3.4

THE USE OF ARTIFICIAL INTELLIGENCE METHODS IN IMPROVED VISIBILITY FORECASTING FOR SURFACE TRANSPORTATION

Ryan Knutsvig
Regional Weather Information Center
University of North Dakota, Grand Forks, North Dakota

1. INTRODUCTION

Low visibilities due to fog can have hazardous affects on human life. Surface transportation is one activity that is most affected by fog. For instance, the National Highway Traffic Safety Administration's (NHTSA) Fatality Analysis Reporting System (FARS) reports that in 1994 the total number of fatal motor vehicle accidents was 36,799. Fog related fatal accidents accounted for 613, or 1.67% of the total number of crashes. With the cost of motor vehicle crashes in 1994 being $150.5 billion (NHTSA), the associated cost due to fog can be assumed to be approximately $2.5 billion in 1994. Thus, fog has damaging affects socially as well as financially.

Fog forecasting focused on surface transportation may help reduce the number of motor vehicle accidents sustained per year. In this study, the feasibility of using artificial intelligence methods to aid in fog forecasting for surface transportation is explored, namely case-based reasoning (CBR) and rule-based reasoning (RBR).

2. ARTIFICIAL INTELLIGENCE

"Artificial intelligence (AI) technology provides techniques for developing computer programs for carrying out a variety of tasks, simulating the intelligent way of problem solving by humans" (Krishnamoorthy and Rajeev, 1996). Many scientific domains have benefited from AI, ranging from medicine to automobile design. RBR (often called expert systems) and CBR are areas of AI that employ ways humans solve problems. RBR reasons about a set of rules, created by the domain expert, to reach a conclusion about a specific problem. It basically attempts to emulate a domain expert. CBR, on the other hand, "remembers" back to cases similar to the current case to find a solution. This method assumes that similar problems have similar solutions and that problems an agent encounters tend to recur (Leake, 1996).

* Corresponding author address: Ryan Knutsvig, University of North Dakota, Regional Weather Information Center, Grand Forks, ND 58202-9007; email: knutsvig@netzero.net.

3. FOG FORECASTING

One of the most challenging meteorological parameters to forecast is the formation of fog. Some current models forecast fog by using regression equations (NGM & AVN MOS), while other models parameterize some of the variables that are necessary for forecasting fog (RUC). In a study by Meyer et. al., verification statistics show an improvement over persistence for NGM model's MOS.

AI is an area of cognitive science that has been utilized somewhat in weather forecasting. There have been a few studies using AI in fog forecasting research. Bjarne Hanson's work (2000a, 2000b, 1998) on CBR in weather forecasting inspired this work this paper presents. He was able to conclude that "querying a large database of weather observations for past weather cases similar to a present case using a fuzzy k-nearest neighbors (fuzzy k-nn) algorithm that is designed and tuned with the help of a weather forecasting expert can increase the accuracy of predictions of cloud ceiling and visibility at an airport" (Hanson, 2000a). Peak and Tag (1989) created an expert system for the prediction of maritime visibility obscuration. Their system (AESOP) has 232 rules and has been designed in terms of nowcasts (0-1 h) and forecasts (1-6 h). They claim AESOP's overall performance is 75% correct and displays considerable forecast skill when compared to 47% for persistence and 41% for random chance.

While fog forecasting has seen reasonable research efforts from the scientific community, current observation systems in place across the U.S. limit the forecasting accuracy that can be achieved. This warrants the need to look to other ways for forecasting fog. The current research efforts described in this paper are designed to create a fog forecasting system that utilizes AI technology and is tailored to the highway transportation industry.

4. WORK IN PROGRESS

Currently, the procedures in place in the CBR system include two main parts: case comparison and case adaptation. In the first part of the system, cases are compared to the current case to determine similarity. In the second part, cases deemed similar are adapted to create a forecast. The two archived datasets used in the system were prepared by the NCDC: "TD-3280", which is the 1984 and later dataset, and “TDF14”, which is the pre-1984 dataset going as far back as 1938 for some sites.

4.1 Case Comparison

As expected, case comparison takes quite a bit longer to complete versus case adaptation with the current algorithms in place. Currently, cases are accessed according to the date of the current case. The archived cases must be within one month of the current case. This is the only filter in place before case comparison. Next, the system compares the past x hours of the current case to the same times of each archived case. The variables compared currently in the system are as follows:

- Surface air temperature
- Surface dew point depression
- Surface wind speed
- Surface wind direction
- Mean sea-level pressure
- Mean sea-level pressure tendency
- Current visibility
- Current precipitation
- Cloud coverage
- Cloud height

Similarity is determined by using fuzzy logic in a manner similar to that used by Hanson (2000a). Each variable has its own membership function, which is dependent upon what is “similar” in regards to fog forecasting for each variable. Figure 1 is an example of a membership function. In fuzzy logic, objects have a membership in a set according to a membership function. Instead of the traditional methods, which state that an object is either a member (1) or not (0), fuzzy logic allows degrees of membership [0,1]. In the CBR system, compared variables are most similar if they achieve a fuzzy membership value of 1 and least similar if their value is 0. The similarity is determined for each hour for each variable. Below is an example of the similarity computed for hours 0 (initial hour) and 1 hour before.

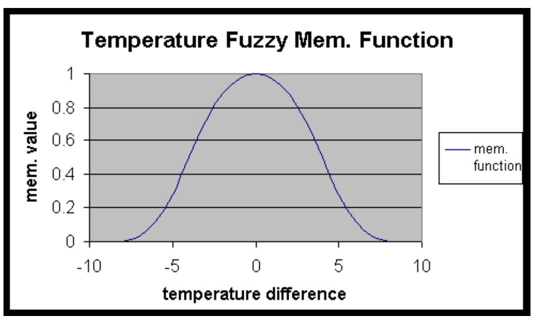

Fig. 1: Fuzzy membership function for temperature.

hour 0		**hour –1**	
Temperature	0.95	Temperature	0.7
Dew Pt. Dep.	0.85	Dew Pt. Dep.	0.8
Wind Speed	0.9	Wind Speed	0.9
Wind Dir.	0.65	Wind Dir.	0.6
MSLP	0.95	MSLP	0.9
MSLP Tend.	0.8	MSLP Tend.	0.45
Visibility	0.8	Visibility	0.8
Precipitation	1	Precipitation	1
Cloud Cover	0.7	Cloud Cover	0.4
Cloud Height	0.75	Cloud Height	0.75
MIN	0.65	**MIN**	0.4

Table 1: Example of similarity calculations for the initial hour and 1 hour prior.

Each variable has its own fuzzy membership in the fuzzy set “similar.” As described by Hanson, the minimum of the similarities is kept; this is called a fuzzy intersection of those fuzzy sets. The variable whose membership value is the lowest can be described as the “weakest link” for that hour since the minimum is kept. This is done for all hours that are used to determine similarity. Currently, the previous 6 hours are used although the system is configured to use a maximum of 24 hours.

The last step in determining a case’s similarity is the use of the “forgetting function” (figure 2) as termed by Hanson. The use of this function helps to show leniency to cases whose earlier hours (hour –4, –5, –6) are not very similar.

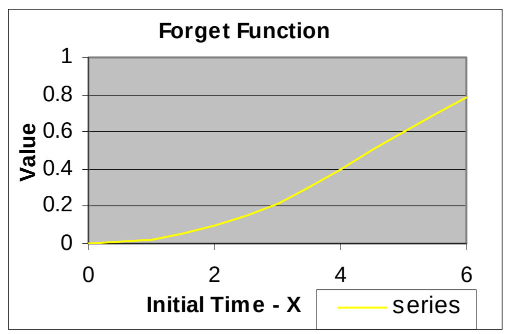

Fig. 2: Forget function for hours 0 through -6.

To calculate the final case similarity, the forget function is “maxed” with the current case similarity in what fuzzy logic calls a fuzzy union. This helps in not using low similarities that earlier hours may have in similarity calculations since the most recent hours are the most important.

hr 0	hr -1	hr -2	hr -3	hr -4	hr –5	hr -6
0.65	0.4	0.45	0.5	0.45	0.25	0.3
0	0.02	0.1	0.22	0.4	0.6	0.78
0.65	0.4	0.45	0.5	0.45	0.6	0.78
MIN = 0.4	Overall Case Similarity = 0.4					

Table 2: In Row 2 are the similarities, in row 3 are the forget function’s values and in row 4 are the “maxed” values.

Once the forget function has determined the similarity for each hour (row 4 in table 2), the overall case similarity is determined by the fuzzy intersection (min) of all the time period's values.

4.2 Case Adaptation

After the overall similarity is determined for each case, there is still the process of adapting those similar cases to the current situation to forecast fog. There are a couple of methods currently investigated. One way is to use the mean of the observations, according to their similarity, to create the forecast output. This method would be useful if the resulting cases are in relatively good agreement. There are a couple of cons associated with this method. First of all, if the system has accessed 10 cases that it deemed "most similar" and some of those cases have fog while others do not, then the mean of the visibility range would not be an accurate representation of the visibility reduction possible. For example, suppose the observations for a given hour across retrieved cases are: 10 miles, 10 mi, 5 mi, 0.25 mi, 0.5 mi, 15 mi, 7 mi, 5 mi, 0.75 mi, and 3 mi. The mean of the observations would be 5.65 miles. A way to sidestep this problem would be to return the earliest and latest time that fog occurred (reducing visibility below 0.6 miles).

A second way of utilizing the retrieved cases would be to return the minimum visibility that was experienced. Also, a probability of visibility dropping below a certain criteria (1 mi, 0.25 mi) could be extracted. In the example above, the minimum visibility was 0.25 and the probability of that happening could be assumed to be 10% since that is what occurred in 10% of the most similar cases. For visibilities < or = 1, the probability would be 30%. This form of information would be helpful, but not ideal. Also, this information could be weighted according to the similarity.

The two adaptation methods described above are methods that could be used for each airport observation site and the area in the vicinity of the airport. But that is not the goal of this research. The next step in this study is to use those forecast sites and the archived data that are available to forecast for highways across the U.S. This poses to be a challenge but there are a few ways that this can be done.

1) Calculate a statistical relationship. Use at least a couple years of data from Roadway Weather Information System (RWIS) sensors, or other densely populated weather sensors along highways, and create a statistical relationship (linear regression) with regards to fog between the airports and RWIS sensors.
2) Use topography to assign airport sites' (historical) data to a grid point of similar topographical influences.
3) Use a set of rules (RBR) to describe when a grid point should use what airport's (historical) data.
4) Do an objective analysis of the data.

5. FUTURE WORK

Future work includes the incorporating of other variables into the CBR system. Other variables considered include (where low level means 950 to 850 mb):

- Soil Moisture
- Low level dew point depression
- Low level wind speed
- Low level wind direction

Other work that is scheduled to be done is an investigation into how RBR can be used to forecast fog across the U.S. A RBR system would need to have rules that are unique for a given uniform area, have rules that cover the entire problem domain, and return a fog probability and/or visibility range. A RBR/CBR hybrid may be beneficial if stand-alone systems do not prove beneficial.

6. REFERENCES

Tag, P.M. and Peak, J.E., 1989: An Expert System Approach for Prediction of Maritime Visibility Obscuration. *Mon. Wea. Rev.* Vol. **117**, No. 12, pp. 2641-2653.

Hansen, B.K. 2000a: Weather Prediction Using Case-Based Reasoning and Fuzzy Set Theory. Master of Computer Science Thesis, Technical University of Nova Scotia, Halifax, Nova Scotia, Canada. http://www.chebucto.ns.ca/~bjarne/thesis/

Hansen, B. K., 2000b: Analog forecasting of ceiling and visibility using fuzzy sets. 2nd Conference on Artificial Intelligence, American Meteorological Society, 1-7. http://www.chebucto.ns.ca/~bjarne/ams2000

Hansen, B. K. and Riordan, D., 1998: Fuzzy case-based prediction of ceiling and visibility. 1st Conference on Artificial Intelligence, American Meteorological Society, 118-123. http://www.chebucto.ns.ca/~bjarne/fuzzy_cbr/

Krishnamoorthy, C.S., and Rajeev, S., 1996: *Artificial Intelligence and Expert Systems for Engineers*. CRC Press.

Leake, D.B.,1996: CBR in Context: The Present and Future. In Leake, D., editor, Case-Based Reasoning: Experiences, Lessons, and Future Directions, AAAI Press/MIT Press, 1996.

Meyer, F., Dagostaro, V., and Miller, D.: NGM-Based MOS Visibility and Obstruction to Vision Guidance for the Contiguous United States. http://www.nws.noaa.gov/om/windwave.htm

National Highway Traffic Safety Administration (NHTSA): The Economic Cost of Motor Vehicle Crashes, 1994. NHTSA Technical Report, U.S. Department of Transportation, NHTSA, Washington, DC 20590 http://www.nhtsa.dot.gov/people/economic/ecomvc1994.html

NHTSA's Fatality Analysis Reporting System (FARS): http://www-fars.nhtsa.dot.gov/