7.4 PROGRESS ON A KNOWLEDGE-BASED INTERNET WEATHER FORECASTING SYSTEM

Harvey Stern*
Victorian Regional Office, Bureau of Meteorology, Australia

1. INTRODUCTION

In a paper presented to the 18th IIPS Conference, Stern (2002) described some early work towards the development of a knowledge-based system for the generation of internet weather forecasts.

The performance of the Stern (2002) system was evaluated (real-time) over one month (November 2001) and found to be superior to forecasts based on persistence and climatology. However, the system's predictions were somewhat inferior to the corresponding officially issued forecasts, especially for days 1 and 2.

For more details, readers may refer to:
http://www.weather-climate.com/internetforecasts.ppt

2. DISCUSSION

A subsequent analysis of the verification data suggested that there were three main reasons for this inadequate performance:

- Firstly, the forecasts were largely a function of the broad-scale synoptic pattern suggested by the NWP models;
- Secondly, the system failed to take into account NWP model forecasts of the thermal pattern; and,
- Thirdly, the system failed to take into account NWP model forecasts of moisture in the troposphere.

The current paper describes how the knowledge-based system has been modified in order to address the aforementioned deficiencies.

It also presents the results of a verification exercise on developmental data. This exercise demonstrates that potential for a significant increase in the accuracy of the forecasts has been achieved.

Corresponding author address: Harvey Stern, Victorian Regional Office, Bureau of Meteorology, Box 1636M, Melbourne, 3001, Australia; e-mail: H.Stern@bom.gov.au

3. MODIFICATIONS

The 2002 system operated by producing its predictions from a restricted number of discrete "forecast sets". The set that was chosen (by the system) was largely determined by the particular synoptic pattern suggested by the selected NWP model.

The 2003 modification utilises regression analysis to allow predictions to be selected from a continuous array of possible forecasts.

The particular form of regression analysis employed is parameter enveloping (Stern, 1994; 1996). Parameter enveloping allows definition of how the various predictors impact upon, or envelope, the influence (on a predictand) of other predictors.

The process of parameter enveloping provides a means to explicitly interpolate between elements in a data set. In its simplest form, it may be illustrated as follows:

- Suppose that one needs to derive a regression equation to forecast a value for predictand F using predictors P and Q.
- Linear regression, if applied to the data, leads to an equation of the form
  $F = a + bP + cQ$ [a, b, c are constants]   (1)
- In order to establish the spectrum of prediction equations (non-linear effects described by some predictors involving the product of one or more other predictors), that might describe the variation in the value of the constants with respect to the time of the year, defined by the sine and cosine of the month, S and C, a new regression equation is derived that includes as predictors all the terms in the product $(1+S+C) \times (1+P+Q)$.
- This leads to a prediction equation of the form:
  $F = a + bP + cQ +$
  $dS + eC + fSP + gSQ + hCP + iCQ$
  [$a, b, c,$ … etc. are constants]   (2)

- Hence, the prediction equation provides "envelopes" which define how the time of the year affects how the other predictors affect the value of the predictand. In the present case, the partial derivative of F with respect to P is $b+fS+hC$

An illustration of this process is presented in Fig 1.
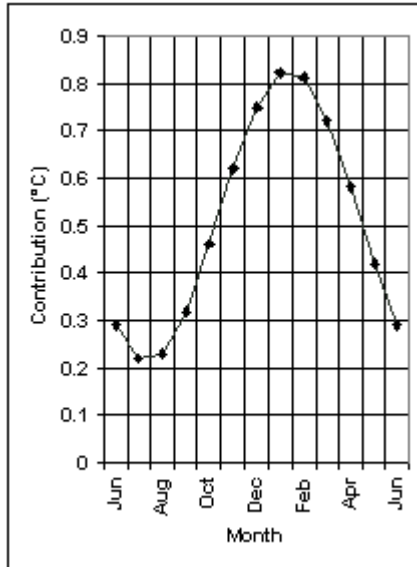
Fig 1 Seasonal variation of the contribution towards the maximum temperature, of a 1°C rise in the 850hPa temperature, for the synoptic type "Moderate NNW Anticyclonic".

4. SPECIFICATIONS

Terms in the *maximum temperature regression equations* include combinations of:
- sine and cosine of the day of the year;
- yesterday's maximum temperature; and,
- NWP model predictions of 850 hPa temperature and precipitable water.

Terms in the *minimum temperature regression equations* include combinations of:
- sine and cosine of the day of the year;
- yesterday's maximum temperature; and,
- NWP model predictions of precipitable water.

Terms in the *probability of precipitation equations*, and also the *amount of precipitation equations*, include combinations of:
- sine and cosine of the day of the year; and,
- NWP predictions of precipitable water at the beginning and end of the forecast period.

In operation, if the solution to the *probability of precipitation equation* is <50%, then, regardless of the solution to the *amount of precipitation equation*, that amount is "set" equal to "nil".

5. DEDUCING PRECIPITABLE WATER FROM THE GLOBAL PREDICTION SYSTEM

Our global prediction system provides forecasts of relative humidity at 700 hPa (RH), and temperature at 850 hPa (T), but not (directly) of total precipitable water (W).

Fig 2 depicts the relationship between precipitable water and the temperature at 850 hPa at Melbourne. It demonstrates that the relationship between the precipitable water of a fully saturated atmosphere and the 850 hPa temperature is approximately:

$$W=20+1.5T \tag{3}$$

Hence, cognisant of the facts that:
- dry air masses are associated with unrepresentatively low humidity at 700hPa (that level usually being above the inversion); and,
- the atmosphere is rarely fully saturated;

one might assert that the relationship between W and T and RH would be approximately:

$$W=(\max(\text{MIN}(RH,90\%),30\%))\times(20+1.5T) \tag{4}$$

Equation (4) may, therefore, be used to deduce the NWP model forecast of total precipitable water.

6. VERIFICATION ON DEVELOPMENTAL DATA

The 2003 modification utilises 40 years of data (1961-2000), made up of 14,610 individual synoptic situations, in its development.

The data is stratified into a set of 50 synoptic types (as defined by Treloar and Stern, 1993; Stern and Parkyn, 1999), utilising NCEP data and a synoptic-typer interface (Dahni, 2003).

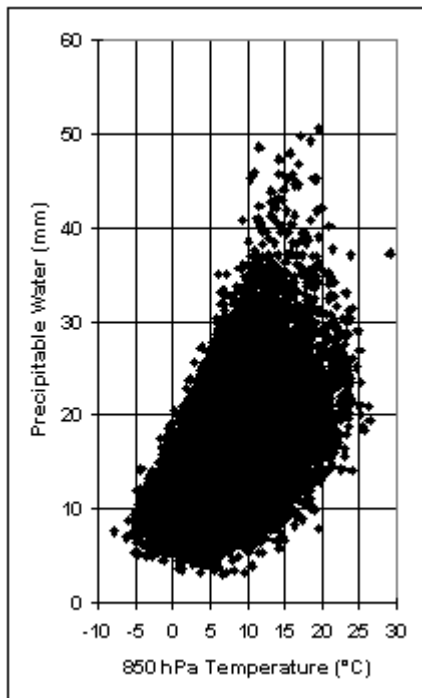Regression analysis is then carried out on data associated with each of the synoptic types.
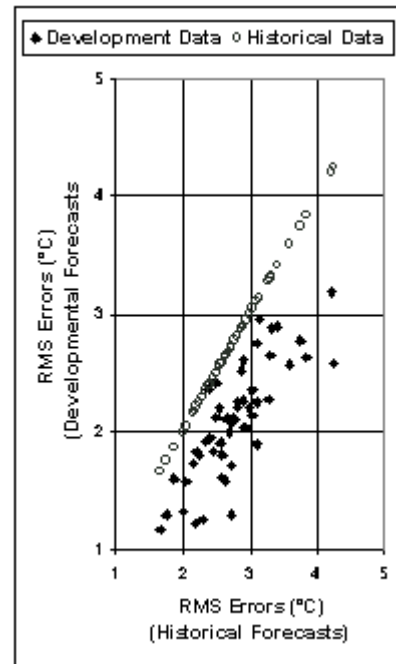

Fig 2 Relationship between precipitable water (mm) and 850 hPa temperature at Melbourne.

Maximum temperature forecasts for Melbourne (generated from the developmental data set) display greater accuracy than that achieved historically by official day-1 predictions, for every one of the synoptic types (Fig 3). This may be deduced because all of the diamonds lie below the "no-improvement" line of circles, thereby suggesting that for all synoptic types, the developmental forecasts are better than the historical forecasts.

However, it should be pointed out that, over the years, there has been an increase in the accuracy of the official day-1 temperature forecasts.

To illustrate, the RMS error of day-1 maximum temperature forecasts over the last 10 years, 1991-2000, was 2.23°C, in comparison with 2.69°C for the entire 1961-2000 period (Dawkins and Stern, 2003).

7. VERIFICATION ON REAL-TIME DATA

It is planned to carry out an evaluation of the performance of the modified system real-time, and over one month, in November 2002, and to compare its performance with that of the "pilot" system in November 2001.

Readers interested in detailed results of this evaluation may refer to:
http://www.weather-climate.com/internetforecasts.html

Fig 3 Accuracy of developmental data forecasts in comparison with historical forecasts, for each synoptic type.

8. CONCLUSION

A description of how Stern's (2002) knowledge-based system was modified has been presented. The result of a verification exercise carried out on developmental data demonstrates potential for a significant increase in forecast accuracy.

9. REFERENCES

Dahni, R. R., 2003: An automated synoptic typing system using archived and real-time NWP model output. *19th Conference on Interactive Information Processing Systems*, *Amer. Meteor. Soc.,* Long Beach, California.

Dawkins, S. S. and Stern, H., 2003: Trends and volatility in the accuracy of temperature forecasts. Submitted to the *7th International Conference on Southern Hemisphere Meteorology and Oceanography*, *Amer. Meteor. Soc.,* Wellington, New Zealand.

Stern, H., 1994: Enveloping predictor parameters in a regression equation - an alternative to model output statistics (MOS). *10th Conference on Numerical Weather Prediction*, *Amer. Meteor. Soc.,* Portland, Oregon.

Stern, H., 1996: Statistically based weather forecast guidance. *Ph. D. Thesis, School of Earth Sciences, University of Melbourne (subsequently published by Bur. Met., Australia, in 1999, as Meteorological Study 43).*

Stern, H. and Parkyn, K., 1999: Predicting the likelihood of fog at Melbourne Airport. *8th Conference on Aviation, Range and Aerospace Meteorology, Amer. Meteor. Soc.,* Dallas, Texas.

Stern, 2002: A knowledge-based system to generate internet weather forecasts. *18th Conference on Interactive Information and Processing Systems*, *Amer. Meteor. Soc.,* Orlando, Florida.

Treloar, A. B. A.. and Stern, H., 1993: A climatology and synoptic classification of Victorian severe thunderstorms. *4th International Conference on Southern Hemisphere Meteorology and Oceanography*, *Amer. Meteor. Soc.,* Hobart, Australia.