MESOSCALE THUNDERSTORM FORECASTING USING RAOB DATA, SURFACE MESONET OBSERVATIONS, AND AN EXPERT SYSTEM

by

John F. Weaver and Roger S. Phillips

NOAA/NESDIS/RAMM Branch*
Colorado State University
Fort Collins, CO 80523

1. INTRODUCTION

Most meteorological forecasts begin with lengthy, often tedious procedures for data assimilation, synthesis and analysis. In recent years, these duties have increased in complexity and scope as more and more meteorological data (e.g., new sensor types and increased spatial and temporal observation density) have been introduced. While one might assume that such additions would naturally lead to better forecasts, this has not always been the case. The problem is that the efficient use of new data can be severely limited by increased information handling requirements. In efforts to address this problem, researchers have turned increasingly to the computer. Research programs such as the Program for Regional Observing and Forecasting Services (PROFS) in Boulder, Colorado have developed efficient, comprehensive interactive computer techniques for data acquisition, assimilation and display. Their well-integrated, menu guided approach offers a great deal of potential for relieving the forecaster of most of the routine data preparation duties required prior to the development of a good weather forecast.

Once the task of data preparation is optimized, a qualified operational meteorologist must be able to prepare an accurate forecast within a relatively limited window in time. However, specific forecast problems (such as severe thunderstorm forecasts) often require highly specialized and complex knowledge. While most meteorologists entering the workforce today are college trained, the knowledge gained in class is necessarily more theoretical than practical and is fairly generalized. Each specific forecast problem still involves an active learning period; one which is often lengthy and difficult. As society makes increasing demands on the modern weather forecast (e.g., increased accuracy, a wider range of parameters, more frequent updates, etc), meteorologists must once again turn to the computer; this time to short cut the learning process. For this application, we are investigating the utility of the emerging technology embodied within expert, or knowledge-based systems.

One of the key components of an expert system is a collection of factual material relating to a specific problem domain. These facts, or rules, are supplied by experts in the domain of interest, and may be viewed as providing a much more sophisticated version of the well-known forecasters "checklist," (as, for example, that designed by Miller 1972 for severe thunderstorm forecasts).

The other primary component of an expert system is known as an "inference engine". It allows for a systematic application of the rules, or its so-called "knowledge base" to the specific problem. Inference engines vary considerably in design, but a true expert system does not rely on a specific solution algorithm. Thus, the inference engine attempts to verify the truth of as many rules in the knowledge base as it can by accessing associated data/information and by asking pertinent questions of the user. Conclusions are reached when no new facts can be substantiated. A confidence factor is usually associated with each conclusion to indicate the degree of consensus among the possible rules. Thus the expert system is able to reach an imperfect, but "best possible" answer in the face of missing data.

A properly designed expert system can be used to oversee each task in the forecast procedure, from gathering the appropriate data sets, to issuing one or more forecast conclusions. It is not intended to replace the forecaster. Instead, the system is employed as a capable and reliable assistant during the analysis phase, and as an expert "consultant" during decision making. Additionally, most expert systems allow the user to interrogate the proceedings at any point, to ascertain why certain questions are being asked and/or why certain decisions were made. This feature represents an on-going training potential for the neophyte.

This paper describes an expert system application called "CONVCTIV" that was developed to address several questions. Among these are: Can we completely and faithfully extract, document, and verify domain-specific meteorological expertise (in this case, severe weather forecasting in northeast Colorado) through the employment of expert system technology? Can captured knowledge be consistently applied by the non-expert within the expert system environment? Does the resultant expert system provide a

* Also affiliated with the Cooperative Institute for Research in the Atmosphere (CIRA).

substantial increase in our abilities to train users within the application domain? Can applications of this nature be accomodated within a personal computer (PC) environment? And finally, is the expert system concept one worth pursuing for application on operational forecast systems of the future?

2. APPLICATION DESCRIPTION

CONVCTIV 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 the particular, domain-specific rules are removed from a working and commercially successful knowledge-based system (KBS). EXSYS is written in the C language and will accomodate an unlimited number of rules. Approximately 5,000 rules may be utilized by a PC operating with 640K bytes of memory, which is the device used in our application development. Rules are expressed in an "If...then..." form with "Else" optional, and confidence factors may be associated with each rule to quantify uncertainty. EXSYS allows a limited exchange of information between a KBS application and external software. External programs can either be invoked from within EXSYS or may be run prior to its start, leaving a data file for the application to access later.

The primary author of this paper supplied the knowledge-base for CONVCTIV, as well as the FORTRAN program for sounding analysis. Briefly, this author has had considerable experience as a professional severe thunderstorm forecaster, including several years at the National Severe Storms Laboratory in Norman, OK in the capacity of forecaster/nowcaster, thunderstorm intercept team co-ordinator, and research scientist. Over the years, he has participated in several field programs in a thunderstorm forecast capacity.

The CONVCTIV rule base was implemented by the second author, fulfilling the role commonly referred to in the vernacular as a "knowledge engineer". This work was highly dependent upon many in-depth discussions between both authors. The sessions were iterative in nature, allowing each subtopic to be separately entered into the knowledge-base, then evaluated as an integral part of the whole.

It was decided to structure CONVCTIV to address two separate but related problems. First, it would attempt to estimate the likelihood of convective activity within the northeast Colorado region (defined as encompassing the area, west-to-east, from the base of the foothills of the Rockies to the Kansas border, and north-to-south, from the Wyoming-Nebraska border to the Palmer Ridge, just north of Colorado Springs). The forecast would be valid for the ensuing 6-8 hours of a given day, and would use conventional weather observing and forecasting information available by 10:00 a.m. (LDT) that same day. Second, and optionally, the system would attempt to determine specific locations at which the chances of convection and/or its anticipated severity would be significantly greater-or-less-than the expected regional average. This optional, mesoscale forecast would utilize PROFS' mesonet observations to provide the finer resolution needed (both in space and time) for such a task and would be utilized during the early afternoon hours to update point-specific, convective potential.

For both utilities, CONVCTIV relies heavily on an external program for performing the relatively complex number crunching associated with objective radiosonde data analysis. This software was originally conceived and developed by the first author. It has subsequently been modified by PROFS for use at their facility and adapted by the RAMMB staff for use on its PC. Through several seasons of forecast testing and verification, it has proven to be a highly effective tool. It performs its analysis through an incremented, one-dimensional updraft model with an empirically derived precipitation drag coefficient appended. Its utility to CONVCTIV is two-fold. First, for the regional scale application, the analysis program uses the 1200 GMT Denver RAOB data with a set of forecast, afternoon surface and dewpoint temperatures (provided interactively through CONVCTIV) to determine the expected convective temperature and airmass instability (positive parcel buoyancy). In the case of the mesoscale application, it uses the individual temperature and dewpoint observations from selected PROFS mesonet stations in northeast Colorado, along with a linear, time-dependent boundary layer mixing function to derive convective temperature and positive buoyancy values for each mesonet location.

CONVCTIV is dependent upon knowledgeable, interactive responses to several queries about both the present and anticipated state of several regional weather variables. Upon initiation, the forecaster-user is requested to provide the expected maximum afternoon surface and dewpoint temperatures representative of the northeast Colorado region. These responses are passed to the sounding analysis package and used to calculate the regional approximations of the convective temperature and relative airmass instability, based upon that morning's Denver sounding. CONVCTIV compares the convective temperature provided by the analysis with the forecast maximum afternoon temperature to determine a relative likelihood of significant convective activity developing in the region. The positive buoyancy parameter is used to determine the degree of severity of any potentially significant convective weather.

A regional convective forecast is selected and fine-tuned, based first upon the information just described and then by applying subjective adjustments which are dependent upon the forecaster-user's response to several queries from CONVCTIV. These have to do with the absence or presence of low-level upslope windflow and with synoptic scale, upper-level convective support or suppression. The possible range of forecast results are:

I. Category 1 - Convective Probability
   a) Chances of convective activity during the forecast period are minimal. The convective temperature exceeds the forecasted maximum surface temperature by more than 2.5 degrees Fahrenheit,
   b) Chances of convective activity during the forecast period are good, most likely in the evening, or
   c) Chances of convective activity during the forecast period are good, most likely in the afternoon.

II. Category 2 - Convective Intensity
   a) Convective activity during the forecast period will be weak,
   b) Convective activity during the forecast period will be moderate, or
   c) Convective activity during the forecast period will be severe.

The decision for Category I is made based upon the relationship between the forecast maximum temperature and the computed convective temperature. The basis for the first choice is self explanatory; the second is made when the two temperatures are within 2.5°F of each other; and the third when the forecast temperature exceeds the convective temperature by more than 2.5°F. Category 2 choices are used in combination with a decision for probable convective activity (b or c in Category 1), based upon the relative degree of atmospheric instability (as expressed by the computed positive buoyancy parameter). As can be seen, uncertainty is addressed through the employment of linguistic variables such as "weak," "moderate" and "severe," rather than utilizing the built-in facility for confidence factors, since this seems to be the currently preferred practice within the NWS.

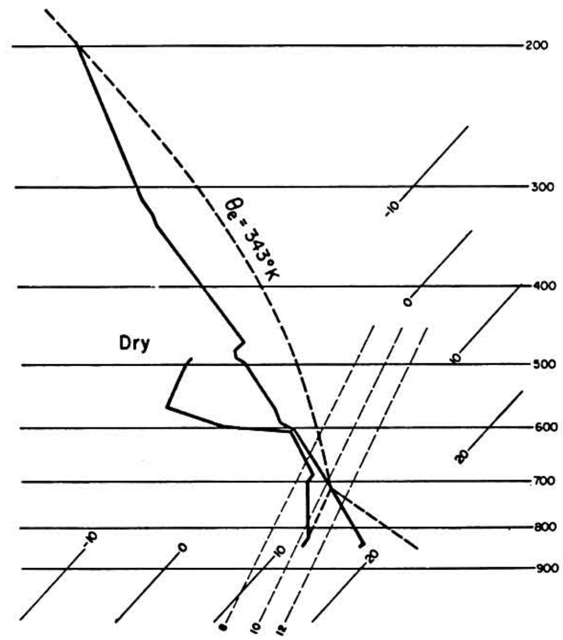
In the cases where CONVCTIV has determined that the chances for convective activity are "good," and before that forecast is presented, the forecaster-user is provided with an option to pursue a mesoscale convective forecast. As stated previously, it is anticipated that this option would be most likely selected in the afternoon, in order to take advantage of the most current mesonet information describing that day's meteorological environment. During a mesoscale forecast, the sounding analysis will use the latest temperature and dewpoint observation from each mesonet site in conjunction with the earlier Denver sounding information to determine the updated convective temperature and positive buoyancy associated with those locations. The mesonet wind speed and direction are used to determine the presence of upslope or downslope conditions at each site. At those locations where the chances for convection are good, the convective severity forecasts are adjusted accordingly. Similarly, the individual mesonet surface heating rates are compared with the average rate over all the non-mountain sites to identify any locations at which convection may initiate sooner or later than the expected regional average time.

The mesonet forecast consists of statements concerning the relative potential for convection and its expected severity, but only at mesonet locations for which those conditions differ appreciably from the regional expectations.

3. EXAMPLE CASE - 2 AUGUST 1986

On August 2, 1986 a major outbreak of severe weather occurred within the region we have labeled northeast Colorado. There were numerous reports of 0.75" to 3.0" hail with major property damage in many cities (Storm Data, 1986). In rural areas, thousands of birds and hundreds of animals were killed or injured, and crop damage totaled well over 100 million dollars.

The weather pattern on August 2nd was a fairly common one for severe thunderstorms on the high plains (Doswell, 1980). A cool front had swept through the state on the previous day, leaving northeast Colorado in low-level easterly (upslope) flow. This pattern tends to advect moist, low-level air westward to the Front Range of the Rockies. The 500 mb analysis revealed moderately strong flow (305° at 35 knots at Denver), with a minor shortwave approaching the region from the northeast.



1. Denver, CO, 1200 GMT radiosonde data plotted on a Skew T - Log P diagram. Height is in millibars. Thick, solid lines are temperature (right), Dewpoint (left). Thicker dashed line is the forecast afternoon parcel lapse rate, and is labeled θ_e=343K. Thin dashed lines are constant moisture values in gms/Kg.

The afternoon temperature/dewpoint values were forecast to be 74°F and 54°F, respectively. When these values are used to define the adjusted afternoon boundary layer, the airmass may be seen to be strongly conditionally unstable. Figure 1 shows the 1200 GMT sounding. The line labeled "θ_e=343°K" shows the parcel lapse rate for the 74°/54° forecast. For this parcel, the Lifted Index (Galway, 1956) has a value of -5 and the sounding analysis package calculated a total positive energy of 2,850 Joules/Kg. On a scale where 1500 Joules/Kg marks the approximate lower boundary for severe thunderstorms, those numbers are particularly ominous.

As discussed above, CONVCTIV begins its regional scale evaluation by ingesting the morning sounding data, and then prompts the user for a forecast afternoon temperature and dewpoint. The user may apply moderate values or extremes; it is entirely discretionary. Once the values are entered, the sounding analysis program uses the forecast dewpoint together with the sounding temperature profile to find the level of free convection (LFC). The convective temperature is calculated by descending dry adiabatically to the surface from this point. If this convective temperature is much lower than the forecast afternoon high, CONVCTIV displays the Category I, selection (a) message and quits. On August 2nd, the convective temperature was nearly equal to the forecast high. This implies fairly late convective development, i.e., the program chose selection (b) in category I.

After determining the relative likelihood and timing of convection, CONVCTIV next polls the user regarding the presence of supportive or suppressive synoptic-scale factors. This is accomplished with the following prompts. (August 2nd choices starred):

An upslope gradient associated with a frontal passage through the front range region is,
 1) absent
* 2) present

Upper-level synoptic support (e.g., PVA, U-L cold air advection, L-L warm air advection, etc.) is:
 1) absent
* 2) weak
 3) moderate
 4) strong

Upper-level synoptic suppression (e.g., NVA, U-L warm air advection, L-L cold air advection, etc.) is:
 1) absent
 2) weak
 3) moderate
 4) strong

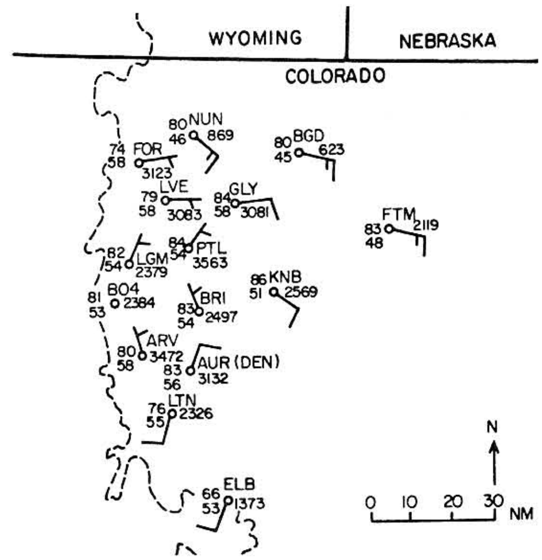
(This last prompt was not issued during the August 2nd case, since the expert system logic already knew the answer, by inference from the user's response to the preceding prompt.)

These decisions are applied through empirically determined adjustments to the total positive buoyant energy calculated earlier (i.e., 2850 Joules/Kg). In this case, the starred answers resulted in a final buoyancy figure of 3050 Joules/Kg. The final decision on convective severity is based on empirically-chosen thresholds of positive buoyant energy. The current version of CONVCTIV specificies <1000 Joules/Kg as weak, 1000-1500 as moderate, and >1500 to be severe convection. Thus, in the case of August 2nd, CONVCTIV specified choice c) under category 2; namely, "convective activity during the forecast period will be severe."

As a final output, the program prints out the mean, density-weighted wind between 700 mb and 300 mb as calculated from the morning RAOB. This parameter is presented to help the forecaster estimate the thunderstorm steering current.

An updated analysis was run using the 2100 GMT PROFS mesonet data. Temperature, dewpoint and derived buoyant energy values are shown in Figure 2 for each station. The regional average buoyancy at this time was determined to be 2614 Joules/Kg. The stations highlighted by CONVCTIV at this time were NUN, BGD, FTM, and ELB (with expected severity of convective activity would be specified as less than the regional average) and FOR, LVE, GLY, PTL, and AUR (with expected severity specified as greater than the regional average). Additionally, conclusions regarding enhanced general convective potential were provided for several of the sites experiencing upslope winds. (These will not be detailed here.)

The data shown in Figure 2 have been modified by an addition to CONVCTIV that was made early in 1987. At this time, the first iteration of an attempt to account for site specific convective climatology was added. This information comes from a study by Klitch, et al (1985), which is a study of high-resolution satellite imagery detailing hourly convective frequencies during the summer of 1982 over the terrain dominated regions of Colorado, eastern Wyoming, and northern New Mexico. At sites where convection was found to be



2. *PROFS mesonet data for 2100 GMT. Winds are plotted in the conventional manner, and in knots. Upper left number in the station models is temperature (°F), lower left is dewpoint (°F), and the number to the right is the buoyant energy estimate in Joules/Kg using the morning Denver sounding together with the individual site values of temperature and dewpoint. Dashed line is continental divide.*

infrequent until later in the day, buoyancy values have been reduced. CONVCTIV now incorporates a closer approximation to these climatically derived frequencies, and applies them two hours in advance to allow for the short range forecast application.

It should be noted that all of the locations showing high buoyancy values experienced severe convection within a few hours of the 2100 GMT analysis. NUN, BGD and ELB had only weak convection in their vicinity.

Longmont, Colorado (LGM) had one of the most severe hail storms of the day, but was not identified by CONVCTIV as having potential for convection above the regional average at any time during the period. There are many possible explanations for this, but we feel the most likely supposition is that, had there been a mesonet site available where the storm formed (approximately 20 NM WNW), that area would have been flagged. We speculate that the storm formed in an area where buoyancies were as high as any shown, and subsequently moved into the LGM area. The lower buoyancies there would have meant a much lower updraft speed, and less ability to support very large hail stones. This might explain why large hail fell right at LGM. On this note, it should also be mentioned that the computed buoyancy at FOR dropped from 3123 to 2768 Joules/Kg between 2100 and 2200 GMT. This occurred as cirrus anvil material cut off the heating in Fort Collins, about one hour before the storm. Perhaps the resulting lowered buoyancies at FOR account for the large hail which fell out of the storm there as well. However, these comments are only offered as reasonable speculation. The authors are looking at these, and other possible explanations in order to make improvements to CONVCTIV in the near future.

4. SUMMARIZING REMARKS

As it now stands, CONVCTIV has been tested for only a few cases, all of which were representative of strong convection. The potential demonstrated by the expert system in these cases has been quite promising, but we are now interested in pursuing a rigorous and more conclusive examination of its behavior under all weather conditions. Current plans are to operate CONVCTIV on a daily basis, excepting weekends, throughout the Summer of 1987 in order to collect suitable performance statistics. We also intend to carefully evaluate empirical classification parameters that have not yet been extensively verified since their introduction into the knowledge base.

In the way of improvements, we currently have in mind the addition of a provision for alerting the user about meteorological incompatibilities in the sounding, mesonet and/or input parameters. For instance, if a mesonet site appears to be heating more slowly than anticipated or is drier than expected (based upon the morning sounding, the morning forecast, and the current mesonet observations), the user would be apprised of such and would be provided the opportunity to adjust certain input parameters, if desirable. A menu would be provided to furnish guidance in such instances.

As interim responses to the questions set forth in the introduction to this paper, the following applies: 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. Give and take discussions serve to flesh out tidbits of information and understanding reposited, and perhaps temporarily dormant, in the mind of a domain expert. Translating that extracted knowledge into a knowledge base of rules is a straightforward and relatively painless process, at least within EXSYS. Verification of pieces or subsets of knowledge proved to be similarly efficient.

The efficacy of consistently applying the knowledge of CONVCTIV 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.

The training benefits of expert systems appear to be self evident but, in our case, have not yet been tested.

We have been suitably impressed with our PC-AT's ability to accomodate EXSYS-CONVCTIV and particularly the external sounding analysis program. Interactive query/response times have not exceeded the mind's comfort zone for attentiveness, at least not yet.

ACKNOWLEDGEMENTS

The authors would like to express their appreciation to Mr. Robert N. Green of this office for his help in establishing communications between a DEC VAX 11/780 (which receives and stores data from PROFS on a routine basis), and the PC-AT on which our expert system application runs. Without this important downlink, our project would not have been possible.

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.

REFERENCES

1. Doswell, C.A., 1980: Synoptic-scale environments associated with High Plains severe thunderstorm. Bull. Amer. Meteor. Soc. 61(11), 1388-1400.

2. Galway, J.G., 1956: The Lifted Index as a prediction of latent instability. Bull. Amer. Meteor. Soc., 37(10), 528-529.

3. 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.

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