Operational Consensus Forecasting: from sites to grids

Chermelle Engel, Beth Ebert and John Bally
BMRC, Melbourne, VIC 3001, Australia

Introduction

Forecasters in Australian regional weather forecasting centres have a wealth of numerical guidance available to them, including regional and global models produced in the Australian Bureau of Meteorology and a subset of the guidance from global models produced in other international centres. Numerical guidance from international centres is routinely received via Global Telecommunications System (GTS) at temporal resolutions of six- to 24-hours with reduced spatial resolution. Hourly numerical guidance is available from a few locally produced, nested regional models.

When guidance is available from a number of different sources, consensus-forecasting techniques that combine the sources have been found to be more accurate on the average than techniques that try to predict the best "model of the day" (Hibon and Evgeniou 2005; Fritsch et al. 2000). The site-based Operational Consensus Forecast (OCF) scheme developed at the Australian Bureau of Meteorology performs a statistical correction of direct model output (DMO) forecasts followed by weighted average consensus. It has been shown to produce objective guidance competitive with subjective official forecasts for daily forecast fields such as maximum and minimum daily air temperatures (Woodcock and Engel, 2005).

Whilst site-based OCF daily forecasts provide useful guidance for some public weather forecasts, fire-weather and aviation forecasts require higher temporal resolution numerical guidance. Aviation forecasters use numerical guidance to assist in the production of Terminal Aerodrome Forecasts (TAFs). TAFs provide localised forecast information up to hourly temporal resolution out to 12 or 24 hours respectively, to the airline industry worldwide. They are produced at regular times every day for each Aerodrome as well as on special request, with the base time of TAF issues varying with station (Aeronautical Services Handbook 2005, Chapter 7). Fire weather forecasters use numerical guidance to assist in the production of public weather and spot fire forecasts. These include predictions for a number of different weather elements around the time of maximum air temperature and other time periods, along with the timing of events such as significant wind direction and speed changes and indications of other events such as rainfall or lightning (Weather Services Handbook 1976, Chapter 5).

An extension of the OCF scheme to hourly temporal resolution has been developed to generate hourly forecasts of screen-level air temperature, dewpoint temperature, relative humidity, mean sea-level pressure derived from the barometric pressure at the station location (QNH), along with 10m wind speed and direction out to 42 hours for approximately 280 Australian sites. Site-based hourly OCF applied to the hours on which all models are available, yielded average reductions in Mean Square Error (MSE) of 60%(air temperature), 47%(dewpoint temperature), 47%(relative humidity), 35% (QNH), 48%(wind speed) and 5% (wind direction) in comparison to the component models for May 2004 with slightly lower reductions for a 30 day period starting mid-February 2004. Given that the available numerical weather prediction (NWP) models are received at a variety of temporal resolutions, a method by which a "blended" site-based hourly OCF forecast could be achieved using a mix of model forecast intervals was developed (Engel, 2005).

There has been recent interest in extending the site-based hourly OCF scheme to produce 5km grid-based objective guidance, to assist in the generation of forecasts at all points across a domain of responsibility. Detailed analysis of the performance of the site-based statistics provides valuable insights into possible approaches for the grid-based system.

Overview of site-based OCF scheme

The site-based OCF scheme is an objective multi-model consensus produced from a weighted average of bias-corrected DMO forecasts. The scheme is based upon simple bias-correction and compositing algorithms. Forecasts are produced through two distinct stages of processing: bias correction and compositing using a weighted average. The bias correction scheme estimates the ***current*** systematic error at each site using a modified mean of error statistics over the past 15 to 30 days (BES; Wonnacott and Wonnacott 1972, section 7.3). The mean absolute error (MAE) of this sample after bias correction is then used to generate the weighting parameters. If a_i is the MAE over the last 15 to 30 days for the i^{th} contributing forecast scheme of n bias-corrected contributors, then it is weighted w_i, where

$$w_i = a_i^{-1}\left(\sum_{j=1}^{n} a_j^{-1}\right)^{-1} \quad \text{where} \quad i, j = 1,2,..,n \quad (1)$$

Model forecast errors can be conceptualised as having two components: systematic error and random error. The OCF algorithm works by removing the systematic errors and minimizing the random error component. Equation (2) demonstrates the propagation of errors experienced, where the n forecasts f with approximated bias $<b>$, and weights w, are broken into three components: the observed value o, systematic error b and random error e.

$$\begin{aligned} OCF &= \sum_{i=1}^{n}\left(w_i\left[f_i - <b>_i\right]\right) \\ &= \sum_{i=1}^{n}\left(w_i\left[o + (b_i - <b>_i) + e_i\right]\right) \\ &= o + \sum_{i=1}^{n}\left(w_i (b_i - <b>_i)\right) + \sum_{i=1}^{n}\left(w_i\, e_i\right) \\ & \text{where} \quad i = 1,2,..,n \qquad (2) \end{aligned}$$

Given that a simple weighted average forms the basis of the compositing algorithm, deficiencies in bias removal and random error characterization can propagate through to the resultant forecast. The second term in equation (2) demonstrates the possible retention of systematic errors. It is therefore important to remove as much systematic error as possible and ensure that the weighting parameters adequately reflect the nature of the random error components. For a description of the method by which a "blended" site-based hourly OCF forecast is achieved using a mix of model forecast intervals refer to (Engel, 2005).

Spatial and temporal variability of error characteristics

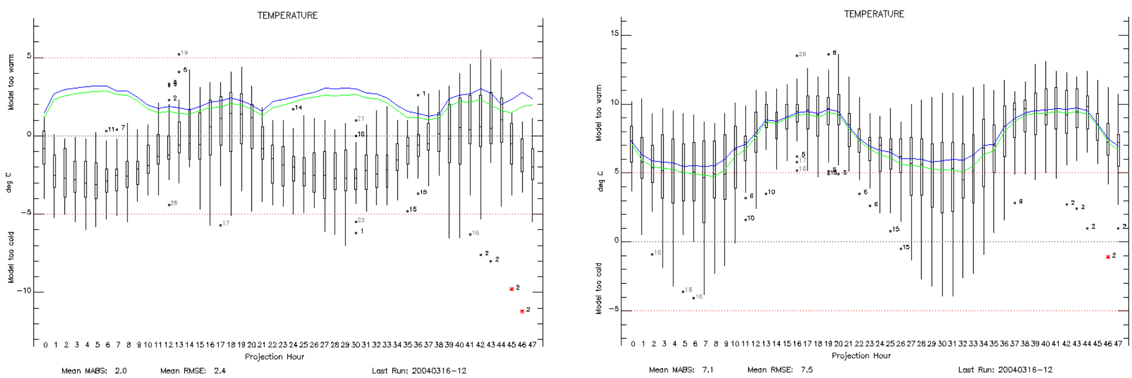

Figure 1: 12.5km MesoLAPS, 12Z based run (Model - Observation) 2m air temperature forecast error Box and Whisker Plots over 15th February to 15th March 2004 for Hobart Airport (left) and Mt. Wellington (right). The light line represents the MAE, whilst the darker the RMSE. Outliers are plotted explicitly, accompanied by the day of the month number. Note the variation in both systematic and random error characteristics across the 48-hour forecasting period and the differences between the two nearby locations.

In Woodcock and Engel (2005), daily forecast errors were found to vary on a site-by-site basis, with systematic and random errors increasing with forecast period. This led to bias corrections and weighting factors being calculated on a site-by-site and day-by-day basis. The systematic and random natures of hourly Numerical Weather Prediction (NWP) forecast errors were found to vary with forecast hour, with a diurnal modulation overlaying the normal error growth with time. Median error values were found to vary in strength in close proximity to each other, showing high spatial variability. The median error values also changed over the forecasting period, with some values becoming increasingly negative and others increasingly positive and some even changing signs. Station specific analysis of the model errors using box and whisker plots revealed location specific diurnal modulations for each of the weather elements investigated. The Hobart Airport and Mt. Wellington station box and whisker plots demonstrate the variation in error characteristics amongst stations (Figure 1).

Relative performance of coarse resolution climate models

In site-based OCF, DMO forecasts are derived at specific sites through interpolation from model nearest layer grid values without any attempt to accommodate differences between grid-point and site elevations or underlying surface variations such as land or sea. NWP grids from the overseas models are made available at spatial resolutions of around 1.25°, whilst local regional models have spatial resolutions of 0.375° or less. After bias correction, the overseas global models perform comparatively well with the local regional models (Figure 2). This is true even though the forecasts are evaluated against highly localised observations that are affected by features that cannot be resolved by even the highest resolution grids.

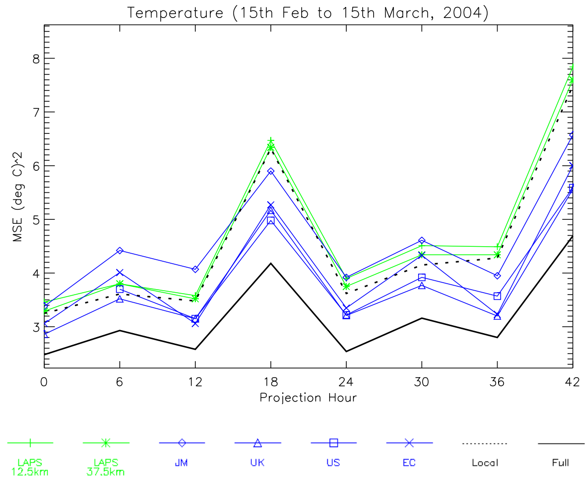

Figure 2: 12Z base run (Model – Observed) Air Temperature MSE statistics over all stations over February 2004. Included in the plot are the bias-corrected NWP models along with the full and local subset composites. Australian models are plotted in light grey and overseas models in dark grey. Due to the redundancy in the local models, the gain due to compositing is lower in comparison to using the full suite. The full composite performs better than that based on local models alone. EC statistics are 24 hours older than indicated, as used in the consensus.

To help understand why this is so, let us visualise atmospheric conditions as having large-scale (ls), mesoscale (ms) and small-scale (ss) components, where the mesoscale and small-scale processes are deviations from the large-scale flow. We can consider a site-specific observation in terms of its large-scale value plus localised deviations, such that o_{ls} is the average observational value over the area covered by a grid-box, o_{ms} is the deviation from the large-scale value required to represent the observational value over a smaller "mesoscale" area and o_{ss} is the further deviation required to obtain the observation at a specific location.

$$o = o_{ls} + o_{ms} + o_{ss}$$

The coarser overseas model forecasts can be considered as having only a large-scale component, where $f_{overseas}$ stands for the overseas forecast.

$$f_{overseas} = f_{ls}$$

Since these model forecasts have no mesoscale or small-scale processes, the error would be as follows, where $e_{overseas}$ stands for the error in the overseas forecast.

$$e_{overseas} = f_{overseas} - o = f_{ls} - \left(o_{ls} + o_{ms} + o_{ss}\right)$$

Therefore the bias in the forecast ($b_{overseas}$), would contain the bias in the large-scale forecast, plus climatological values for both the mesoscale and small scale weather components.

$$b_{overseas} = \overline{f_{overseas} - o} = \overline{f_{ls} - o} - \overline{o_{ms}} - \overline{o_{ss}}$$

Thus, the process of bias-correction can be seen to add climatological information about mesoscale and small-scale weather processes into the forecast.

$$f_{overseas} - b_{overseas} = (f_{ls} - \overline{f_{ls} - o}) + \overline{o_{ms}} + \overline{o_{ss}}$$

Since the local models run at spatial resolutions high enough to resolve mesoscale weather phenomena, the bias correction procedure could be considered as correcting both large-scale and mesoscale biases, whilst adding in small-scale climatological information. The fact that coarser overseas models compare so favourably with the high-resolution local models in terms of MAE and MSE statistics indicates that either the overseas models generate more accurate predictions of large-scale weather patterns, or predictions of mesoscale weather phenomena from the higher resolution models display less skill than climatology.

Implications for grid-based OCF

The high spatial and temporal variability found in model error characteristics for site-based OCF, indicate that each grid-box will need to be processed individually on an hour-by-hour basis. The implication that meso- and small-scale deviations are heavily represented by climatological information in site-based OCF indicates that scale-separation techniques may be useful for grid-based OCF. There may be a limit to the amount of accuracy achievable on the smaller scales.

An outline of the intended grid-based OCF scheme will be included as part of this presentation. The difficulties faced in the development of the scheme will be highlighted along with approaches currently under consideration.

References

Engel, C. 2005. Hourly Operational Consensus Forecasts. *BMRC Research Report,* in press.

Fritsch, J.M., Hilliker, J., Ross, J. and Vislocky, R.L. 2000. Model consensus. *Wea. Forecasting,* 15, 571-582.

Hibon, M. and Evgeniou, T. 2005. To combine or not to combine: selecting among forecasts and their combinations. *Int. J. Forecasting*. 21, 15-24.

Woodcock, F. and Engel, C. 2005. Operational Consensus Forecasts. *Wea. Forecasting,* 20, 101–111.

Wonnacott, T. H. and Wonnacott, R.J. 1972. *Introductory Statistics*. Wiley, 510 pp.

1997: Weather and Oceanographic Services Handbook[1]. *Bureau of Meteorology*, Ch 5, 30-1.

2005: Aeronautical Services Handbook[1]. *Bureau of Meteorology*, 70-84.

[1] Available from the Bureau of Meteorology, GPO Box 1289, Melbourne, Australia 3001.