Reprinted from the 12th Conf. on Weather Analysis
and Forecasting, October 2-6, 1989, AMS, Monterey, CA

J3.2

MESOSCALE THUNDERSTORM FORECASTING USING AN EXPERT SYSTEM

by

John F. Weaver and Roger S. Phillips

NOAA/NESDIS/RAMM Branch
Colorado State University
Ft. Collins, CO 80523

1. Introduction

Most meteorological forecasts begin with tedious, repetitious procedures for data assimilation, synthesis and analysis. The analysis phase is usually followed by the application of one or more procedural checklists (either objective or subjective) to assure some degree of forecast consistency. These two steps in the forecast routine, while often not requiring a particularly high degree of skill, can involve a great deal of time and patience. The consequence is that the final phase -- the proficient human interpretation of results -- is often relegated to a relatively few moments of attention just prior to deadline.

In recent years, more and more meteorological data (in the form of new sensor types, increased spatial and temporal resolution of observations, etc.) have been introduced. While one might assume that such additions would lead to better forecasts, this has not always been the case. The problem is that the efficient use of these data can be severely limited by the increased information handling requirements. To solve this problem, forecasters have turned increasingly to the computer for more efficient data acquisition, assimilation and presentation. Automated systems have begun to relieve the forecaster of many routine duties necessary to produce a good weather forecast.

Even when the task of data handling is optimized, the forecaster is still faced with the need to prepare an accurate forecast within a limited time-window. However, each forecast problem often requires specialized knowledge, and even though most meteorologists entering today's work force are college trained, the knowledge gained in class must necessarily be somewhat generalized. In an effort to help solve this portion of the problem, research is becoming more directed toward the merits of expert system technology.

In a meteorological application, an expert system consists of three parts: 1) a suite of rules, provided by an "expert" in the problem of interest, 2) a so-called "inference engine" which is capable of moving logically through the rules to reach conclusion, and 3) data, from which the forecast decisions can be made. A properly constructed expert system can be used to monitor the individual tasks within a procedure, from extracting the applicable data sets, to the issuing of one or more forecast "conclusions." The forecaster remains the focal point of the system, answering specific questions as they are presented, and eventually formulating one or more decisions or forecasts based upon the results and their associated confidence parameters.

This paper addresses an expert system developed to assist with convective weather forecasts. The prototype version of this system was described in Weaver and Phillips (1987). The current paper picks up from where its antecedent left off. It recaps the basic system, details capabilities installed since 1987, and describes the authors' continuing education in applying an expert system to meteorological forecasting.

2. System Description

Our system, named CONVEX, was developed within an expert system "shell" called EXSYS from a company by the same name in Albuquerque, New Mexico. Conceptually, a shell is what remains when particular, domain-specific rules are removed from a working knowledge-based system (KBS). Approximately 5,000 rules may be incorporated into EXSYS utilizing a PC operating with 640 K bytes of memory (a PC/AT was used with CONVEX). Rules are expressed in an "if ... then ... " format with "else" optional, and confidence factors may be associated with the "conclusions" to quantify uncertainty. EXSYS allows for an exchange of information between the KBS application and external software.

CONVEX is designed to operate on two spatial scales, both of which are meso-beta. The first is what we call the "regional" scale, and incorporates the area within the heavy black outlined area in figure 1 (top). The second is a so-called "sub-regional" scale. This scale subdivides the primary region

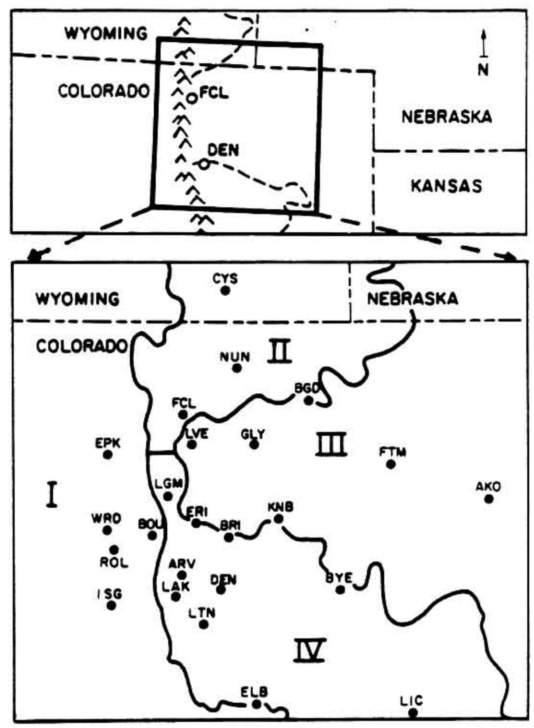

FIGURE 1. Map illustrating forecast area discussed in text. Top part of figure shows regional forecast area. Long/short dashed lines are state borders. Carets show approximate location of the so-called 'Front Range' of the Rockies. Dashed lines are the Cheyenne Ridge and Palmer Lake Divide.
Bottom part of figure is a blow-up of the regional area, showing various sub-regions, as well as locations of PROFS mesonet sites as discussed in text.

into four, approximately equal, areas based on terrain characteristics (figure 1, bottom). The forecast problem for both spatial domains is the same -- to forecast the timing and intensity of convective weather within a three to eight hour forecast time frame. It utilizes weather data available by 1000 Local Daylight Time (LDT) that same day.

3. Regional Application

The primary data source for both applications is the local (in this case Denver, DEN) radiosonde observation. The system relies heavily upon an external program for performing a relatively complex analysis of the sounding data. On the first pass, the expert system simply obtains the boundary layer mixed dew point from the current sounding data utilizing the sounding analysis program. This value is then presented on-screen as a potential forecast afternoon dew point. The forecaster is given the options of 1) accepting this value, 2) providing a subjectively preferable value, or 3) branching to a decision tree which helps derive a forecast dew point based upon an analysis of additional synoptic information independent of that available to the host computer.

After soliciting an afternoon forecast temperature from the user, that value and the dew point are passed to the sounding analysis module, which calculates such parameters as expected afternoon positive buoyancy, convective temperature, approximate maximum hail size, and mean cloud layer winds. In the event that the lower layers of the sounding are found to be too dry to allow deep convection, tests are made for potential mid-level moisture and buoyancy. On the high plains, such conditions may lead to the onset of severe microbursts. The sounding analysis module is designed to estimate an approximate maximum microburst wind gust speed when appropriate.

Once calculated, the convective parameters are passed back to the CONVEX knowledge-base for further analysis. If the forecast buoyancy value is large enough to allow convection, but not large enough to assure severe thunderstorms, CONVEX will ask a series of questions related to available synoptic support and/or suppression. The answers to these questions are factored into the final decision regarding severity. CONVEX also compares the forecast maximum temperature to the calculated convective temperature to determine a relative time for convective initiation.

4. Sub-regional application.

The sub-regional scale application differs from the previous one in that it incorporates data from a mesoscale surface network operated by the Program for Regional Observing and Forecast Systems (PROFS). Mid-morning surface temperature and dew point data from the appropriate mesonet sites are averaged and used to estimate a forecast afternoon temperature and dew point for each sub-region by the sounding module. These numbers are presented to the forecaster, who is then asked to either accept them or override them with new values input from the keyboard. These decisions are made on a sub-region-by-sub-region basis. When the forecast values have been decided upon, the session proceeds as in the regional application.

One final difference between the two applications is in the estimation of "time of activity." In the sub-regional application, a bias is introduced to account for differences in timing as defined by climatological results. This information comes from a study by Klitch, et. al., (1985), which is a study of high-resolution satellite imagery detailing hourly convective frequencies within the terrain-dominated regions of eastern Colorado, Wyoming, and New Mexico. For example, the study shows that deep convection seldom occurs before late afternoon, or early evening in sub-region 3. This bias is built into the section which estimates timing of convective onset.

5. Format of Final Forecast.

The results from CONVEX for both the regional area and each of the four sub-scale areas are presented in the same format. The range of possible results are:

I. Category 1 - Convective Probability

a. Chances of convective activity during the forecast period are minimal. (In this case, a forecast potential microburst gust speed may be included if appropriate).
b. Convection is likely during the afternoon, or
c. Convection is likely during the evening.

II. Category 2 - Convective Intensity

a. Convective activity will be "non-significant,"
b. Convective activity will be "significant,"
c. Convective activity will be "severe."

For forecasts from this application to be correct, a "severe" thunderstorm must exhibit at least one of the following: a tornado, hail with diameter greater than or equal to 0.75," or surface winds of 50 kts or greater. A "significant" storm will have at least one of the following: hail between 0.25" and 0.74," or surface winds between 35 and 49 knots.

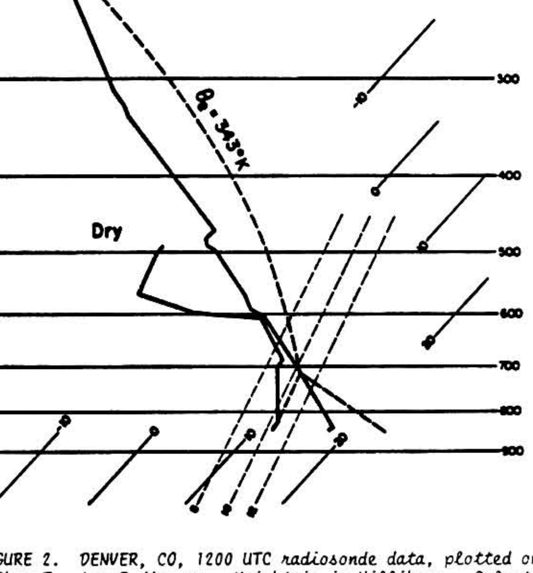

FIGURE 2. DENVER, CO, 1200 UTC radiosonde data, plotted on a Skew T - Log P diagram. Height is in Millibars. Selected temperature line segments shown skewed to P-axis. Thin dash lines are constant moisture values in gm/Kg ... thicker dash line is the parcel lapse rate corresponding to θ_e =343K. The plotted parameters (solid lines) are temperature (right) and dewpoint (left).

6. Application to a Case Study.

In a previous paper (Weaver and Phillips, 1987), the authors discussed the results of applying their expert system to a severe hailstorm case which occurred in Colorado on August 2, 1986. As a point of reference we shall reevaluate this case using the updated version, CONVEX.

Figure 1 shows the 1200 UTC, DEN sounding. The line labeled "θ_e = 343K" shows the saturated lapse rate for a parcel whose mixed temperature and dew point correspond to surface values of 74F and 54F, respectively. A dew point of 54F was quite common August 2nd in sub-regions 2, 3, and 4. A mixed dew point of 54F corresponds to a convective temperature of 74F. Actual afternoon temperatures were somewhat warmer than that; however, for this parcel, the regional version of CONVEX forecast "severe" convective activity, beginning "late" in the day, with a calculated positive buoyancy of 1,993 Joules/kg. On a scale where 1,500 Joules/Kg marks the approximate lower boundary for severe activity, those numbers are particularly significant. Potential hailstone size for this parcel was estimated at 0.78". This would have represented a reasonable regional forecast on August 2nd.

We next applied CONVEX to the various sub-scale regions. Table I shows individual input parameters and results for each sub-region. In all cases, the "forecast" temperature and dew point values used were approximately those which actually occurred on the case day. Sub-region 1 corresponds to the foothills and front range of the northern Colorado Rocky Mountains. Climatologically, convective activity forms earlier here than in the other sub-regions (Klitch, et. al., 1985). Also, thunderstorms seem to be somewhat less severe. This tendency might be caused by the mountainous regions having less absolute moisture (on average) due to the higher elevation, and/or the fact that the extremely low population density in the area precludes accurate reporting of severe activity. In any case, notice that on August 2nd the estimated buoyancy in sub-region 1 is substantially less than all of the others. Indeed, convection began just after "noon LDT." Only one severe weather event was reported; 0.75" hail, 30 miles northwest of FCL. (For all severe reports referred to herein, see Storm DATA, 1986).

TABLE 1. CONVEX RESULTS

VARIABLE	REGIONAL	SUBREGION 1	SUBREGION 2	SUBREGION 3	SUBREGION 4
forecast T (F)	N/A	63.9	78.6	81.7	77.1
T (F) used	80	60	76	81.7	80
forecast DP (F)	47.2	37.9	40.4	42.5	42.2
DP (F) used	54	37.9	54	54	55
timing	early	before 2 PM	after 4 PM	before 5 PM	before 4 PM
probability (%)	/////	/////	/////	/////	/////
of "severe"	100%	10%	100%	100%	100%
"significant"	0%	90%	0%	0%	0%
Buoyancy (J/Kg)	2680	1481	2684	2866	2924
Hailsize (in.)	1.5	0.3	1.1	1.7	0.5

The first major activity of the day occurred just east of DEN, in sub-region 4. Beginning at 1416 LDT and continuing through 2055 LDT, there were a total of 16+ reports of hail ranging in size from 0.75" to 1.25" in diameter. Notice that the CONVEX derived forecast for this sub-region was quite accurate, both in timing and in the severity of the weather which occurred. Similarly, severe activity in sub-regions 2 and 3 began at 1645 and 1623 LDT, respectively, and included hail of diameters 0.75" - 1.5" (sub-region 2), and 0.75" - 3.0" (sub-region 3). Additionally, a small tornado was reported in sub-region 2 at 1950 LDT.

7. Concluding Remarks.

A preliminary test of the regional section of CONVEX was conducted during the summer of 1988. Also, a small number of test cases were run through the sub-regional application when it had been completed. Unfortunately, no objective tests could be made, since CONVEX was under development at the time, and upgrades were being made constantly during the test period. Nevertheless, subjective results from both sections were generally as encouraging as the example case presented above. The regional application seemed to predict correctly in a clear majority of cases, and the sub-regional section seemed to provide the specialized, sub-scale differentiation that was expected of it. In particular, the averaging of mesonet observations seemed to be a much more representative way of looking at the evolving convective environment than previous efforts based on single sites (e.g., Weaver and Phillips, 1987).

The authors are looking forward to participation in two separate tests of CONVEX during the summer of 1989. In the first, a modified version of CONVEX (into which a flash flooding algorithm has been incorporated) will be run routinely at the Cheyenne, WY National Weather Service Warning and Forecast Office. This experience may facilitate our current objective of adapting CONVEX to address specific convective forecasting problems in other areas, outside of the Front Range of the Rocky Mountains.The second test of the system will be as a part of the Environmental Research Laboratory (Boulder, CO) real-time test of several expert systems designed to forecast convective weather -- a test which has been designated "Shootout '89." We are hoping that both of these field experiences will help us develop a better feel for what constitutes the "optimum" in human interaction. This is the area of our deepest concern, both philosophically and practically. We conscientiously set about **not** to develop a model, but to bring about what might be called a reliable "consultant." Thus, we are forced to place a heavier burden upon the forecast/user than might be deemed practical by others.

As stated in our previous paper, "The expert system building convention of developing a knowledge base (through the collective and cooperative efforts of both the expert and a knowledge engineer) seems to be a viable method of extracting and documenting a comprehensive collection of learned expertise about a meteorological subject. Translating that extracted knowledge into a knowledge base of rules is a straightforward and relatively painless process, at least within EXSYS.

The efficacy of consistently applying the knowledge of CONVEX is clearly seen. We liken this utility to that of providing an automated checklist of things to do in a stepwise and logical sequence. A user is not allowed to forget an important parametric input although a graceful escape option is provided if an informed response is not possible; "unknown" or "I don't know" are acceptable responses when such is the case. EXSYS (and most other expert system shells we know about) will attempt to find the best conclusion with the information available."

ACKNOWLEDGEMENTS

This research is funded in part by NOAA/ERL Grant No. NA-85-RAH-05045, and the Cooperative Institute for Research in the Atmosphere at Colorado State University. The authors also wish to thank Mrs. Debbie Eckdahl for preparation of the manuscript.

REFERENCES

Klitch, M.A., J.F. Weaver, F.P. Kelly, and T.H. Vonder Haar,1985: Convective Cloud Climatologies Constructed from Satellite Imagery. Mon. Wea. Rev., 113 (3), 326-337.

Storm Data, 1986, Vol. 28 (8): NOAA/NESDIS publication, National Climatic Data Center, Asheville, NC.

Weaver, J.F. and R.S. Phillips, 1987: Mesoscale Thunderstorm Forecasting Using RAOB Data, Surface Mesonet Observations, and an Expert System. Proc., Symposium on Mesoscale Analysis and Forecasting Incorporating Nowcasting, Vancouver, BC, CANADA, 17-19 August, 327-330.