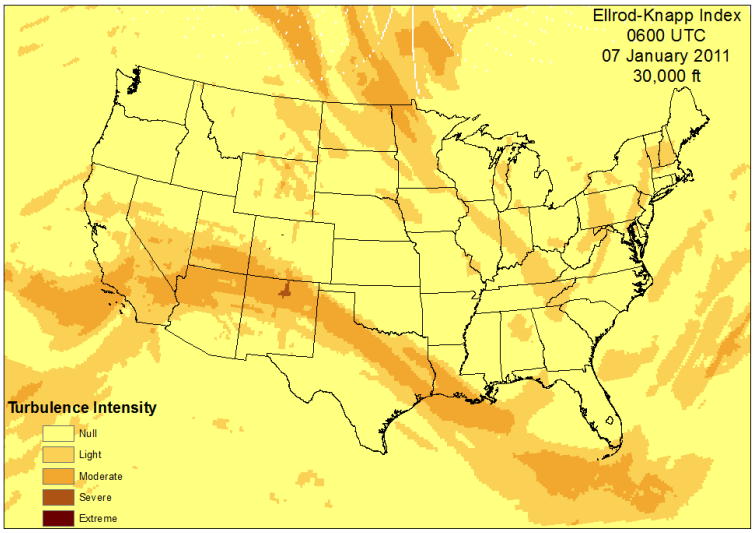
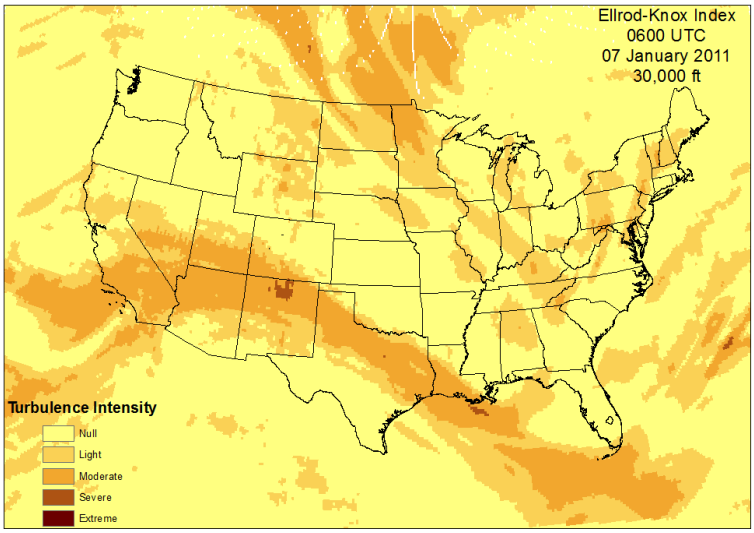


Figure 6.8: Synoptic maps for the January 2011 outbreak at (a) 1800 UTC on 06 January 2011, (b) 0000 UTC on 07 January 2011, (c) 0600 UTC on 07 January 2011, (d) 1200 UTC on 07 January 2011, and (e) 1800 UTC on 07 January 2011. In the left column, 300 hPa heights, windspeeds, and divergence are shown. Heights are shown by the solid black contours, divergence is shown by the purple contours, and wind speeds are illustrated by the shaded area. In the right column, 700 hPa to 500 hPa Petterssen frontogenesis, heights, and temperatures are illustrated. Heights are shown by the solid black contours, temperatures are illustrated using the dashed blue contours, and mean frontogenesis is shown in the shaded area. (Source: Storm Prediction Center)

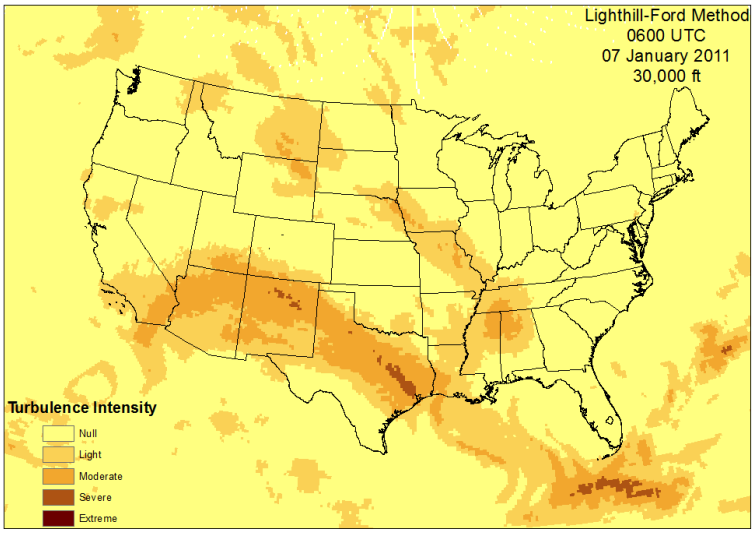
(a)


Ellrod-Knapp Index
0600 UTC
07 January 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

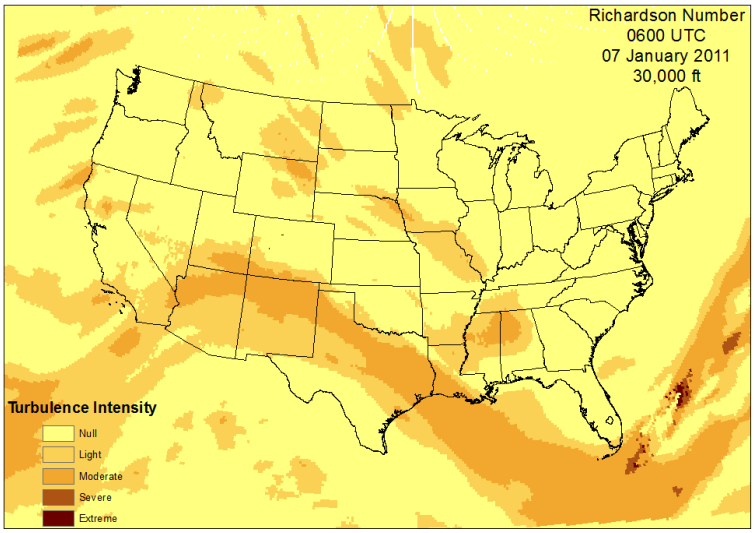
(b)


Ellrod-Knox Index
0600 UTC
07 January 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

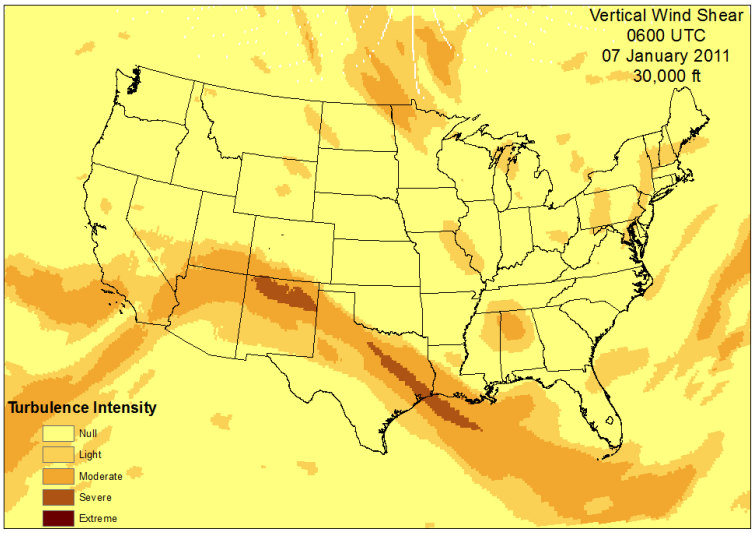
(c)


Lighthill-Ford Method
0600 UTC
07 January 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

(d)


Richardson Number
0600 UTC
07 January 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

(e)


Vertical Wind Shear
0600 UTC
07 January 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

(f)

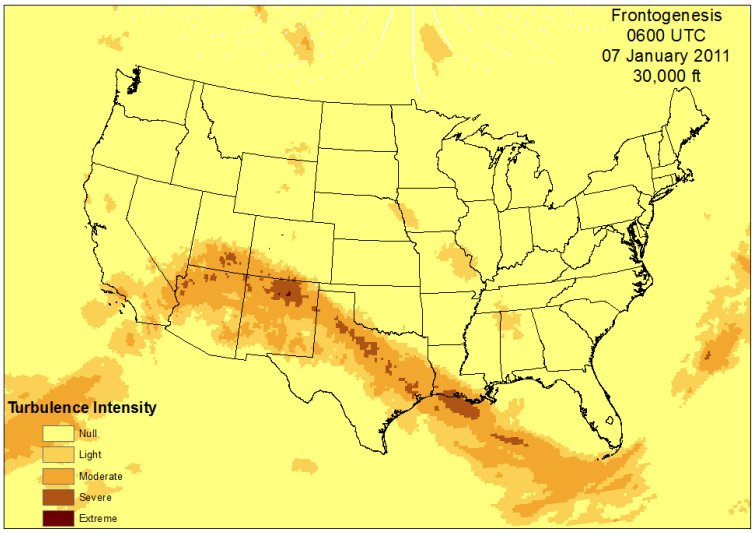
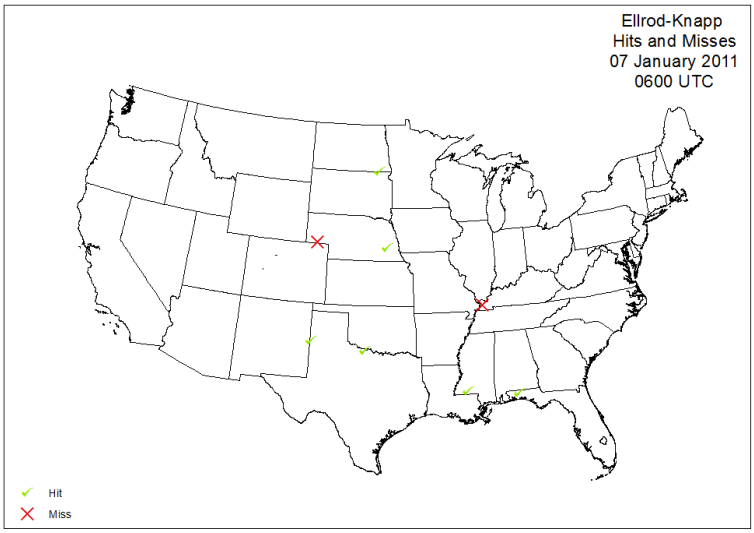
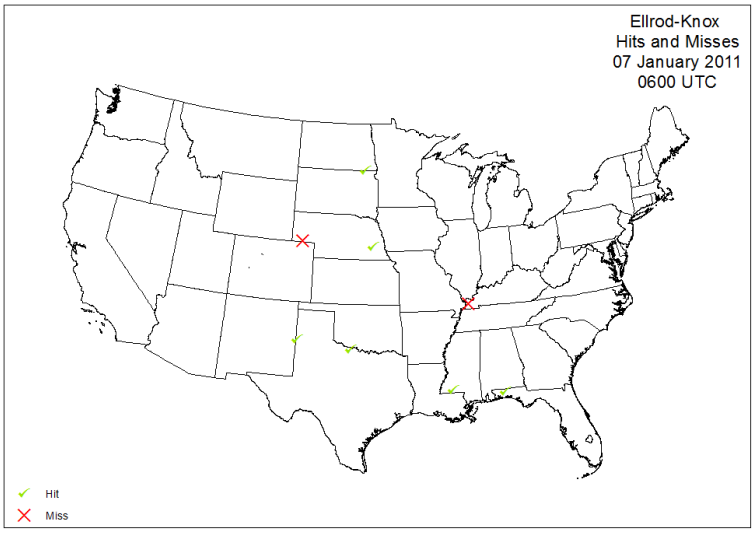

Frontogenesis
0600 UTC
07 January 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

Figure 6.9: Turbulence forecasts using (a) the Ellrod-Knapp index, (b) the Ellrod-Knox index, (c) the Lighthill-Ford method, (d) the Richardson number, (e) vertical wind shear, and (f) frontogenesis for 0600 UTC on 07 January 2011 for the altitude of 30,000 feet (9.1 km).

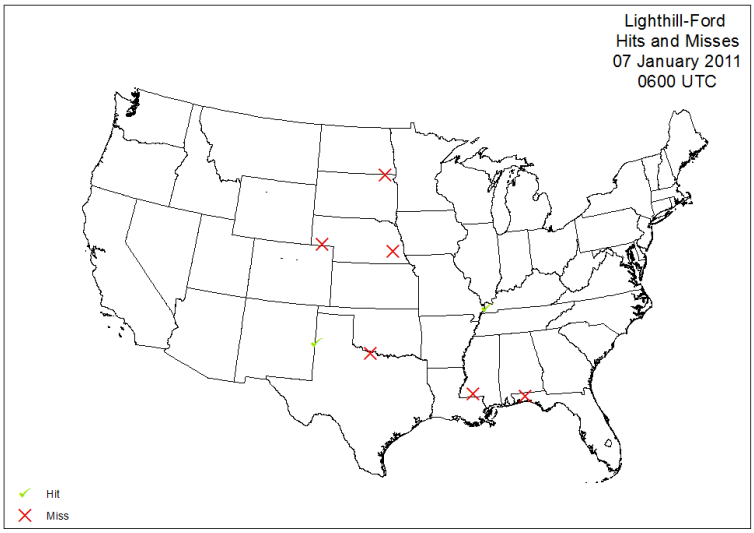
(a)


Ellrod-Knapp
Hits and Misses
07 January 2011
0600 UTC
Hit
Miss

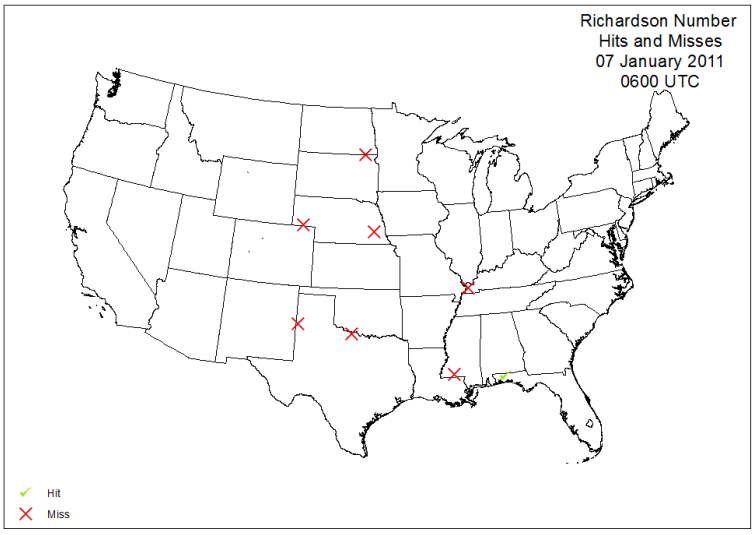
(b)


Ellrod-Knox
Hits and Misses
07 January 2011
0600 UTC
Hit
Miss

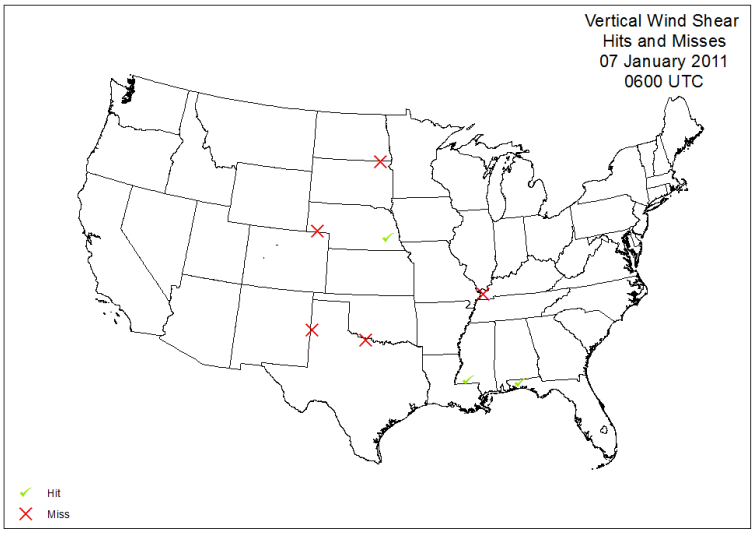
(c)


Lighthill-Ford
Hits and Misses
07 January 2011
0600 UTC
Hit
Miss

(d)


Richardson Number
Hits and Misses
07 January 2011
0600 UTC
Hit
Miss

(e)


Vertical Wind Shear
Hits and Misses
07 January 2011
0600 UTC
Hit
Miss

(f)

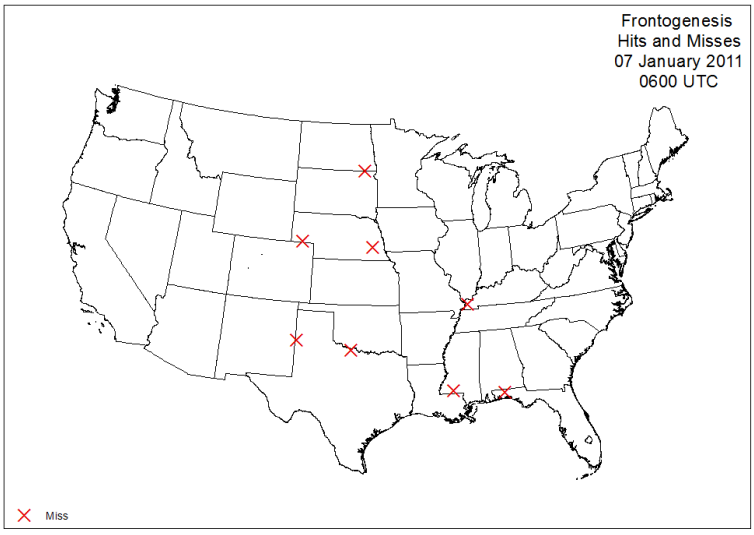

Frontogenesis
Hits and Misses
07 January 2011
0600 UTC
Miss

Figure 6.10: Turbulence forecast hits and misses using (a) the Ellrod-Knapp index, (b) the Ellrod-Knox index, (c) the Lighthill-Ford method, (d) the Richardson number, (e) vertical wind shear, and (f) frontogenesis for 0600 UTC on 07 January 2011.

Table 6.9: Forecast verifications from the January 2011 turbulence outbreak.

(a)

Ellrod-Knapp	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 256	*False Alarm* 597
Forecasted No Turbulence	*Misses* 124	*Correct Rejection* 1522

(b)

Ellrod-Knox	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 279	*False Alarms* 766
Forecasted No Turbulence	*Misses* 101	*Correct Rejections* 1353

(c)

Lighthill-Ford	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 181	*False Alarms* 460
Forecasted No Turbulence	*Misses* 199	*Correct Rejections* 1659

(d)

Richardson Number	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 189	*False Alarms* 467
Forecasted No Turbulence	*Misses* 191	*Correct Rejections* 1652

(e)

Vertical Wind Shear	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 250	*False Alarms* 739
Forecasted No Turbulence	*Misses* 130	*Correct Rejections* 1380

(f)

Frontogenesis	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 130	*False Alarms* 149
Forecasted No Turbulence	*Misses* 250	*Correct Rejections* 1970

Table 6.10: Forecast verification statistics for the January 2011 outbreak.

Turbulence Index	POD_y	POD_n	POFD	FAR	HSS	TSS	CSI
Ellrod-Knapp	0.674	0.718	0.282	0.700	0.259	0.392	0.262
Ellrod-Knox	0.734	0.639	0.361	0.733	0.217	0.373	0.243
Lighthill-Ford	0.476	0.783	0.217	0.718	0.202	0.259	0.215
Richardson Number	0.497	0.780	0.220	0.712	0.213	0.277	0.223
Vertical Wind Shear	0.658	0.651	0.349	0.747	0.186	0.309	0.223
Frontogenesis	0.342	0.930	0.070	0.534	0.305	0.272	0.246

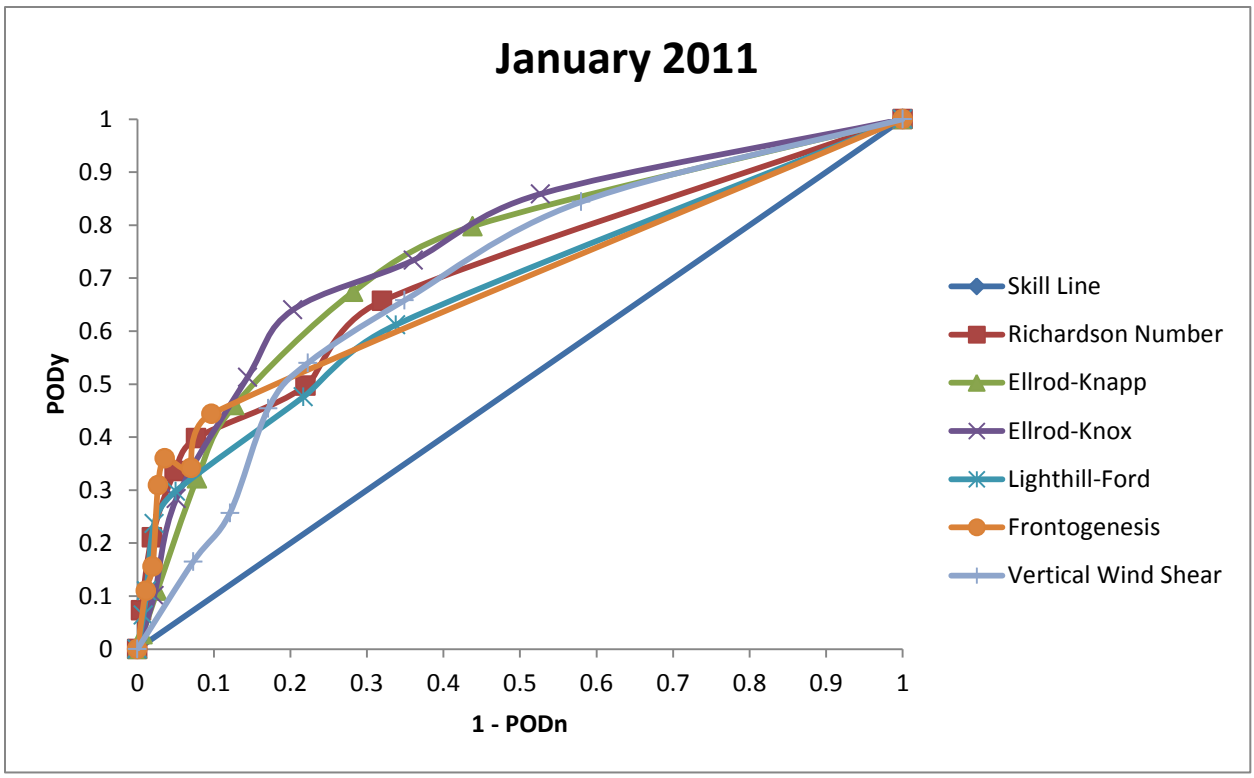


Figure 6.11: ROC curve for the turbulence indices during the January 2011 outbreak.

Table 6.11: AUC values for the six turbulence indices during the January 2011 outbreak.

Turbulence Index	AUC
Ellrod-Knox	0.758
Ellrod-Knapp	0.740
Richardson Number	0.706
Vertical Wind Shear	0.696
Frontogenesis	0.681
Lighthill-Ford	0.674

Table 6.12: Forecast verifications for moderate-or-greater (MOG) turbulence from the January 2011 turbulence outbreak.

(a)

Ellrod-Knapp	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 39	*False Alarms* 171
Forecasted No MOG Turbulence	*Misses* 70	*Correct Rejections* 2219

(b)

Ellrod-Knox	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 53	*False Alarms* 342
Forecasted No MOG Turbulence	*Misses* 56	*Correct Rejections* 2048

(c)

Lighthill-Ford	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 17	*False Alarms* 58
Forecasted No MOG Turbulence	*Misses* 92	*Correct Rejections* 2332

(d)

Richardson Number	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 33	*False Alarms* 118
Forecasted No MOG Turbulence	*Misses* 76	*Correct Rejections* 2272

(e)

Vertical Wind Shear	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 45	*False Alarms* 399
Forecasted No MOG Turbulence	*Misses* 64	*Correct Rejections* 1991

(f)

Frontogenesis	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 32	*False Alarms* 72
Forecasted No MOG Turbulence	*Misses* 77	*Correct Rejections* 2318

Table 6.13: Moderate-or-greater turbulence forecast verification statistics for the January 2011 outbreak.

Turbulence Index	POD_y	POD_n	POFD	FAR	HSS	TSS	CSI
Ellrod-Knapp	0.358	0.928	0.072	0.814	0.198	0.286	0.139
Ellrod-Knox	0.486	0.857	0.143	0.866	0.152	0.343	0.118
Lighthill-Ford	0.156	0.976	0.024	0.773	0.155	0.132	0.102
Richardson Number	0.303	0.951	0.049	0.781	0.214	0.253	0.145
Vertical Wind Shear	0.413	0.833	0.167	0.899	0.100	0.246	0.089
Frontogenesis	0.294	0.970	0.030	0.692	0.269	0.263	0.177

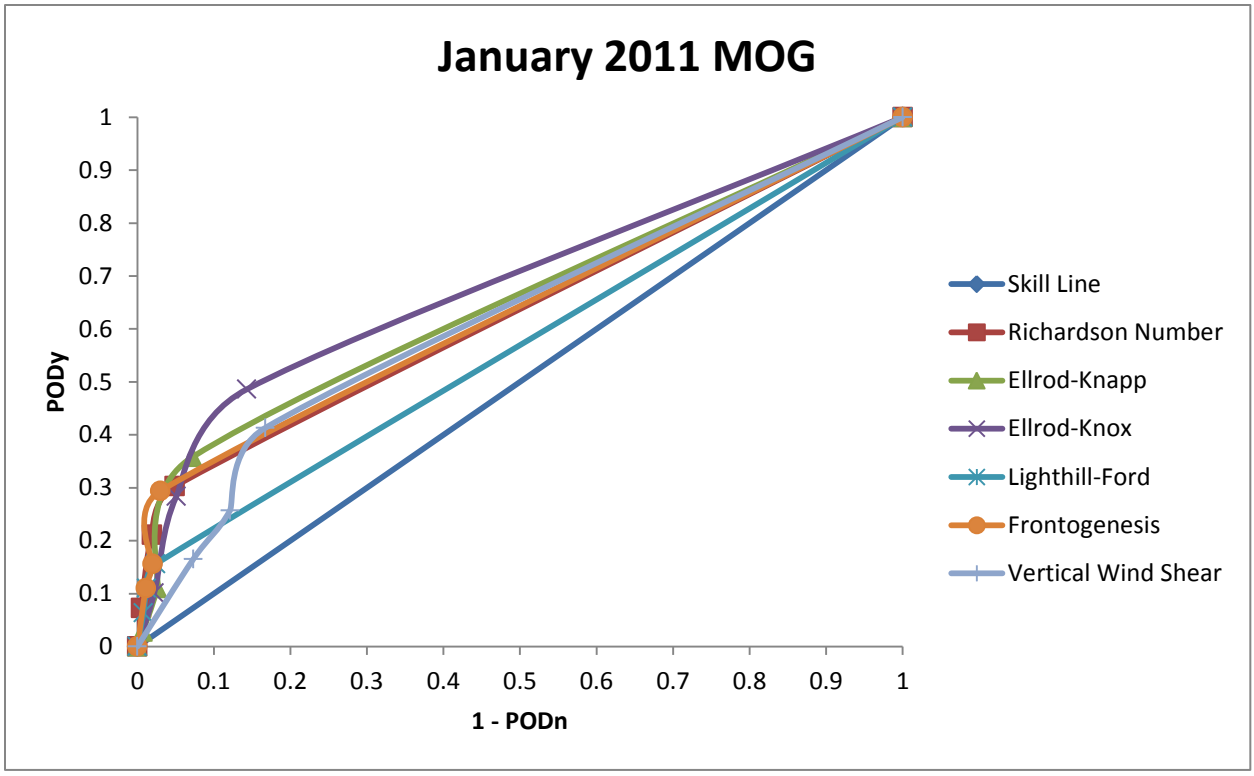


Figure 6.12: ROC curve for the turbulence indices at the moderate-or-greater scale during the January 2011 outbreak.

Table 6.14: AUC values for the six turbulence indices at the moderate-or-greater scale during the January 2011 outbreak.

Turbulence Index	AUC
Ellrod-Knox	0.679
Ellrod-Knapp	0.642
Frontogenesis	0.632
Richardson Number	0.629
Vertical Wind Shear	0.620
Lighthill-Ford	0.566

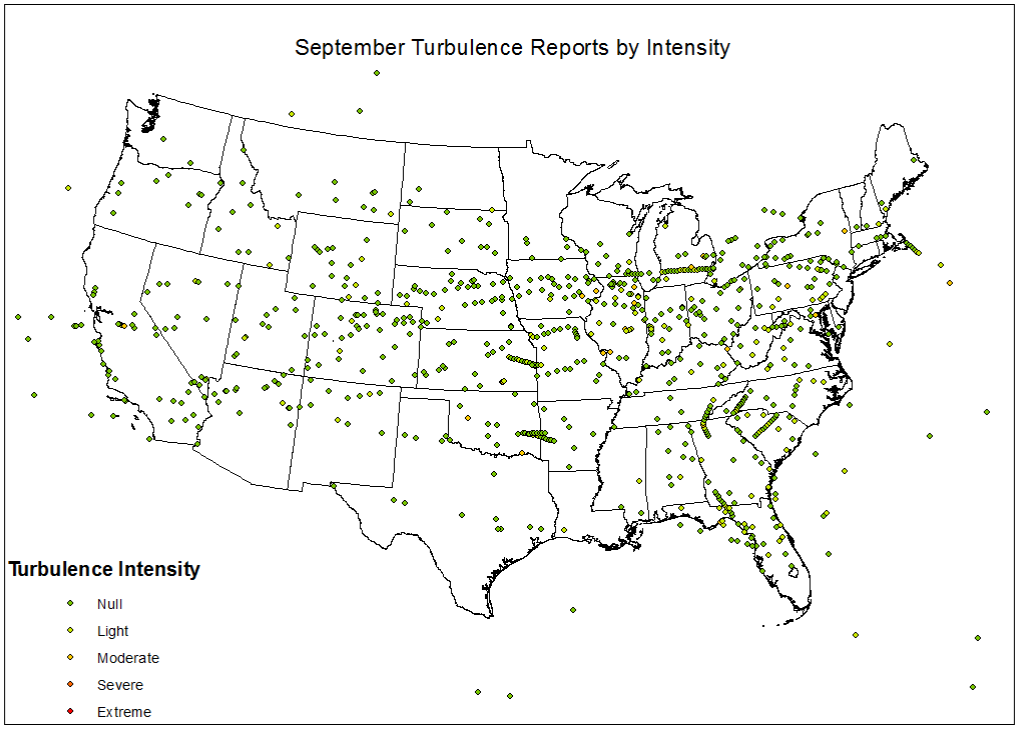
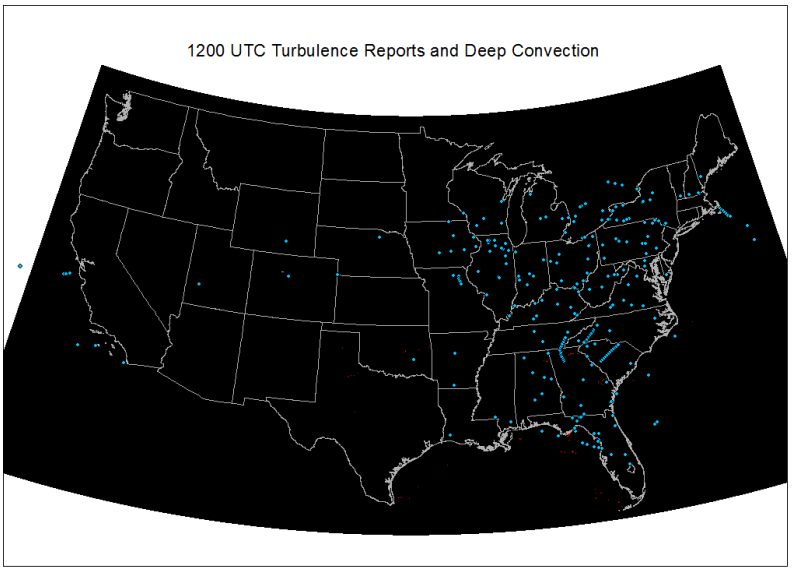


Figure 6.13: EDR data from the September 2011 turbulence outbreak, sorted by turbulence intensity.

Table 6.15: Turbulence reports by intensity for the September 2011 outbreak.

	Null	Light	Moderate	Severe
All Turbulence Reports	442	134	49	0

(a)


1200 UTC Turbulence Reports and Deep Convection

(b)

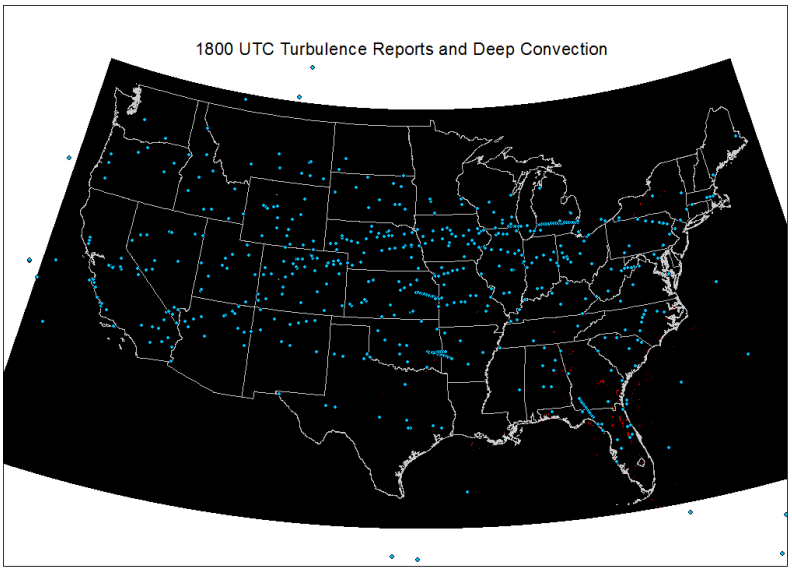
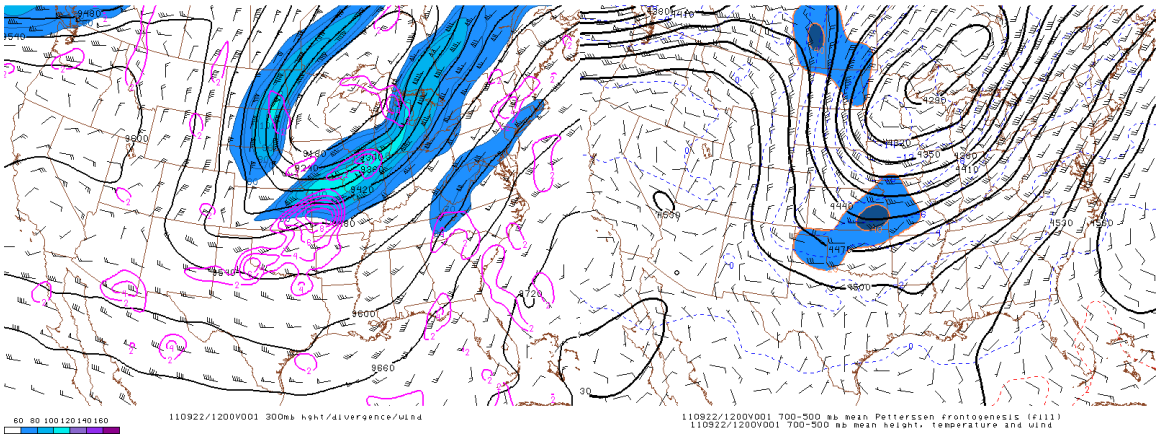

1800 UTC Turbulence Reports and Deep Convection

Figure 6.14: Turbulence reports from September 22 overlaid with convection greater than 50 dBZ. In this figure, the blue dots are EDR observations, and the red dots are areas with convection greater than or equal to 50 dBZ.

(a)



(b)

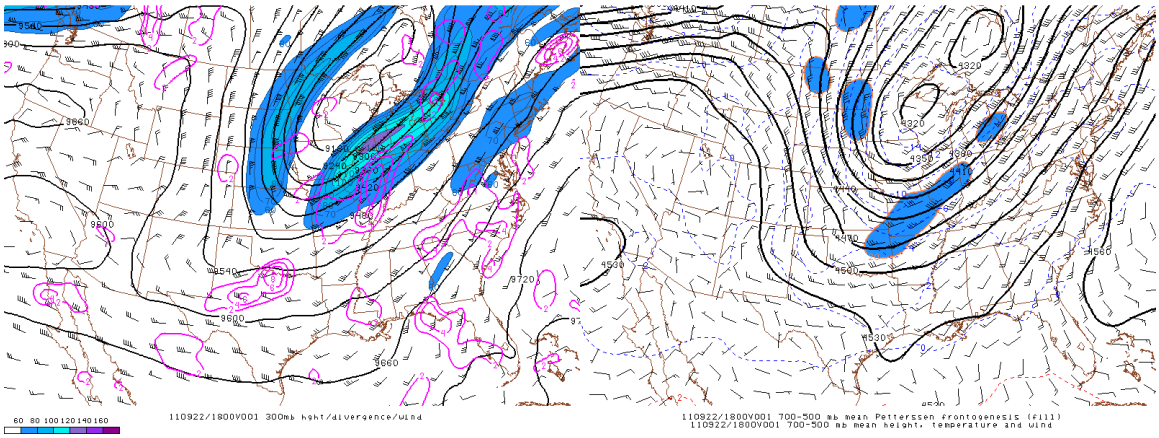
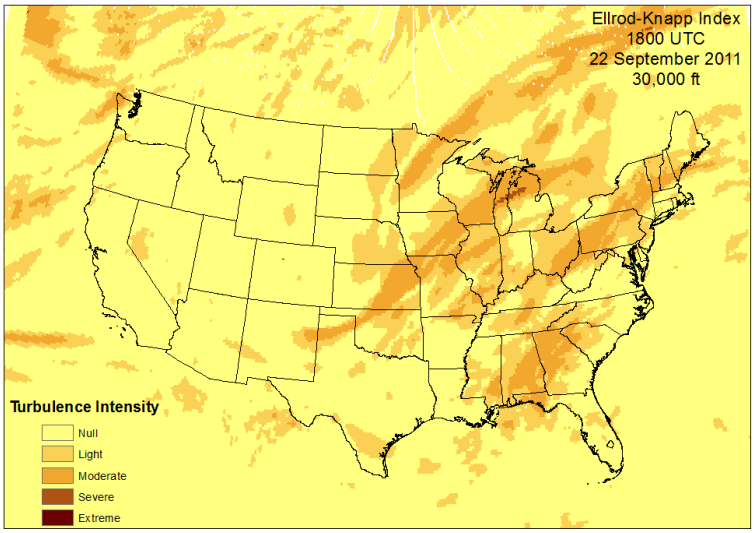
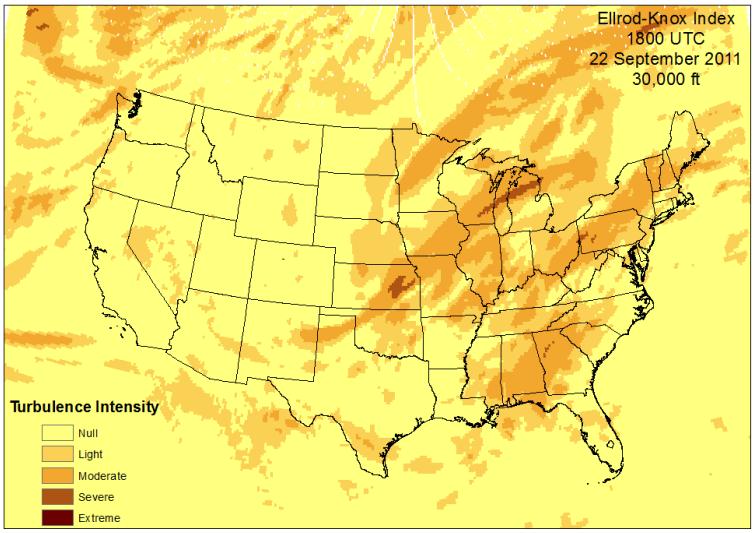


Figure 6.15: Synoptic maps for the September 2011 outbreak at (a) 1200 UTC on 22 September 2011 and (b) 1800 UTC on 22 September 2011. In the left column, 300 hPa heights, windspeeds, and divergence are shown. Heights are shown by the solid black contours, divergence is shown by the purple contours, and wind speeds are illustrated by the shaded area. In the right column, 700 hPa to 500 hPa Petterssen frontogenesis, heights, and temperatures are illustrated. Heights are shown by the solid black contours, temperatures are illustrated using the dashed blue contours, and mean frontogenesis is shown in the shaded area. (Source: Storm Prediction Center)

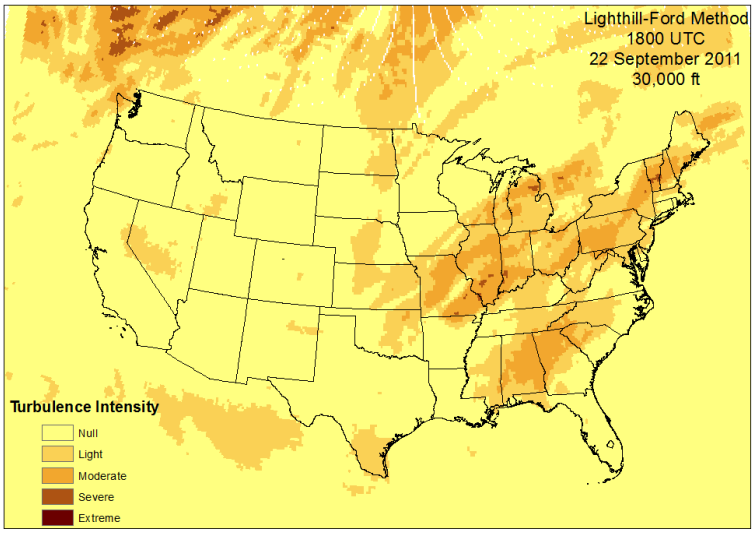
(a)


Ellrod-Knapp Index
1800 UTC
22 September 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

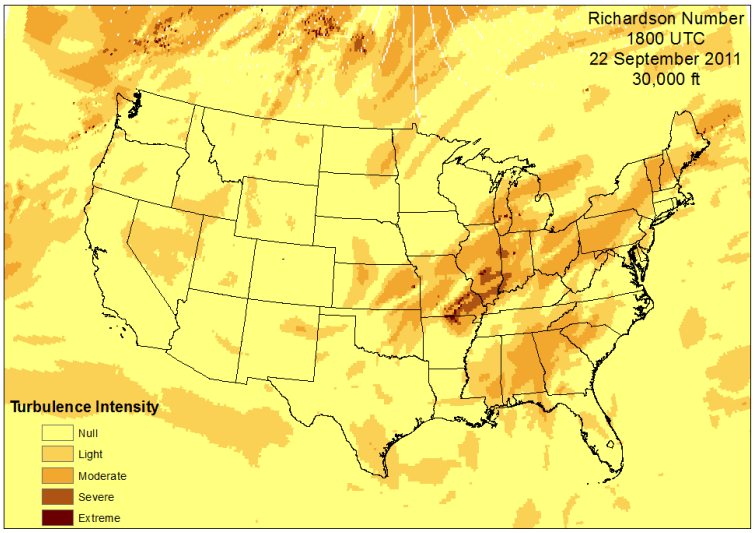
(b)


Ellrod-Knox Index
1800 UTC
22 September 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

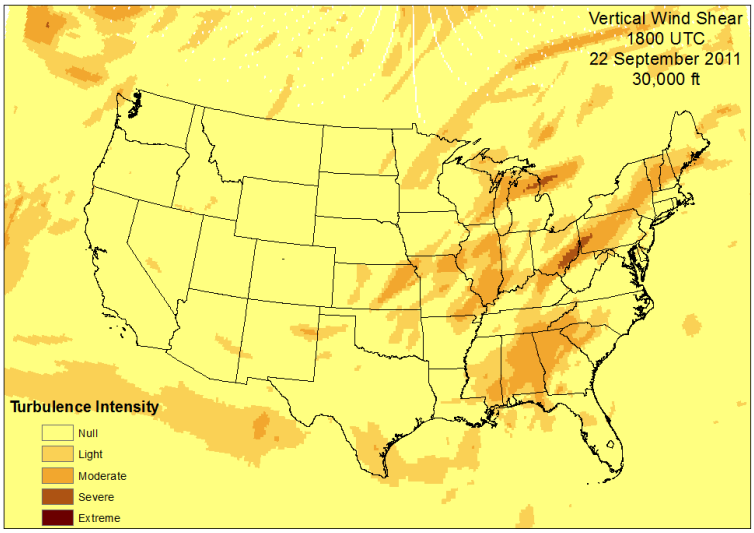
(c)


Lighthill-Ford Method
1800 UTC
22 September 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

(d)


Richardson Number
1800 UTC
22 September 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

(e)


Vertical Wind Shear
1800 UTC
22 September 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

(f)

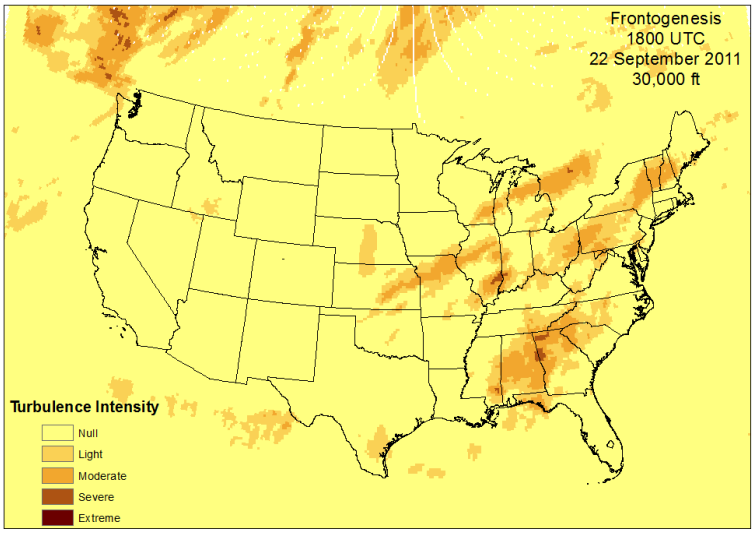
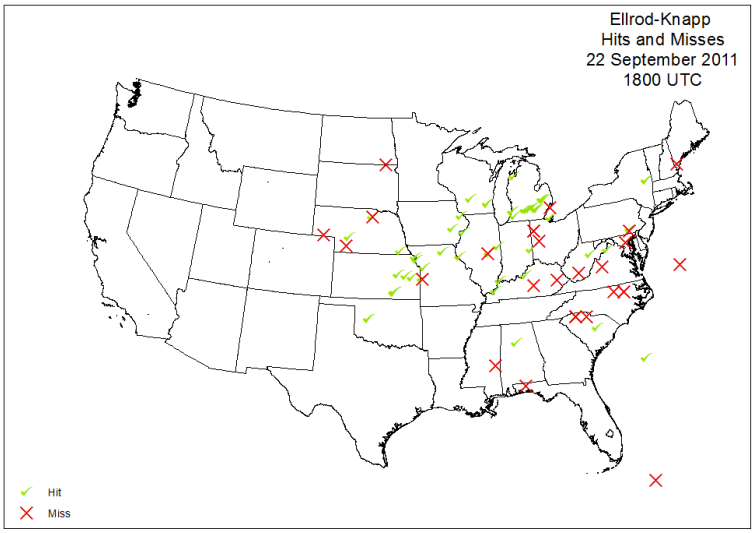
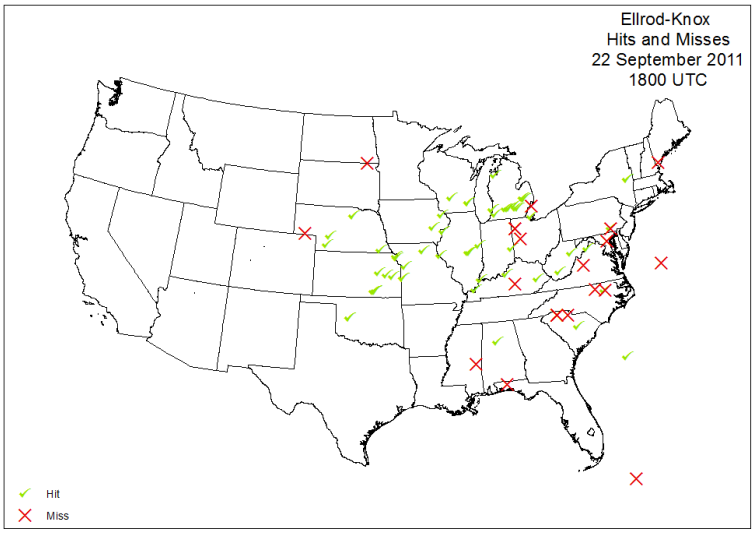

Frontogenesis
1800 UTC
22 September 2011
30,000 ft
Turbulence Intensity
Null
Light
Moderate
Severe
Extreme

Figure 6.16: Turbulence forecasts using (a) the Ellrod-Knapp index, (b) the Ellrod-Knox index, (c) the Lighthill-Ford method, (d) the Richardson number, (e) vertical wind shear, and (f) frontogenesis for 1800 UTC on 22 September 2011 for the altitude of 30,000 feet (9.1 km).

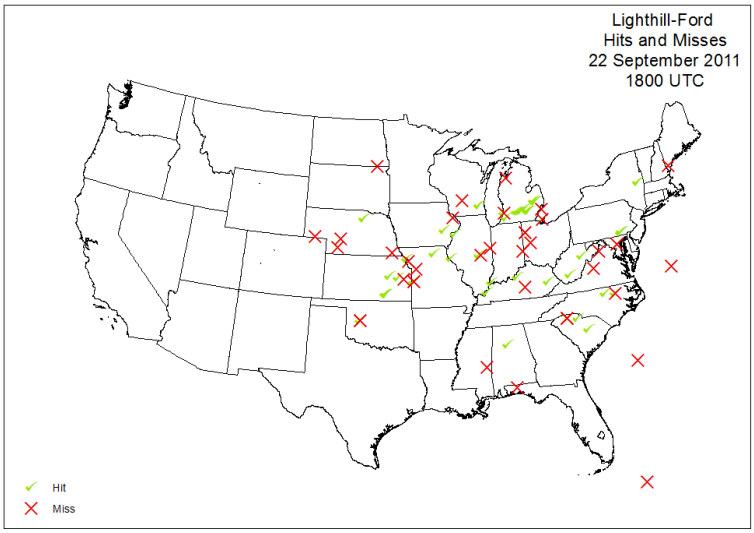
(a)


Ellrod-Knapp
Hits and Misses
22 September 2011
1800 UTC
Hit
Miss

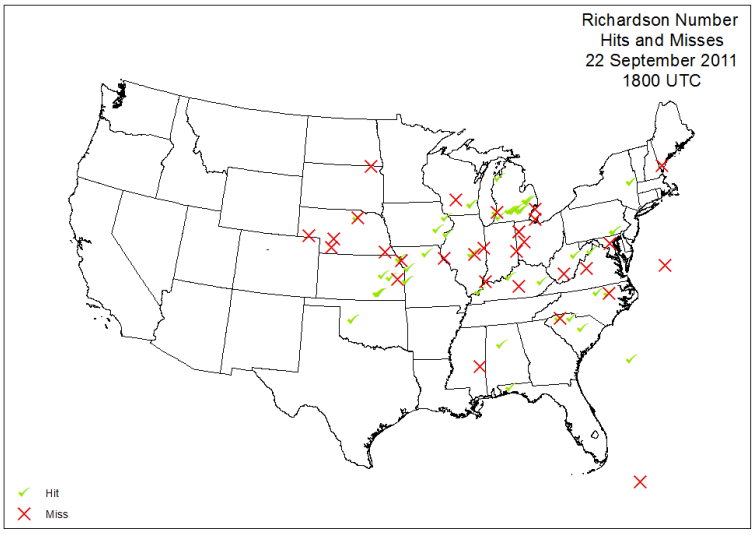
(b)


Ellrod-Knox
Hits and Misses
22 September 2011
1800 UTC
Hit
Miss

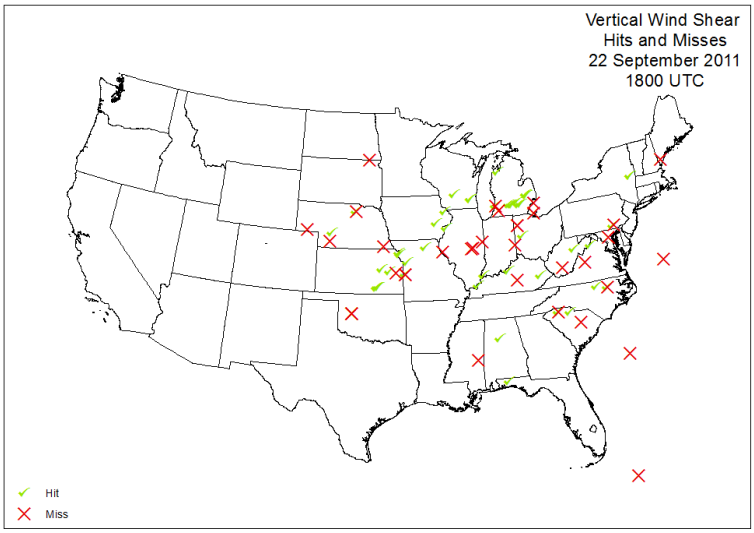
(c)


Lighthill-Ford
Hits and Misses
22 September 2011
1800 UTC
Hit
Miss

(d)


Richardson Number
Hits and Misses
22 September 2011
1800 UTC
Hit
Miss

(e)


Vertical Wind Shear
Hits and Misses
22 September 2011
1800 UTC
Hit
Miss

(f)

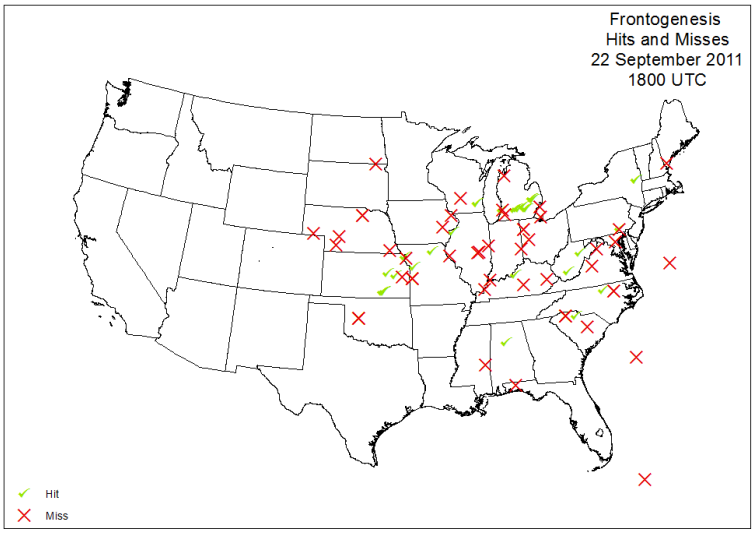

Frontogenesis
Hits and Misses
22 September 2011
1800 UTC
Hit
Miss

Figure 6.17: Turbulence forecast hits and misses using (a) the Ellrod-Knapp index, (b) the Ellrod-Knox index, (c) the Lighthill-Ford method, (d) the Richardson number, (e) vertical wind shear, and (f) frontogenesis for 1800 UTC on 22 September 2011.

Table 6.16: Forecast verifications from the September 2011 turbulence outbreak.

(a)

Ellrod-Knapp	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 123	*False Alarms* 221
Forecasted No Turbulence	*Misses* 60	*Correct Rejections* 221

(b)

Ellrod-Knox	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 135	*False Alarms* 252
Forecasted No Turbulence	*Misses* 48	*Correct Rejections* 190

(c)

Lighthill-Ford	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 107	*False Alarms* 128
Forecasted No Turbulence	*Misses* 76	*Correct Rejections* 314

(d)

Richardson Number	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 115	*False Alarms* 164
Forecasted No Turbulence	*Misses* 68	*Correct Rejections* 278

(e)

Vertical Wind Shear	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 104	*False Alarms* 210
Forecasted No Turbulence	*Misses* 79	*Correct Rejections* 232

(f)

Frontogenesis	Observed Turbulence	Observed No Turbulence
Forecasted Turbulence	*Hits* 60	*False Alarms* 75
Forecasted No Turbulence	*Misses* 123	*Correct Rejections* 367

Table 6.17: Forecast verification statistics for the September 2011 outbreak.

Turbulence Index	POD_y	POD_n	POFD	FAR	HSS	TSS	CSI
Ellrod-Knapp	0.672	0.500	0.500	0.642	0.137	0.172	0.304
Ellrod-Knox	0.738	0.430	0.570	0.651	0.126	0.168	0.310
Lighthill-Ford	0.585	0.710	0.290	0.545	0.272	0.295	0.344
Richardson Number	0.628	0.629	0.371	0.588	0.223	0.257	0.331
Vertical Wind Shear	0.568	0.525	0.475	0.669	0.077	0.093	0.265
Frontogenesis	0.328	0.830	0.170	0.556	0.171	0.158	0.233

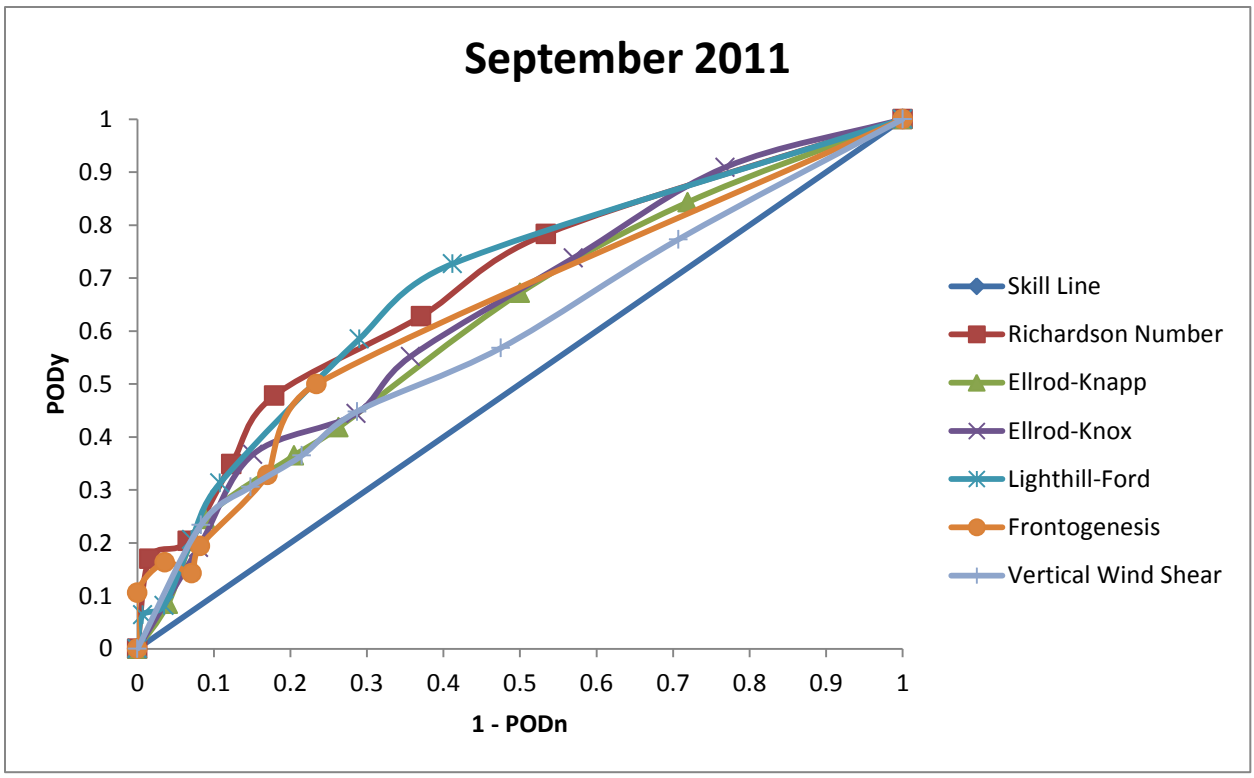


Figure 6.18: ROC curve for the turbulence indices during the September 2011 outbreak.

Table 6.18: AUC values for the six turbulence indices during the September 2011 outbreak.

Turbulence Index	AUC
Lighthill-Ford	0.687
Richardson Number	0.686
Ellrod-Knox	0.639
Frontogenesis	0.636
Ellrod-Knapp	0.622
Vertical Wind Shear	0.590

Table 6.19: Forecast verifications for moderate-or-greater (MOG) turbulence from the September 2011 turbulence outbreak.

(a)

Ellrod-Knapp	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 19	*False Alarms* 111
Forecasted No MOG Turbulence	*Misses* 30	*Correct Rejections* 465

(b)

Ellrod-Knox	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 22	*False Alarms* 157
Forecasted No MOG Turbulence	*Misses* 27	*Correct Rejections* 419

(c)

Lighthill-Ford	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 9	*False Alarms* 40
Forecasted No MOG Turbulence	*Misses* 40	*Correct Rejections* 536

(d)

Richardson Number	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 16	*False Alarms* 70
Forecasted No MOG Turbulence	*Misses* 33	*Correct Rejections* 506

(e)

Vertical Wind Shear	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 16	*False Alarms* 120
Forecasted No MOG Turbulence	*Misses* 33	*Correct Rejections* 456

(f)

Frontogenesis	Observed MOG Turbulence	Observed No MOG Turbulence
Forecasted MOG Turbulence	*Hits* 9	*False Alarms* 36
Forecasted No MOG Turbulence	*Misses* 40	*Correct Rejections* 540

Table 6.20: Moderate-or-greater turbulence forecast verification statistics for the September 2011 outbreak.

Turbulence Index	POD_y	POD_n	POFD	FAR	HSS	TSS	CSI
Ellrod-Knapp	0.388	0.807	0.193	0.854	0.111	0.195	0.119
Ellrod-Knox	0.449	0.727	0.273	0.877	0.080	0.176	0.107
Lighthill-Ford	0.184	0.931	0.069	0.816	0.114	0.114	0.101
Richardson Number	0.327	0.878	0.122	0.814	0.152	0.205	0.134
Vertical Wind Shear	0.327	0.792	0.208	0.882	0.065	0.118	0.095
Frontogenesis	0.184	0.938	0.063	0.800	0.126	0.121	0.106

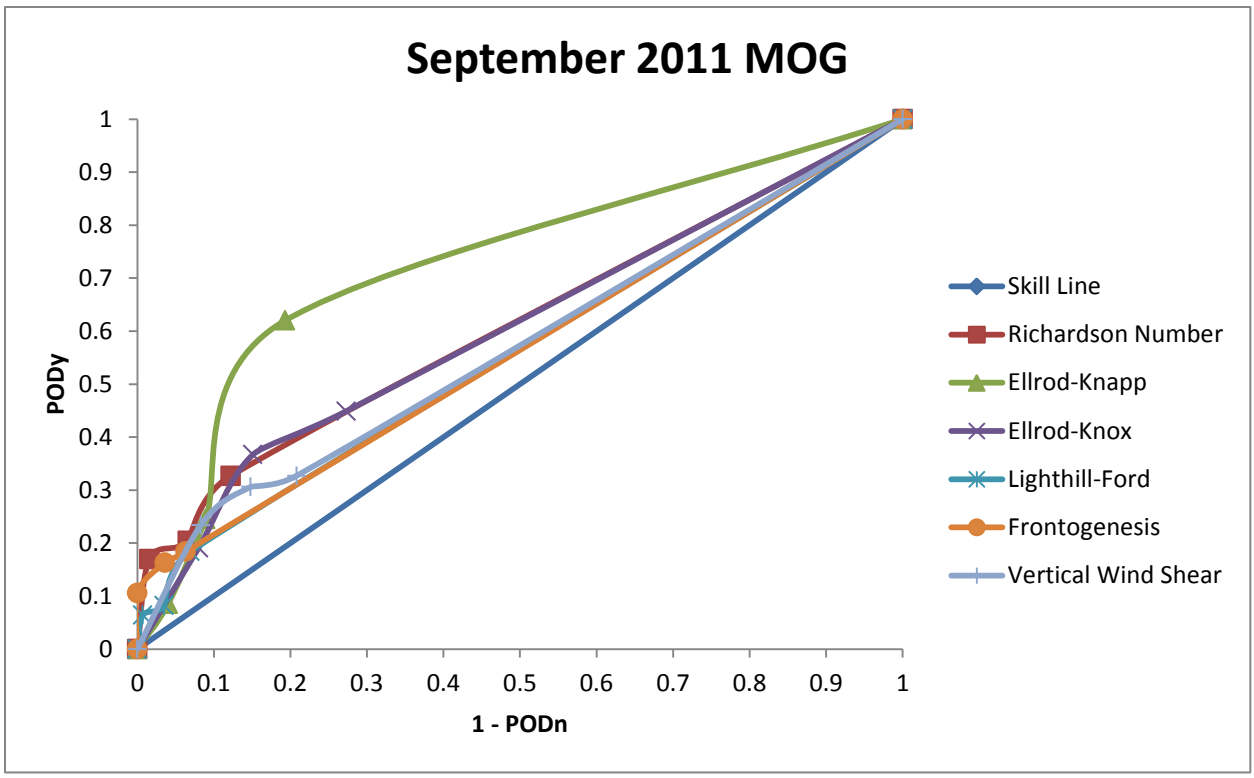


Figure 6.19: ROC curve for the turbulence indices at the moderate-or-greater scale during the September 2011 outbreak.

Table 6.21: AUC values for the six turbulence indices at the moderate-or-greater scale during the September 2011 outbreak.

Turbulence Index	AUC
Ellrod-Knapp	0.708
Richardson Number	0.608
Ellrod-Knox	0.604
Vertical Wind Shear	0.572
Frontogenesis	0.564
Lighthill-Ford	0.558

CHAPTER 7

CONCLUSION

Despite significant advances in technology over the decades, clear-air turbulence remains a serious threat to airlines, as well as an unsolved problem in the atmospheric sciences. Throughout the years, many indices have been developed to try and predict where regions of CAT may develop. These indices range from simple meteorological parameters such as frontogenesis (Baughman 1946) to complex diagnostics such as the GTG index (Sharman 2006) or the Lighthill-Ford method (Knox 2008). Even with the development of several CAT prediction indices and diagnostics, there is no single method that is best for predicting episodes of turbulence.

This thesis was designed to explore three main questions regarding CAT. First, was the Lighthill-Ford method effective at predicting episodes of CAT in several case studies? Next, was the Lighthill-Ford method complementary to other turbulence indices? Last, what were the synoptic conditions like during these outbreaks of CAT?

In this thesis, these questions were answered using three case studies of CAT outbreaks, which were examined using six different turbulence forecast indices: the Ellrod-Knox index, the Ellrod-Knapp index, the Lighthill-Ford method, the Richardson number, vertical wind shear, and frontogenesis. These forecasts were compared to actual EDR observations of turbulence, and were then verified by calculating skill statistics and producing ROC curves.

For all turbulence events, the Ellrod-Knox index had the highest POD_y values in all three cases, while the Ellrod-Knapp index was never far behind. While the Ellrod-Knox index always had the highest POD_y, it often also had the highest values for both the POFD and the FAR, due to the large regions of forecasted turbulence that this index created. In terms of predictions of null turbulence events, frontogenesis and the Lighthill-Ford method had the highest POD_n values in all three cases. This was due to the fact that, out of the six turbulence indices, both of these diagnostics typically had the smallest regions of forecasted turbulence, meaning that most of the null reports of turbulence were in regions where turbulence was not expected.

In terms of the more complex skill statistics for all turbulence events, there were not any clear patterns between the cases. For the December case, frontogenesis had the highest HSS value, vertical wind shear had the highest TSS value, and the Richardson number had the highest CSI. In the January case, once again, frontogenesis had the highest HSS, but the Ellrod-Knapp index had the largest TSS and CSI values. The Lighthill-Ford method had the best scores for all three skill statistics in the September case study. The scores for the HSS and the TSS may be somewhat less reliable than other skill statistics, because they can be biased by a large number of correct rejections.

In terms of the area under the curve, which is perhaps the best metric to illustrate an index's forecast skill because it is a combination of the hit rate and the false alarm rate, the Ellrod-Knox index and the Ellrod-Knapp index proved to have the most forecast skill during the first two CAT outbreaks. The Lighthill-Ford method and frontogenesis had the least skill, in terms of AUC, during these two cases. During the September case, the Lighthill-Ford method had the most forecast skill, while the Ellrod-Knox index and the Ellrod-Knapp index fell into the bottom half of the forecast skill rankings. Perhaps the success of the Lighthill-Ford method

during this last case is due to the convection that was occurring on this day. The Lighthill-Ford method may outperform other turbulence indices during convective events (Trier et al. 2012).

For the moderate-or-greater turbulence events, once again, the Ellrod-Knox index had the highest POD_y values in all three cases. Similar to the previous case, this index also had some of the highest values for both the POFD and the FAR, due to the large regions of forecasted turbulence that this index predicted. In terms of predictions of null turbulence events, yet again, frontogenesis and the Lighthill-Ford method had the highest POD_n values in all three cases. This is due to the fact that, out of the six turbulence indices, both of these diagnostics typically had smaller regions of forecasted turbulence.

In terms of the more complex skill statistics for moderate-or-greater turbulence events, the Richardson number appeared to do well at forecasting turbulence. For the December case, the Richardson number had the highest values for the HSS and the CSI, while the Ellrod-Knox index had the largest TSS. In the January case, the Richardson number had the highest CSI, as well as the second highest HSS value. Frontogenesis had the largest HSS, and the Ellrod-Knox index had the largest TSS for this case. The Richardson number had the largest scores for all three skill statistics in the September case study.

Based on the area under the curve for the moderate-or-greater turbulence events, the Lighthill-Ford method had the least skill at forecasting turbulence. This turbulence index had the lowest AUC value during all three case studies. In most of the cases, the Lighthill-Ford method's AUC was barely above the "no forecast skill" value of 0.500. The Ellrod-Knox index performed the best during the December and January cases, while the Ellrod-Knapp index had the highest AUC during the September case.

These results differ from McCann et al. (2012), who found that calculating the Lighthill-Ford method using ULTURB outperformed all other turbulence indices at all intensity levels. Both McCann et al. (2012) and this thesis used 13-kilometer model data to produce turbulence forecasts. McCann et al. (2012) used RUC data, and this thesis used WRF-RR data, which is the replacement for the RUC. To verify the ULTURB forecasts, McCann et al. (2012) used PIREPs, which are known to be more error-prone, instead of the EDR data that was used to verify the forecasts in this thesis. This difference in the observational data that was used to verify turbulence forecasts may account for some of the difference in the results of these two studies. The success of ULTURB may be due to its computation of the Lighthill-Ford method. Perhaps including the gravity wave physics (that are contained in the ULTURB program but not in Sharman's formulation used in this thesis) when computing the Lighthill-Ford method plays a major role in its success as an effective turbulence index.

In summary, based on overall forecast skill, Sharman's formulation of the Lighthill-Ford method did not perform well during these cases. In all but one case, the September 2011 outbreak for all turbulence events, the Lighthill-Ford method had either the worst or the second worst AUC value out of all six forecast indices. In the September all-turbulence case, the Lighthill-Ford method had the highest AUC value out of all six indices. This method often had one of the worst hit rates; however, this method was effective at predicting regions of null turbulence, due to the fact that this index often had one of the highest POD_n values. The Lighthill-Ford method did not perform well overall in terms of the advanced skill statistics like the HSS, the TSS, or the CSI. The only time that the Lighthill-Ford method had the highest values for these statistics was during the September 2011 case for all turbulence intensities. Perhaps calculating the Lighthill-Ford method using ULTURB, rather than Sharman's

formulation, would have produced better overall results for this index during the three case studies.

All forecast indices, including the Lighthill-Ford method, predicted turbulence in similar regions of the United States, and along similar large-scale atmospheric features. The Ellrod-Knox method, the Ellrod-Knapp method, and the Richardson number typically predicted larger regions of turbulence than did the Lighthill-Ford method. The Lighthill-Ford method typically forecasted larger regions of turbulence than did frontogenesis and vertical wind shear. There also appears to be a strong correlation between the spatial extent of a turbulence forecast and the number of hits, correct rejections, and the false alarm rate. Indices with a large area of forecasted turbulence, such as the Ellrod-Knox index, typically had many forecast hits, as well as a high false alarm rate. Forecast indices that predicted smaller regions of turbulence, such as the Lighthill-Ford method or frontogenesis, typically had more correct rejections and low false alarm rates.

To examine the synoptic setup during each turbulence outbreak, the locations of forecasted turbulence and EDR turbulence reports were compared to archived weather maps. The synoptic setups for all three case studies contained troughs, strong jet streams, and regions of frontogenesis. In all three cases, turbulence was forecasted to occur in troughs, in areas of frontogenesis, and in the vicinity of the jet stream. The plots of observed turbulence events showed that turbulence was mostly occurring in troughs, with many of these events occurring in or near the jet stream. Some of the observed turbulence events occurred in regions of diagnosed frontogenesis, but the relationship between turbulence, troughs, and jet streams was stronger. The Ellrod-Knox index typically had the most forecast hits within the troughs, as well as being one of the only indices to correctly forecast turbulence in the divergent flow associated with the

beginning of a ridge during the December case. During these cases, the Lighthill-Ford method performed well in the middle of troughs. In the January case, it was the only turbulence index to correctly forecast a turbulence event over Kentucky, which happened to be located in the middle of a deep trough at the time. These three cases support the long-standing relationship between atmospheric features like troughs and jet streams to aircraft turbulence.

In the future, this thesis could be expanded upon in several ways. First, several more case studies of CAT outbreaks could be examined using these six forecast indices, to see if similar results are obtained. It would also be interesting to look at CAT outbreaks under different synoptic conditions. The three CAT cases that were analyzed in this thesis had troughs, so it would be interesting to look at CAT outbreaks that occurred in anticyclonic situations. Next, the newly created ULTURB method could be used to examine CAT outbreaks (McCann et al. 2012). It would be interesting to see how this method of using Lighthill-Ford compares to Sharman's method of using Lighthill-Ford, as well as comparing ULTURB to the other turbulence indices. The ULTURB method may show more skill at predicting turbulence, due to the gravity wave physics that it contains. While the ultimate turbulence diagnostic is still a scientific dream, perhaps this thesis has played a small role in the quest for that diagnostic.

REFERENCES

Baughman, E.E., Jr., 1946: Turbulence with a stable lapse rate. *Bull. Amer. Meteor. Soc.*, **27**, 459-462.

Benjamin, S., and S. Sahm, 2011: Rapid Refresh (RR). [Available online at http://rapidrefresh.noaa.gov/].

Bluestein, H.B., 1993: *Observations and Theory of Weather Systems*. Vol. II. *Synoptic-Dynamic Meteorology in Midlatitudes*, Oxford University Press, 594 pp.

Buldovskii, G.S., S.A. Bortnikov, and M.V. Rubinshtejn, 1976: Forecasting zones of intense turbulence in the upper troposphere. *Meteor. Gidrol.*, **2**, 9-18.

Chambers, E., 1955: Clear air turbulence and civil jet operations. *J. Roy. Aeronaut. Soc.*, **59**, 613-628.

Colson D., and H.A. Panofsky, 1965: An index of clear-air turbulence. *Quart. J. Roy. Meteor. Soc.*, **91**, 507-513.

Dutton, J., and H.A. Panofsky, 1970: Clear air turbulence: A mystery may be unfolding. *Science*, **167**, 937-944.

Ellrod, G.P., and D.L. Knapp, 1992: An objective clear-air turbulence forecasting technique: Verification and operational use. *Wea. Forecasting*, **7**, 150-165.

Ellrod, G.P., and J.A. Knox, 2010: Improvements to an operational clear-air turbulence diagnostic index by addition of a divergence trend term. *Wea. Forecasting*, **25**, 789-798.

Ellrod, G.P., P.F. Lester, and L.J. Ehernberger, 2003: Clear air turbulence. In *Encyclopedia of the Atmospheric Sciences*, eds. J.R. Holton, J. Pyle, and J.A. Curry, Academic Press, New York, 393-403.

Endlich, R.M., 1964: The mesoscale structure of some regions of clear-air turbulence. *J. Appl. Meteor.*, **3**, 261-276.

Ford, R., 1994: Gravity wave radiation from vortex trains in rotating shallow water. *J. Fluid Mech.*, **281**, 81-118.

Frehlich R., and R. Sharman, 2004: Estimates of upper level turbulence based on second order structure functions derived from numerical weather prediction model output. Preprints, *11th Conf. on Aviation, Range, and Aerospace Meteorology,* Hyannis, MA, Amer. Meteor. Soc., CD-ROM, P4.13.

Jolliffe, I.T., and D.B. Stephenson, 2003: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley and Sons.

Kaplan, M.L., and Coauthors, 2004: Characterizing the severe turbulence environments associated with commercial aviation accidents. A real-time turbulence model (RTTM) designed for the operational prediction of hazardous aviation turbulence environments. NASA CR-2004-213025, 54 pp.

Knox, J.A., 1997: Possible mechanism of clear-air turbulence in strongly anticyclonic flows. *Mon. Wea. Rev.*, **125**, 1251-1259.

Knox, J.A., D.W. McCann, and P.D. Williams, 2008: Application of the Lighthill-Ford theory of spontaneous imbalance to clear-air turbulence forecasting. *J. Atmos. Sci.*, **65**, 3292-3304.

Knox, J.A., D.W. McCann, and P.D. Williams, 2009: Reply to Plougonven et al.'s "Comments on 'Application of the Lighthill-Ford theory of spontaneous imbalance to clear-air turbulence forecasting'". *J. Atmos. Sci.*, **66**, 2511-2516.

Kronebach, G.W., 1964: An automated procedure for forecasting clear-air turbulence. *J. Appl. Meteor.*, **3**, 119-125.

Lane, T.P., R.D. Sharman, S.B. Trier, R.G. Fovell, and J.K. Williams, 2012: Recent advances in the understanding of near-cloud turbulence. *Bull. Amer. Meteor. Soc.*, **93**, 499-515.

Lindborg, E., 1999: Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? *J. Fluid Mech.*, **388**, 259-288.

Marroquin, A., 1998: An advanced algorithm to diagnose atmospheric turbulence using numerical model output. Preprints, *16th Conf. on Weather Analysis and Forecasting*, Phoenix, AZ, Amer. Meteor. Soc., 79-81.

Massachusetts Water Resources Authority, 2005: ROC Curve Preparation. [Available online at http://www.mwra.state.ma.us/harbor/enquad/pdf/2005-20.pdf].

McCann, D.W., 2001: Gravity waves, unbalanced flow, and aircraft clear air turbulence. *Natl. Wea. Dig.*, **25**, 3-14.

McCann, D.W., J.A. Knox, and P.D. Williams, 2012: An improvement in clear-air turbulence forecasting based on spontaneous imbalance theory: The ULTURB algorithm. *Meteor. Apps.*, **19**, 71-78.

Plougonven, R., C. Snyder, and F. Zhang, 2010: Comments on "Application of the Lighthill-Ford theory of spontaneous imbalance to clear-air turbulence forecasting." *J. Atmos. Sci.*, **66**, 2506-2510.

Ray, P.S. (Ed.), 1986: *Mesoscale Meteorology and Forecasting*. Amer. Meteor. Soc., Boston, MA, 793 pp.

Reap, R.M., 1996: Probability forecasts of clear-air turbulence for the contiguous U.S. National Weather Service Office of Meteorology Tech. Procedures Bull. 430, 15 pp.

Schwartz, B., 1996: The quantitative use of PIREPs in developing aviation weather guidance products. *Wea. Forecasting*, **11**, 372-384.

Sharman, R., C. Tebaldi, G. Wiener, and J. Wolff, 2006: An integrated approach to mid and upper-level turbulence forecasting. *Wea. Forecasting*, **21**, 268-287.

Takacs, A., L. Holland, R. Hueftle, B. Brown, and A. Holmes, 2005: Using in situ eddy dissipation rate (EDR) observations for turbulence forecast verification. Report to the FAA Aviation Weather Research Program. [Available online at http://www.rap.ucar.edu/research/verification/FAA_reports/E5.pdf].

Trier, S.B., R.D. Sharman, and T.P. Lane, 2012: Influences of moist convection on a cold-season outbreak of clear-air turbulence (CAT). *Mon. Wea. Rev.*, doi:10.1175/MWR-D-11-00353.1, in press.

Wolff, J.K., and R.D. Sharman, 2008: Climatology of upper-level turbulence over the contiguous united states. *J. Appl. Meteor. Clim.*, **47**, 2198-2214.

Woodcock, F., 1976: The evaluation of yes/no forecasts for scientific and administrative purposes. *Mon. Wea. Rev.*, **104**, 1209-1214.