270 COMPUTER PROGRAMS WITH THE CAPACITY TO READ AND MANIPULATE DATA WITHIN WEB DOCUMENTS AS A VEHICLE TO THE AUTOMATIC GENERATION OF MORE ACCURATE WEATHER AND CLIMATE FORECASTS

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

1. INTRODUCTION

1.1 *Background*

Stern *et al.* (2011) explored how new technologies might be harnessed to integrate material from various sources on the web to generate new products.

Berners-Lee (2010) nominates linked data as “a great example of (the web's) future promise (and suggests that) today's web is quite effective at helping people publish and discover documents (but that) our computer programs cannot read or manipulate the actual data within those documents".

He observes that “as this problem is solved, the web will become much more useful, because data about nearly every aspect of our lives are being created at an astonishing rate (and that) locked within all these data is knowledge about how to cure diseases, foster business value and govern our world more effectively".

He notes that “scientists are actually at the forefront of some of the largest efforts to put linked data on the web”.

1.2 *Purpose*

It is the primary goal of the work presented here to develop within computer programs the aforementioned “capacity to read and manipulate the actual data within web documents”.

By this means, the value of the data is fully realised via the automatic generation of a broad range of more accurate weather and climate forecasts and other products.

The specific purpose of the present paper is to provide an update on the work described by Stern *et al.* (2011), in particular that work dedicated to developing a seamless (across time scales - Day 1 to 14, monthly, seasonal, decadal) weather and climate forecasting framework.

*Corresponding author address: Harvey Stern, Victorian Regional Office, Bureau of Meteorology, Box 1636, Melbourne, 3001, Australia; e-mail:

2. ‘REAL TIME’ TRIAL

2.1 *History*

A "real time" trial of a methodology utilised to generate Day-1 to Day-7 forecasts, by mechanically integrating (that is, combining) judgmental (human) and automated predictions, has been ongoing since 20 August 2005.

Since 20 August 2006, forecasts have also been generated for beyond Day-7 (out to Day-10). Since 18 January 2009, forecasts have also been generated out to Day-14.

2.2 *Results*

The verification data demonstrate how the combining process improves the accuracy of the ‘raw’ official forecasts for Melbourne – temperature predictions (Figure 1) in 71 of the 77 months (August 2005 to December 2011); and, rainfall predictions (Figure 2) in 58 of the 77 months, albeit to a lesser extent during the recent exceptionally wet period.

The data reflect the greater potential for an increase in accuracy in temperature forecasts during the summer months (when the variability of temperature is larger) than during the winter months (Figure 3).

The data also reflect the greater potential for an increase in accuracy in rainfall forecasts during the summer months (when the variability of rainfall is larger) than during the winter months (Figure 4).

The annual cycle is evident from the 5-month moving averages shown in Figures 1 and 2, with the improved skill at forecasting both rainfall and temperature peaking during the summer (more changeable) time of the year, with one exception. That exception is during the recent wet period, when the 'combined forecasts system' actually made the rainfall forecasts worse - the unusual weather pattern was identified well by the human forecasters, but not by the 'combined forecasts system'.

There is evidence of an overall increase in the accuracy of the official forecasts. This is illustrated by:

(1) The extent to which the mean square error in the temperature forecasts was able to be decreased by the combining process (from 1.00 deg C, on average, during the first three

years, to 0.48 deg C, on average, during the most recent three years; and,

(2) The extent to which the % correct rain/no rain forecasts was able to be increased by the combining process (from 6.30%, on average, during the first three years, to 2.76%, on average, during the most recent three years.

That the 'combined forecasts system' was unaltered during the period, whilst at the same time the extent to which the system was able to improve upon the 'raw' official forecasts diminished, is evidence that there was a ‘real’ increase in the skill displayed by the official forecasts.

Regarding the verification data for the Melbourne experimental Day 1-14 forecasts, positive correlation coefficients between forecast and observed minimum and maximum temperature, and amount and probability of precipitation, indicate the presence of some worthwhile skill out to Day-10 for all four forecast elements - (but also with some limited skill beyond) (Figure 5).

Accompanying the Day 1-14 forecasts are both a monthly and a seasonal climate outlook. Preliminary correlation coefficients (forecasts for the coming month in Melbourne versus observed for June 2009 to December 2011) are +0.09 (rainfall), +0.06 (min temp) and +0.03 (max temp).

3. FUTURE DEVELOPMENT

Regarding future development, worthy of consideration is a seamless (across time scales) framework whereby predictions are generated for key population centres (e.g. the eight State and Territory capitals, Adelaide, Brisbane, Canberra, Darwin, Hobart, Melbourne, Perth, Sydney, plus Broome - representing northern Western Australia, Alice Springs - representing Central Australia, and Cairns - representing northern Queensland); on a regional forecast district basis; and, on a State by State (& Territory) basis.

4. REFERENCES

Berners -Lee, T. 2010: Long Live the Web: A Call for Continued Open Standards and Neutrality. *Scientific American Magazine, December 2010.*

Stern, H., Campbell, B., Efron, M. and Cornall-Reilly, J., 2011: Towards the effective communication of weather and climate information - harnessing new technologies to integrate material from various sources on the web to generate new products. *9th Conference on Artificial Intelligence and its Applications to the Environmental Sciences, 23-27 January 2011, Seattle WA.*

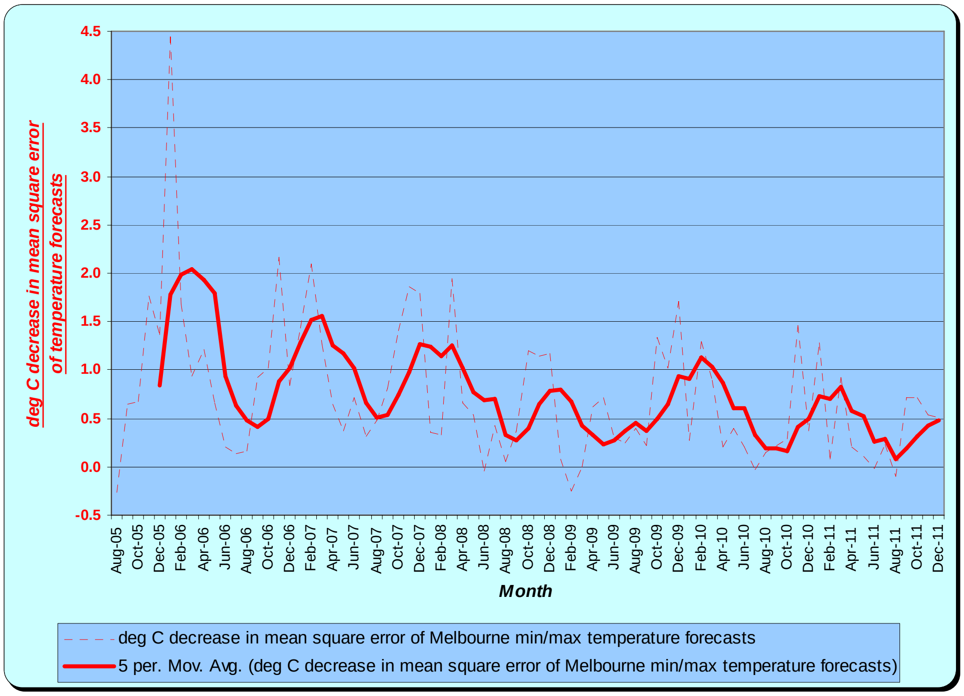

Figure 1 The extent to which the combined forecasts system was able to improve upon the official forecasts (August 2005 to December 2011) at predicting temperature.

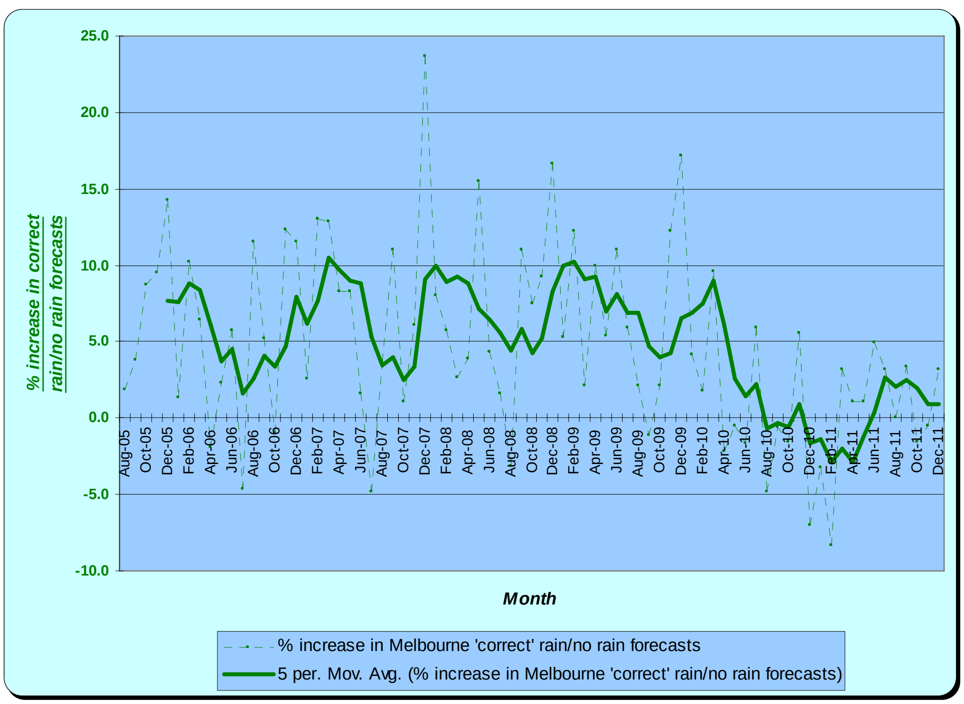

Figure 2 The extent to which the combined forecasts system was able to improve upon the official forecasts (August 2005 to December 2011) at predicting rainfall.

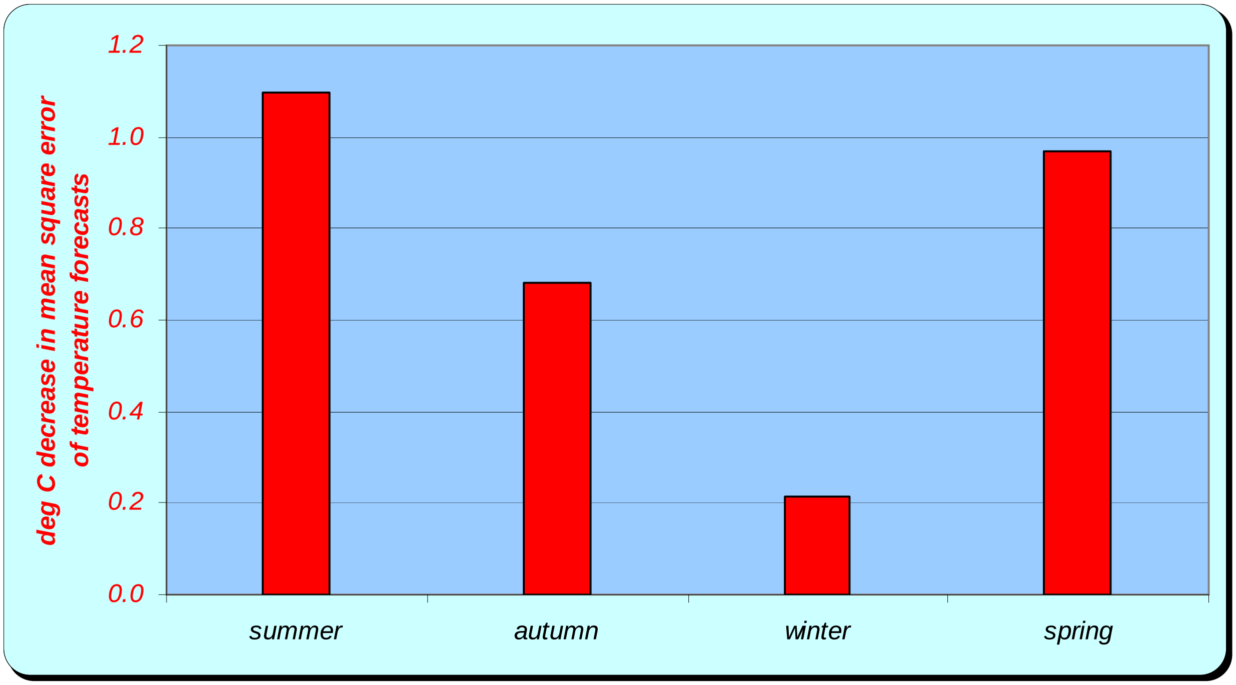

Figure 3 Seasonal variation in the extent to which the combined forecasts system was able to improve upon the official forecasts (August 2005 to December 2011) at predicting temperature.

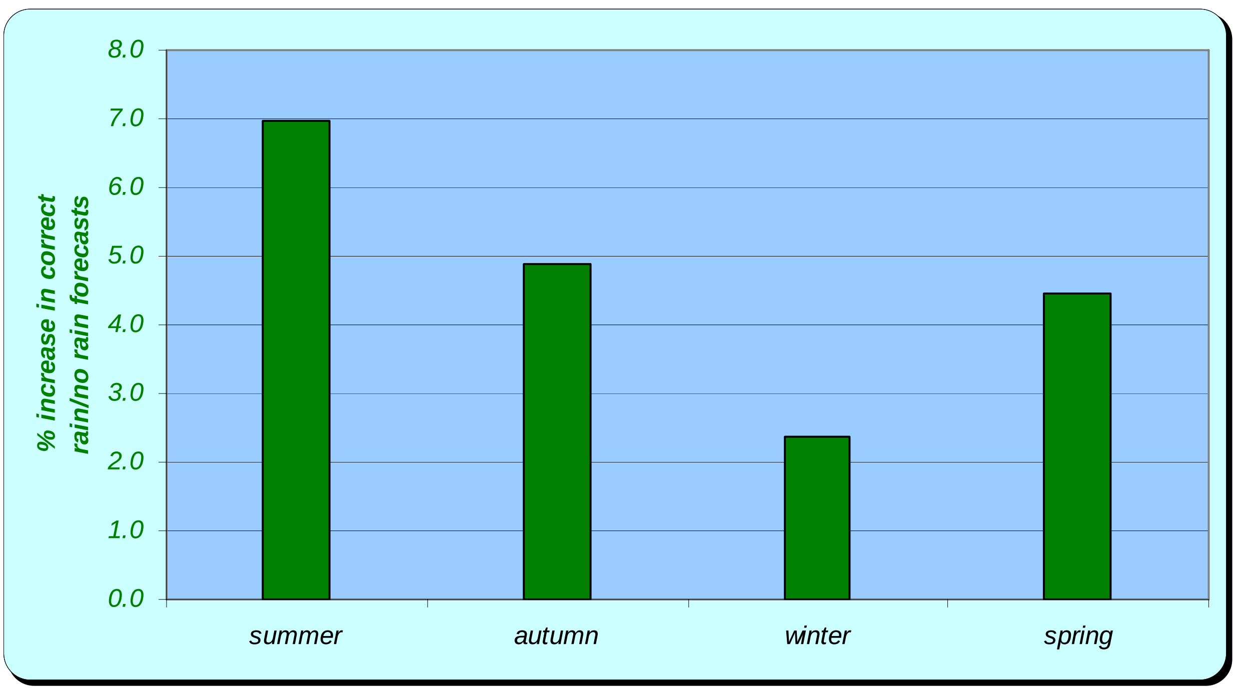

Figure 4 Seasonal variation in the extent to which the combined forecasts system was able to improve upon the official forecasts (August 2005 to December 2011) at predicting rainfall.

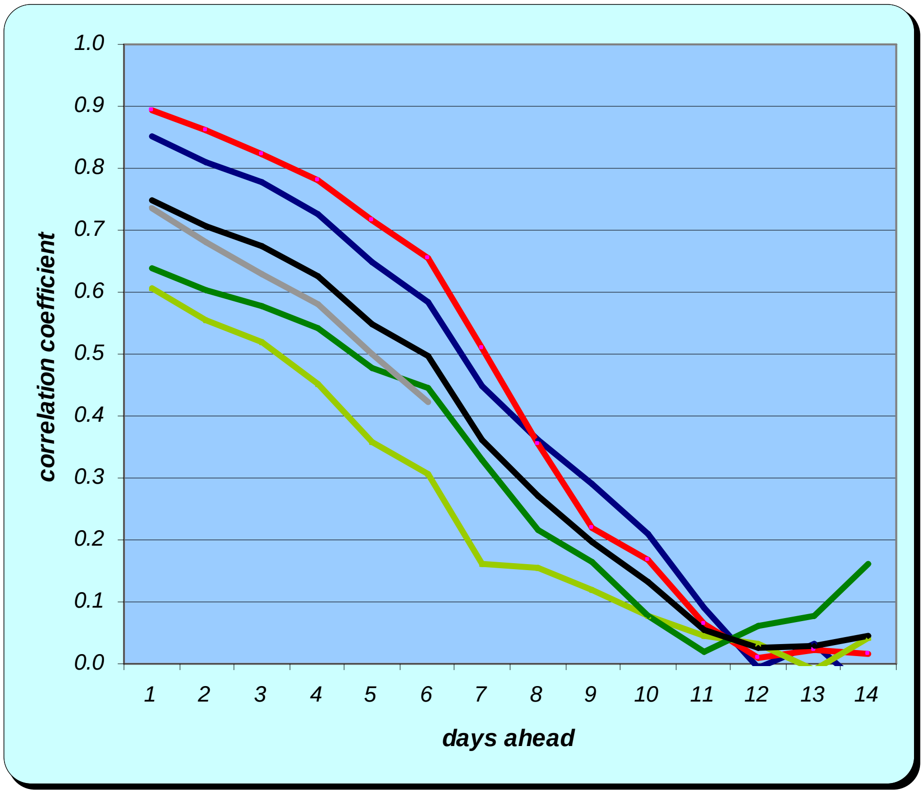

Figure 5 Positive correlation coefficients between forecast and observed minimum and maximum temperature (blue and red lines, respectively), amount and probability of precipitation (lime and green), and the mean correlation coefficient for all four elements for the combined forecasts system (black), indicate the presence of some worthwhile skill out to Day-10 - (but also with some limited skill beyond). The mean correlation coefficient for the official forecasts is in grey.