12.6 STATISTICAL WIND FORECASTING FOR ARCTIC LOCATIONS USING RECURSIVE PARTITIONING AND REGRESSION TREES

Andrew Giles*
EC, Edmonton, AB CANADA

1. INTRODUCTION

The Canadian Arctic is an area without many manned observations or upper air reports. As such, numerical models over this part of the world do not have a lot of surface or upper air data to assimilate. Figure 1 shows a map of the surface (red) and upper air observing stations (green). The yellow triangles indicate the locations of Clyde, NU (to the right on the map) and Paulatuk, NT.

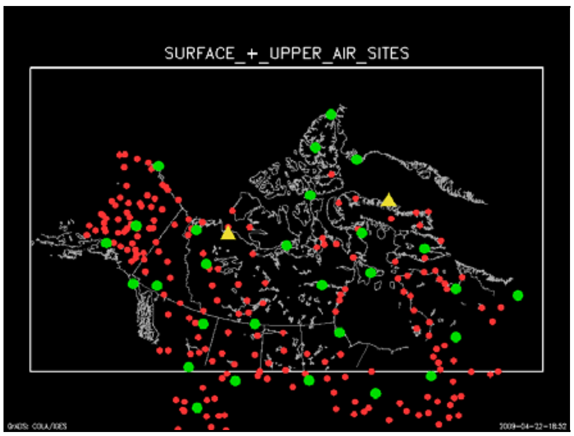

Figure 1 Surface and upper air observing sites

The Canadian Updateable Model Output System [UMOS] (Wilson and Vallée, 2002) has been running at the Canadian Meteorological Center (CMC) since 12 UTC, 14 September 2000 and is the main statistical processing system for the output of the Canadian GEM model. UMOS uses forecast model information and observed data in the development of the equations used in the UMOS forecast. UMOS is used as the principal statistical guidance for the SCRIBE forecast system (Verret et al, 2000). SCRIBE is now the principal tool used by all Environment Canada meteorologists across the country to produce their public forecasts.

2. RECURSIVE PARTITIONING AND REGRESSION TREES

Forecasters have noted over the past several years that SCRIBE / UMOS has difficulty forecasting winds over the Canadian Arctic. A project was put in place by the Hydrometeorological and Arctic Lab in Edmonton, AB to see if use of a technique known as Recursive Partitioning and Regression Trees could improve wind forecasts. This technique has successfully been used in producing lightning forecasts for Canada (Burrows et al, 2005) and synoptic map classification (Cannon et al, 2002).

The statistical package RPART was used in the production of the forecasts. The RPART technique involves training the forecast system using one set of data, with matched forecast and observed parameter pairs. In this case, observed wind data (speed and direction) from the winter months during 2005 and 2007 were used to train the system. A regression tree results from the training process. This nonparametric algorithm minimizes predictand variance through a series of decisions that use predictor threshold values to cluster groups of similar training data into a set of "terminal nodes." At the end of the tree branches are homogeneous groups, whose values are averages of the data in that group. Separate regression trees are produced for wind speed and wind direction.

The following figure shows a simple example of a regression tree – decisions resulting in the lower right node enhanced.

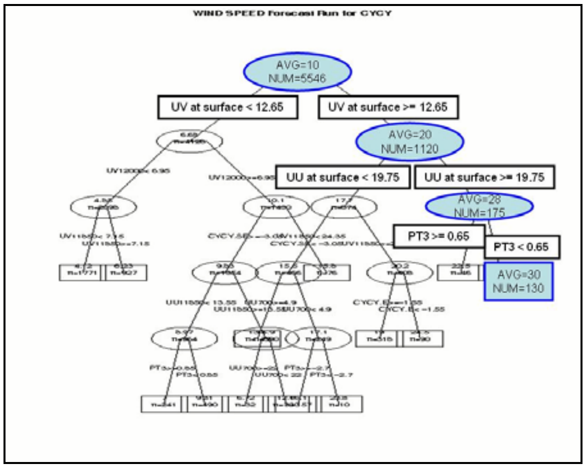

Figure 2 Simple tree structure for Wind Speeds in Clyde, NU

3. FORECAST SYSTEM

The system has been set up to run twice daily, upon reception of 00Z and 12Z forecast model data. Using the model parameters determined to be useful in the regression trees, a forecast is produced for about 20 sites across the Canadian Arctic. Forecast of wind speed and direction out to 48 hours are made available to the forecasters with the Prairie and Arctic Storm Prediction Centre and Canadian Meteorological Aviation Centre – West in Edmonton around 06Z and 18Z.

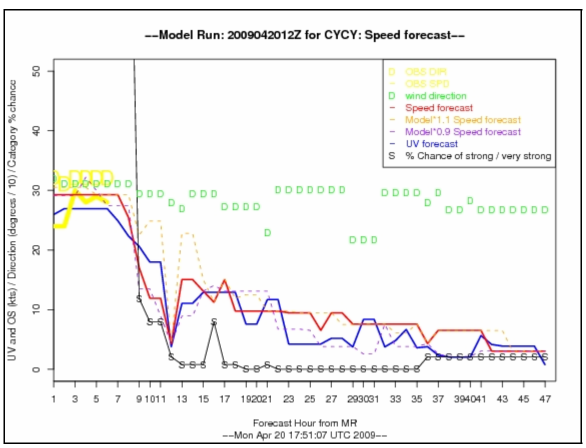

Figure 3 Example of forecast winds for Clyde, NU from 12Z April 20th, 2009

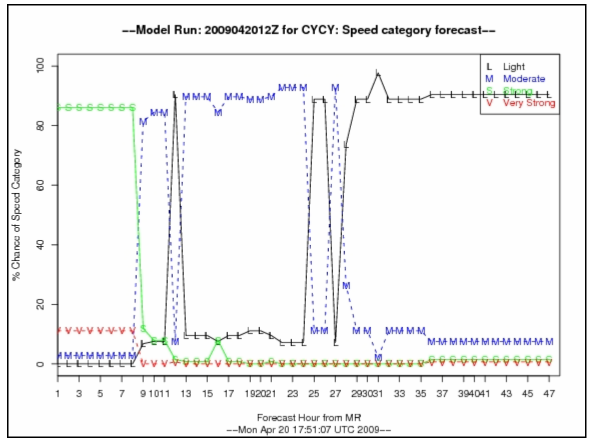

Figure 4 Forecast wind speed category for Clyde, NU from 12Z April 20th, 2009

4. Verification

Here is a scatter plot of wind speed forecasts vs. observed values for Clyde, NU for 2008. Note that there are more significant over-forecasts of wind speed than under-forecasts.

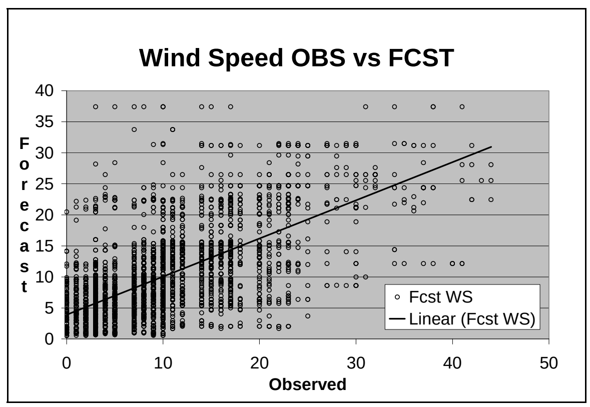

Figure 5 Forecast vs. Observed wind speed scatter plot: Clyde, NU : 2008

The following chart shows wind speed forecasts and observations for Clyde, NU for a period in the spring of 2008.

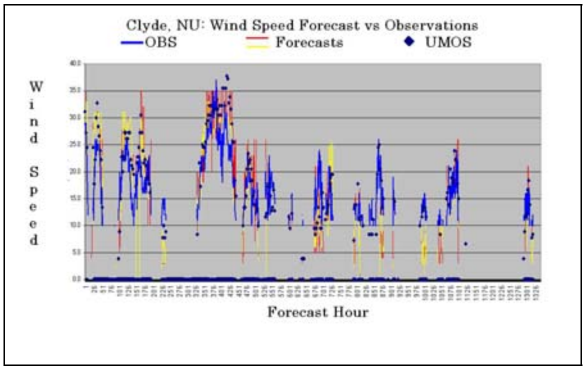

Figure 6 Wind Speed Forecast vs. Observations for Clyde, NU : Spring 2008

Verification of the UMOS forecasts was received from CMC in Dorval, QC. Verification of the RPART technique winds shows some improvement over the UMOS wind forecast, using the RMSE error scores.

		FORECAST HOUR			
	Forecast	**6**	**9**	**12**	**15**
Clyde, NU	***UMOS***	*6*	*7.1*	*7.5*	*7.1*
	RPART	5.73	6.3	6.62	6.23
	Forecast	**6**	**9**	**12**	**15**
Paulatuk, NT	***UMOS***	*7.9*	*8.9*	*9.5*	*9.3*
	RPART	6.24	7.02	6.82	6.99

Table 1 RMSE error scores - wind speed for Clyde, NU and Paulatuk, NT

5. Conclusions

Wind speed forecasts for various locations in the Canadian Arctic are being produced daily using a technique that involves some improvement over UMOS forecasts for Clyde, NU and Paulatuk, NT for 2008.

This forecasting tool is designed to assist forecasters in production of their forecasts, and not replace the UMOS forecast data.

Further verification is now underway for additional locations. Expansion to a larger number of sites is being considered.

6. References

Burrows, W.R., C. Price and L. J. Wilson, 2005: Warm season lightning probability prediction for Canada and the northern United States. *Weather and Forecasting,* 20, 971-988.

Cannon A. J., P. H. Whitfield and E. R. Lord, 2002: Synoptic map pattern classification using recursive partitioning and principal component analysis. *Monthly Weather Review* 130: 1187–1206.

Verret R., G. Babin, D. Vigneux, J. Marcoux, J. Boulais, R. Parent, S. Payer and F. Petrucci, 1995: SCRIBE an Interactive System for Composition of Meteorological Forecasts. *Preprints, 11th International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, AMS, Dallas, Texas, January 15-20, 1995, 56-61.

Wilson, L. J. and M. Vallée, 2002: The Canadian updateable model output statistics (UMOS) system: Design and development tests. *Weather and Forecasting*, **17,** 206–222.