Some aspects of the verification of weather forecasts for Melbourne, Australia

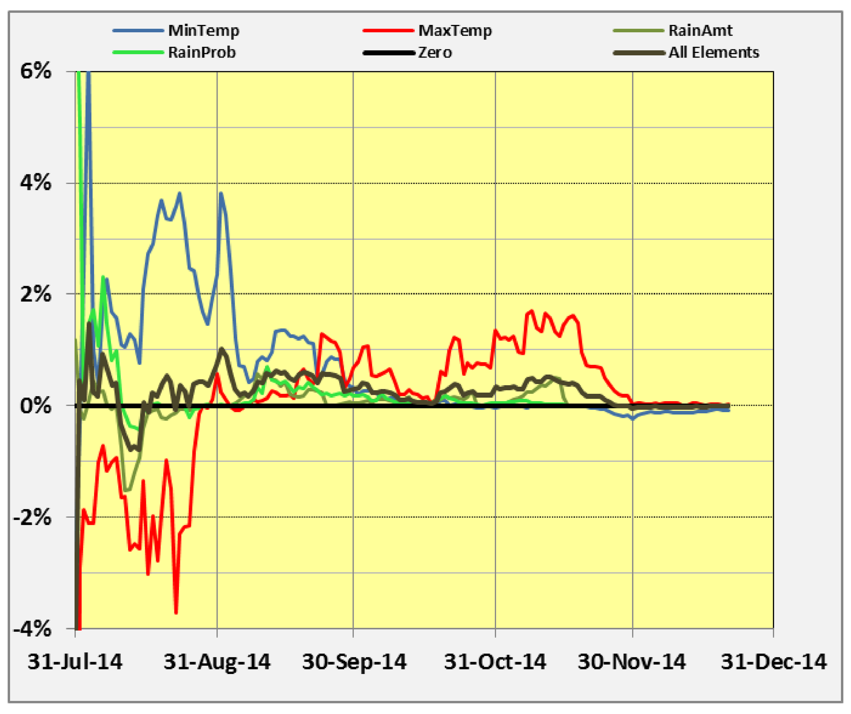
Harvey Stern[1] & Noel Davidson[2]

[1]*School of Earth Sciences, University of Melbourne, Parkville, Australia, e-mail: hstern@unimelb.edu.au*
[2]*Centre for Australian Weather and Climate Research, Bureau of Meteorology, Melbourne, Australia, e-mail: n.davidson@bom.gov.au*

Background. The authors have recently completed a piece of work exploring trends in the skill of day-to-day weather prediction at lead times of 1 to 14 days for Melbourne, Australia:
[http://www.weather-climate.com/MelbourneForecastAccuracy.pdf]
The system that was used to establish these trends was, in part, based upon an algorithm that generates local weather forecasts by statistically interpreting the Global Forecasting System (GFS) NWP model output. It was considered that it would be interesting to assess what might be achieved using the output of other global NWP models. Preliminary results are presented about what has been achieved during a six month trial (Jul-14 to Dec-14). The trial involved applying an algorithm to statistically interpret the output of the ECMWF (EC) NWP control models in terms of day-to-day local weather out to Day-32 (*refer to the summary figure opposite*). In addition, an account of what has been achieved by averaging the EC and GFS based output is also presented. The EC control models are those applied in the EC ensemble prediction system:
[http://old.ecmwf.int/about/corporate_brochure/leaflets/EPS-2012.pdf]

The algorithm used to interpret the EC model uses fewer predictors than that used to interpret the GFS model. This may be why some forecasts derived from the EC model output are less accurate than corresponding forecasts derived from the GFS model output.

Summary Figure. Accumulated (Jul-14 to Dec-14) ***Percent Variance of the Observations Explained*** (***PVOE***)* by the day-to-day EC Day 15-32 predictions of minimum temperature (blue), maximum temperature (red), rainfall amount (dark green), & rainfall probability (light green). Predictions are generated by application of an interpretive algorithm to the model output. As the data base grows, the accumulated skill displayed by the various sets of predictions trends towards zero.



*Where ***ACC***, the ***Anomaly Correlation Coefficient***, represents the correlation coefficient between the observed & forecast departure from the seasonal normal:

$$PVOE = (ACC^2)*(|ACC|/ACC)$$

Figure 1 shows that the algorithm utilised to interpret the GFS model output yields more accurate Week-1 (Day 1-7) minimum temperature predictions (***PVOE***=60.4%) than does the algorithm utilised to interpret the EC model output (***PVOE***=56.8%), more accurate maximum temperature predictions (68.7% vs 59.7%) but less accurate rainfall amount predictions (25.8% vs 29.0%) and rainfall probability predictions of similar accuracy (34.6% vs 34.7%).

Overall, across all four elements, the GFS *PVOE* is 47.4%, whilst the EC *PVOE* is 45.1%.

Combining, by averaging the forecasts associated with the two models, increases the accuracy of forecasts for three of the weather elements to a level superior to those of either model - for minimum temperature to 62.9% from 60.4% & 56.8%, for rainfall amount to 30.5% from 25.8% & 29.0% and for rainfall probability to 37.2% from 34.6% & 34.7%). However, the GFS maximum temperature forecasts are more accurate than the combined forecasts (68.7% vs 67.0%).

Overall, across all four elements, combining increases skill to 49.4% from 47.4% & 45.1%.

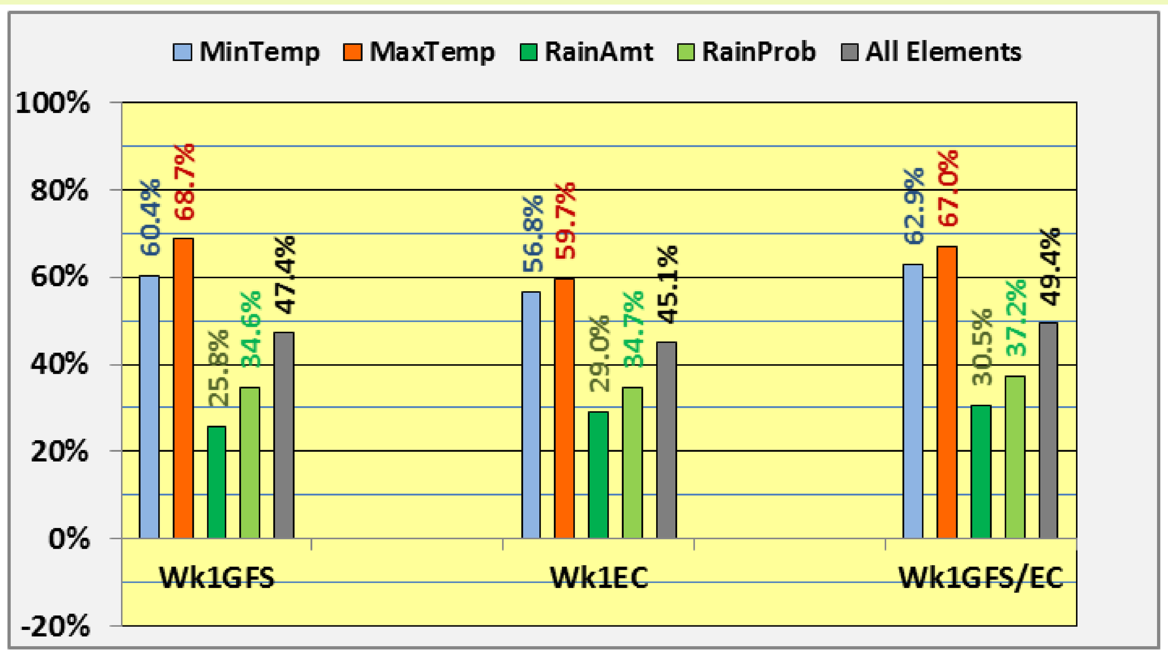


Figure 1 ***PVOE*** for Week-1 (Day 1-7) predictions of minimum temperature (blue columns), maximum temperature (red columns), rainfall amount (dark green columns) and rainfall probability (light green columns), and overall (grey columns). The predictions are generated by the application of an interpretive algorithm to the output of the GFS model (left-hand columns), the EC model (middle columns). The performance of a set of forecasts based upon the average of the aforementioned predictions is illustrated by the set of columns on the right.

Figure 2 shows that the algorithm utilised to interpret the GFS model output yields more accurate Week-2 (Day 8-14) predictions of minimum temperature predictions (***PVOE***=6.9%) than does the algorithm utilised to interpret the EC model output (***PVOE***=5.9%), more accurate rainfall amount predictions (1.5% vs 0.3%) and more accurate rainfall probability predictions (1.7% vs 1.3%) but less accurate maximum temperature predictions (9.6% vs 9.9%).

Overall, across all four elements, the GFS *PVOE* is 4.9%, whilst the EC *PVOE* is 4.4%.

Combining increases the accuracy of the predictions to a level superior to that of either model - for minimum temperature to 10.1% from 6.9% & 5.9%, for maximum temperature to 14.3% from 9.6% & 9.9% and for rainfall probability to 2.4% from 1.7% & 1.3%. However, the GFS rainfall amount predictions are more accurate than the combined forecasts (1.5% vs 1.3%).

Overall, across all four elements, averaging increases skill to 7.0% from 4.9% & 4.4%.

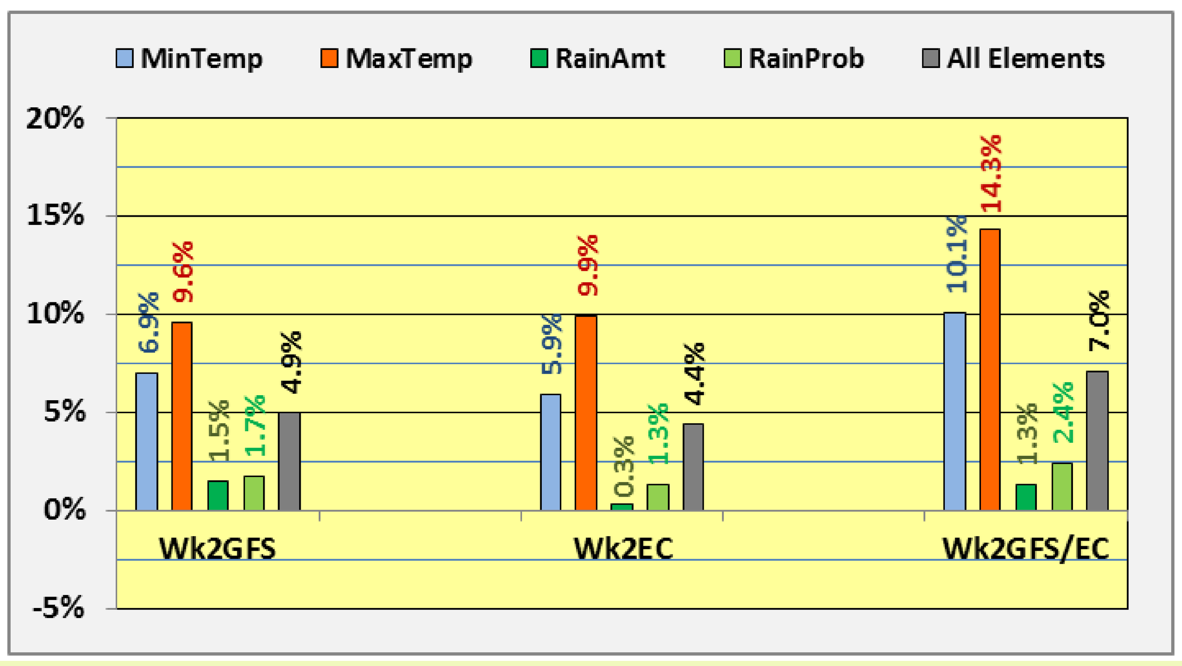


Figure 2 As per Figure 1, but for Week-2 (Day 8-14) predictions.

Figure 3, in summarising the outcome of the verification across Weeks 1&2 (Day 1-14) shows that combining increases the accuracy of the predictions to a level superior to that of either model for all four of the weather elements – for minimum temperature to 35.4% from 31.5% & 27.1%, for maximum temperature to 39.8% from 36.9% & 31.3%, for rainfall amount to 11.9% from 9.5% & 9.0% and for rainfall probability to 17.1% from 14.9% & 13.0%.

Overall, across all four elements, combining increases skill to 26.0% from 23.2% & 20.1%.

Figure 3 also shows that the EC Day 15-32 predictions display little forecasting skill.

Indeed, the minimum temperature *PVOE* is slightly negative (-0.1%). This reflects the fact that the correlation coefficient between the observed and forecast departure of minimum temperature from the seasonal normal, that is, the *ACC*, is also slightly negative.

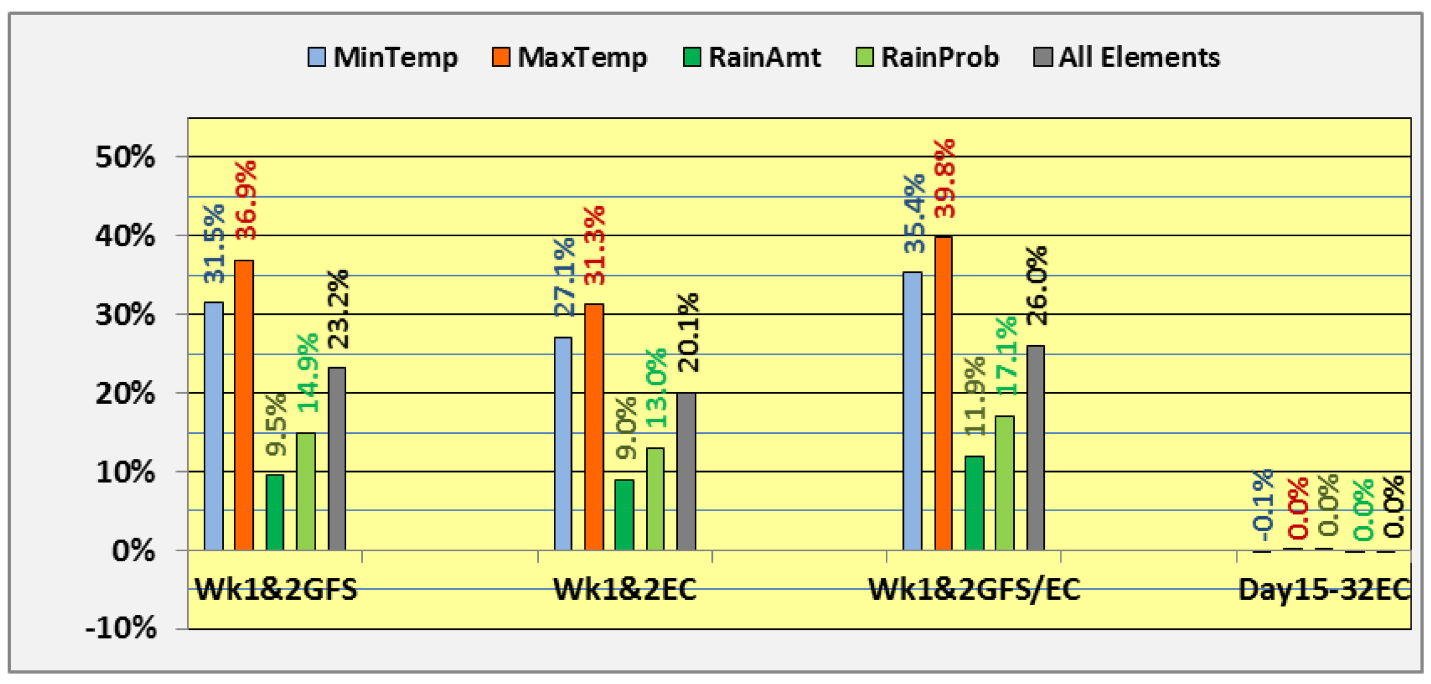


Figure 3 As per Figure 1, but for Weeks 1 & 2 (Day 1-14) predictions and also for the EC Day 15-32 predictions.

Figure 4 shows that combining increases the overall accuracy of the predictions to a level superior to that of the official forecasts for rainfall amount to 35.7% from 26.9% and for rainfall probability to 41.4% from 35.1%. The accuracy of the combined minimum temperature predictions is about the same as that of the corresponding official predictions (both 62.9%) as also is that of the combined maximum temperature predictions (67.0% vs 66.8%).

Overall, across all four elements, combining increases skill to 51.7% from 47.9%.

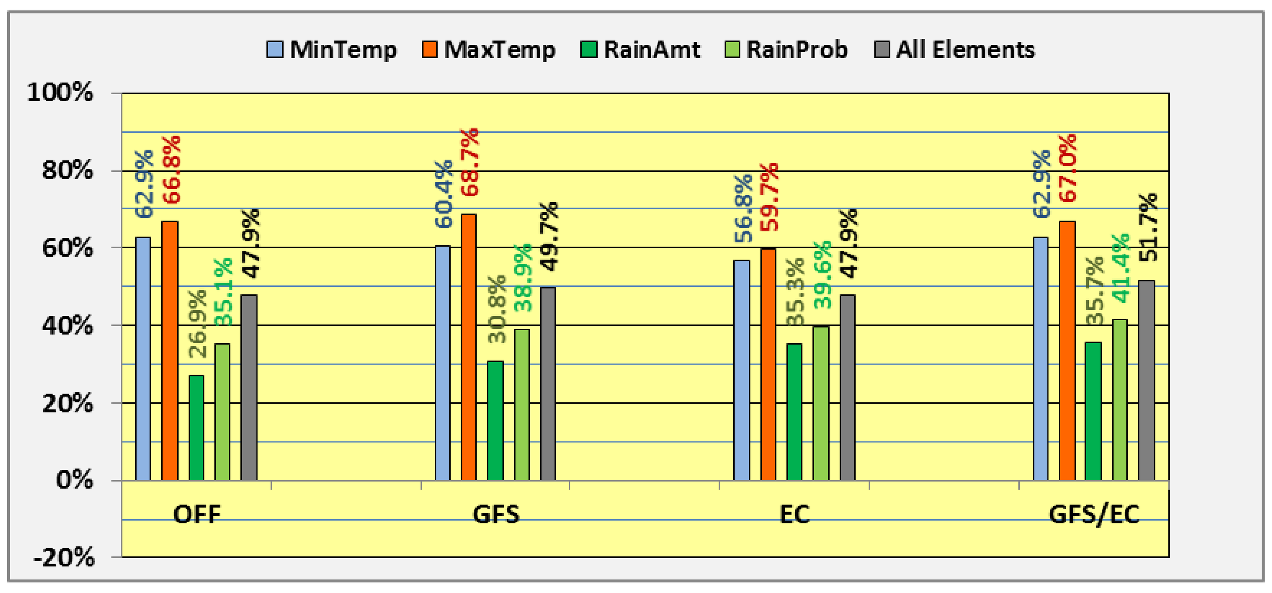


Figure 4 Percent variance explained by official (OFF) predictions of minimum & maximum temperature (Day 1-7), and rainfall amount & probability (Day 1-6). Note that official forecasts of rainfall amount & probability are only issued out to Day-6 at this time. Comparisons with those generated by the application of an interpretive algorithm to the output of the GFS model, the EC model and an average of those of the GFS and EC models are also shown.

SUMMARY

1. **The results suggest that an ensemble approach to weather forecasting, applied via a process of combining forecasts from various sources, increases the accuracy of day-to-day weather predictions.**

2. **During the (albeit) short trial, very long range day-to-day weather forecasts for Melbourne, Australia, derived by interpreting the output of the EC control models, displayed little forecasting skill.**