FINAL REPORT for COMET Cooperative Project -- Improved frost forecasting through coupled artificial neural network time series prediction techniques and a frost deposition model

University: Iowa State University

Name of University Researcher Preparing Report: William A. Gallus, Jr.

NWS Office: DMX (Johnston, IA)

Name of Primary NWS Collaborator: Karl A. Jungbluth

Cooperative Project, UCAR Award No. S01-32791

1. Date: October 16, 2003

1 SUMMARY OF ORIGINAL PROPOSED SCOPE OF WORK:

The project originally proposed to develop a time-series prediction system for parameters needed as input by a frost deposition model (road temperature, air dew point, winds), by training a neural network on meteorological observations and model output. In addition, the project's objectives also included a comparison of RWIS observations with those of standard ASOS/AWOS instruments to determine the quality of the data. Plans also included working with the NWS office so that they could use the neural network system to improve forecasting of other parameters.

2 WORK ACCOMPLISHED, CHANGES IN SCOPE, PROBLEMS ENCOUNTERED

A) WORK ACCOMPLISHED:

i) ANN-forecasting system

After gathering Model Output Statistics (MOS) data from the NGM model for the three year period (1996-1998) matching our frost observation record period, we also collected an appropriate data set of RWIS measurements. We accessed the RWIS archive valid over this time period, and interpolated the data to a standard 20 minute time interval, performed some quality control of the data, and flagged those time periods when gaps between observations exceeded 4 hours, or could interfere with important time periods when frost was possible. The corrected and usable data from both MOS and RWIS observations, were organized into a very large database which was used to train the automated neural network (ANN). The training set used 85% of all of the data, with the remaining 15% saved as a test set. 181 possible inputs to the model were tested. It was decided that the best approach would be to limit the inputs to the 8-10 found to be most important for each parameter. All of this work was performed by a graduate student, Bradley Temeyer, supported under the project. Programs were developed to automatically collect the data in real-time with a regular 20 minute resolution for future applications.

Mr. Temeyer worked with Dr. Eric Bartlett, the designer of the automated neural network software package we used, and tested several different neural network configurations to predict roadway surface temperature, air temperature, air dew point, and wind speed, the four variables needed as input into the frost prediction model. Despite the artificial neural network's ability to predict multiple outputs with one model, they can only predict at one point in time, so it was therefore determined that the best way to predict the necessary parameters with a time frequency of 20 minutes was to create separate neural network models for each weather parameter. Thus, for the four weather parameters that were needed as input into the frost prediction model, where predicted values are needed every 20 minutes over a 24 h period, a total of 288 neural network models were needed. We tested many configurations to determine the best one for each parameter at each time. When validation was done over the entire training and test set, the ANN predictions correlated well with the RWIS observations and generally performed better than the NGM-MOS output alone. However, for the test set, the ANN predictions of temperature and wind speed were found to be better than MOS only for the first few hours of the prediction period (corresponding to forecasts out 6-9 hours), with MOS forecasts having smaller root mean square errors (RMSEs) later in time. The ANN had smaller RMSEs for wind speed, but the smaller average errors were achieved by always forecasting light winds, even when strong winds occurred. MOS does not include a forecast for roadway temperature, and thus the ANN system provided much better forecasts during daylight hours when compared to the use of the 2 meter MOS air temperature as a proxy for road temperature. At nighttime, the ANN showed a small improvement over the crude estimate based on MOS. In general, the errors with the ANN system were better than those obtained when 3 mesoscale models were run with two different sets of initial conditions over Iowa for the MDSS project in the winter of 2003. As in that project, dew point errors are much larger than temperature errors.

The ANN output was then tested in the frost deposition model for the winters of 1996-98, 2001-02 and 2002-03. The verification statistics for the winters of 1996-98 were provided by the DOT. During this time, yes and no observations of frost were made by maintenance personal looking out the windows of their vehicles on the way into work. Therefore, great uncertainty is present about the quality of the verification statistics. The ANN-based predictions of frost for this period were not particularly good. False alarm rates were very high (generally .90 - .98 at most stations tested), and probability of detection tended to be in the .2 to .6 range. Although the lack of skill in the forecasts was likely partly related to poor verification data, it was observed that the ANN system often predicted an air dew point exceeding the air temperature. This is a physically impossible condition, and it stems from the independent nature of the ANN models forecasting each parameter.

For the winter of 2001-02, frost verification data was available from an Iowa DOT-sponsored project in Ames, Iowa, where ISU students closely monitored a bridge in the region during the early morning hours. These yes and no frost observations are believed to be highly accurate, although frost is a very microscale phenomena, again complicating interpretation of verification data. For this period, the ANN predictions were very accurate, nearly matching the forecasts made after the fact when RWIS data from Ames were fed to the frost deposition model. Thus, for this period, the ANN forecast approached the "best" value that might be expected. False alarm rates, though still high, were around .56. The probability of detection was .57. Even more promising, the forecasts were much more skillful than those of the private forecasting firm that had been contracted by the Iowa DOT to forecast frost during this winter.

In the winter of 2002-03, the Ames frost verification project was expanded using AURORA funding, so that 3 bridges were monitored. This again complicated the picture concerning verification. Data for verification were not available until the end of the COMET project, and thus verification statistics are preliminary. It appeared that forecast skill was much less in 2002-03 than in the previous winter, with the FAR rising to .79, and the POD falling to .30. Nonetheless, the threat score, a combined measure of skill, was still comparable to that of the private firm the previous winter. No other verification data was yet available from 2002-03 to which the ANN could be compared.

In summary, the ANN forecasting system showed promise, and the performance during the 2001-02 winter was very encouraging. The fact the system earned much better scores than a contracted forecasting agency was also considered a success. However, it should be pointed out that some other winters exhibited much worse performance of the ANN system. The independent nature of the predictions for different parameters allows for some unphysical situations to be predicted. An improved ANN system in the future could predict the dew point depression instead of temperature and dew point as separate quantities. This would avoid the problem of unrealistic forecasts, and may help improve performance.

ii) RWIS-ASOS comparison:

Detailed comparisons of RWIS and AWOS/ASOS observations were performed. During the COMET project, a separate initiative was established at ISU to combine all mesoscale data into one convenient mesonetwork, known as the Iowa Environmental Mesonet. Several meetings occurred at ISU which included NWS-DMX personnel, Iowa Dept. of Transportation personnel, and other interested parties from state agencies. The establishment of the IEM greatly assisted our ability to compare RWIS observations to more standard meteorological observations of surface conditions, in particular, ASOS readings.

A thorough comparison was performed for several weather parameters by averaging RWIS and ASOS or AWOS data over an entire year. These comparisons suggest that RWIS temperature readings have a high bias when wind speeds are light, with the worst problem at the lightest wind. This behavior is what would be expected because RWIS sensors are not aspirated, and ASOS and AWOS are aspirated. In addition, RWIS winds were found to typically read a few knots less than AWOS/ASOS at the stations examined (where RWIS sensors were within 10 miles of AWOS or ASOS sensors). Although both platforms are designed to record wind at the same elevation (10 m), many RWIS platforms are sited in ditches which may shelter the instruments somewhat. At twelve stations, average wind speeds were computed for RWIS and ASOS instruments as a function of wind direction (octants). The impact of siting clearly be seen, with average wind speed differences often varying from around zero at a particular site for a given wind direction to roughly 3 knots for other directions (a roughly 40% decrease in magnitude from the ASOS values). For most directions at all sites, the RWIS values are less than the ASOS. Other comparisons of different parameters (such as dew point) as a function of hour of the day, or of wind speed, were performed as well. It appears that systematic differences can be quantified which should allow for calibration to improve use of the combined dataset. We are preparing a paper on the comparison to be submitted in the next few months (probably to the Bulletin of the Atmospheric Sciences), and some of the results were discussed at COMET/FHWA Meetings in Washington, D.C. in September 2002, and in Albany, NY in August 2003.

In addition to the detailed comparisons discussed above, a web site was set up that allows direct comparisons between the different measuring systems in real time, and it can be accessed at http://mesonet.agron.iastate.edu/compare/. Some results from these comparisons were published in the Preprints of the 18th International Conference on IIPS in Orlando Florida, held in January 2002.

iii) extension of ANN applications to the NWS

The graduate student also explored an additional use of neural networks designed to assist the DMX office in forecasting. A neural network predictive system for warm season QPF was designed. For this neural network system, the technique of Hall et al. (Weather and Forecasting, 1999) was followed, with a neural network being trained on data collected from the Eta model and soundings valid for both Omaha,Nebraska and Davenport, Iowa over the warm seasons of 1998-2001. Over 30 variables were used as input parameters in the model. Results using just the Omaha (OMA) data were very encouraging, so that the student added over 400 cases from Davenport (DVN) to expand the number of patterns used by the neural network. The expansion was determined to be necessary since early results implied some memorization or overparameterization might have been occurring. Results using the OMA and DVN data together indicated that the neural network system still performed better than either the operational Eta model, or 10 km Eta simulations in predicting 24-hour rainfall, particularly for heavier amounts that are generally unpredictable in grid point models, but the combining of the data did seem to harm the results some (errors were larger than using just OMA alone). Results from just the Omaha training were presented at the 18th International Conference on IIPS in Orlando Florida in January 2002. A discussion of all results was given at the United States Weather Research Program Science Symposium in Boulder, CO in April, and at the World Weather Research Programme's QPF Conference held in September 2002in Reading, United Kingdom. The neural network system based on the combined data was run in real-time during summer 2002, and results were made available to the National Weather Service via a web page. In this real time test period, the system was altered to allow inputs directly from the Eta model (instead of a mixture of model and rawinsonde output). This allowed for forecasts to be made at grid points across the state of Iowa. This output was also made available via the web page. The summer 2002 test, however, indicated a high bias in the neural network predictions. It appears that some calibration would be necessary to improve the forecasts.

B) CHANGES TO SCOPE:

There were no changes to scope of this project. We benefitted from having a running start on the neural network modeling and by the fact the IEM project allowed for easier than expected comparison of RWIS observations to ASOS and AWOS.

C) PROBLEMS ENCOUNTERED:

No serious problems were encountered in conducting the research, apart from some Windows OS-related computer crashes that occasionally caused data losses.

3) DIVISION OF LABOR

The primary method of interaction during this project occurred through occasional meetings, phone calls and emails with the DMX-NWS office and the Iowa DOT. The vast majority of the work, because it involved very site-specific neural network software, took place at the university. The DOT's involvement extended beyond the 2.5 years of this project, since they had provided us with the 1996-98 frost observations prior to the start of the COMET project.

4) RECOMMENDATION FOR FUTURE WORK RELATED TO THIS PROJECT:

The creation of a neural network forecasting system is very labor-intensive. Much of the project period was required simply to round up all necessary data, quality control it, and enter it into the appropriately-formatted input tables needed by the ANN software. Although extensive testing was performed to find the best neural network models for each parameter, there are enough degrees of freedom to the problem that additional improvements would be likely if more experimenting could be done with the ANN software. A future project that could benefit both the DOT and the NWS would be one in which the forecasting skill of a combined MOS-ANN system could be evaluated. In addition, it may be possible to run an ensemble of MOS and ANN systems. The mean would likely provide the most valuable input to the frost deposition model. An ongoing project at ISU has hinted that a mesoscale model such as MM5 likely has more skill than the NGM-MOS that was used to train our ANN. It would be very interesting to test the performance of an ANN system that used MM5 output for training instead of NGM-MOS. When this project began, there were only 3 years of RWIS data that could be used for training. That period has now nearly doubled. With all ANN systems, the more data to train on, the better the forecast skill. In summary, our project clearly showed that artificial intelligence systems such as neural networks may offer valuable improvements in forecasting, but obviously there is much room for improvement.

5) BROADER RECOMMENDATIONS:

Although we were able to follow closely our original scope of work and did not run into major problems, we still found it difficult to emphasize one core project that would benefit both the DOT and NWS partners significantly. In fact, of the 5 groups that received funding, we were the only group whose main focus was not basically the creation of a mesonet (merger of DOT data with NWS and other data sources). If one thinks about it, the creation of the mesonet is about the only project that seems to benefit both partners roughly equally. It should be pointed out that even in our project, we included a task involving mesonet applications -- the comparison of RWIS data with ASOS and AWOS. Probably the only reason our project was not like the other 4 was the fact that the Iowa DOT has a long history of being at the cutting edge of initiatives, so that we had already had access to their RWIS data prior to the COMET RFP and the establishment of the Iowa Mesonet was already well underway (probably the farthest along of any such effort in the nation). I believe any future combined COMET/FHWA projects will likewise run into a problem whereby unless the project is simply manipulating multiple data sources, the main research thrust will either primarily assist the DOT, or primarily assist the NWS, but not both. Traditional COMET projects benefit the NWS. I would suggest that it be understood for future collaborations where the FHWA is involved that there may be very little interaction and direct benefit to the NWS. For our project, the data comparison was believed to help the NWS, and we also let them know that they would be welcome to use the ANN system. However, the system was obviously designed to forecast road frost, something the NWS is not permitted to do. Thus, at best, the NWS would only ever be able to use the software system in a secondary manner.