1.5 STATISTICAL ANALYSIS OF VARIABLES ASSOCIATED WITH CONVECTIVE INITIATION ALONG THE SOUTHERN PLAINS DRYLINE

Michael P. Griesinger* and Christopher C. Weiss
Texas Tech University, Lubbock, Texas

1. INTRODUCTION

One of the most prominent weather features on the southern Great Plains during the spring and early summer months is the dryline. The dryline is a mesoscale boundary layer feature that serves as a delineator between moist maritime tropical air originating from the Gulf of Mexico and dry continental tropical air originating from the high terrain of the desert southwest United States and the Mexican Plateau (Schaefer 1986). The density difference between the two air masses, which leads to the development of a solenoidal circulation, along with the pressure gradient force associated with a trough of low pressure usually concurrent with the dryline, leads to the confluence of the surface winds (Ziegler and Rasmussen 1998). This convergence of the surface winds is one of the main reasons why the dryline serves as a focus for deep moist convection during the spring on the southern Great Plains (Rhea 1966; Doswell 1982; Ziegler and Rasmussen 1998; and Weiss and Bluestein 2002).

The study by Rhea (1966) found that of the drylines identified in the research only 70% were associated with deep moist convection somewhere along their length, leading to the frequently asked questions by forecasters every spring on the South Plains of will the dryline be active, and if it is, where will it be active. The goal of this research is to answer the questions of where and will a dryline be active by using data that is readily available to an operational forecaster. After data collection a multitude of variables were calculated and inserted into a regression model to determine an equation for forecasting the likelihood of convective initiation (CI) along the dryline.

Section 2 will include data sources used and an explanation of the regression method utilized can be found in section 3. Section 4 reveals and explains the variables in the CI probability equation. Two case studies utilizing the CI probability equation will be presented in section 5, followed by a brief summary and conclusions in section 6.

* *Corresponding author address:* Michael P. Griesinger, NOAA/National Weather Service, Peachtree City, GA 30269; e-mail: michael.griesinger@noaa.gov.

2. METHODOLOGY

There were two types of data used for this research, surface and upper air. For surface data, the West Texas Mesonet (WTM) was utilized. The WTM is a collection of 48 observation sites across West Texas that are both east and west of the Caprock Escarpment and are spaced no more than 35 km apart (Fig. 1; Schroeder et al. 2005). The higher spatial and temporal resolution of the WTM compared to the synoptic observations leads to a better placement of the dryline and a more accurate calculation of the convergence (Griesinger 2006).

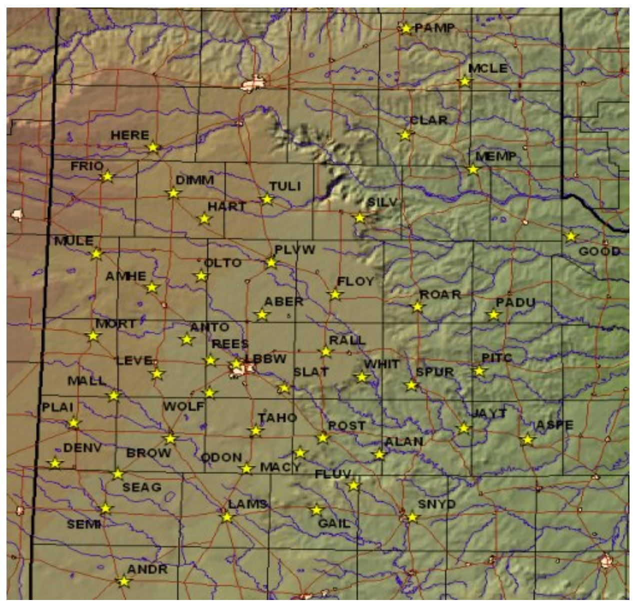

Figure 1. Map showing current 48 WTM sites.

For upper air data two radiosonde locations were used, Midland (MAF) and Amarillo, (AMA) TX. Up to five soundings were created each day at specific locations within the WTM by doing a simple linear interpolation between conditions at AMA and MAF. If both locations were in the moist air at both 1200 and 0000 UTC, then composite soundings were generated by interpolating in both time and space to get soundings at desired times and locations. If either AMA or MAF was dry at the surface at 0000 UTC, then composite soundings were created by interpolating in space only with the 1200 UTC conditions at each site. The generated composite soundings were then completed by inserting an observation from a WTM site for the sounding surface conditions. Composite sounding profiles generated by both methods were then corrected by following the dry adiabat associated with the surface temperature until it intersected the original composite sounding temperature. Moisture was also corrected by following the mixing ratio associated with the surface dewpoint to the level where the temperature correction was no longer made. The corrections made led to an inverted-v boundary layer profile.

Cases for this research were drawn from days that had a dryline in West Texas during the years of 2004 and 2005. Three criteria were used in selecting cases for this study. First, a dryline had to be present within the WTM domain. For a moisture gradient to be considered a dryline, a 10^{o} C per 100 km dewpoint gradient needed to be present within the WTM domain. If a dryline was determined to be present in West Texas during the day, it then needed to stay within the WTM domain during the entire day. Finally, if the first two requirements were met, the soundings at 1200 UTC for AMA and MAF needed to be present and moist at the surface in order for the dryline to be included within the data set used in the regression model.

If the dryline met all of these requirements, then calculation locations were selected by picking WTM sites that were directly out ahead of the dryline. One location was selected per north-south row of counties in the WTM, resulting in, at most, five variable locations for each dryline. The calculations were made either 30 minutes before the first precipitating convection appeared on radar, or the time when the peak temperature occurred at a location when convection was not present. The above methodology yielded 22 drylines, with a total of 97 observation points on these drylines. Of the 97 observations, 23 dryline segments had CI occur while the other 74 did not.

3. STEPWISE LOGISTIC REGRESSION

After dryline cases were selected, composite soundings generated, and variables calculated, stepwise logistic regression was utilized to fit the data to a model. Logistic regression places a restraint on the response variable, which is that it must be binary (Ryan 1997). For this research, that means a location had a response value of “1” if CI occurred along the chosen dryline (determined by the presence of a radar echo), and a “0” if CI did not happen. The binomial nature of the response variable results in the regression equation and curve being an exponential. The logistic regression equation for a single predictor is as follows:

$$y = \frac{\exp(\beta_o + \beta_1 X)}{1 + \exp(\beta_o + \beta_1 X)} \qquad (1)$$

where β_o and β_1 are coefficients that must be estimated and X is the predictor variable, which has no restrictions.

Since multiple predictor variables were used, stepwise regression was needed to test the significance of each variable. To accomplish this, an intercept-only model (IOM) must first be created using only the response variables to generate the first model to be tested against. The IOM results in a climatological probability for CI in the two years studied (23/97). Once the IOM was created, multiple steps of forward selections and backward eliminations were conducted to improve the models ability to predict CI along the dryline.

Forward selection was used to add variables to the model to see which variable improves the fit of the data to a logit curve the most. This is accomplished by adding each variable not already in the model to the current equation of the step to see which variable improves the fit of the data to a logistic curve the most. In S-Plus, Mallow’s Cp statistic is used to test the effectiveness of the equation at fitting the model. The variable addition that lowers the Cp value the most gets added to the equation. When a variable is added, a backward elimination test is conducted by removing each variable in the equation individually to see how the Cp statistic changes. As long as the removal of a variable does not lead to a decrease in the Cp statistic they are all kept and the procedure moves on to the next forward selection. Once the addition of a variable no longer decreases the Cp statistic, the process is done and a final equation for predicting CI is generated. A summarization of the results from each step can be found in table 1, while the resulting equation can be found in section 4.

4. CI PROBABILITY EQUATION

After running the stepwise logistic regression procedure using data collected from all dryline segments, the following equation for the logistic regression model was found:

$$Z = -\beta_o - \beta_1 * dz - \beta_2 * (T_{700} - T_{500}) + \beta_3 * (T_{850} - T_{500}) - \beta_4 * SDD \qquad (2)$$

Where: $\beta_o = 3.194$ $\beta_1 = 0.046\ m^{-1}$
$\beta_2 = 1.384\ ^{o}C^{-1}$ $\beta_3 = 0.957\ ^{o}C^{-1}$

$\beta_4 = 0.238\ ^{o}C^{-1}$.

In (2), dz is the 500 hPa height difference between AMA and MAF (AMA-MAF) at 1200 UTC, ($T_{700} - T_{500}$) is the 700 to 500 hPa lapse rate, ($T_{850} - T_{500}$) is the lapse rate between 850 and 500 hPa, and SDD is the surface dewpoint depression. Inserting the answer from (2) into (1) results in the following probability equation for CI along the dryline:

$$P(CI)=\frac{\exp(Z)}{1+\exp(Z)} \qquad (3)$$

In (2), dz is a measure for the amount of background synoptic lift present, as a large value for dz is usually coincident with strong mid-level troughing over the SGP. Although dz is subtracted from the equation, it is a positive forcer for CI, since the 500 hPa height at AMA is usually less than that at MAF, making the height difference negative. In all of the cases where convection occurred along the dryline the value for dz was at least 20 m.

Perhaps the most compelling result from the logistic regression procedure is the negative relationship between the 700 to 500 hPa lapse rate and CI along the dryline. The negative correlation means a smaller lapse rate is more conducive to CI than a large one. It is hypothesized that the reason for this relationship lies with entrainment. Houze (1993) noted that entrainment increases for narrower and stronger updrafts. Since dryline updrafts already tend to be narrow, a stronger updraft can exacerbate the effects of entrainment on dryline updrafts, causing them to dry out before the level of free convetion (LFC) is attained. Since strong updrafts are associated with large lapse rates, a smaller 700 to 500 hPa lapse rate would lead to slightly weaker updrafts, potentially leading to less entrainment of dry environmental air into a moist updraft, which would allow a surface based moist air parcel to reach its LFC before drying out.

The other lapse rate included in the CI model was the lapse rate between 850 and 500 hPa. Unlike the 700 to 500 hPa lapse rate, the 850 to 500 hPa lapse rate has the expected positive correlation with CI. Considering the 850 to 700 hPa layer, which represents the principle difference from the previous predictor, a surface based parcel is usually unsaturated, which means the entrainment of environmental air is not evaporating cloud condensate, the effect of which is hypothesized to increase the cooling and stabilization in the 700 to 500 hPa layer. The lack of evaporation means entrainment is not as detrimental to a moist parcel, and the stronger updrafts associated with larger 850 to 500 hPa lapse rates can help a surface based moist air parcel attain its LFC and initiate deep moist convection.

The final variable included in the CI probability equation is the SDD. As expected, there is a negative correlation between CI and the SDD (after the variance of the previous predictors is accounted for), resulting in a smaller SDD being more conducive to CI. A small SDD is associated with more low-level moisture, a lower lifted condensation level (LCL), and less environmental convective inhibition. All of these factors result in increased chances for a surfaced-based parcel to be able to reach its LFC.

5. CASE STUDIES

1. 23 May 2006 Dryline

The dryline of 23 May 2006 was not associated with deep moist convection, but was able to lift parcels to the LCL, as evidenced by a visible satellite image (Fig. 2). The dryline propagated to the eastern edge of the WTM array. Therefore, the sites at Memphis, Paducah, Guthrie, Jayton, and Snyder, TX were used for CI prediction locations. On this day all of the CI probabilities (Table 2) stayed below 5 %, which coincides well with the lack of deep convection. The low dz value, caused by upper level ridging, along with the large SDDs are the main factors that led to the low CI probabilities along the dryline.

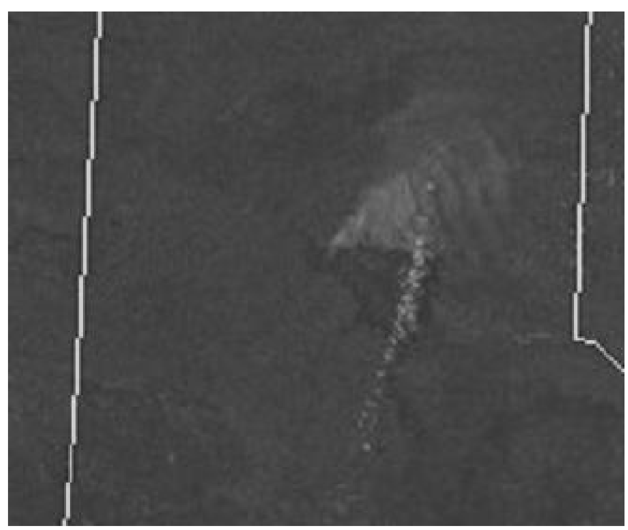

Figure 2. Visible satellite image for 23 May 2006 valid at 2245 UTC. The feature to the west of the dryline is likely a smoke plume from a fire in Palo Duro Canyon.

2. 28 May 2006 Dryline

The dryline of 28 May 2006 was associated with CI along its entire length in the WTM domain. Convection was first initiated at 2030 UTC near Memphis, TX, while the last storms initiated at 2230 UTC in the vicinity of Snyder, TX. The strongest storms on this day were those in the southeast TX Panhandle (Fig. 3), where scattered reports of wind damage and hail occurred.

The values for CI probabilities on this day (Table 3) were all larger in comparison to 23 May 2006. The occurrence of convection coincides well with the higher CI probabilities. One of the reasons the CI probabilities are higher on the 28th is the 30 m height difference between AMA and MAF seen on this day. Although the trough at 500 hPa stayed north of the South Plains, it came close enough to the TX Panhandle to provide some background synoptic lift to the WTM domain, especially in the north. The reader is also directed to the values for the 700 to 500 hPa lapse rate. It should be noted that the lapse rates in this layer are smaller on 28 May 2006, when CI occurred, than on 23 May 2006, when convection was not present, consistent with the relationship between CI and the 700 to 500 hPa lapse rate found with the data from 2004 and 2005 used to create the prediction equation. Although none of the CI probabilities are particularly large, the largest value of 38.45 % at Memphis is larger than the predicted probability from climatology of 24 % and is coincident with the location of the strongest storm.

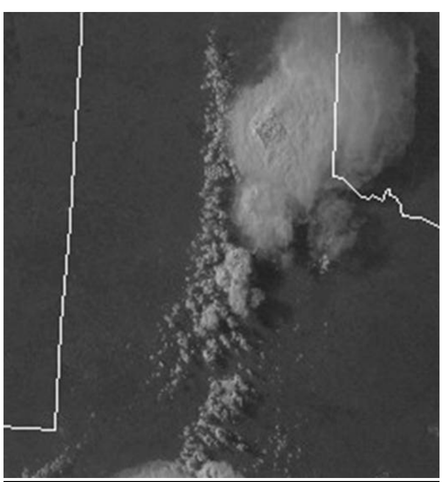

Figure 3. Visible satellite image valid at 2345 UTC on 28 May 2006.

6. SUMMARY AND CONCLUSIONS

The purpose of this research was to use observational tools available to the operational forecaster on the SGP of the U.S., such as atmospheric soundings and surface observations, to predict CI along the dryline. To accomplish this goal dryline cases from the spring months of 2004 and 2005 were compiled. In particular, drylines that stayed within the confines of the WTM array were analyzed. Drylines that were chosen for this research were then categorized as CI if the dryline initiated deep moist convection, or NCI if convection was not present along the dryline. For this study, CI was deemed to occur when a precipitation echo was noted on radar along the dryline.

From the data collected during the springs of 2004 and 2005, a stepwise logistic regression procedure was executed to generate a forecast equation for CI along the dryline. It includes the 500 hPa height difference between AMA and MAF, the 700 to 500 hPa and 850 to 500 hPa lapse rates, and the SDD. It is hypothesized that the negative correlation between CI and the 700 to 500 hPa lapse rate (after the variance by dz is explained) is the result of increased entrainment resulting from stronger updrafts present with larger lapse rates. Results from two case studies show the potential usefulness of the CI probability equation, as the highest probabilities calculated were all on 28 May 2006, when convection initiated along the dryline. Additionally the largest probability was also associated with the strongest convection.

7. References

Doswell, C. A., III, 1982: The operational meteorology of convective weather. Vol. I: Operational mesoanalysis. NOAA Tech. Memo. NWS NSSFC-5, 186 pp.

Griesinger, M. P, 2006: Statistical analysis associated with convective initiation along the dryline. M. S. thesis, Dept. of Geosciences, Texas Tech University, 118 pp.

Houze, D. A., Jr., 1993: *Cloud Dynamics*. Academic Press, San Diego, CA. 573 pp.

Rhea, J. O., 1966: A study of thunderstorm formation along drylines. *J. Appl. Meteor.*, **5**, 58-63.

Ryan, T. P., 1997: *Modern Regression Methods.* John Wiley & Sons, Inc., New York, NY, 515 pp.

Schaefer, J. T., 1986: The dryline. *Mesoscale Meteorology and Forecasting,* P. S. Ray, Ed., Amer. Meteor. Soc., 549-572.

Schroeder, J. L., W. S. Burgett, K. B. Haynie, I. Sonmez, G. D. Skwira, A. L. Doggett, and J. W. Lipe, 2005: The West Texas mesonet: A technical overview. *J. Atmos. Oceanic Technol*., **22**, 211-222.

Weiss, C. C., and H. B. Bluestein, 2002: Airborne pseudo-dual Doppler analysis of a dryline-outflow boundary intersection. *Mon. Wea. Rev.,* **130,** 1207-1226.

Ziegler, C. L., and E. N. Rasmussen, 1998: The initiation of moist convection at the dryline: Forecasting issues from a case study perspective. *Wea. Forecasting*, **13**, 1106-1131.

Table 1: Summary of Cp values at each step of the logistic regression procedure. Bold and italicized values represent variables selected at each step.

	Step 1 Cp	**Step 2 Cp**	**Step 3 Cp**	**Step 4 Cp**	**Step 5 Cp**
Step Cp value	99	97.425	97.7974	101.7778	120.8609
CIN (J kg^{-1})	97.1677	96.5077	96.3193	101.6191	121.0772
Convergence (s^{-1})	100.6986	98.7955	99.1658	103.0885	122.1805
dz; 500 hPa AMA to MAF 12 UTC height difference	***92.6008***	--	--	--	--
Surface dewpoint depression (SDD; ^{o}C)	100.3553	99.3537	99.7029	***98.6794***	--
600 hPa dewpoint depression (^{o}C)	99.2719	99.0018	99.7968	103.544	122.7852
LCL height (m)	100.3278	99.3394	99.7271	98.7213	122.821
dz/dt (at 500 hPa; m/12 hr)	100.4028	99.145	99.6222	103.146	122.2619
700 hPa AMA to MAF 12 UTC height difference	99.2624	99.2598	99.4764	103.3825	122.8035
850 hPa dewpoint depression (^{o}C)	100.6326	99.3732	99.3066	100.2572	122.7929
700 hPa dewpoint depression (^{o}C)	96.1838	96.9076	97.2691	103.1663	121.994
500 hPa dewpoint depression (^{o}C)	100.9437	99.0943	99.2207	103.6318	122.7441
T(850 hPa) – T(700 hPa) (^{o}C)	99.2741	97.0874	96.096	103.4882	122.6519
T(850 hPa) – T(500 hPa) (^{o}C)	100.3408	99.4189	***96.0914***	--	--
T(700 hPa) – T(500 hPa) (^{o}C)	93.3762	***96.0857***	--	--	--
T (700 hPa) (^{o}C)	100.6919	99.2997	99.4214	103.6395	122.7671
T (500 hPa) (^{o}C)	100.0889	98.0192	99.4888	103.5906	122.7955
T (surface) (^{o}C)	100.5938	99.4248	99.5382	101.3055	122.8303
T (850 hPa) (^{o}C)	100.9832	98.5886	98.21769	103.5733	122.8077

Table 2: Values for variables in equation 2 and the probability of CI for 23 May 2006.

	MEMP	PADU	PITC	JAYT	SNYD
dz	-20	-20	-20	-20	-20
T(700)-T(500)	22.95	22.675	22.4	22.125	21.85
T(850)-T(500)	36.31327	36.2844	37.01747	35.69344	37.03157
SDD	22.66	21.94	25.27	23.11	25.17
CI %	0.92	1.55	2.05	1.43	4.45

Table 3: Values for variables in equation 2 and the probability of CI for 28 May 2006.

	MEMP	ROAR	WHIT	JAYT	SNYD
dz	-30	-30	-30	-30	-30
T(700)-T(500)	21.45482	22.58882	22.88387	22.075	22.68902
T(850)-T(500)	36.67627	38.8213	39.12737	38.09644	38.95457
SDD	21.12	21.67	22.08	21.11	22.28
CI %	38.45	25.44	21.61	28.4	22.58