A Fuzzy Logic Based Analog Forecasting System for Ceiling and Visibility

Bjarne Hansen, Meteorologist
Cloud Physics and Severe Weather Research Division
Meteorological Research Branch
Meteorological Service of Canada
Dorval, Québec

Workshop on Fog Remote Sensing and Modeling, UQAM,
14 - 15 June 2005, Montreal, Quebec
Outline

Introduction

- Ceiling and visibility prediction
- Fuzzy logic
- Analog forecasting / $k$-nearest neighbors

Combining all of the above

Operational application: WIND-3
Ceiling and Visibility Prediction

Critical airport forecasts for: planning, economy, and safety.
Ceiling height and visibility prediction demands precision in near-term and on local scale:

- Ceiling height, when low, accurate to within 100 feet.
- Visibility, when low, accurate to within 1/4 mile.
- Time of change of flying category should be accurate to within one hour.

Safety concern

“Adverse ceiling and visibility conditions can produce major negative impacts on aviation - as a contributing factor in over 35% of all weather-related accidents in the U.S. civil aviation sector and as a major cause of flight delays nationwide.”

Fuzzy Logic Definition

“Fuzzy logic a superset of Boolean logic dealing with the concept of partial truth – truth values between ‘completely true’ and ‘completely false’. It was introduced by Dr. Lotfi Zadeh of UCB in the 1960’s as a means to model the uncertainty of natural language.”

Analog Forecasting / \( k \)-nearest neighbors\(^1\)

- A basic statistical learning technique
- Analog forecasting contrasts with linear regression, two are complementary
- Linear regression \( \rightarrow \)
- Solutions based on line which best discriminates between two classes
- Generally accurate but evidently locally wrong where effects are non-linear

Analog Forecasting / $k$-nearest neighbors

- 1 nearest neighbor classifier
- Solutions based on single nearest neighbor
- Generally more accurate than linear regression, but locally more unstable

Analog Forecasting / $k$-nearest neighbors\footnote{1}

- 15 nearest neighbor classifier →
- Solutions based on majority of 15 nearest neighbors
- Generally more accurate than linear regression, and less locally unstable than 1 nearest neighbor

Analog forecasting / $k$-nn complements Linear Regression

Compared to the linear model approach (basis of most statistical systems for C&V prediction):

1. The $k$-nearest neighbors technique has a relative lack of structural assumptions about data.

“The linear model makes huge assumptions about structure and yields stable but possibly inaccurate predictions. The method of $k$-nearest neighbors makes very mild structural assumptions: its predictions are often accurate but can be unstable.”

2. $k$-nn is computationally expensive, but newly practical.

Both points borne out in ceiling and visibility prediction system…

**Ceiling and visibility articles since 1970**

**Using Multiple Linear Regression (MLR) and Multiple Discriminant Analysis (MDA)**


**Conditional climatology without MLR or MDA**


Moore’s Law

The empirical observation that at our rate of technological development, the complexity of an integrated circuit, with respect to minimum component cost will double in about 24 months. ¹

Operational Application: Prediction System: WIND-3

WIND: “Weather Is Not Discrete”

Consists of three parts:

- **Data** – weather observations and model-based guidance;
- **Fuzzy similarity-measuring algorithm** – small C program;
- **Prediction composition** – predictions based on selected C&V percentiles in the set of $k$ nearest neighbors, $k$-nn.

**Data: what current cases and analogs are composed of**

- Past airport weather observations: 190 airports, 30 years of hourly obs, time series of ~ 300,000 detailed observations;
- Recent and current observations (METARs);
- Model based guidance (knowledge of near-term changes, e.g., imminent wind-shift, onset/cessation of precipitation).
## Data: Past and current observations, regular METARs

<table>
<thead>
<tr>
<th>Type</th>
<th>Attribute</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>temporal</td>
<td>date</td>
<td>Julian date of year (wraps around)</td>
</tr>
<tr>
<td></td>
<td>hour</td>
<td>hours offset from sunrise/sunset</td>
</tr>
<tr>
<td>cloud ceiling</td>
<td>cloud amount(s)</td>
<td>tenths of cloud cover (for each layer)</td>
</tr>
<tr>
<td>and visibility</td>
<td>cloud ceiling height</td>
<td>height in metres of $\geq 6/10$ths cloud cover</td>
</tr>
<tr>
<td></td>
<td>visibility</td>
<td>horizontal visibility in metres</td>
</tr>
<tr>
<td>wind</td>
<td>wind direction</td>
<td>degrees from true north</td>
</tr>
<tr>
<td></td>
<td>wind speed</td>
<td>knots</td>
</tr>
<tr>
<td>precipitation</td>
<td>precipitation type</td>
<td>nil, rain, snow, etc.</td>
</tr>
<tr>
<td></td>
<td>precipitation intensity</td>
<td>nil, light, moderate, heavy</td>
</tr>
<tr>
<td>spread and temperature</td>
<td>dew point temperature</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td></td>
<td>dry bulb temperature</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>pressure</td>
<td>pressure trend</td>
<td>kiloPascal $\times$ hour$^{-1}$</td>
</tr>
</tbody>
</table>
## Data: Past and current observations

E.g., over 300,000 consecutive hourly obs for Halifax Airport, quality-controlled.

<table>
<thead>
<tr>
<th>YY/MM/DD/HH</th>
<th>Ceiling</th>
<th>Vis</th>
<th>Wind Directn</th>
<th>Wind Speed</th>
<th>Bulb 10's deg</th>
<th>Dry 30's m</th>
<th>Dew deg C</th>
<th>MSL kPa</th>
<th>Station Press kPa</th>
<th>Cloud press</th>
<th>Amount tenth</th>
<th>Weather</th>
</tr>
</thead>
<tbody>
<tr>
<td>64/1/2/0</td>
<td>15</td>
<td>24.1</td>
<td>14</td>
<td>16</td>
<td>-4.4</td>
<td>-5.6</td>
<td>101.07</td>
<td>99.31</td>
<td>10</td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>64/1/2/1</td>
<td>13</td>
<td>6.1</td>
<td>14</td>
<td>26</td>
<td>-2.2</td>
<td>-2.8</td>
<td>100.72</td>
<td>98.96</td>
<td>10</td>
<td>ZR-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/2</td>
<td>2</td>
<td>8.0</td>
<td>11</td>
<td>26</td>
<td>-1.1</td>
<td>-2.2</td>
<td>100.39</td>
<td>98.66</td>
<td>10</td>
<td>ZR-F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/3</td>
<td>2</td>
<td>6.4</td>
<td>11</td>
<td>24</td>
<td>0.0</td>
<td>-0.6</td>
<td>100.09</td>
<td>98.36</td>
<td>10</td>
<td>ZR-F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/4</td>
<td>2</td>
<td>4.8</td>
<td>11</td>
<td>32</td>
<td>1.1</td>
<td>0.6</td>
<td>99.63</td>
<td>97.90</td>
<td>10</td>
<td>R-F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/5</td>
<td>2</td>
<td>3.2</td>
<td>14</td>
<td>48</td>
<td>2.8</td>
<td>2.2</td>
<td>99.20</td>
<td>97.50</td>
<td>10</td>
<td>R-F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/6</td>
<td>3</td>
<td>1.2</td>
<td>16</td>
<td>40</td>
<td>3.9</td>
<td>3.9</td>
<td>98.92</td>
<td>97.22</td>
<td>10</td>
<td>R-F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/7</td>
<td>2</td>
<td>2.0</td>
<td>20</td>
<td>40</td>
<td>4.4</td>
<td>4.4</td>
<td>98.78</td>
<td>97.08</td>
<td>10</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/8</td>
<td>2</td>
<td>4.8</td>
<td>20</td>
<td>35</td>
<td>3.9</td>
<td>3.3</td>
<td>98.70</td>
<td>97.01</td>
<td>10</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/9</td>
<td>4</td>
<td>4.0</td>
<td>20</td>
<td>29</td>
<td>3.3</td>
<td>2.8</td>
<td>98.65</td>
<td>96.96</td>
<td>10</td>
<td>R-F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/10</td>
<td>6</td>
<td>8.0</td>
<td>20</td>
<td>35</td>
<td>2.8</td>
<td>2.2</td>
<td>98.60</td>
<td>96.91</td>
<td>10</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/11</td>
<td>8</td>
<td>8.0</td>
<td>20</td>
<td>32</td>
<td>2.8</td>
<td>2.2</td>
<td>98.45</td>
<td>96.77</td>
<td>10</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/12</td>
<td>9</td>
<td>9.7</td>
<td>23</td>
<td>29</td>
<td>2.2</td>
<td>1.7</td>
<td>98.43</td>
<td>96.75</td>
<td>10</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64/1/2/13</td>
<td>9</td>
<td>11.3</td>
<td>23</td>
<td>32</td>
<td>1.7</td>
<td>1.1</td>
<td>98.37</td>
<td>96.69</td>
<td>10</td>
<td>F</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

...
Data: Computer model based guidance

Predictions of weather elements related to C&V, e.g. temperature, dewpoint, wind, weather, dp/dt.

1. Any available model output can be used.
Algorithm: Collect Most Similar Analogs, Make Prediction

Archive search is like contracting hypersphere centered on present case. Axes measure differences weather elements between compared cases. Distances determined by fuzzy similarity-measuring functions, expertly tuned (for first approximation), all applied together simultaneously.

Basic idea, key to \( k \)-nn

~ \( 10^6 \) points

weather states ↔ ordered points in 12-D weather sequences ↔ generally continuous loci
weather variables tend to flow in certain directions

Forecast ceiling and visibility based on outcomes of most similar analogs.

Spread in analogs helps to inform about appropriate forecast confidence.
MSC aviation weather service reorganization
Two Canadian Meteorological Aviation Centres
CMAC-West in Edmonton, CMAC-East in Montreal
Products: TAFs, GFAs, SIGMETs, AIRMETs

CMAC-W
97 TAFs
39 forecasters
6-7 operational desks
CMAC-E
83 TAFs
33 forecasters
5-6 operational desks

For more information, contact Steve Ricketts, Manager CMAC-W,
steve.ricketts@ec.gc.ca

New opportunities
- Develop software to assist forecasters to handle data, increase situational awareness, and write TAFs
- Increase follow-up on verification statistics
- Develop new products
Prediction

Probabilistic forecast: 10 %ile to 50%ile cig. and vis. from analogs
CSI IFR, February-April 2005

CSI = hits / (hits + misses + false alarms), IFR flying category ⇒ Ceiling < 1000 feet or Visibility < 3 miles.

Statistics are comprehensive for 190 Canadian airports for period from February - April 2005.

- CVG-3 statistics from WIND forecasting system for ~350,000 24-hour forecasts made hourly.

WIND system forecasts ceiling and visibility using analog forecasting (data-mining and fuzzy logic).

- Data consists of current METARs, climatology (hourly obs from 1971-2004), and GEM-based MOS guidance (mainly for the 6-24 hour projection period) from CMC.
- For more details, visit: http://collaboration.cmc.ec.gc.ca/science/arma/bjarne/wind3
Questions?

E-mail: bjarne.hansen@ec.gc.ca
Webpage: www.cmc.ec.gc.ca/rpn/hansen