3.4 A NEURAL NETWORK TO AID FORECASTERS IN IDENTIFYING SIGNIFICANT WEATHER EVENTS

Ron Holmes[1] *and Rich Grumm*
National Weather Service, State College
Dr. George Young
The Pennsylvania State University

1. INTRODUCTION

Significant weather events such as major winter storms, severe weather outbreaks, cold episodes, and heat waves, can have large impacts on the economy and the welfare of the nation. Ross and Lott (2006) reported that during the period from 1980-2005 the U.S. sustained over $500 billion in overall inflation adjusted damages/costs due to significant weather (Fig 1). The mission of the National Weather Service is to protect life and property and a large part of this endeavor is to identify when significant weather will impact the nation and its economy. Forecasters are trained to recognize weather patterns associated with significant weather and tasked to issue watches, warnings, and advisories for such weather. Artificial Neural Networks (ANN) can also be trained to recognize patterns in meteorological data and classify weather systems. The purpose of this paper is to demonstrate how an ANN was trained using climatic anomalies to assist forecasters in identifying when and what types of significant weather will affect their area of responsibility.

2. METHODS

2.1 Identification of events using NCDC Climate Data

National Climatic Data Center (NCDC) Cooperative Observations (COOP) were used to identify significant events in this study. These observations contain 24 hour maximum and minimum temperature, precipitation, snowfall, and snow depth ending at 7:00 AM local time. To help identify significant events it was important to first build a climatology for stations using the NCDC data. The standard 30-year normals computed by NCDC were not used as a basis for this study for several reasons. First, not every station has 30-year normals computed and it is conceivable that this study could be adapted for other locations. Secondly, it was determined there was large variance in the normals computed by NCDC. Therefore a 21-day centered mean and standard deviation was computed for each day of the year for each station. Hart and Grumm (2001) showed that a smoothed, 21-day mean, climatology to be more advantageous than using a monthly mean. The smoothed climatology reduces large daily temperature fluctuations due to transient synoptic scale systems and facilitates identification of truly anomalous days. This was done for a 30 year period (1970-2000) for all COOP locations across Pennsylvania and the results were stored in a MySQL relational database.

Identification of heat and cold events were determined by using a threshold of Standard Deviations (STD) above or below the mean. A heat event was defined as maximum temperature exceeding 2 STD above normal and a cold event was defined as the min temperature below 2 STD below normal. The time of year when these criteria are met is important. A +2 STD above normal day in mid-July would be categorized as oppressive and thus significant. Conversely, a +2 STD above normal day in March may be regarded as pleasant and therefore not significant. Nonetheless it was important to include events that met the criteria for all seasons because not doing so would severely limit the number of training cases available to the neural net. Additionally, Hart and Grumm (2001) showed that normalized anomalies are advantageous to use because they account for seasonal variations and can be compared without regard to time of year.

---

[1]*Corresponding author address*: Ron Holmes, National Weather Service, 328 Innovation Blvd., State College, Pa. 16803

For precipitation the mean and standard deviations were computed using only those days when precipitation greater than 0.01 inches was observed. This mitigates the effects of averaging in days with no precipitation which would significantly lower the mean. Since precipitation does not follow a normal distribution, percentile ranks were used to test for significant precipitation amounts. The 95th percentile was used to identify heavy precipitation days when it rained. This value usually was at least an inch (25.4mm) or more precipitation in 24 hours for a station. Therefore dates with an inch or more of precipitation were recorded as heavy precipitation events. Dates were also limited to the cool months (Sept – April) to focus on large scale synoptic systems.

Although one inch is on the low end of the heavy precipitation spectrum and it normally does not cause significant hydrologic problems, it was used because raising the bottom limit to 1.5 or 2.0 inches would dramatically decrease the number of events. This would reduce the training data for the algorithm leading to poor predictive results (Witten and Frank, 2005). For snowfall events a minimum of 6 inches was used to record heavy snow dates because that is the lower boundary for winter storm watches and warnings over much of Central Pennsylvania where this study was focused.

It is important for the training data set to include an equal number of null cases for each significant event type in order for the neural net to distinguish between an event and a non-event. Therefore an equal number of dates (chosen at random) were recorded for each event type to include neutral or less significant days. Null cases for heat events were defined as maximum temperatures between 0 and 1 STD above normal and for cold events defined as minimum temperatures between 0 and 1 STD below normal. For heavy precipitation events null cases were defined by dates having no precipitation. For heavy snow events amounts < 5 inches were identified as null or neutral events.

2.2 Global Reanalysis Data

The National Centers for Environmental Protection (NCEP) Global Reanalysis (GR) dataset (Kalnay et al. 1996) was used to extract upper atmospheric variables and calculate mean and standard deviation for them. This data set is a 2.5 x 2.5 km grid at 17 pressure levels covering North America from 1948 to present. It is available 4 times per day at 0000, 0600, 1200, and 1800 UTC. Similar to Hart and Grumm (2001) this study used the 21-day centered mean and standard deviation climatology for each grid point in the GR data to calculate standardized anomalies for variables such as column precipitable water, temperature, height, U-wind and V-wind component, mean sea-level pressure, and thickness.

Since this study was focused on Central Pennsylvania dates when significant weather occurred at the Cooperative Observer site in State College (STCP1) were recorded. Significant events spanning the period 1948-2004 were identified based on the above criteria because GR data was available for that period. A total of 477 heat events, 415 cold events, and 190 heavy precipitation events were identified. Because heavy snow does not occur that often, using a single station to identify heavy snow days was too limiting. Therefore it was decided to use a number of COOP stations (7) surrounding State College to identify heavy snow days. Using this method 701 dates of heavy snow were identified. For these matching dates various atmospheric anomalies were calculated from the GR data for a grid point over State College. The anomalies were used as predictors for the significant weather event being predicted. Predictors for heat events, cold events, heavy precipitation, and heavy snow were determined using a combination of experience and data mining tools.

***3.* ARTIFICIAL NEURAL NETWORK**

3.1 Overview of an Artificial Neural Network

An Artificial Neural Network (ANN) is a mathematical construct, modeled after the human brain, to perform pattern recognition and classification (Reed and Marks, 1998). Like the human brain, an ANN learns through repetition. Training data is fed into the ANN and its output is compared to actual observations or some pre-classified known data. Differences between the predicted output and the known data are calculated as error and this error is used as a corrective influence to adjust weights in the network. The weights are thought of as the long term "memory" of the ANN and are similar to the coefficients in a linear regression equation. However the ANN excels over linear regression because it can map any non-linear function.

Similar to linear regression, in which the goal is to choose the best coefficients to minimize the squared error, the goal of learning in the ANN is to keep adjusting the weights to minimize the output error. The method of adjusting the weights via this feedback mechanism is called back-propagation (Bryson and Ho, 1969) and is a form of gradient descent search in weight space. That is, each weight is defined by a point on the error surface and that surface is assumed to have a minimum value at some location. The tangent of a point on the error surface is the slope or partial derivative of the surface with respect to each weight. Adjusting the weights by an amount proportional to this slope down the error surface (gradient descent) to some global minimum causes the network to reduce its error (Russell and Norvig, 1995). There is a chance however that the network may converge on some local minima and not reach the global minima. This possibility can be reduced by including some type of search algorithm such as simulated annealing or genetic search to find the global minima (Reed and Marks, 1998).

3.2 Construct of an Artificial Neural Network

The construction of the ANN for this study followed the standard construct presented in many artificial intelligence books. The Java programming language was used to build a feed-forward back propagation ANN with a set of input nodes, one hidden layer consisting of varying nodes, and an output layer. Reed and Marks (1998) indicate that only one hidden layer is needed to represent a continuous function mapping from *N* inputs to *M* outputs, therefore higher order constructs consisting or 2 or more hidden layers were not implemented. The activation function of each of the nodes was the sigmoid function:

$$f(u) = 1/(1+e^{-u}).$$

which was used because the classes being predicted were either 1, the event occurred, or 0, the event did not occur. The sigmoid function in effect "squashes" large values of input to between 0 and 1. It is also differentiable which is needed by the back propagation algorithm. The number of nodes to use in the hidden layer was determined via methods described by Heaton (2005). This method resulted in a network that was robust and able to generalize well without the problem of overfitting. Generally 3 to 6 hidden layer nodes were used depending on the significant weather type for which the ANN was being trained. Likewise learning rate and momentum terms were adjusted to produce a minimum of error on the training data set. Most commercial AI applications have automated methods to fine tune the network such as weight decay for efficient pruning and optimization. Future enhancements to the ANN will include these more advanced features.

3.3 *Training the Artificial Neural Net*

Training data was built using anomaly values calculated from the GR data for the dates on which a significant event occurred or was neutral. This data set was made for a single point, State College, but the same process could be adapted for any location across Pennsylvania. Each significant event type had its own training data set which also included a binary value that indicated whether the event was significant (1) or was a neutral case (0). The number of inputs used to classify each event type was dependant on the significant weather type. Anomaly fields used as inputs to the ANN were chosen based on meteorological experience. For example, it is well known that high precipitable water values are an important ingredient to heavy precipitation (Junker and Schnieder 1997 and Hart and Grumm 2000) and that low-level temperature and precipitable water anomalies are important during heat waves (Lipton et al. 2005).

A data mining tool was also used to help identify key predictors. The Waikato Environment for Knowledge Analysis

(WEKA) workbench (Witten and Frank, 2005) was used to build decision trees that helped identify key predictors for each significant weather type. For example, one decision tree indicated that low level U and V wind anomalies are the most influencing factor in distinguishing between Synoptic and Frontal Maddox type heavy precipitation signatures (Maddox, 1979). This example shows the power of leveraging new data mining applications with meteorological data sets to extract emerging patterns and salient features hidden in the data.

Output of the ANN was used to classify a significant event type and consisted of 2 or 3 nodes depending on the event type to be classified. For example, Maddox (1979) identified several scenarios for heavy precipitation and categorized them as Synoptic, Frontal, or MesoHigh type events. The heavy precipitation events identified in this study were limited to the cool months so the ANN was trained to distinguish between Synoptic or Frontal Maddox type heavy rain events and days with light rain (neutral events). This required 3 output nodes. Likewise, 3 output nodes were used to classify hot, cold, or normal days.

All training and testing was performed via a process called stratified 10-fold cross validation. Witten and Frank (2005) outline this procedure in which data for each class is collected in equal amounts and randomized so that each class is properly represented in both training and test sets. The data set is split into 10 equal parts (stratification). Some of the data (10%) is held out as test data while the neural net is trained on the remaining 90% of data. Then, the process is repeated 10 times using successive partitions of the data set (cross-validation). With this method all data is eventually used for both training and testing but in any one fold of a *k*-fold training session the training and testing data are held separate. This is important because a classifier must not use test data that was used during the training process as a measure of future performance.

After the number of input nodes was determined training began by initializing the weights of the nodes to random values between -0.5 and +0.5. Then sample to sample, or on-line, training was performed with each training instance fed forward through the network. The output prediction was compared to the actual class type (1 for a significant event or 0 for a neutral event) which resulted in some amount of error. This error was then back-propagated through the network to adjust the weights. This process continued for each training instance and was repeated for 10,000 epochs (cycles) for each partition of the 10-fold cross validation process.

4. RESULTS

Separate ANN's were built and trained according to the steps outlined above to classify different significant event types for a single point over State College. Once an acceptable level of error was reached the weights were saved to a file. The ANN was taken out of training mode and then employed in an operational setting by using the weights calculated during training. The North American Model (NAM) was used as input to the ANN. Various upper air fields from the NAM Grib files were extracted for a grid point over State College and converted to standardized anomalies using the NCEP GR climatology previously built by Hart and Grumm (2001). The same anomaly fields that the ANN was trained on to classify an event were used as input from the operational NAM model. Classification output from the ANN was then converted to a graphic using GrADS (Doty et al. 1995) for easy on-line interpretation every model run (6 hours).

4.1.a Case example - *Cold Event Sept 21, 2006*

The first cold event to hit Central Pennsylvania occurred during the early morning hours of Sept 21, 2006. This was not an extremely cold event (minimum temperatures were in the mid 30s to low 40s over Central Pa) and most likely fell on the low end of the spectrum that the ANN was trained to recognize. However it showed some signal going into the event. The predicted temperature class from the ANN using the 12 UTC 20 Sept NAM anomaly forecasts for a grid point over State College is depicted in Fig 2. It shows the ANN predicted a cold event from 0000 UTC to 0900 UTC 21 Sept. Global reanalysis data for 1200 UTC 21 Sept showed a strong surface high pressure over the region with -1 to -2 standard deviations (STD) below normal 500 mb height and 925 mb temperature anomalies (Fig 3). Additionally the 925 mb V-wind anomaly showed northerly winds over the area.

4.1.b Case example - *Cold Event Oct 12th to 15th, 2006*

Another weak cold event settled over Pennsylvania from 12 Oct to 15 Oct as a deep upper level trough strengthened over the Great Lakes on 00 UTC 13 Oct (Fig 4). By 0000 UTC 14 Oct the trough had expanded over the eastern seaboard and mean sea level pressure anomalies, 500 mb height anomalies, and 925 mb temperature anomalies were -1 to -2 STD below normal (Fig 5). Output from the ANN using NAM anomaly forecasts from the 1800 UTC 10 Oct run indicated initially a warm event (which was currently underway on 10 Oct) would switch to neutral and then enter a cold period beginning 1800 UTC 12 Oct (Fig 6). The ANN output from the 0000 UTC 12 Oct NAM run indicated this event would end around 0000 UTC 15 Oct (Fig 7) and the GR data did show a weakening of the upper level cold anomalies by this time (Fig 8). The COOP temperature graph for State College (STCP1) during this period (Fig 9) showed afternoon highs and morning lows around -1 STD below normal during this period. Again, this is on the low end of the spectrum that the ANN was trained to recognize however this also indicates it is robust enough to distinguish a marginal cold event and a neutral event.

4.2 Case example - *Heat Event Oct 4th, 2006*

A short-lived warm event was forecast over Pennsylvania for 4 Oct, 2006 as warm air flowed northeast ahead of a strong short wave dropping southeast across the Great Lakes. The ANN output using the 1800 UTC 30 Sept NAM run indicated a warm event during the day beginning around 0900 UTC 4 Oct to 06 UTC 5 Oct (Fig 10). The morning low at State College on 4 Oct was 53F which was +1 STD above normal and the high temperature reached 76F during the afternoon which was around +1.5 STD above normal (see Fig 9).

4.3 *Case example - Heavy Precipitation Oct 17th and Oct 19th, 2006*

Two episodes of heavy rain occurred over Central Pennsylvania in mid Oct, 2006. Heavy rain fell from 17 Oct to 18 Oct affecting much of the state with the heaviest rain across the western third of PA including the State College area (Fig 11). After a brief respite the heavy rain picked up again on 19 Oct with the heaviest rain falling across northern and central PA. The 24 hour total at State College from 1200 UTC 19 Oct to 1200 UTC 20 Oct was 2.31 inches (Fig 12). Figure 13 shows the ANN output from the 0600 UTC 16 Oct Nam run. It shows the 2 precipitation events separated by a brief period of no precipitation. The GR analysis (Fig 14) shows a +3 STD above normal 850 mb V-wind anomaly poking into central PA with an axis of +2 STD above normal Precipitable water anomaly along the spine of the Appalachians.

The 36 hour NAM forecast valid 1200 UTC 19 Oct showed a flattening ridge across the Ohio valley with a weak area of low pressure extending north from the Tennessee valley and along the Appalachian mountains (Fig 15). At this time a narrow tongue of above normal precipitable water extended into the Ohio valley. By 0000 UTC 20 Oct precipitable water values had increased dramatically as a strong southwest flow at low levels increased. Meanwhile surface pressure continued to deepen in response to the upper level trough which sharpened as it moved across the Ohio valley (Fig 16). Heavy rain spread north along the inverted trough and surface frontal boundary as it moved into western PA overnight. Precipitable water anomalies reached +2 to +3 STD above normal over central and eastern PA while a nose of high 850 mb V-wind anomaly of +3 STD above normal moved into southern PA (Fig 17). This is a classic Synoptic type heavy precipitation pattern as determined by Maddox (1979). An attempt was made to train the ANN to distinguish between Synoptic and Frontal Maddox type heavy precipitation patterns. Figure 18 shows the output of this ANN from the 0600 UTC 17 Oct, 2006 NAM run which indicated both heavy precipitation events would be comprised of the Synoptic type pattern.

5. CONCLUSIONS

An Artificial Neural Network was trained to recognize significant weather events using GR data and COOP observations. The ANN was put into service during the fall season, 2006 and initial results are encouraging. The ANN has shown the ability to identify weak heat and cold events so it is reasonable to expect it will also be able to recognize stronger significant events. It also appears to recognize days when it will rain vs. dry days and shows promise at distinguishing between Synoptic and Frontal heavy rain signatures. As of fall 2006, it has only been trained for temperature, precipitation, and snow events. It has yet to be tested on a real-time heavy snow event. Further work will focus on recognizing severe weather potential. We hope this kind of tool can be helpful to operational forecasters in the early detection of significant weather events.

6. REFERENCES

Bryson, A. E. and Ho, Y. C. (1969). Applied optimal Control, Blaisdell, New York

Doty, B.E. and J.L. Kinter III, 1995: Geophysical Data Analysis and Visualization using GrADS. Visualization Techniques in Space and Atmospheric Sciences, eds. E.P. Szuszczewicz and J.H. Bredekamp. (NASA, Washington, D.C.), 209-219.

Lipton, K., R. Grumm, R. Holmes, P.Knight, and J.R. Ross, 2005: Forecasting Heat waves using climatic anomalies. *Pre-prints 21st Conference on Wea. and Fore. and the 17th Conference on Numerical Weather Prediction*, AMS, Washington, DC.

Hart R., and R.H. Grumm, 2000: Anticipating heavy rainfall events: Forecast aspects. Preprints, Symposium on Precipitation Extremes: Prediction, Impacts, and Responses, Alburquerque, New Mexico, Amer. Meteor. Soc. 271-275.

Hart, R. and R. H. Grumm, 2001: Using Normalized Climatological Anomalies to Rank Synoptic Scale Events Objectively. *Monthly Weather Review*., **129**, 2426-2442.

Heaton, Jeff. “Introduction to Neural Networks with Java”, First Edition, Heaton Research, Chesterfield, MO, 63017-4976.

Junker, N.W. and R.S Schnieder, 1997: Two case studies of quasi-stationary convection during the 1993 great midwest flood. *National Weather Association Digest*., **21**,5-13.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc*., **77**, 437-471.

Lott, Neal and Ross, Tom (2006) Tracking and Evaluating U.S. Billion Dollar Weather Disasters, 1980-2005. AMS Conference Pre-print January, 2006

Maddox,R.A., C.F Chappell, and L.R. Hoxit. 1979: Synoptic and meso-alpha aspects of flash flood events. *Bull. Amer. Meteor. Soc*.,**60,**115-123.

Reed, Russell, D. and Marks, Robert, J. (1998) “Neural Smithing – Supervised Learning in Feed Forward Artificial Neural Networks”, MIT Press, Cambridge, Massachusetts.

Russell, Stuart and Norvig, Peter (1995) “Artificial Intelligence – A Modern Approach”, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 07458

Witten, Ian, H. and Frank, Eibe (2005) "Data Mining: Practical machine learning tools and techniques", 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

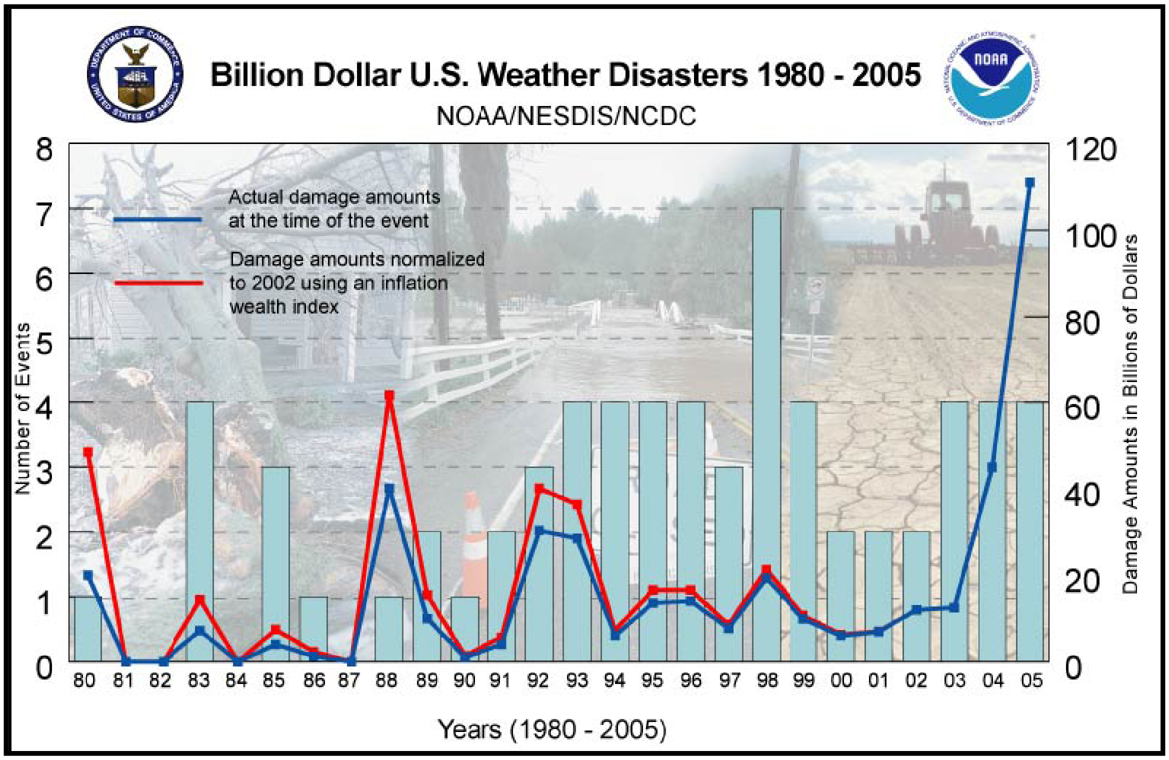

Figure 1. Number of significant weather events and damage in billions of dollars from 1980 to 2005. Used by permission, courtesy Neal Lott and Tom Ross.

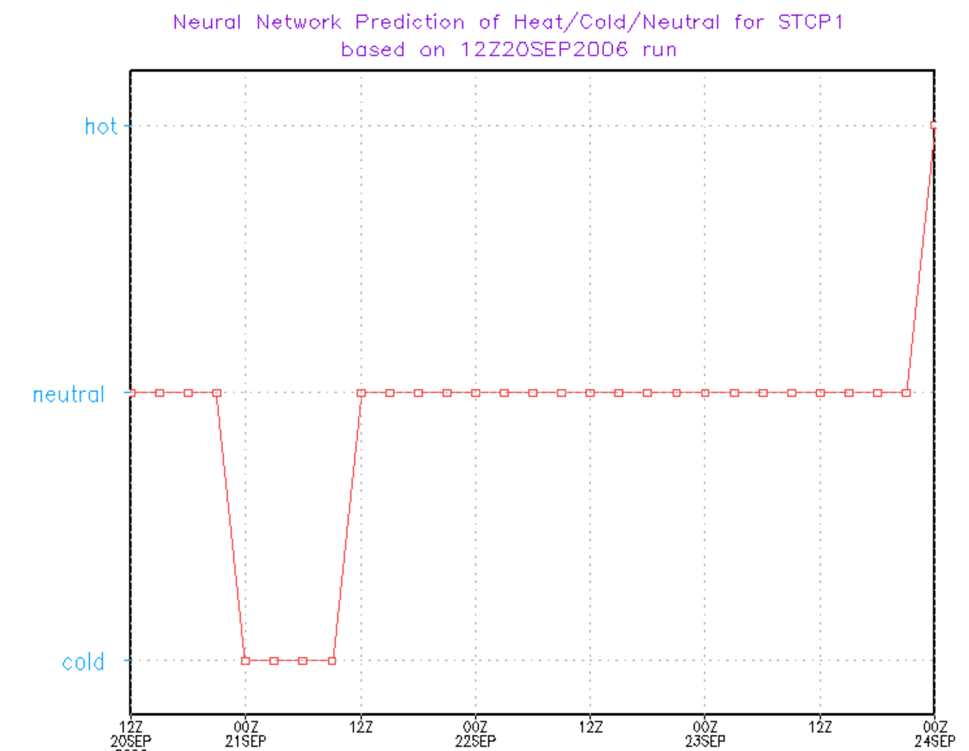

Fig 2. Output of Neural Network showing cold event from 00 UTC to 09 UTC Sept 21, 2006 using NAM anomaly forecasts based on 12 UTC Sept 20, 2006 run.

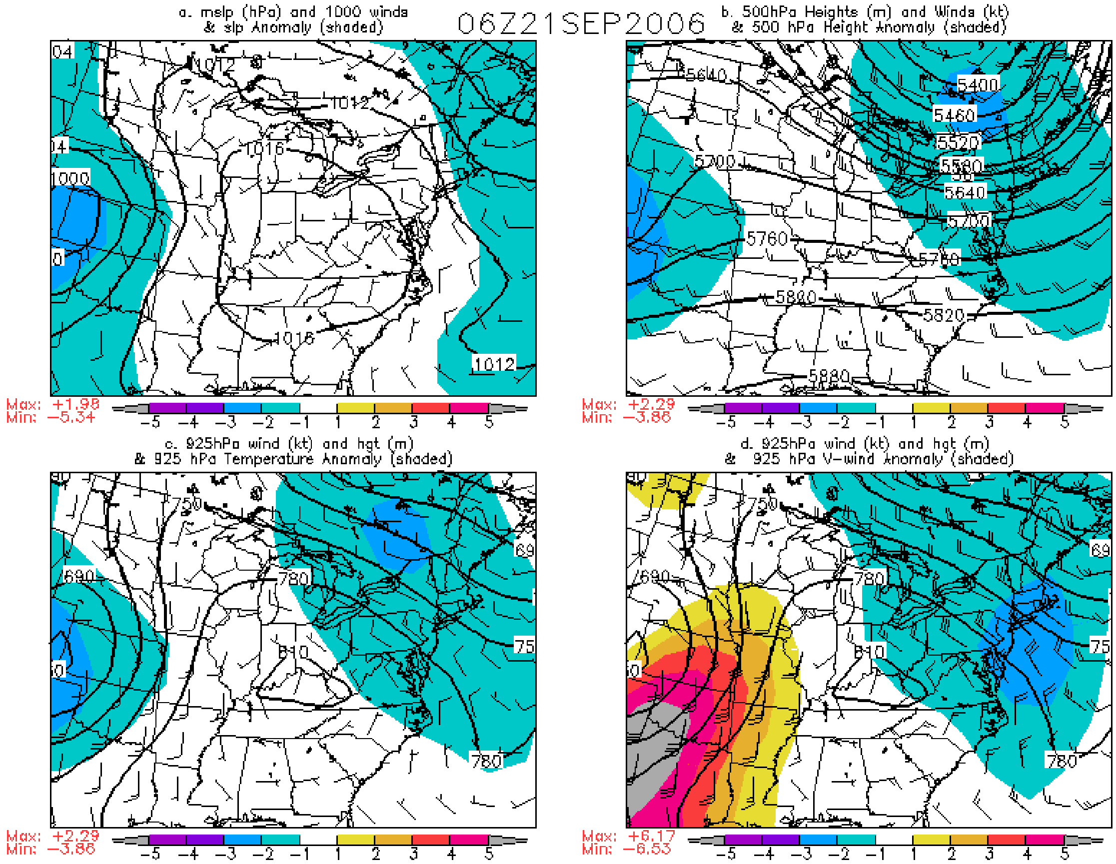

Fig 3. Global reanalysis of a) Mean Sea Level Pressure (MSLP), 1000 mb Winds, and MSLP Anomaly, b) 500 mb Heights, Winds, and Height Anomaly, c) 925 mb Heights, Winds, and Temperature Anomaly, and d) 925 mb Heights, Winds, and V-wind Anomaly for 06 UTC Sept 21, 2006.

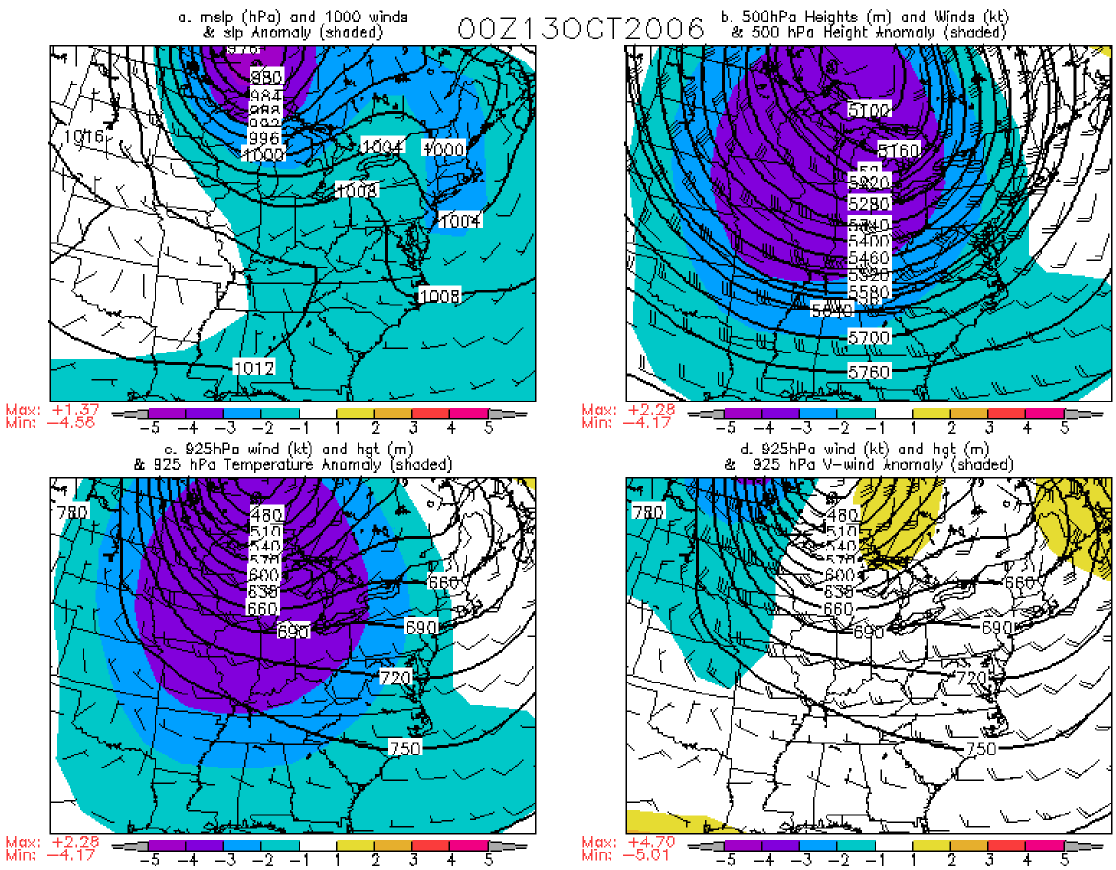

Fig 4. As in Fig 3 except for 00 UTC Oct 13, 2006

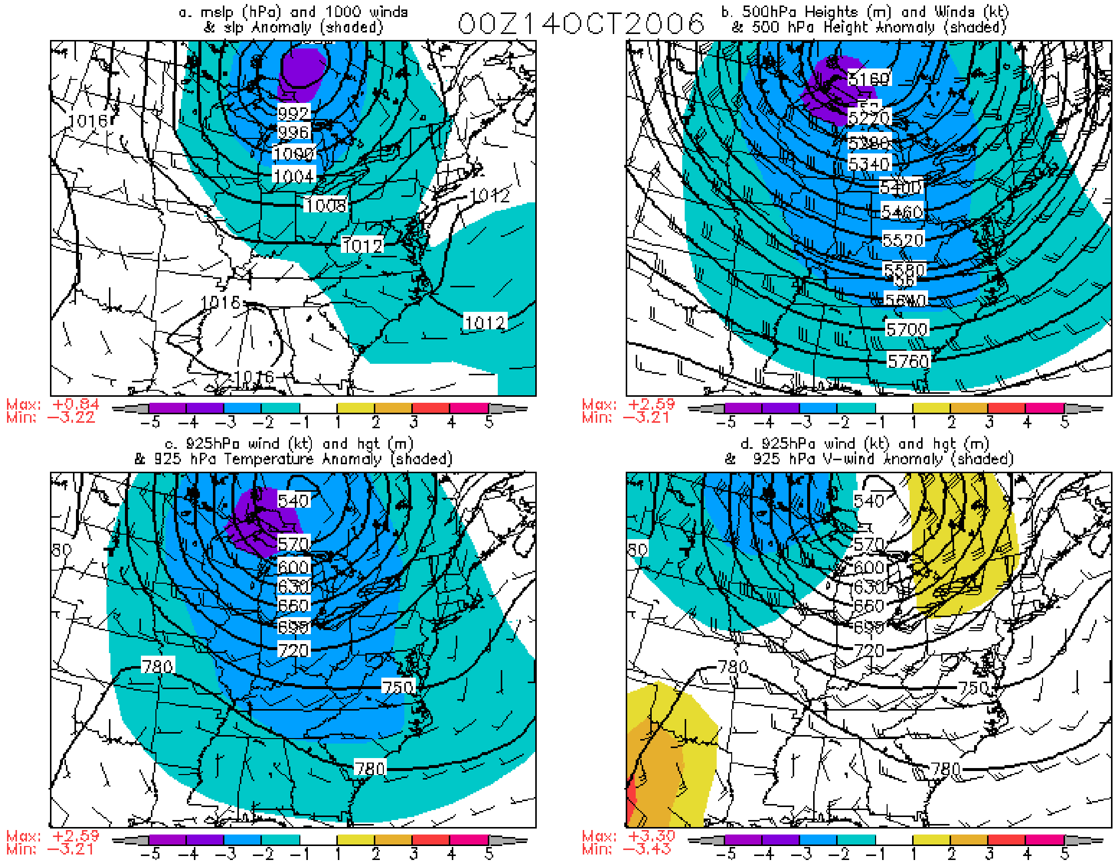

Fig 5 As in Fig 3 except for 00 UTC Oct 14, 2006

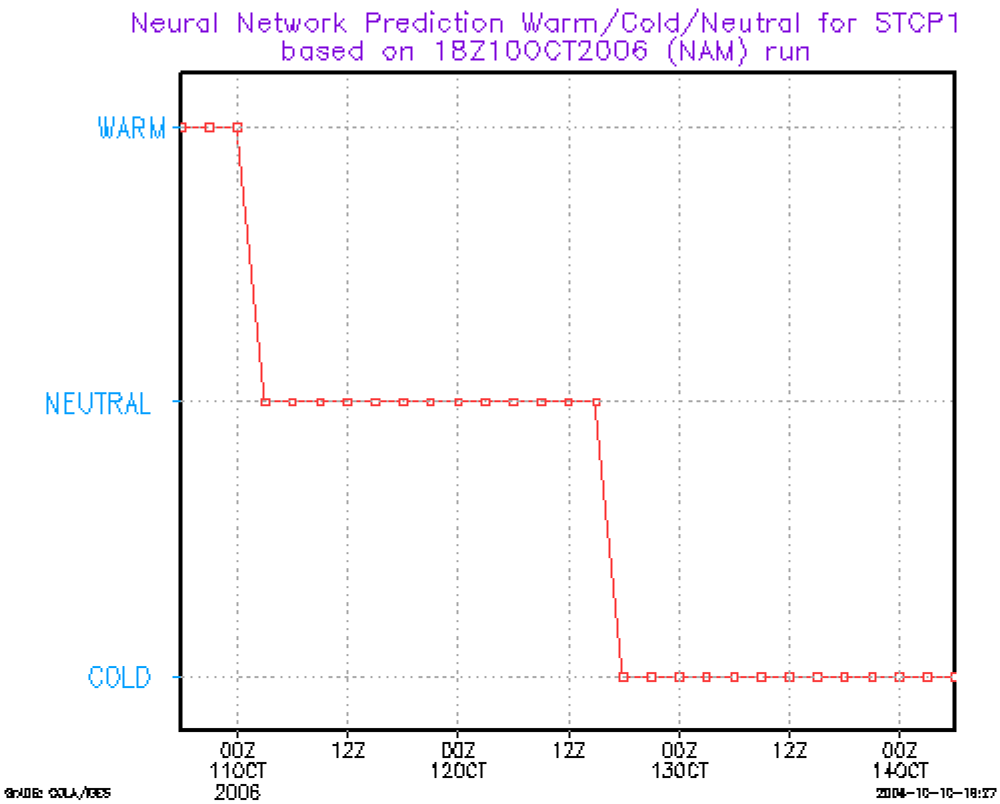

Fig 6. As in Fig 2 except for 18 UTC Oct 10, 2006 NAM run.

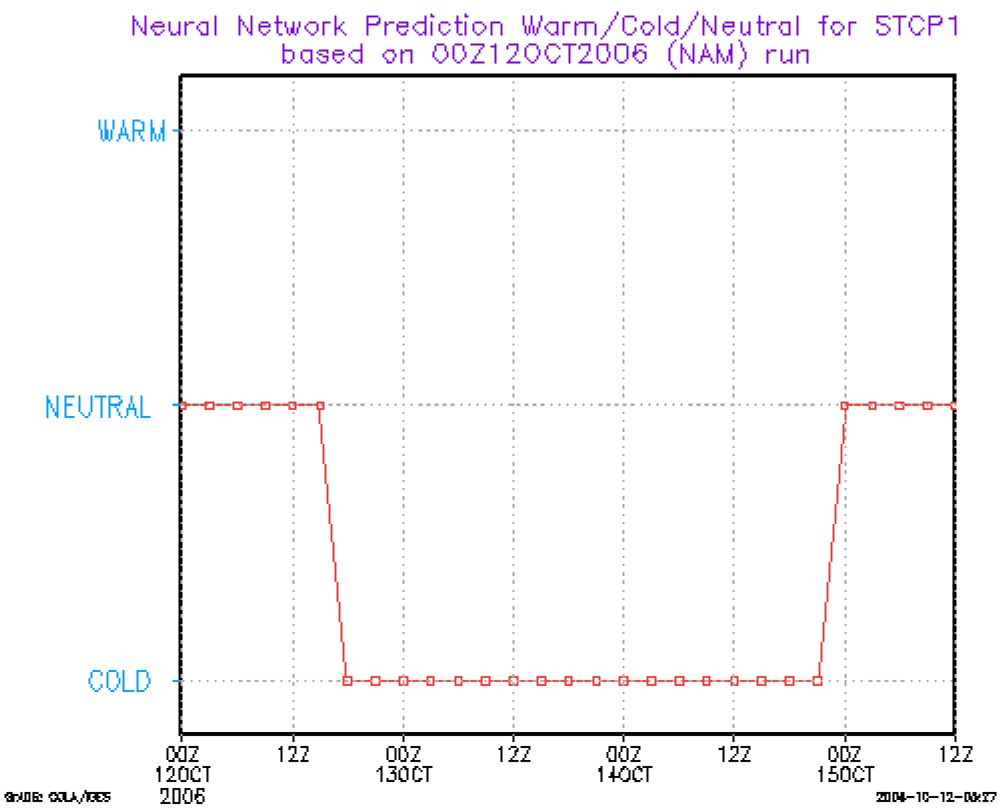

Fig 7. As in Fig 2 except for 00 UTC 12 Oct, 2006 NAM run.

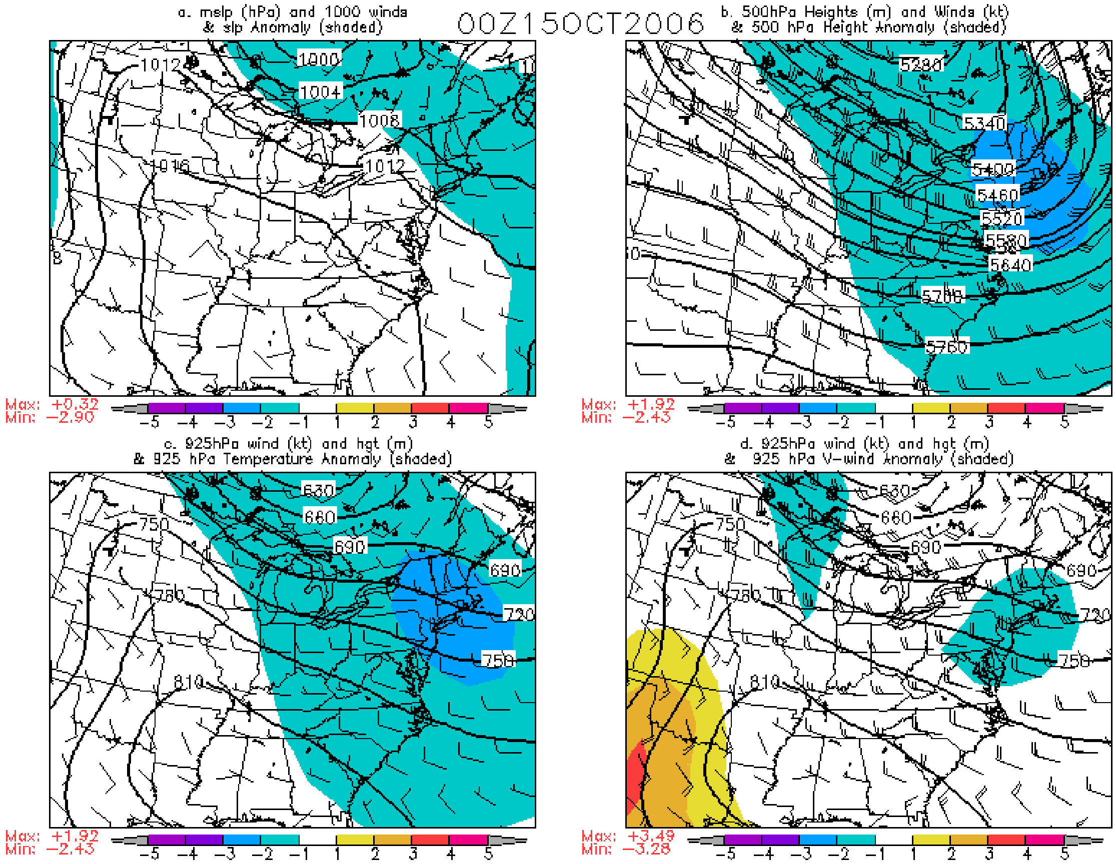

Fig 8. As in Fig 3 except for 00 UTC 15 Oct, 2006.

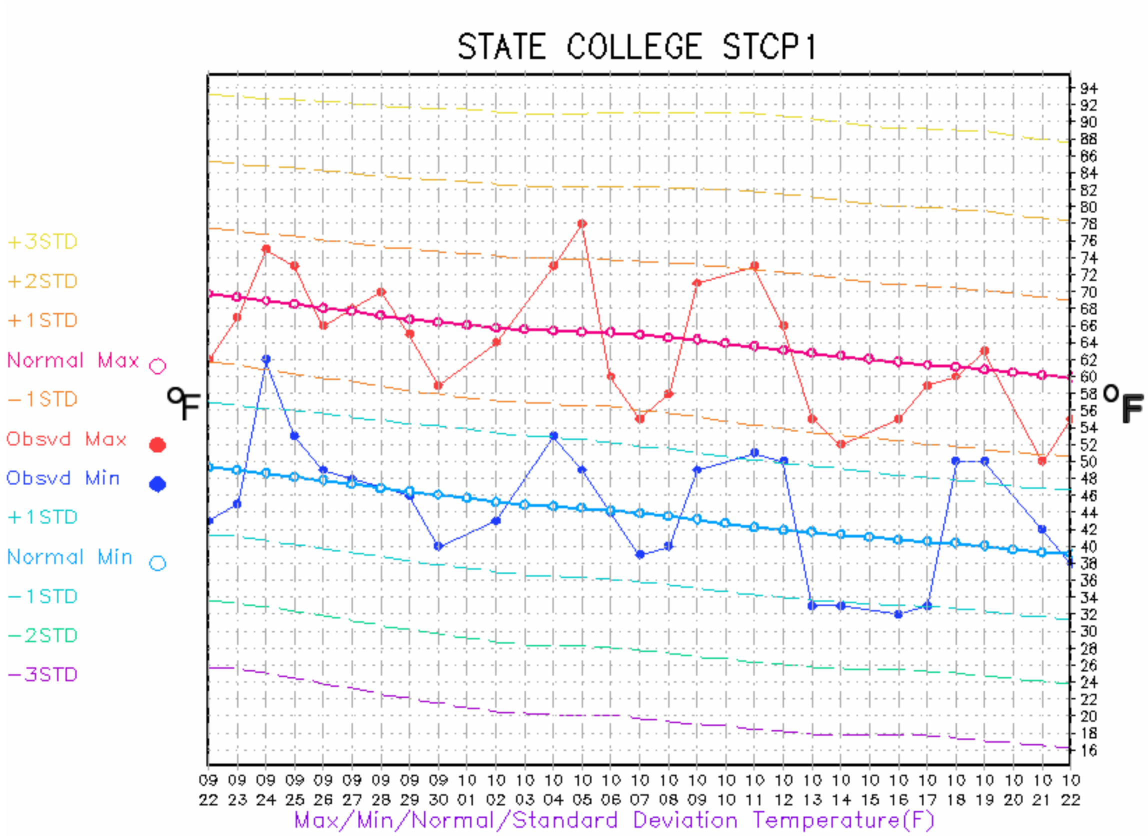

Fig 9. Observed max and min temperature at State College, Pa (STCP1) showing a cold period Oct 13th to 17th with below normal high temperatures and morning low temperatures around -1 standard deviation below normal. Note: temperature based on 24 hour measurement from 700 am local time.

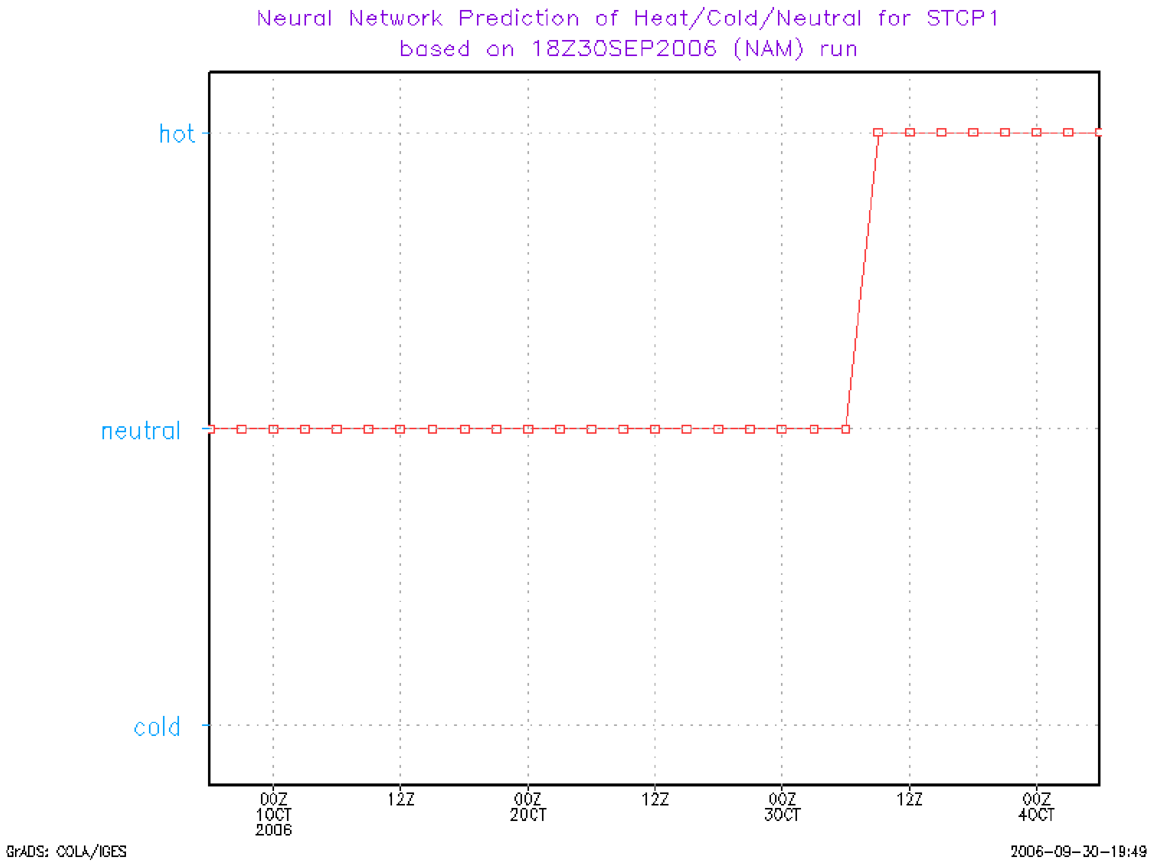

Fig 10. As in Fig 2 except 18 UTC 30 Sept, 2006 NAM run.

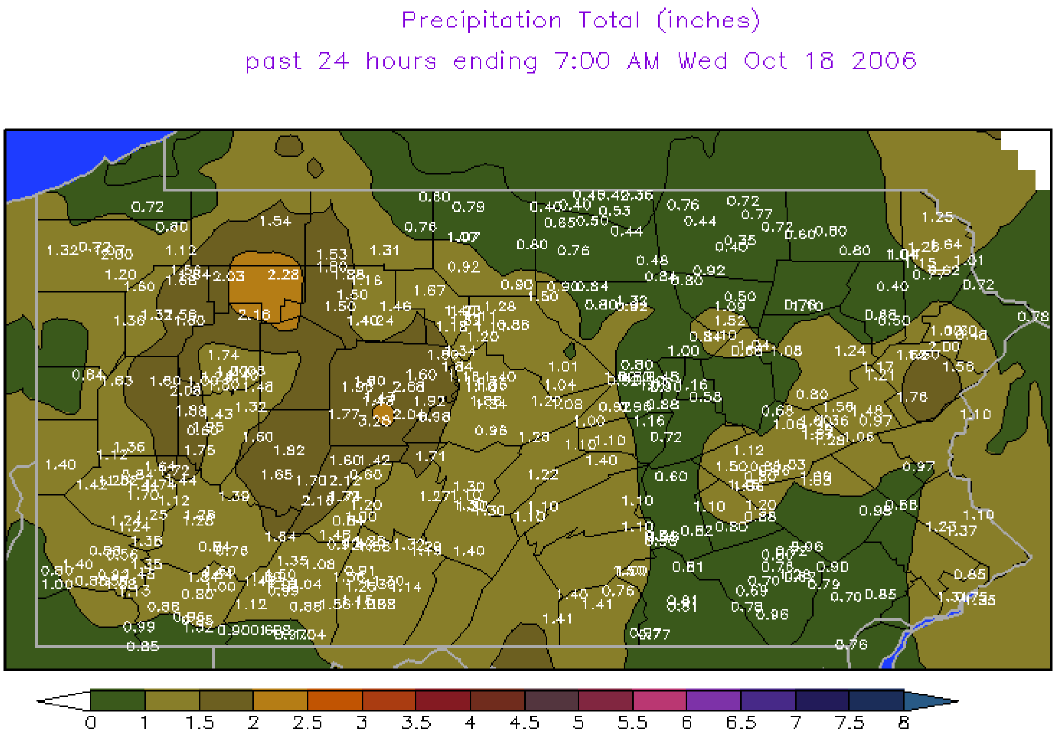

Fig 11. Cooperative Observer 24 hour total rainfall (inches) from 12 UTC 17 Oct to 12 UTC 18 Oct, 2006.

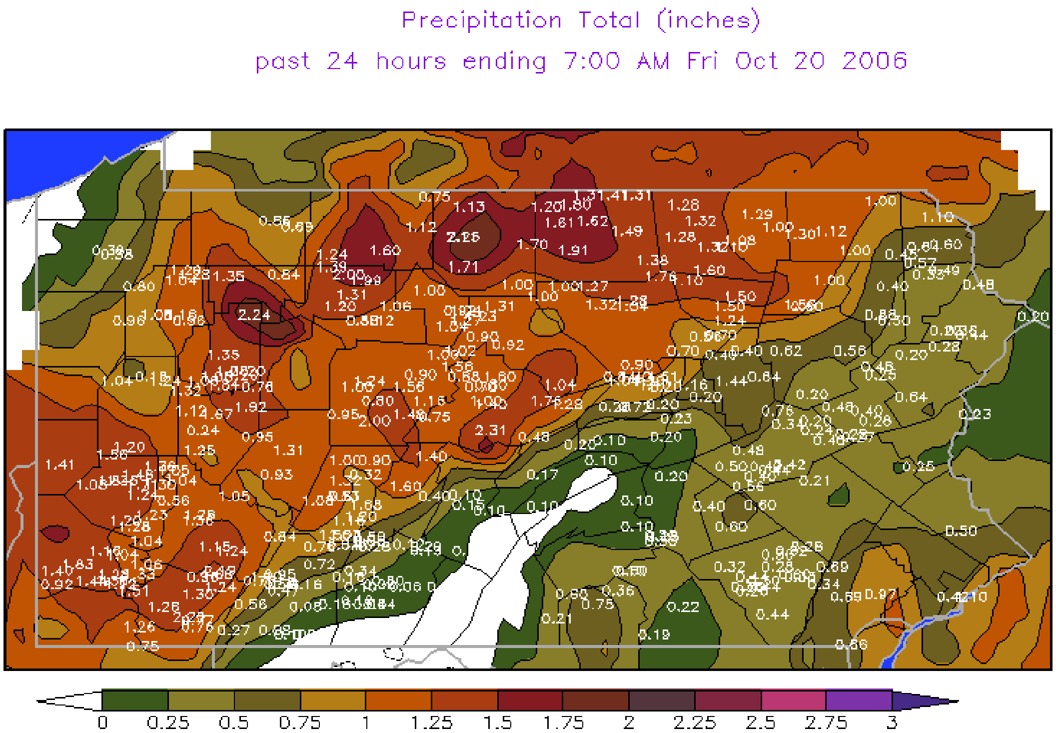

Fig 12. As in Fig 11 except for from 12 UTC 19 Oct to 12 UTC 20 Oct, 2006.

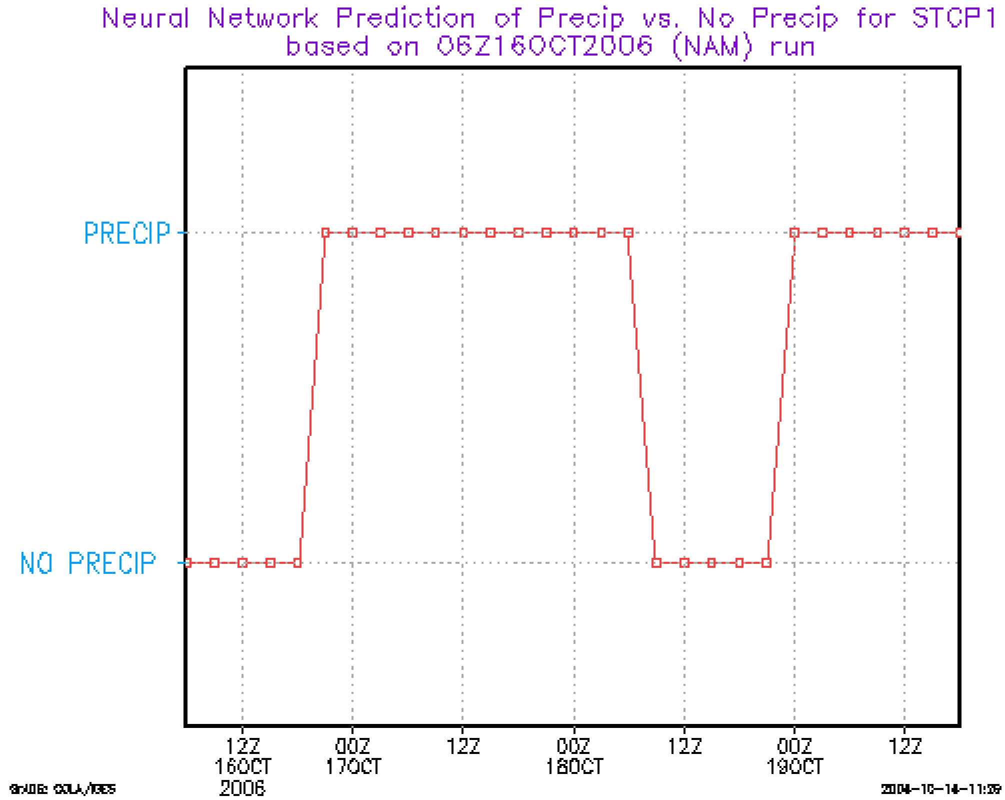

Fig 13. Output of Neural Network from 06 UTC 16 Oct, 2006 NAM run indicating 2 precipitation events separated by a period of no precipitation.

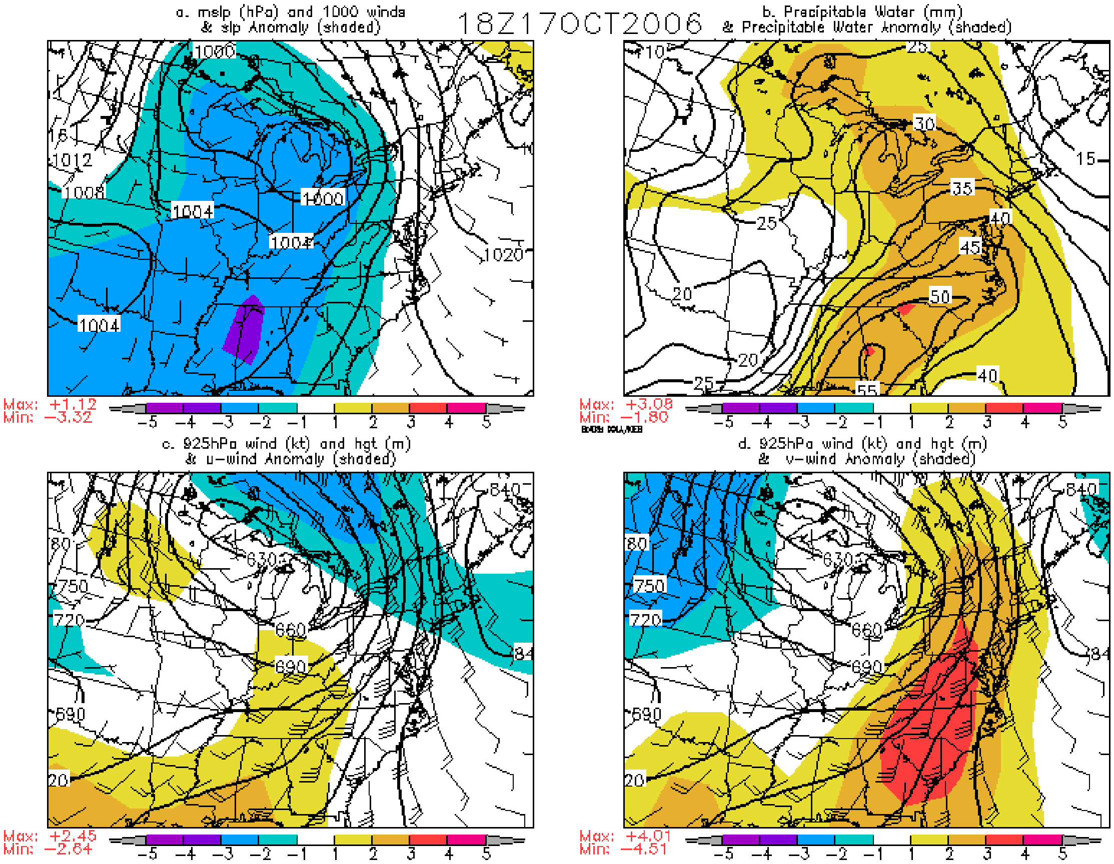

Fig 14. Global reanalysis of a) Mean Sea Level Pressure (MSLP), 1000 mb Winds, and MSLP Anomaly, b) Precipitable Water and Anomaly, c) 925 mb Heights, Winds, and U-wind Anomaly, and d) 925 mb Heights, Winds, and V-wind Anomaly for 18 UTC Oct 17, 2006.

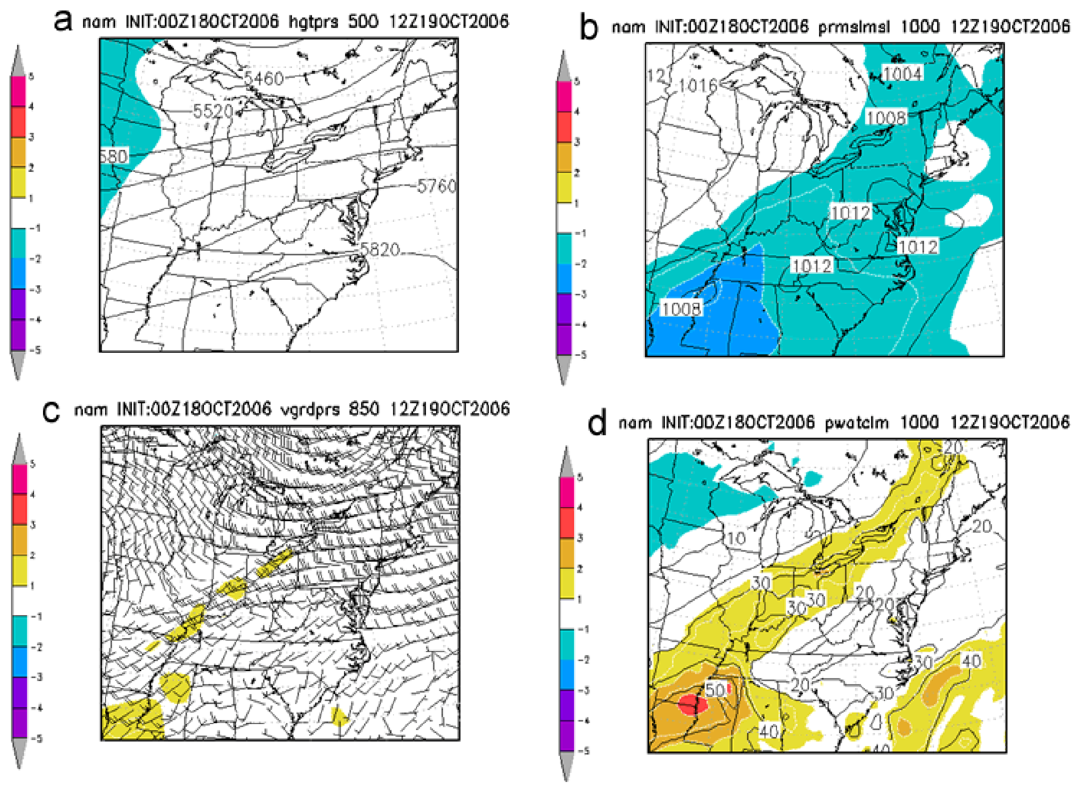

Fig 15. NAM 36 hour forecast for a) 500 mb Height and Anomaly, b) Mean Sea Level Pressure and Anomaly, c) 850 mb Wind and V-wind Anomaly, and d) Precipitable Water and Anomaly valid 12 UTC 19 Oct, 2006.

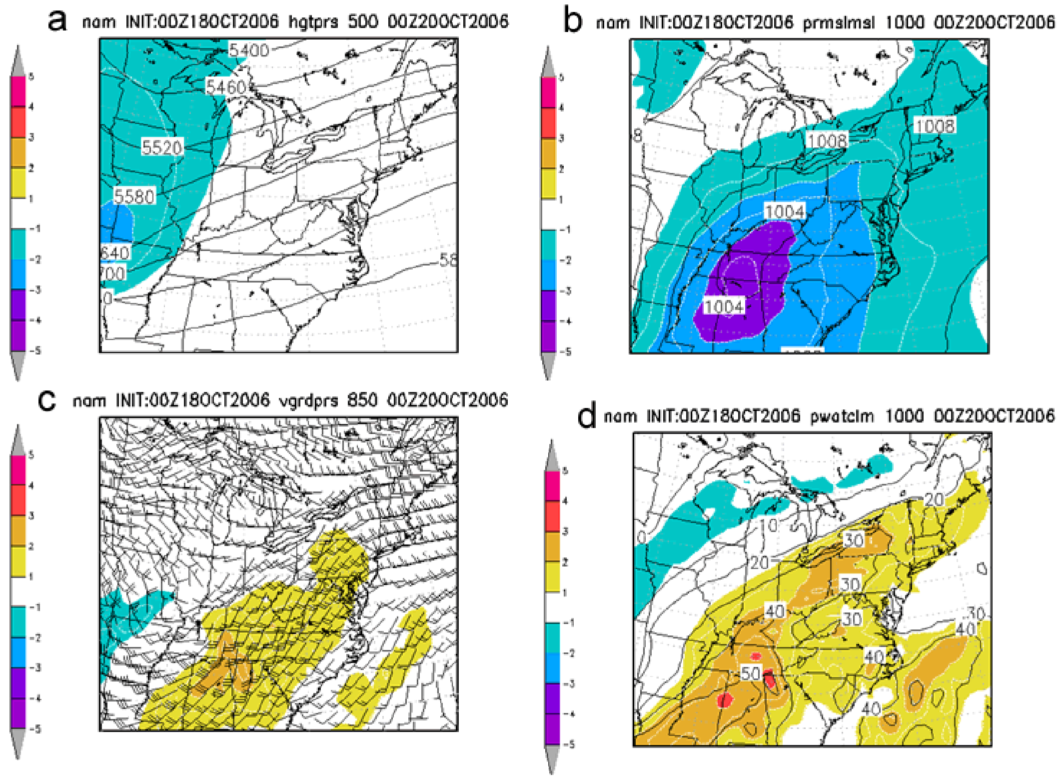

Fig 16. As in Fig 15 except 48 hour forecast valid 00 UTC 20 Oct , 2006

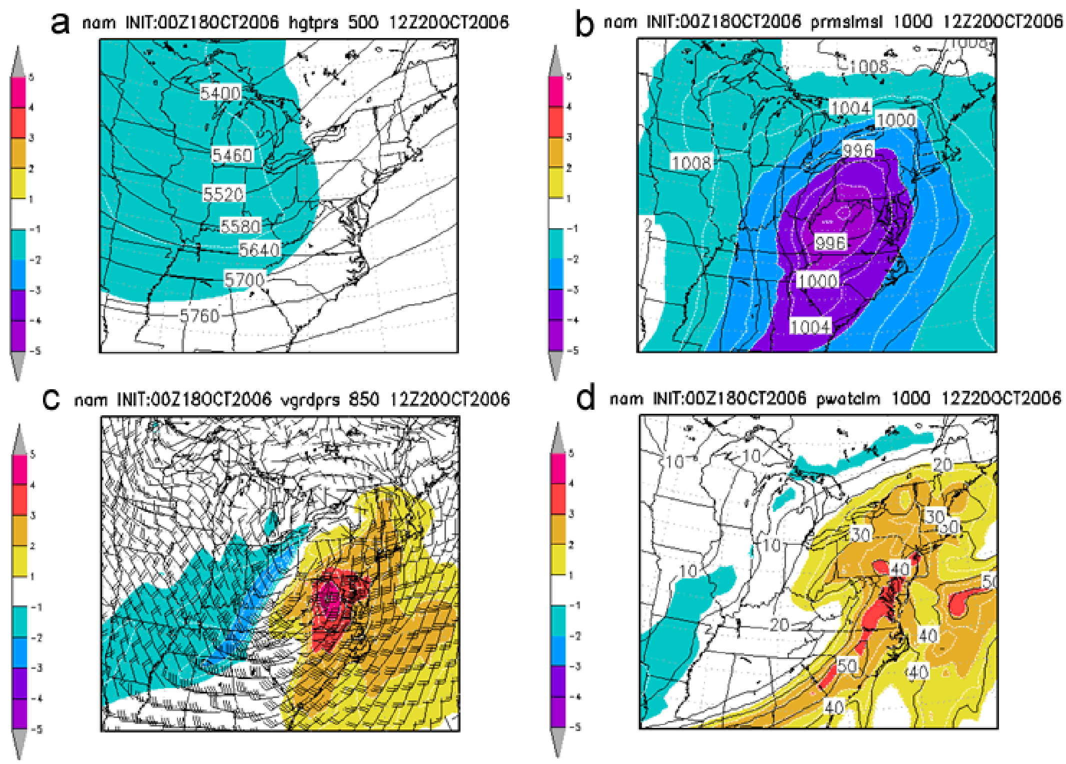

Fig 17. As in Fig 15 except 60 hour forecast valid 12 UTC 20 Oct, 2006

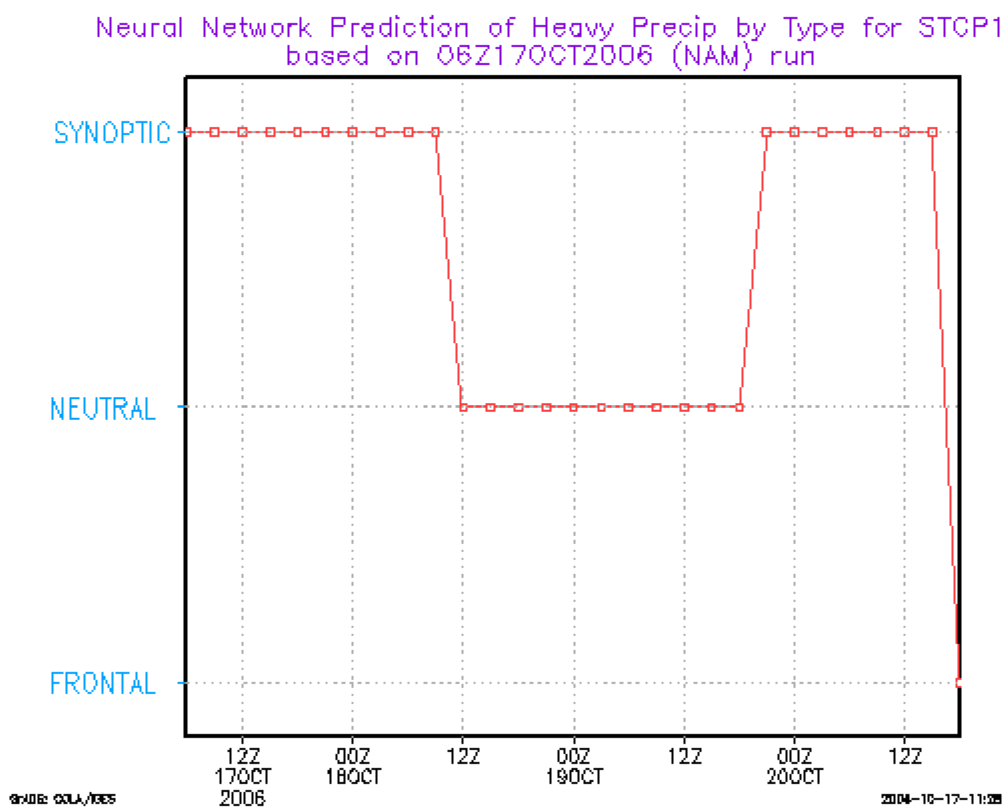

Fig 18. Output from Neural Network from 06 UTC 17 Oct, 2006 NAM run showing 2 heavy precipitation events for Oct 17-18 and Oct19-20 with a dry period in between. Heavy precipitation signature categorized as Synoptic or Frontal Maddox type.