COMPARING THE SKILL DISPLAYED BY TWO STATISTICAL SCHEMES THAT INTERPRET THE ECMWF ENSEMBLE PREDICTION SYSTEM CONTROL MODEL AND THE NCEP GLOBAL FORECAST SYSTEM (GFS) MODEL

Harvey Stern*
School of Earth Sciences,
University of Melbourne, Victoria, Australia

1. INTRODUCTION

Weather forecasters have access to a number of numerical weather prediction (NWP) models and a range of statistical systems to interpret their output.

2. BACKGROUND

Stern and Davidson (2015) provide a comprehensive analysis of long-term trends in the skill of weather prediction for Melbourne, Australia, at lead times of 1 to 14 Days, and an update (to 31 October 2017) is given at Figure 1.

Figure 1 illustrates how, for most forecast variables and for most lead times, current performance levels of skill are very close to being as high as they ever have been.

For example, regarding the percentage variance about the seasonal norm for Day-1 forecasts explained, minimum temperature predictive skill has lifted from about 55% to 85%, maximum temperature forecast skill has lifted from about 55% to 90%, Quantitative Precipitation Forecast (QPF) skill has lifted from 30% to 60%, whilst Probability of Precipitation (PoP) forecast skill has lifted from 35% to 50%.

3. PURPOSE

The primary purpose of the current paper is to compare the relative skill displayed by statistical systems used to generate predictions for Melbourne, Australia, when applied to the output of two NWP models.

The two NWP models subjected to this evaluation are the *ECMWF Ensemble Prediction System Control Model* and the *NCEP Global Forecast System (GFS) Model*.

The predictions evaluated are, for the *Melbourne City* observation site:

- Estimates of the inter-diurnal change in minimum and maximum temperature (*Temp*); and,
- Estimates of the inter-diurnal change in the amount and probability of precipitation (*Precip*).

In addition to, for the Melbourne Airport observation site:

- Estimates of the inter-diurnal change in 9am and 3pm wind direction and speed (*Wind*).

4. ANALYSIS

4.1 Combining predictions

Figure 2 illustrates how combining forecasts derived from two different NWP models increases the overall (Day 1-10) level of skill (% inter-diurnal variance explained by the predictions) to a level superior to that of either of the models taken separately.

Figure 3 is similar to Figure 2, but for aims to illustrate the results for specific lead times between Day-1 and Day-14, and also for Week-1 predictions taken as a whole, Day 8-10 predictions taken as a whole, and also for Day 11-14 predictions taken as a whole.

Combining predictions is shown to lead to forecasts of greater overall skill (*Temp*, *Precip*, & *Wind*) than the individual sets for most lead times.

4.2 Comparison with official predictions

Figure 4 is also similar to Figure 3, but additionally aims to compare the overall skill (*Temp* & *Precip*) displayed by the official forecasts for Days 1 to 7, with the others.

Combining predictions is shown to lead to forecasts of greater overall skill (*Temp, Precip*) than the official sets for most lead times between Day-1 & Day-7 (the exception being Day-1), but note that the raw Day-1 to Day-7 ECMWF & GFS predictions involve the official *Temp.* and *Precip.* forecasts in their generation, and that the official wind forecasts were not readily available for this evaluation.

Figure 5 and Figure 6 are similar to Figure 4, but respectively evaluate the skill displayed by the *Temp* & *Precip* forecasts.

* *Corresponding author address: Harvey Stern, School of Earth Sciences, University of Melbourne, Parkville, Victoria, Australia, 3010;* e-mail: hstern@unimelb.edu.au

Figure 7 evaluates the skill displayed by the ECMWF & GFS *Wind* forecasts. The *Wind* forecasts arising from the two models' output combined clearly outperform those of either model taken separately.

7. SUMMARY

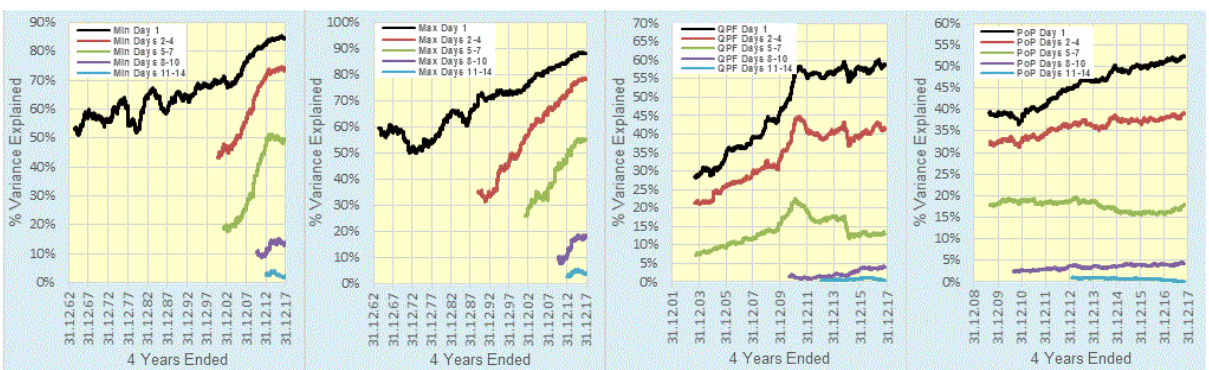
For most forecast variables and for most lead times, current performance levels of skill are very close to being as high as they ever have been.

Combining forecasts derived from two different NWP models increases the overall level of skill to a level superior to that of either of the models taken separately and also leads to forecasts of greater overall skill than the official sets for most lead times

7. REFERENCE

Stern, H. and Davidson, N. E. (2015), Trends in the skill of weather prediction at lead times of 1–14 days. *Q.J.R. Meteorol. Soc*., 141: 2726–2736. doi:10.1002/qj.2559

(a) Variance about the seasonal norm explained:



(b) Inter-Diurnal Variance explained:

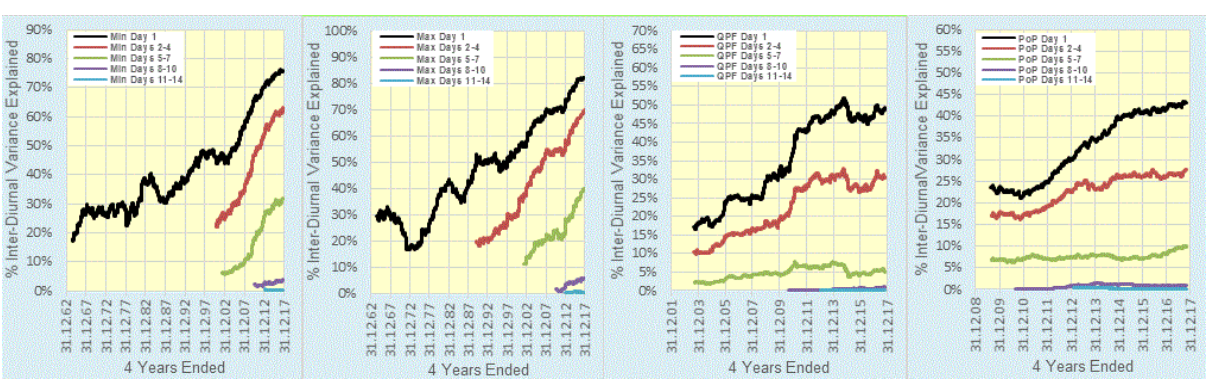


FIGURE 1

Long-term trends in the skill of weather prediction for Melbourne, Australia, at lead times of 1 to 14 Days. The variance about the seasonal norm explained, Figure 1 (a), is the correlation coefficient between the sets of departures from normal, of forecast values and observed values, squared. The inter-diurnal variance explained, Figure 1 (b), is the correlation coefficient between the sets of changes from one day to the next, of forecast values and observed values, squared.

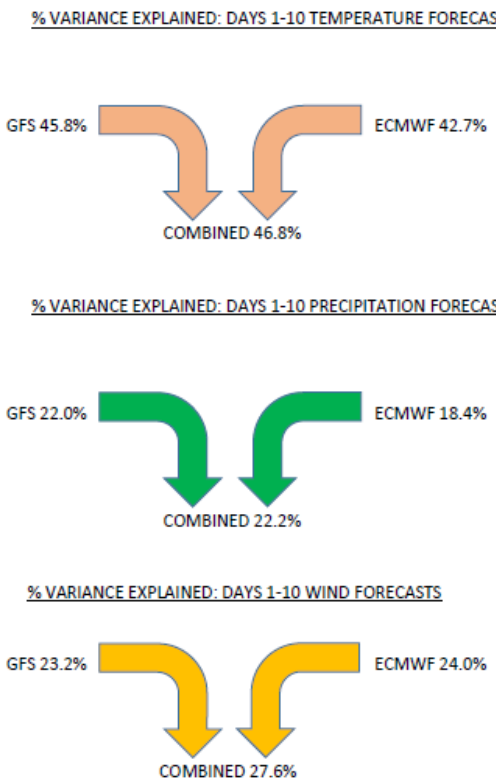


FIGURE 2

An illustration of how combining forecasts derived from two different NWP models increases the level of skill (% inter-diurnal variance explained) to a level superior to that of either of the individual models taken separately.

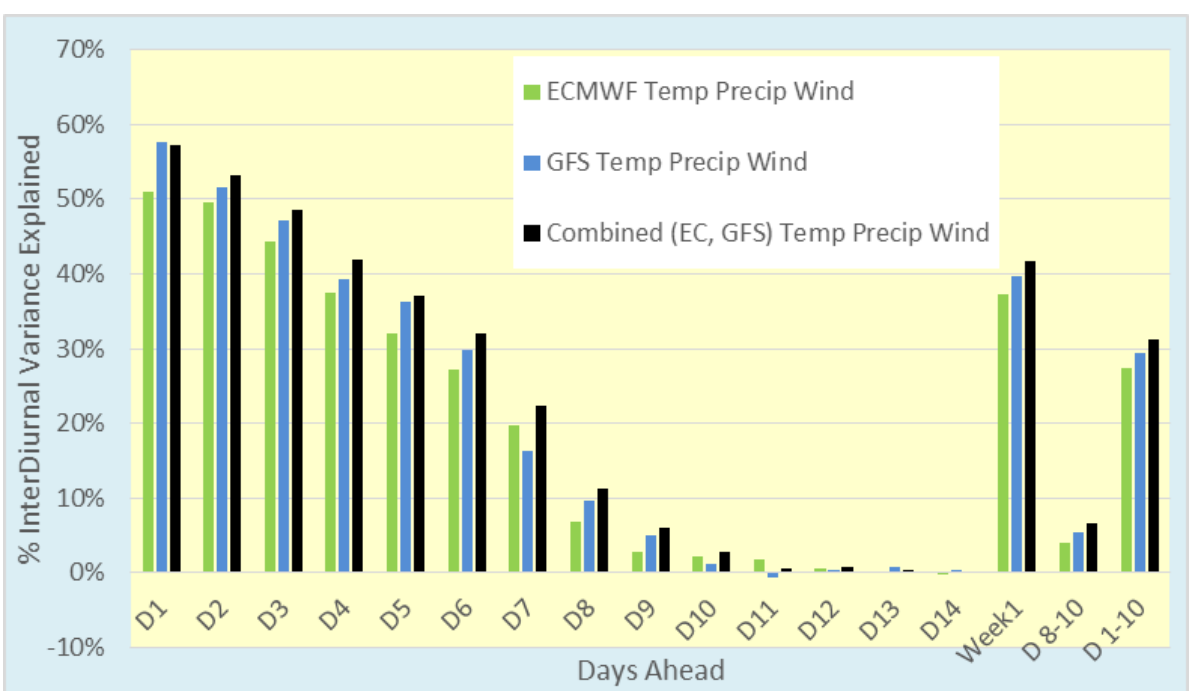


FIGURE 3

Another illustration (as per Figure 2) of how combining forecasts derived from two different NWP models increases the level of overall (*Temp*, *Precip*, & Wind) forecast skill (% inter-diurnal variance explained) to a level superior to that of either of the individual models *taken* separately., but for specific lead times Day-1, Day-2, Day-3, - - - - - , Day-14 and also for Week-1, Days 8-10 and Days 1-10.

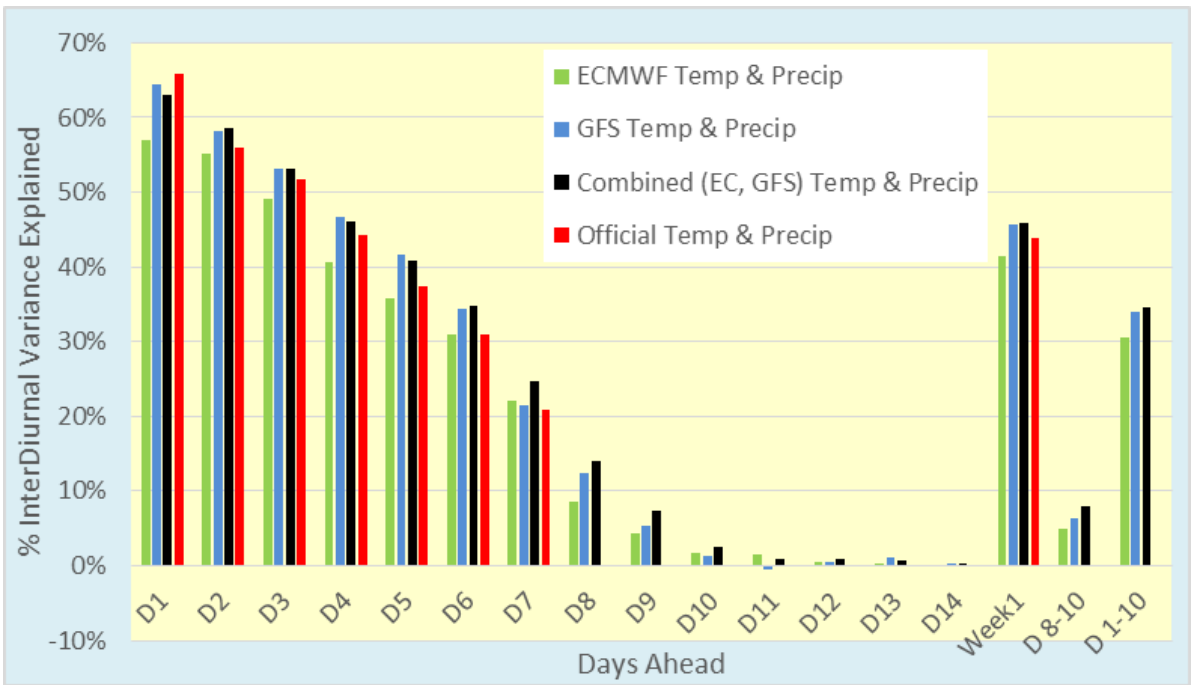


FIGURE 4

As per FIGURE 3, but with the additional aim of comparing the overall (*Temp & Precip*) skill displayed by the official forecasts with that displayed by the two models, and also the two models combined.

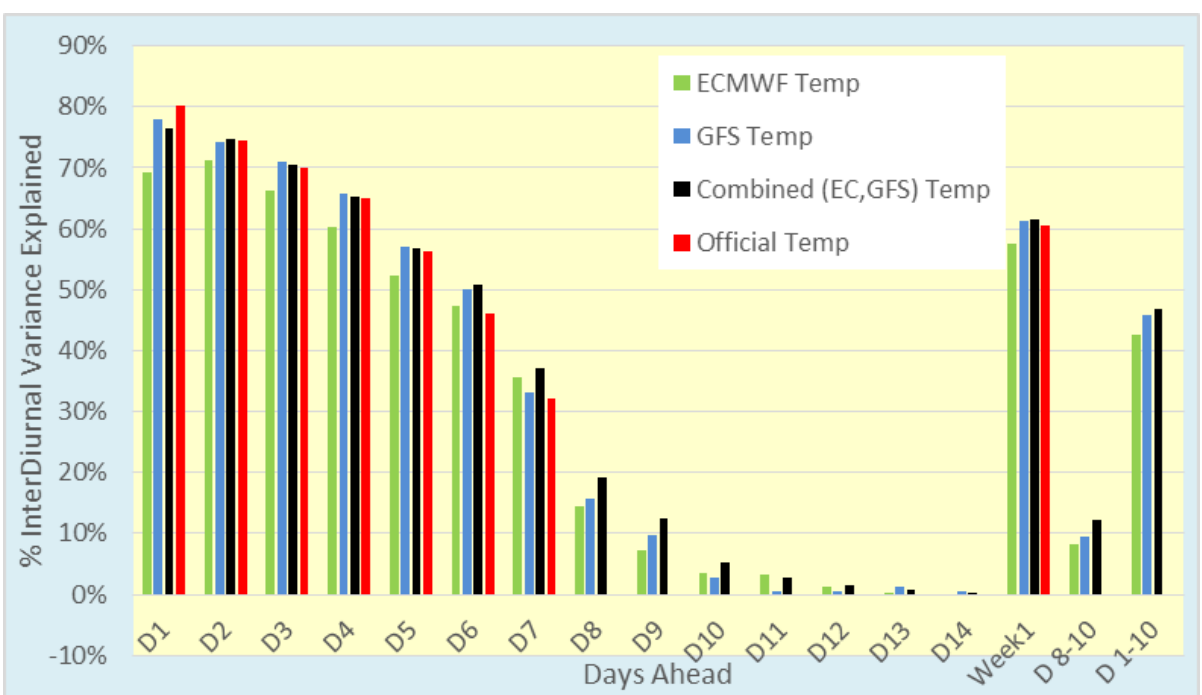


FIGURE 5

As per FIGURE 3, but with the specific aim of comparing the *Temp* skill displayed by the official forecasts with that displayed by the two models, and also the two models combined.

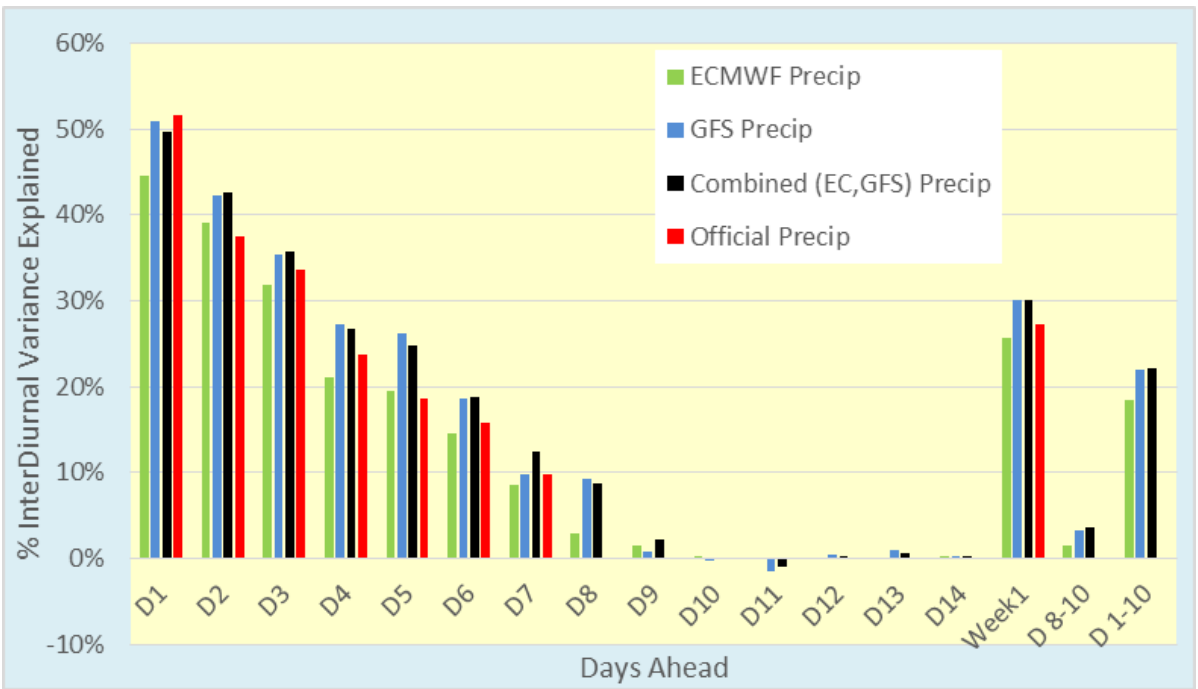


FIGURE 6

As per FIGURE 3, but with the specific aim of comparing the *Precip* skill displayed by the official forecasts with that displayed by the two models, and also the two models combined.

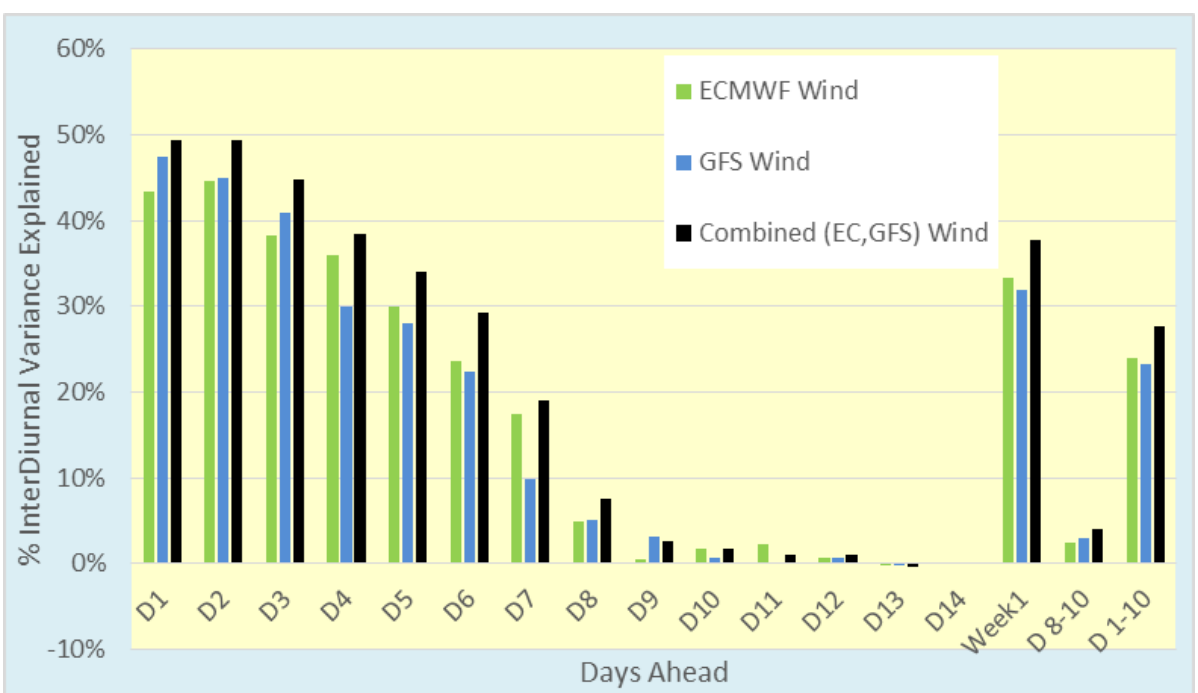


FIGURE 7

As per FIGURE 3, but with the specific aim of comparing the *Wind* skill displayed by the displayed by the two models individually, and also the two models combined.