Neural Network versus Multiple regression: Methods of estimating VIL thresholds for severe hail based on storm environment

Jeffrey T. Davis
National Weather Service Office Amarillo, Texas

1. Introduction

As the National Weather Service deploys its new Weather Surveillance Radar 1988 (WSR-88D), more meteorologists will be introduced to vertically integrated liquid water (VIL) data. Until recently, only about a dozen Weather Service Offices have had the opportunity to use VIL operationally. Over the years, studies (Beasley, 1986; Devore, 1983; Garner, 1989) have shown VIL to be a reliable radar-derived product in predicting and identifying severe thunderstorms. Some studies (Stewart, 1991; Teague, 1990) have attempted to use VIL to predict tornadoes and downburst winds, but Beasley (1986) found VIL to be a better indicator of severe hail (hailstone diameter .75 inch or greater).

Even though the use of VIL has improved severe thunderstorm detection at several Weather Service Offices (Garner, 1989; Teague, 1990), VIL is not a flawless predictor of severe weather. Beasley (1986) found VIL values associated with hailstorms to vary monthly, weekly, and daily, making VIL a difficult radar-derived product to use in the detection of severe thunderstorms. Beasley's research not only found average VIL values in hailstorms to vary significantly by season, but that absolute VIL could vary significantly by region in Oklahoma. These findings led Beasley to develop a set of seasonally and geographically normalized regression equations for estimating hailstone size.

In a similar study to that of Beasley (1986), Wagenmaker (1992) found hailstorm VIP-5 heights varied significantly month-to-month, week-to-week, and even day-to-day. Thus, instead of evaluating hailstorm VIP-5 heights on a month-by-month, or seasonal climatology, he compared VIP-5 heights with certain environmental parameters. Wagenmaker's study showed that VIP-5 height thresholds for hail producing thunderstorms can provide useful information when normalized with respect to certain characteristics of an individual thunderstorm environment. Furthermore, when not considering potential problems with the calculation of VIL, Teague (1990) pointed out that the range in VIL values associated with severe weather on the same day might be the result of a changing airmass. A change of airmass throughout the day would significantly alter the temperature or humidity of the environment, resulting in a wide range of VIL threshold values.

As Wagenmaker (1992) suggested, the weekly, monthly, seasonal, and even the diurnal variations in hailstorm VILs are likely related to the wide range of atmospheric environments in which severe thunderstorms occur. Therefore, a seasonally or even geographically fixed VIL threshold algorithm would be unable to adjust to changes in hailstorm VILs brought about by the change in environment from day to day. For instance, consider a cool season type airmass in place over the Texas and Oklahoma panhandles on a given day in May. In this case, a warm season attuned VIL algorithm would likely be unreliable and predict too high of a VIL threshold for that type of airmass. Given this, VIL will only be evaluated with respect to the environment before determining potential storm severity thresholds. In this paper two different methods of estimating VIL thresholds for severe hail will be discussed. One method will utilize a neural network, while the other will rely on a multiple regression analysis. Both methods use VIL, the freezing level height, and the total totals index for predictors of hailstone size. Several hailstorm cases will be used to demonstrate and compare both the neural network and multiple regression methods.

2. Data

The data set used in this study was prepared by Garner (1989). All reports of hail from March 15, 1989 to May 31, 1989 within 125 miles on Amarillo, Texas were compare to RADAP II maps of VIL. The reports of hail were compared to the maximum VIL value corresponding to the location of the hail report. For all cases, the value of VIL used occurred at or no more than 20 minutes prior to the time of the hail report.

All freezing level heights and total totals used in this study were taken from the Amarillo, Texas 12z and 00z soundings. For all hail reports occurring between 18z and 06z, the 00z sounding was used. The freezing level height and total totals from the Amarillo 12z sounding were used for all other times of the day.

3. Neural Network versus Multiple regression

Analog versions of back-propagating neural networks can handle numeric problems which would otherwise be processed with regression analysis. The independent variables used in regression become the inputs to the neural network and the dependent variable becomes the output. There are a few key differences between neural networks and multiple regression, however. First, a neural network is highly non-linear so it can resolve many problems that traditional regression methods cannot. A neural network can also provide a tight data fit (shown in figure 1) that cannot be achieved with regression. Moreover, a neural network needs fewer cases to learn than regression methods as long as the cases represent the entire range of expected parameters. Lastly, a neural network can generalize (the ability to do well with new cases for which it has not been trained) better than regression. For those interested in learning more about neural networks, refer to McCann (1992) and Leblang (1991).

4. Sensitivity Tests

One approach to test a neural network and multiple regression equation is to look at their responses to simple input. If the responses are not realistic and in agreement with similar studies, then the network and equation are probably not reliable.

For the first test, the freezing level height (12,000 ft) and the total totals index (50) were kept constant while VIL varied from 35 kg/m^2 to 75 kg/m^2. Figures 2 and 3 show that the output of hailstone size from both the neural network and the regression equation increases as VIL increases. This is in agreement with observations and other studies (Beasely, 1986; Teague, 1990).

Similarly, in a second test, the freezing level height was varied from 950 ft to 18,000 ft while holding the total totals index (50) and VIL (50 kg/m^2) constant. Figures 4 and 5 show that as the freezing level increases in height the hailstone size decreases. In general, this is a valid solution since observations indicate that warmer thunderstorm environments require higher VIL thresholds for severe hail. Therefore, with all other factors like vertical wind shear, buoyant energy, precipitation drag, and dynamic pressure effects etc. set equal, a given VIL threshold of 50 kg/m^2 would correspond to smaller hailstone sizes as the thunderstorm environment becomes warmer (as freezing level increases in height).

For the last test, the total totals varied from 35 to 70 while holding the freezing level height at 12,000 ft and VIL at 50 kg/m^2. As figures 6 and 7 show, the hailstone size increases as the total totals index increases in value. Again, this is a reasonable response.

5. Case studies

Selection of the following six cases for this study was based on the availability of recorded VIL and hail data. Details of the atmospheric environments in which these hail events occurred were not available. Because of the limited data set, no conclusive results can be drawn from this study. Rather, the main purpose of this study was to demonstrate two different ways of analyzing VIL and hail data, and in doing so it is necessary to look at a few hail events.

The six cases selected from the Texas Panhandle occurred from June 1, 1989 to June 4, 1989. As illustrated in table 1, the neural network provided better estimates of hailstone size for all of the six cases. The multiple regression analysis generally underestimated hailstone diameters for the most of the cases studied.

6. Discussion

In this study it was demonstrated that by using a few current thermodynamic parameters for a given weather situation, forecasters may effectively predict a VIL threshold needed for severe hail on that particular day. This study also looked at two different types of methods which can be used to estimate a "VIL value of the day" that will yield severe hail. From the limited data set, it was found that the neural network provided better estimates of hailstone size for all six of the cases studied.

Although no conclusive results can be drawn from this study, future work should consider using a neural network. The highly non-linear nature and generalizing capabilities of a neural network make them suitable for this type of study. Future work at the Amarillo National Weather Service Office should also consider using other thermodynamic or even synoptic forcing parameters recorded over several severe weather seasons.

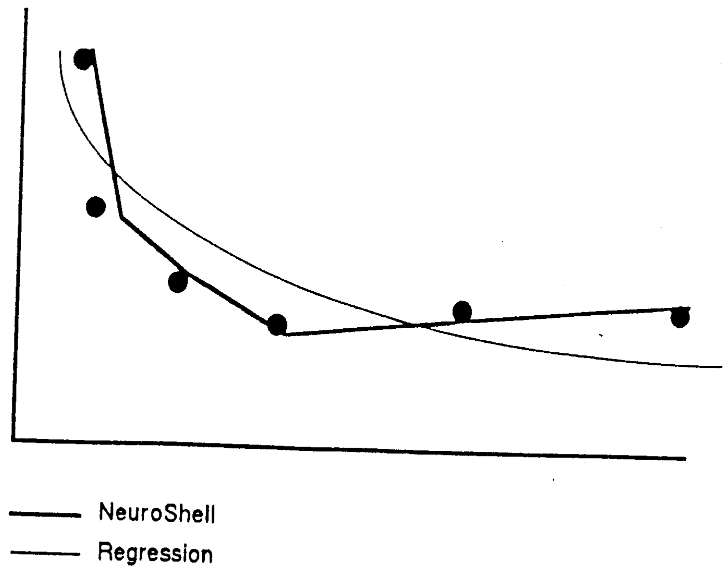

Figure 1. (from Neuroshell, 1990)

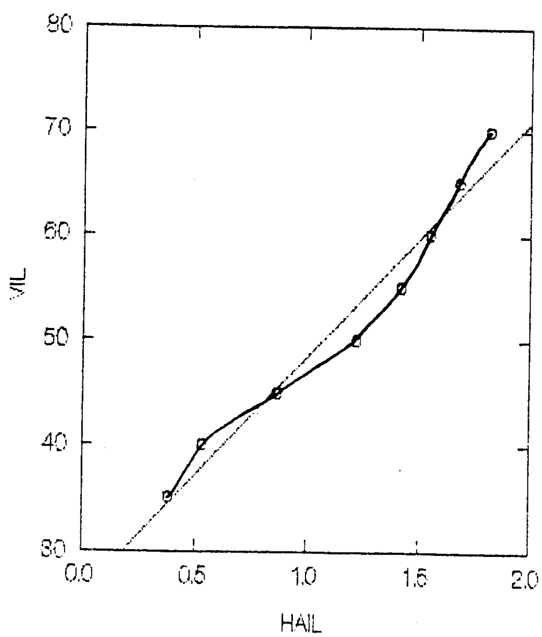

Figure 2. Neural Network

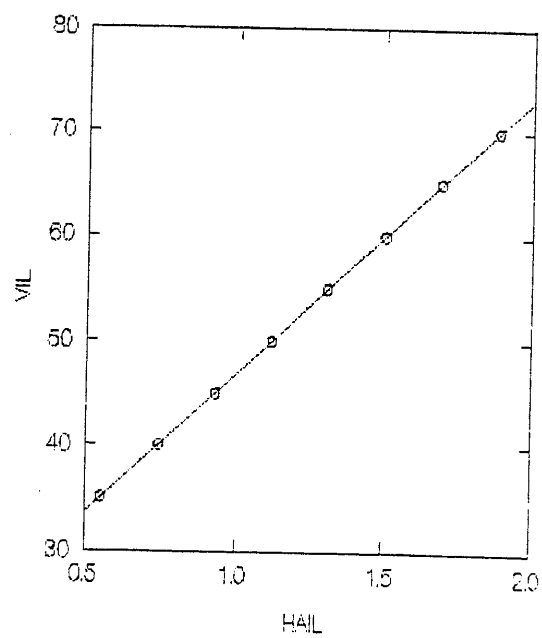

Figure 3. Regression Analysis

DATE	Freezing Level Height	Total Totals Index	VIL	Observed Hail	Regression Hail	Neural Network Hail
6/1/89	13,400ft	50	70	1.75	1.83 (+.08)	1.73 (-.02)
6/3/89	13,800ft	47	55	1.75	1.22 (-.53)	1.35 (-.40)
6/3/89	13,800ft	47	50	1.50	1.03 (-.47)	1.20 (-.30)
6/3/89	13,800ft	47	60	1.75	1.60 (-.15)	1.70 (-.05)
6/4/89	11,900ft	50	65	1.75	1.69 (-.06)	1.80 (+.05)
6/4/89	11,900ft	50	60	1.00	1.51 (+.51)	1.20 (+.20)

TABLE 1.