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

Australian Government
Bureau of Meteorology

**Corresponding author address:* Harvey Stern, Bureau of Meteorology, Box 1636, Melbourne, 3001, Australia; e-mail: h.stern@bom.gov.au

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". Berners-Lee 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". 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.

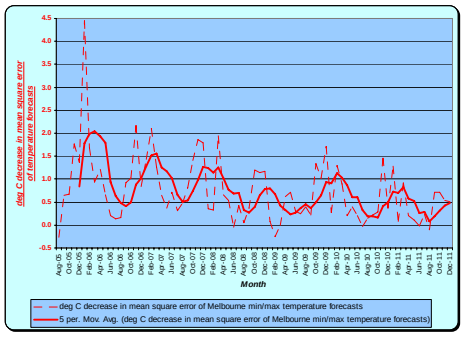

Figure 1 The extent to which the combined forecasts system was able to improve upon the official forecasts (August 2005 to December 2011) at predicting temperature.

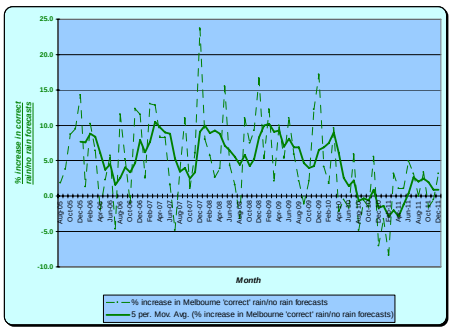

Figure 2 The extent to which the combined forecasts system was able to improve upon the official forecasts (August 2005 to December 2011) at predicting rainfall.

Increasing forecast accuracy

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.

There is evidence of an overall increase in the accuracy of the official forecasts. This is illustrated by the extent to which both the temperature and rainfall forecasts were able to be improved by the combining process. 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.

The accuracy of the Day 1-14 predictions

The graphic below displays (very) preliminary verification data for Melbourne experimental Day 1-14 forecasts. Positive correlation coefficients indicate the presence of some worthwhile skill out to Day-10 for all four forecast elements - minimum and maximum temperature, amount and probability of precipitation (but also with some limited skill beyond).

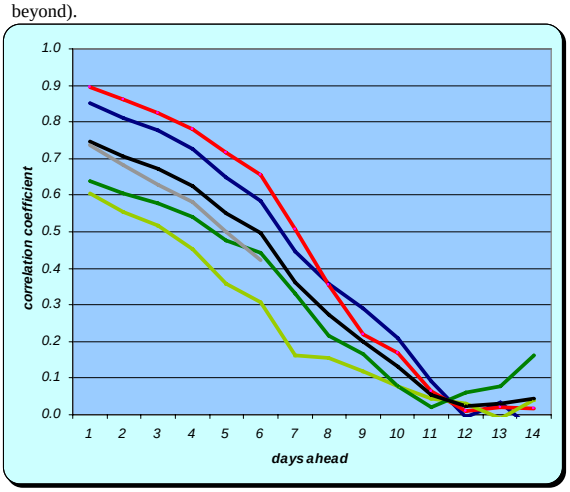

Figure 5 Correlation coefficients for forecast and observed minimum and maximum temperature (blue and red lines, respectively), amount and probability of precipitation (lime and green), and the mean correlation coefficient for all four elements (black), with the mean correlation coefficient for the official forecasts shown in grey.

Monthly predictions

Accompanying the Day 1-14 forecasts are both a monthly and a seasonal climate outlook.

Preliminary correlation coefficients (monthly forecasts versus observed between June 2009 and December 2011) are +0.09 (rainfall), +0.06 (min temp) and +0.03 (max temp).

Regarding future development, worthy of consideration is a seamless (across time scales) framework whereby predictions are generated for key population centres:

-namely, 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); and also,

- on a regional forecast district basis; and, on a State by State (& Territory) basis.

Seasonal variation in accuracy of temperature forecasts

The graphic below reflects 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.

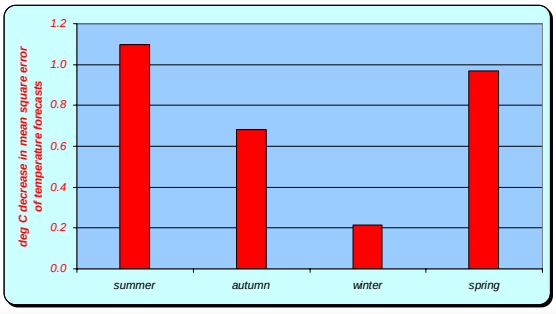

Figure 3 Seasonal variation in the extent to which the combined forecasts system was able to improve upon the official forecasts (August 2005 to December 2011) at predicting temperature.

Seasonal variation in accuracy of rainfall forecasts

The graphic below reflects 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.

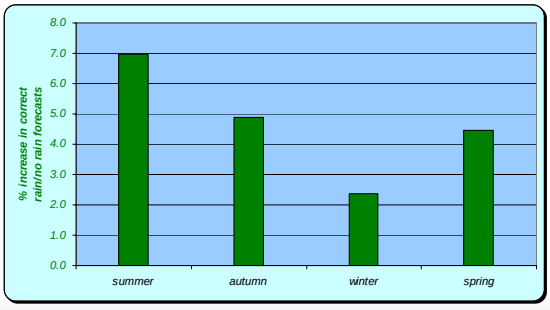

Figure 4 Seasonal variation in the extent to which the combined forecasts system was able to improve upon the official forecasts (August 2005 to December 2011) at predicting rainfall.