Anticipating Lightning Probabilities Using the Perfect Prog Technique
October 2009

Chris Outler
NWSFO Medford, OR

I - Introduction

Anticipating lightning potential during the dry summer months across much of the Western U.S can be challenging, and due to the convective dependence on topography, is often difficult to forecast in advance. Furthermore, due to the Western United States susceptibility to extreme fire danger, effective techniques to forecast lightning during periods of critical fire weather are highly advantageous to fire weather operations, emergency management, and the public who may be affected. This report will highlight the development and operational usage of a lightning probability product generated from statistical analysis and use of the perfect prog technique to develop numerical lightning guidance for 101 RAWS locations in the Medford, Oregon county warning area (CWA). In addition, a summary of the results will be provided illustrating both the advantages and limitations of the study.

II - Background

The Medford CWA includes all of Southwestern Oregon and far Northern California, a region rich with mountainous and varied topography. Coastal beaches are found on the western edge, the Cascade Range bisects the area from north to south, the Siskiyou Mountains dominate most of the Northern California zones, and an elevated basin lies across the eastern extent of the region. These diverse topographical features result in numerous complex climate zones within the CWA itself.

Since the climatic zones are so diverse, the forcing mechanisms which most commonly initiate convection can vary dramatically by region, therefore, for the purposes of this study, the CWA has been broken down into eight regions of distinct climate zones (Figure 1).

These zones include the coastal western edge, where convection is least common in the summer months and more common in winter and spring as cold air aloft moving over warmer Pacific waters results in very steep lapse rates and isolated thunderstorm activity. The Coast Range, as well as the Umpqua and Rogue River Basins feature rugged mountainous terrain with a few broader valleys of low elevation. Typically these areas are susceptible to some occasional thunderstorm activity in the summer months, however, it is usually supported by some synoptic forcing as storms organize and drift off the higher terrain.

The Cascades and Siskiyou zones feature a chain of elevated mountains, which are susceptible to diurnal thunderstorm activity due to the high elevation and orographic forcings. Finally, the Klamath Basin, Eastside, and Modoc zones are all elevated basins with broad valleys and plateaus as well as some mountain ridges. Since these locations are downwind of the Cascade crest, stabilizing maritime influences are blocked by the terrain, and sites in these zones observe the most frequent thunderstorm activity in the CWA (Figure 2).

It is important to note that a lightning occurrence for the purposes of this study was determined by the number of cumulative lightning strikes occurring in a six hourly period (base times 12z, 18z, 0z, and 6z), occurring within a grid oriented 25 miles east or west, 25 miles north or south, and centered on a selected RAWS site. Then the number of lightning strikes were categorized to highlight the significance of the lightning event and broken down into three categories, one strike, ten strike, and one hundred strike events. Additionally, since this study was focused more on warm season convective weather for fire weather purposes, only strikes occurring between May and October were used in this study.

One of the greatest difficulties with this project is the lack of thunderstorm activity observed in the Medford CWA. Widespread thunderstorm events in Northern California and Southern Oregon are a rarity, and having sufficient lightning strike data to supply an accurate statistical correlation can be difficult, especially when dealing with coastal sites, or sites within some proximity to the coastline. Even across the easternmost zones, which are convectively most active, complications can arise when attempting to correlate large lightning outbreaks (100+ CG Strikes) with model data, due to the fact that some sites have had either very few 100+ lightning strike events in a given 6 hourly period, or none at all. Therefore, for sites that suffered from these problems, only probabilities for a single CG strike and for 10 CG strikes in a 6 hourly period were calculated, while 100+ CG strike probabilities were neglected due to unreliable statistical correlations.

The ultimate goal of this project is to better understand regional convective forcing mechanisms which drive thunderstorm activity across the region. Through statistical analysis of model data from the 101 RAWS sites and observed lightning data for the 2000-2008 period, correlations of atmospheric variables can be determined using the perfect prog technique. Once the statistical guidance is available, it will be used operationally and implemented into AWIPS to aid convective forecasting, fire weather forecasts, and spot forecasts.

III - Methodology

The development of this project required utilization of many resources including model and lightning data archives, programming, databasing, and statistical analysis software. The final result provides statistical lightning probabilities using the perfect prog technique, which assumes the regression equations perfectly predict the state of the atmosphere and neglects any potential model biases. So if guidance is correct, the forecast probability should perform well. If guidance is flawed, the forecast probabilities will be flawed as well.

Model data was provided by downloading the appropriate gridded binary (GRIB) data from the North American Regional Reanalysis (NARR) page (http://nomads.ncdc.noaa.gov) where GRIB model data is archived. Model data from the year 2000 through 2008, months of May through October, and broken into 6 hourly data periods were downloaded and used for this study. Then, a C program code was written which decoded the GRIB data and converted it into an ASCII database of meteorological variables. Additionally, a file was created containing the latitude and longitude of each of the RAWS sites, so the NARR data could be subseted to be location specific.

Next, lightning data was provided which included an archived database of all lightning strikes for both Northern California and Oregon. Then using the file containing the exact latitude and longitude for all the RAWS sites, lightning data was appended to the model data for each site if it occurred within a grid oriented 25 miles east or west, 25 miles north or south, and centered on the selected RAWS site. The resulting lightning occurrences near a specific RAWS site were then appended to the end of each 6 hourly record in each specific model database, so that the model data and time of lightning occurrence matched. Finally, one strike, ten strike, and hundred strike events were categorized in binary form and appended each record as well. The resulting database for each site included 6 hourly model data, total nearby lightning occurrences in each 6 hourly time frame, and whether the occurrence was a one+ strike, ten+ strike, or hundred+ strike event.

Now with the databases complete, each file was converted to a comma separated value (.csv) format, which was imported into SYSTAT, which is a statistical analysis software package which was used for this study (www.systat.com). SYSTAT was important in determining the best meteorological predictors of lightning for each RAWS site, which was then used to create a site specific regression equation. Statistics for the best predictors of lightning were recorded for each RAWS site, and for each categorical lightning threshold (one strike, ten strike, hundred strike). Then, the best predictors were generalized by zone, and used to create scripts which systematically computed regression coefficients for each RAWS site. The resulting output was then saved into text files for later reading in WordPad.

Finally, using a locally created C program named LTGPARAM.C, regression equations were extracted for each site. This program decoded the output from SYSTAT logistic regression analysis and compiled an equation database for each RAWS site. With the completed equation files, the statistics could be run against the NCEP model forecast using the Perfect Prog approach to estimate the 6 hourly lightning probabilities. These probabilities were then made locally available on AWIPS for aid in operational and fire weather forecasting (Figure 3).

IV - Analysis

The results of the statistical analysis indicated that, as expected, convective indices generally performed best at anticipating lightning potential, especially when considering the probability of multiple strike events. Therefore, all zones included Lifted Index, CAPE, and CIN in the development of the regression equations.

Moisture fields were a trouble spot in the statistical analysis for this study, as in some sites they proved useful while in others they did not. Correlations were attempted using relative humidity of multiple layers of the troposphere, as well as the surface, but did not seem to strengthen the regression equation to any significant degree. Attempts were also made to calculate Theta-E at 700 mb; however, it too proved unsuccessful at enhancing the regression equation. This is likely because in operational meteorology, locating Theta-E ridges often highlights the greatest conditional thunderstorm threat and largest axis of energy, but Theta-E by itself is not as useful. The only fields which did prove successful were precipitable water, and observed precipitation, which both featured a positive correlation with observed lightning.

Of the temperature fields, tropopause temperature, 1000 mb temperature, and 2 meter temperature, all were used to strengthen the regression equations. Lapse rates for many different levels were tested as well to determine any convective dependencies; however, they did not improve the correlations. This could be explained by the high frequency of warm, summer days with very steep lapse rates across the region, but due to the low moisture content of the atmosphere that is so dominant in the area during the summer, most zones remain devoid of convection. In other parts of the country, lapse rates may be much more useful in statistical analysis, but since convection is such a rarity across Southern Oregon and Northern California, it becomes much less useful.

Height fields were also examined but proved less useful than some of the other parameters listed below. Of exception were both the height of the planetary boundary layer, and the height of the tropopause, but individual layer heights lacked any strong correlation with lightning days. In future studies, it may be useful to investigate height tendencies, and see if periods of rapid height rises or falls are associated with convection, but that was not a focus for this study.

Of the vertical motion fields, vertical velocities in both the lower and upper troposphere proved useful for anticipating lighting potential. The vertical velocities at both the 800-600 mb level, and the 400-300 mb level, were frequently listed as strong predictors of lightning during the statistical analysis using SYSTAT. These fields were especially critical for the Coast Range sites, as well as those in the Umpqua and Rogue Basin’s, where synoptic scale forcing is often required to sustain convection drifting off the higher terrain.

Lastly, some zones featured a tendency to favor specific wind components in convective development. Specifically, sites in Siskiyou County favored a U component to the wind in the lower 6000 feet of the atmosphere, while favoring a southerly V component above. This is likely a result of the complex topography across Northern California and localized areas of surface convergence which result from these wind patterns. These tendencies were included in the top predictors used to develop the regression equation for sites in this zone. Similar relationships were found and included within the Cascades, Klamath Basin, and Modoc county sites. Therefore, it appears the more elevated mountain zones have less dependency on synoptic scale ascent, and more reliance on mesoscale processes to organize diurnal convection.

V – Conclusion

Since its development, the regression equations are run daily against NCEP model forecasts, with the statistical results stored into an operational guidance product available on the local Medford AWIPS for use in forecasting lightning probability. Since it is a new, experimental product, further work will be required to improve its performance, and verification studies will be needed to monitor this products accuracy. Unfortunately, this product was not available most of the summer of 2009 since it was still being developed. Thus, the summer of 2010 will likely provide many opportunities to verify the accuracy of the statistical guidance.

In future studies, it would be helpful to focus on some of the problems which were encountered with the development of this product. Most importantly, the period of study needs to be broadened, as 8 years of lightning data was still insufficient to provide concrete results that one can trust, especially within coastal zones where lightning frequency is low. Future work should expand on the current time range and season, to provide a more thorough record for developing statistics.

Additionally, future development will require more reliance on moisture fields, which showed poor correlations in this study. Certainly moisture plays a key role in convective activity even through the statistical analysis did not pick up on it. Perhaps auto-correlations or relationships with similar variables are to blame for some of the low correlations that were represented in the study. Future statistical analysis will need to investigate why there is such variance in moisture fields and convective predictability among the different RAWS sites. Furthermore, other meteorological fields such as layer dewpoints, Theta-E, and lapse rates could be tested for positive correlations with lightning frequency.

Continued work into verifying the accuracy, as well as expanding on the project itself will lead to the continued improvement of this product. In the fire seasons to come, this guidance should prove valuable to forecasters, fire weather operations, and the public who are affected. Hopefully, this product will allow forecasters to better anticipate the atmospheric conditions leading to lightning outbreaks when fire danger is at critical levels.

VII – Figures

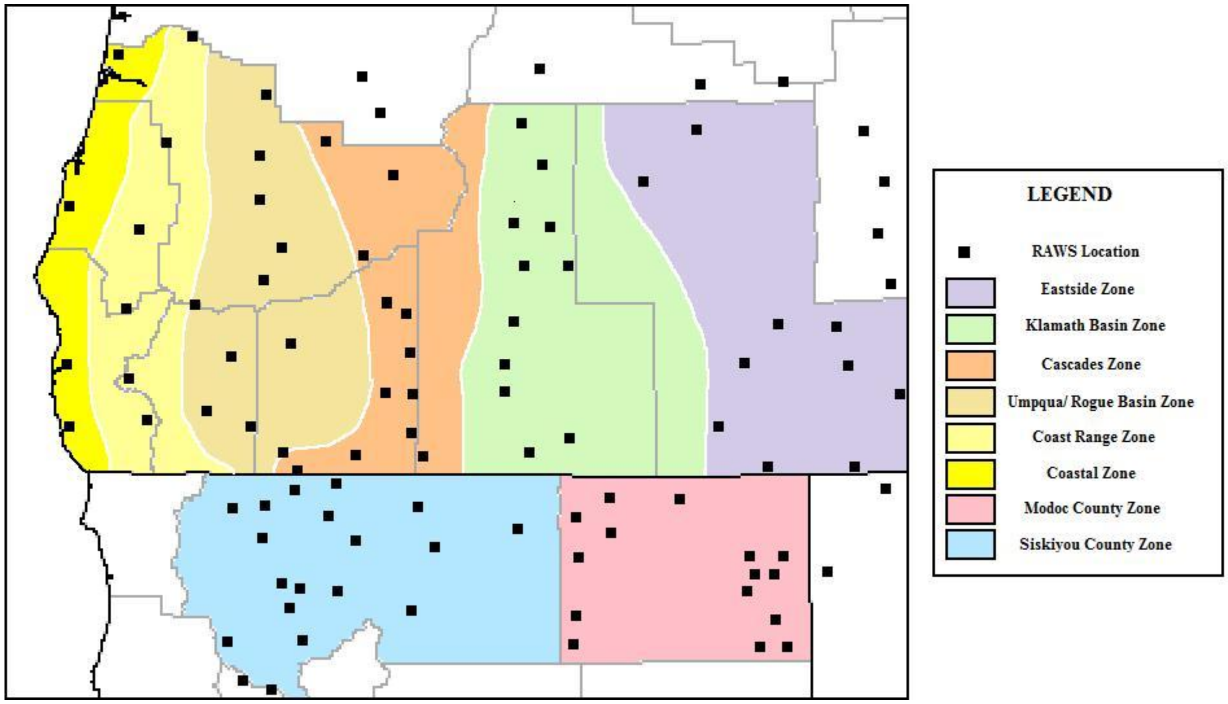

*Figure 1:*RAWS locations and Climatic Zones

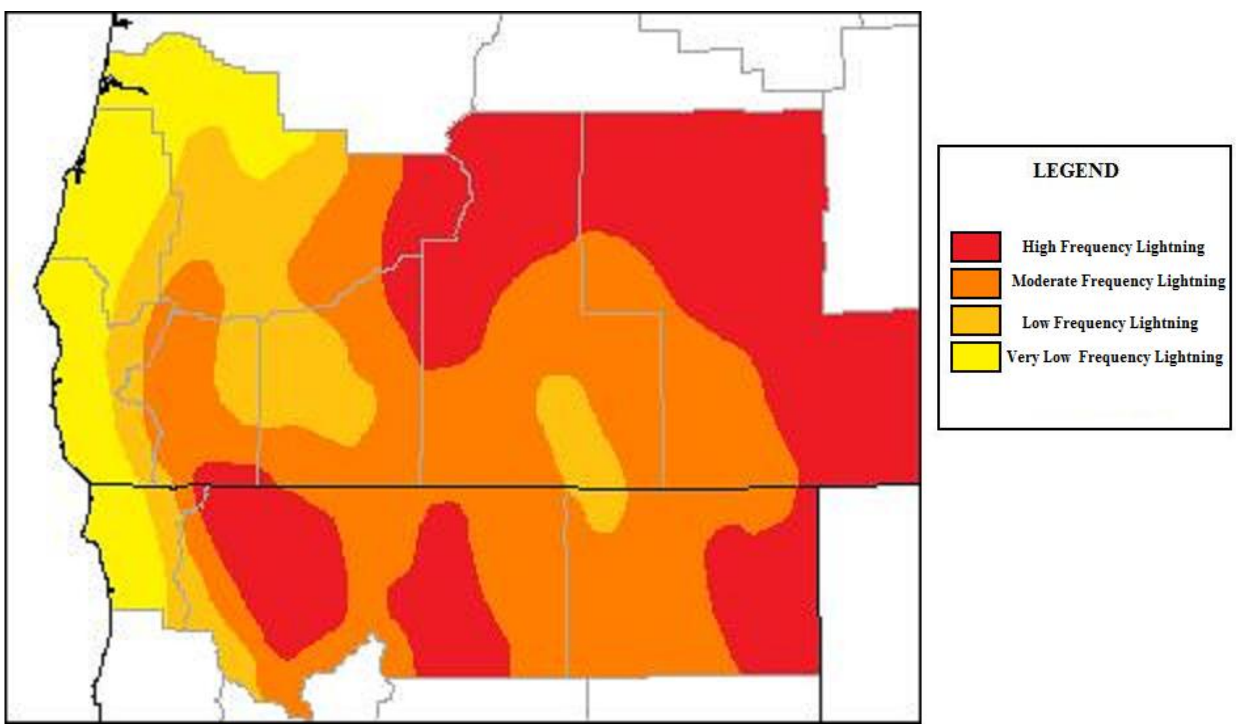

Figure 2: Lightning Frequency in the Medford CWA

```
ZCZC PDXFWGCIN SDC
TTAA00 KMFR DDHHMM
MOS FWG GUIDANCE FOR Cinnamon  08/10/1200Z
Parm  10  10  11  11  11  11  11  11  11  11  12  12  12  12  12  12  12  12  13
     18Z 21Z 00Z 03Z 06Z 09Z 12Z 15Z 18Z 21Z 00Z 03Z 06Z 09Z 12Z 15Z 18Z 21Z 00Z
P01    1       1       2       0       2       4       7       8       5      14
P10    0       0       2       0       0       1       7       8       1       2
S01    0       1       1       1       1      13       7       6       4      28
S10    0       0       1       0       0       5       6       6       2      13

Tmp   71  77  79  69  56  51  46  58  74  77  76  66  54  51  48  53  71  73  68
Ext           80              45              79              46              75

Dpt   45  44  46  51  51  48  44  48  44  42  46  50  51  49  47  50  50  51  52
Hum   39  31  31  53  83  89  93  69  34  29  34  56  90  93  96  89  47  46  57
Ext           24              98              29              96              46

WDr  248 240 268 283   6 145  47  72 239 267 270 279 282 288 345 280 249 265 272
WSpd   3   4   3   4   2   3   3   3   4   7   8   5   2   2   1   2   5   7   6
Max            8               4               8               2               7

GDr  227 238 239 239 313  39   6  38 244 248 247 239 269 307 302 298 244 248 243
GSpd   7  11  11  12   5   2   4   3   9  20  23  15   8   3   4   4  11  20  17
Max           20               8              23               8              20

MOS FWG EXTENDED GUIDANCE FOR Cinnamon  08/10/1200Z
Parm  13  13  13  13  13  13  13  14
     03Z 06Z 09Z 12Z 15Z 18Z 21Z 00Z
P01       15      17      16      14
P10        8      17       4       4
S01       14       9       1       0
S10       10       9       1       0

Tmp   62  54  49  46  50  61  64  64
Ext               45              67

Dpt   52  52  49  46  49  48  48  45
Hum   70  93 100 100  96  62  56  50
Ext              100              41

WDr  280 278 279 293 275 248 263 274
WSpd   5   2   2   1   2   3   6   6
Max                2               7

GDr  250 276 312 295 285 236 249 243
GSpd  13   8   3   5   5   9  15  17
Max                8              20
```

Figure 3: Example of Statistical Output. P01 and P10 represent the probability of .01 inches and .10 inches of precipitation respectively. S01 and S10 represent the probability of 1 lightning strike and 10 lightning strikes respectively. Lightning strike probability is for a 6 hour period.