ESTABLISHING THE LIMITS OF PREDICTABILITY AT MELBOURNE, AUSTRALIA USING A KNOWLEDGE BASED FORECASTING SYSTEM AND NOAA'S LONG RANGE NWP MODEL.

Harvey Stern*

Bureau of Meteorology, Melbourne, Vic., 3001, Australia

ABSTRACT

The work of Lorenz suggests that there is a 15-day limit to day-to-day predictability of the atmosphere. Ongoing increases in the accuracy of NWP model output continue to be evident. It was, therefore, considered appropriate to assess whether or not there may now be scientific justification to prepare day-to-day forecasts for the day 8 to day 15 period, with a view to providing a "link" between the day 1 to day 7 forecast and the three-month Seasonal Climate Outlook. A knowledge based forecasting system was utilised to *objectively* interpret the output of NOAA's long range (16 days) numerical weather prediction model statistically in terms of local weather at Melbourne (maximum temperature, minimum temperature, probability of precipitation, and amount of precipitation) in order to rigorously establish current limits of predictability. The results presented suggest that, for the first time, we have emerging evidence that there may now be some forecast skill out to Lorenz's suggested 15-day limit, particularly for temperature.

1. INTRODUCTION

The WMO (World Meteorological Organisation) weather research programme, THORPEX (The Observation, Research and Predictability Experiment), responds to the challenges associated with accelerating improvements in the skill of high-impact weather forecasts that (firstly) reduce and mitigate weather disasters and (secondly) increase the benefits provided by improved forecasts…(because) the term "high-impact weather forecasts" also emphasizes (those) societal and economic benefits resulting from advances in meteorological science…

THORPEX addresses the influence of sub-seasonal time-scales on high-impact forecasts out to two weeks, and thereby aspires to bridge the "middle ground" between medium range weather forecasting and climate prediction…" (Shapiro and Thorpe, 2004).

The work presented in the current paper, the primary motivation of which is to investigate whether or not there may now be scientific justification to prepare day-to-day forecasts out to two weeks, may be viewed as a small contribution towards the realisation of that aspiration.

This investigation involves *objectively* interpreting the output of a long-range Numerical Weather Prediction (NWP) model in terms of local weather at Melbourne and evaluating the forecasts of temperature and precipitation so generated.

2. BACKGROUND

Weather forecasts provided to the public are steadily improving (Figure 1(a)) and, during the late 1990s, the present author (Stern, 1998, 1999) presented the results of an experiment to establish the limits of that predictability. The experiment involved verifying a set of forecasts for Melbourne (Australia) out to 14 days. These forecasts were based upon a subjective interpretation of the ensemble mean output of the NCEP[1] NWP model.

The verification data suggested that, during the late 1990s, routinely providing or utilising day-to-day forecasts beyond day 4 would have been inappropriate. However, the data also suggested that it might have been possible to provide some useful information about the likely weather up to about a week in advance for some elements and in some situations.

Shortly after Stern's (1998) presentation (in April of that year), the Australian Bureau of Meteorology's (ABM) Victorian Regional Forecasting Centre (RFC) commenced a formal (official) trial of day-to-day forecasts for Melbourne out to day 7.

There have been considerable advances in NWP modelling since then, and also in associated techniques for statistically interpreting the NWP model output utilising objective methods. Stern (2004a) has recently demonstrated that the skill displayed by the trial maximum temperature forecasts is superior to that of climatology (even) at day 7 (Figure 1(b)).

In the light of the skill displayed by the official trial forecasts, the ABM recently commenced routinely issuing a forecast out to day 7 to the public each evening. Predictions for days 5, 6, and 7 are couched in general terms.

The ABM routinely issues its three-month Seasonal Climate Outlook (SCO) to the public on about the middle day of each month prior.

The work of Lorenz (1963, 1969a&b, 1993) suggests that there is a 15-day limit to day-to-day predictability of the atmosphere. Ongoing increases in the accuracy of NWP model output continue to be evident. It was, therefore, considered appropriate to now

---

[1]National Center for Environmental Prediction

repeat Stern's (1998, 1999) experiment. This was done to enable assessment of whether or not there may now be scientific justification to prepare day-to-day forecasts for the day 8 to day 15 period, with a view to providing a "link" between the day 1 to day 7 forecast and the three-month SCO. Incidentally, Stern (1999) also investigated whether, even though the day 5 to day 14 forecasts for the *individual* days displayed little skill at that time, they might provide an indication of *overall* weather conditions during the day 5 to day 14 period. Analysis of the data revealed that there was some basis for an affirmative response to this question. However, the relationship between the forecast weather and observed weather was not strong (level of significance between 0.2% and 13%). Interestingly, with increases in the accuracy of NWP model output, the South African Weather Service now provides such a product that presents a depiction of the likely weather during Week 2.

3. A LONG RANGE GLOBAL FORECASTING SYSTEM

The National Oceanic and Atmospheric Administration's (NOAA) Global Forecasting System long range (GFSlr) NWP model provides output that includes forecast data every 12 hours from forecast hour 192 (day 8) to 384 (day 16) on a 2.5 degree latitude/longitude grid covering the globe. The data is updated 4 times per day. Anecdotal evidence suggested that the model displayed skill (refer to *Acknowledgements*).

The current paper presents a preliminary study of the skill displayed by forecasts derived from 100 twice-daily "runs" of the GFSlr model (base analyses between 12UTC on 5 August, 2004, and 00UTC on 13 November, 2004).

4. INTERPRETING MODEL OUTPUT IN TERMS OF LOCAL WEATHER

Over recent years, the author (Stern 2002, 2003, 2004b&c) has been involved in the development of a knowledge based weather forecasting.

Operating under the “perfect prog” assumption, the knowledge based system uses a range of forecasting aids to interpret NWP model output in terms of such weather parameters as precipitation amount and probability, maximum and minimum temperature, fog and low cloud probability (Stern and Parkyn, 2001), thunderstorm probability (Stern, 2004c), wind direction and speed, and swell (Dawkins, 2002).

The knowledge based forecasting system was utilised to *objectively* interpret the output of the GFSlr model statistically in terms of local weather at Melbourne (maximum temperature, minimum temperature, probability of precipitation, and amount of precipitation) in order to rigorously establish current limits of predictability. By contrast, Stern's (1998, 1999) experiment was based upon *subjective* interpretation of model output.

5. CONSISTENCY OF OUTPUT

Examination of the output of "runs" of the GFSlr model reveals a modest, but useful, level of consistency of output from one "run" to the next. Indeed, there were a number of occasions when consistent advance notice was given by the GFSlr model to the prediction of a number of unusual events. For example, Figure 2 shows that the warm day on 20 September, 2004, was anticipated well in advance by the GFSlr model, as interpreted by the knowledge based system.

However, there is some jerkiness in the forecasts from one "run" to the next – for example, whilst the warm day was anticipated 13.5 and 14 days in advance, it temporarily disappeared at 13 days - and this would render them unsatisfactory, were a decision made to issue the forecasts to the public.

One approach to address this "jerkiness" may be to regard the individual forecasts as members of an ensemble. For example, the output of the GFSlr model's four most recent runs may then be averaged.

To illustrate, the output of the GFSlr model's 850 hPa temperature and 700 hPa relative humidity are among the data that are input into the knowledge based system in order to obtain a weather forecast.

Firstly, regarding the 850 hPa temperature, Figure 3(a) depicts the mean and standard deviation (uncertainty) of the Melbourne 850 hPa temperature that were forecast by the 20 October, 2004, "runs" - 00UTC, 06UTC, 12UTC and 18UTC. It also shows a line of best-fit based on the standard deviation (sd) data, which suggests that there is an overall increase in uncertainty associated with the forecasts of 850 hPa temperature as one moves from day 1 (Oct-21) to day 16 (Nov-5).

Secondly, regarding the 700 hPa relative humidity, Figure 3(b) depicts the mean and standard deviation (uncertainty) of the Melbourne 700 hPa relative humidity that were forecast by the 20 October, 2004, "runs" - 00UTC, 06UTC, 12UTC and 18UTC. It also shows a line of best-fit based on the standard deviation (sd) data, which suggests that there is an overall increase in uncertainty associated with the forecasts of relative humidity as one moves from day 1 (Oct-21) to day 16 (Nov-5).

Another approach to address the "jerkiness" may be to average the most recent interpretations of the output of the GFSlr model (instead of the output, itself). Figure 3(c) demonstrates how for Day 5 and beyond, such an approach might lead to an improvement in the forecasts of maximum temperature.

To illustrate the skill displayed by the maximum temperature forecasts close to the event, Figure 4 compares departures from normal of observed and forecast maximum temperatures 8 and 8.5 days in advance.

6. INDEPENDENCE OF FORECAST DATA SETS

There are 22 forecast data sets, each comprising predictions of:

- Minimum temperature (*Min*);
- Maximum temperature (*Max*);
- Quantity of Precipitation Forecast (*QPF*); and,
- Probability of Precipitation (*PoP*).

Four of these data sets correspond to the official ABM forecasts for 1, 2, 3, and 4 days ahead.

Three of these data sets correspond to official trial forecasts for 5, 6, and 7 days ahead.

Fifteen of these data sets correspond to the forecasts based on the interpretation of the output of the GFSlr NWP model for 8, 8.5, 9, 9.5 … 15 days ahead.

However, the elements of each of these 22 forecast data sets are not truly independent. This lack of independence arises from the fact that weather patterns often persist for several days.

Now, Figure 5 shows that the overall % variance explained by the *Max*, *Min*, *QPF* and *PoP* Equations:

```
(Observed departure from normal²)=

a + b(Observed departure from normal a number of days before)

where a and b are constants          (1)
```

suggests that persistence of weather patterns is confined largely to day 1.

One may, therefore, deduce that consecutive data elements are not truly independent (this is particularly so for minimum and maximum temperature), whilst beyond day 1, the data elements do appear to be fairly independent[3].

The numbers of degrees of freedom utilised to establish confidence limits in the analyses that follow are all, therefore, reduced to half of what they would have been,

---

[2]In this context, the QPF is regarded as √(forecast *Precipitation Amount*), the normal *QPF* is regarded as the √(mean daily precipitation for a particular month), whilst the observed *Precipitation Amount* is either √(observed precipitation), if precipitation is observed, or 0 (if precipitation is not observed). Furthermore, in this context, the normal *PoP* is regarded as the monthly proportion of days with precipitation, whilst the observed *PoP* is either 100%, if precipitation is observed, or 0% (if precipitation is not observed).

[3] However, a cycle of diminishing amplitude appears to be present with weak negative correlation evident with data from days 3 to 7, weak positive correlation from days 8 to 10, and weak negative correlation from days 11 to 15.

had all the data elements been truly independent, because, in the present analysis, only every second day's data may be regarded as independent.

7. ANALYSIS OF RESULTS

7.1 *Minimum temperature*

Figure 6 depicts three best-fit 2nd order polynomial curves[4], about the regression coefficients 'b' ('b' may be regarded as a measure of predictability) in the Equations derived on data from the 22 *Min* forecast data sets:

(Observed *Min* departure from normal)=

a + b(Forecast *Min* departure from normal a number of days in advance)

where a and b are constants (2)

The three curves are:

- Top curve: regression coefficients 'b';
- Middle curve: 75% lower confidence limit for regression coefficient 'b'; and,

[4] Other curves, including higher order polynomials and a sigmoid curve, were also considered, but the relatively simple 2nd order polynomial curve appeared to best depict the relationship.

- Bottom curve: 95% lower confidence limit for regression coefficient 'b'.

The curves show that:

- It is more likely than not that there is skill at forecasting minimum temperature out to 15 days ahead;
- It is three times more likely than not that there is skill at forecasting minimum temperature out to 14 days ahead; and,
- One can be 95% confident that there is skill at forecasting minimum temperature out to 11 days ahead.

The 'b's represent the proportion of the forecast departure from normal to utilise, should one wish to achieve optimal forecast skill. Hence, by way of example, for forecasts for 8 days ahead, although the significance of the skill is high (at the 95% level), the magnitude of that skill is not - from the 2^{nd} order polynomial for coefficient 'b' one observes that the proportion of the forecast departure from normal to utilise is only about 0.45. To illustrate, should the prediction for day 8 be for a minimum temperature that is 10 deg C above normal, for optimal forecast skill one should predict a minimum temperature that is 4.5 deg C above normal, assuming an intercept of zero.

7.2 *Maximum temperature*

Figure 7 depicts three best-fit 2^{nd} order polynomial curves, about the regression coefficients 'b' in the Equations derived on data from the 22 *Max* forecast data sets:

(Observed *Max* departure from normal)=

a + b(Forecast *Max* departure from normal a number of days in advance)

where a and b are constants (3)

The three curves are:

- Top curve: regression coefficients 'b';
- Middle curve: 75% lower confidence limit for regression coefficient 'b'; and,
- Bottom curve: 95% lower confidence limit for regression coefficient 'b'.

The curves show that:

- It is more likely than not that there is skill at forecasting maximum temperature out to 15 days ahead;
- It is even three times more likely than not that there is skill at forecasting maximum temperature out to 15 days ahead; and,
- One can be 95% confident that there is skill at forecasting maximum temperature out to 12 days ahead.

As with the case for minimum temperature, the 'b's represent the proportion of the forecast departure from normal to utilise, should one wish to achieve optimal forecast skill.

Figure 7 shows that, for maximum temperature forecasts for 8 days ahead, the optimal proportion of forecast departure from normal to utilise (about 0.50) is slightly higher than the corresponding value for minimum temperature.

Furthermore, Figure 7 shows that the optimal proportion of forecast departure from normal to utilise for day 7 is about 0.55. This value is greater than the corresponding value derived by Stern (2004a) for day 7 (0.511) using 1998-2003 data from the official trial. One may interpret this to be suggesting that there has been an improvement in the accuracy of the official trial forecasts since the 1998-2003 period.

7.3 *Quantitative Precipitation Forecast (QPF)*

Figure 8 depicts three best-fit 2nd order polynomial curves, about the regression coefficients 'b' in the Equations derived on data from the 22 *QPF* forecast data sets:

(Observed *Precipitation Amount* departure from normal)=

a + b(*QPF* departure from normal a number of days in advance)

where a and b are constants (4)

The three curves are:

- Top curve: regression coefficients 'b';
- Middle curve: 75% lower confidence limit for regression coefficient 'b'; and,
- Bottom curve: 95% lower confidence limit for regression coefficient 'b'.

The curves show that:

- It is more likely than not that there is skill at forecasting precipitation amount out to 11 days ahead;
- It is three times more likely than not that there is skill at forecasting precipitation amount out to 9 days ahead; and,
- One can be 95% confident that there is skill at forecasting precipitation amount out to 7 days ahead.

7.4 *Probability of Precipitation (PoP)*

Figure 9 depicts three best-fit 2^{nd} order polynomial curves, about the regression coefficients 'b' in the Equations derived on data from the 22 *PoP* forecast data sets:

(Observed *PoP* departure from normal)=

a + b(Forecast *PoP* departure from normal a number of days in advance)

where a and b are constants (5)

The three curves are:

- Top curve: regression coefficients 'b';
- Middle curve: 75% lower confidence limit for regression coefficient 'b'; and,
- Bottom curve: 95% lower confidence limit for regression coefficient 'b'.

The curves show that:

- It is more likely than not that there is skill at forecasting *PoP* out to 12 days ahead;
- It is three times more likely than not that there is skill at forecasting *PoP* out to 10 days ahead; and,
- One can be 95% confident that there is skill at forecasting *PoP* out to 8 days ahead.

8. COMPARISON WITH ABM TRIAL OF FORECASTS OUT TO DAY 7

An analysis of the variance explained by the official ABM forecasts for 1, 2, 3, and 4 days ahead, and the official trial forecasts for 5, 6, and 7 days ahead, was carried out on 2000-2003 data. The results of this analysis were compared with a corresponding analysis of forecasts between 1 and 15 days ahead during the 100-day trial conducted in 2004.

Figure 10 depicts the percentage variance explained by the *Min*, *Max*, *QPF* and *PoP* components of the 2000-2003 and 2004 sets of forecasts (temperature and precipitation components combined).

PoP was not included in the depiction for 2000-2003 because *PoP* data was not available for that period. For this reason, a double weighting is given to the 2000-2003

QPF data in order that overall equal weighting be given to the temperature and precipitation components of the forecasts.

Figure 10 shows that the skill (as measured by the percentage variance explained) declines steadily from about 50% for Day 1, to about 15% for Day 7, and that the characteristics of that decline are similar for both the 2000-2003 and 2004 sets of forecasts (notwithstanding that they are not strictly comparable on account of the 2000-2003 *PoP* data not being available).

Figure 10 also shows that the skill displayed by the forecasts (for all lead times between Day 1 and Day 7) is slightly greater for the 2004 forecasts than for the 2000-2003 forecasts, reflecting the ongoing trend towards increasing forecast skill.

Figure 10 also shows that the skill continues to decline (albeit at a slower rate) from Day 7 to Day 10, at which point only 5% of the variance is explained. For forecasts from Day 10.5 to Day 15, the skill averages about 1.4%.

A legitimate question to ask is:

Is a forecast that explains only a small amount of the variance useful to a client?

The answer, in this era of active amelioration of weather-related risks, is "yes", because provided the client is able to activate such risk reduction measures[5], even a low level of skill can be taken advantage of. To this end, Stern and Dawkins (2004) show how weather derivatives may be used as a vehicle to realise the skill inherent in seasonal forecasts whilst Dawkins and Stern (2004) show how weather derivatives may be used to manage weather risk during sporting events.

Figure 11 depicts the percentage variance explained by the temperature and precipitation components of the forecasts taken separately.

---

[5] Some clients may not be in a position to activate risk reduction measures.

Figure 11 shows that the levels of skill displayed by both the temperature and precipitation components of the forecasts (for all lead times between Day 1 and Day 7) are slightly greater for the 2004 forecasts than for the 2000-2003 forecasts (reflecting the ongoing trend towards increasing forecast skill).

Figure 11 also shows that the skill of the temperature forecasts (as measured by the percentage variance explained) declines steadily from about 75% for Day 1, to about 20% for Day 7, and to about 8% for Day 10. For forecasts from Day 10.5 to Day 15, the skill averages 2.8%.

Figure 11 also shows that the skill of the precipitation forecasts (as measured by the percentage variance explained) declines steadily from about 30% for Day 1, to about 8% for Day 7. For forecasts from Day 8 to Day 10, the skill averages only 0.7%, whilst it is negligible (0.01%) from Day 10.5 to Day 15.

8. CONCLUSION

Analysis of the data suggests that application of the knowledge based system to the interpretation of the Global Forecasting System long range model output yields a set of day-to-day weather predictions that display a modest, but nevertheless potentially useful, level of skill, especially at predicting temperature.

This outcome appears to justify the emergence on the web of extended-period day-to-day forecasts.

Furthermore, even a modest level of forecast skill may be applied to financial market instruments, such as weather derivatives, in order to ameliorate weather-related risk. It may, therefore, be justifiable to prepare such forecasts with a view to using them to ameliorate that risk, and also with a view to providing a "link" between the short-term forecasts and the three-month Seasonal Climate Outlook.

The significance of the results presented herein is that, for the first time, we have emerging evidence there may now be skill out to Lorenz's (1963, 1969a&b, 1993) suggested 15-day limit of day-to-day predictability of the atmosphere, particularly for temperature.

Acknowledgements. To Stuart Coombs, of the Bureau of Meteorology's Regional Forecasting Centre (Victoria), who inspired this work (Stuart alerted the author to anecdotal evidence that the output of the NOAA GFSlr NWP Model displayed considerable skill, and that, on occasions, it had predicted significant events even towards the end of the forecast period (day 16)), to Robert Dahni and Terry Adair, of the Bureau of Meteorology's Data Management Group, for providing some historical forecast verification data, and to reviewers Robert Seaman and Ian Mason, for their helpful comments.

9. REFERENCES

Dawkins, S. S., 2002: A web-based swell and wind forecasting tool. *9th Conference of the Australian Meteorological and Oceanographic Society, Melbourne, Australia, 18-20 Feb., 2002.*

Dawkins, S. S., and Stern, H., 2004: Managing weather risk during major sporting events: the use of weather derivatives. *2nd International Workshop on Climate, Tourism and Recreation, Orthodox Academy of Crete, Kolimbari, Crete, Greece, 8-11 Jun., 2004. Published in Berichte des Meteorologischen Institutes der Universitat Freiburg Nr. 12. A. Matzarakis, C. R. de Freitas and D. Scott (Eds.) Advances in Tourism Climatology. Freiburg, November 2004. pp 166-173.*

Lorenz, E. N., 1963: Deterministic, non-periodic flow. *J. Atmos. Sci.,* **20**, 130-41.

Lorenz, E. N., 1969a: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.,* **26**, 636-46.

Lorenz, E. N., 1969b: The predictability of a flow which possesses many scales of motion. *Tellus,* **21**, 289-307.

Lorenz, E. N., 1993: The essence of chaos. *University of Washington Press*.

Shapiro, M. A. and Thorpe, A., 2004: THORPEX: A global atmospheric research programme for the beginning of the 21st century. WMO Bulletin, **54**, No.3, July 2004.

Stern, H., 1998: An experiment to establish the limits of our predictive capability. *14th Conference on Probability and Statistics / 16th Conference on Weather Forecasting and Analysis, Phoenix, Arizona, 11-16 Jan., 1998, Amer. Meteor. Soc.*

Stern, H., 1999: An experiment to establish the limits of our predictive capability for Melbourne. *Aust. Meteor. Mag.,* **48**, 159-167.

Stern, H., 2002: A knowledge-based system to generate internet weather forecasts. *18th Conference on Interactive Information and Processing Systems, Orlando, Florida, 13-17 Jan., 2002, Amer. Meteor. Soc.*

Stern, H., 2003: Progress on a knowledge-based internet forecasting system. *19th Conference on Interactive Information and Processing Systems, Long Beach, California, 9-13 Feb., 2003, Amer. Meteor. Soc.*

Stern, H., 2004a: Using verification data to improve forecasts. *Australian Meteorology and Oceanography Society 2004 Annual Conference, Brisbane, Queensland, Australia 5-9 Jul., 2004.*

Stern, H., 2004b: Incorporating an ensemble forecasting proxy into a knowledge based system. *20th Conference on Interactive Information and Processing Systems, Seattle, Washington, 11-15 Jan., 2004, Amer. Meteor. Soc.*

Stern, H., 2004c: Using a knowledge based system to predict thunderstorms. *International Conference on Storms, Storms Science to Disaster Mitigation, Brisbane, Queensland, Australia 5-9 Jul., 2004.*

Stern, H. and Parkyn, K., 2001: A web-based Melbourne Airport fog and low cloud forecasting technique. *2nd Conference on Fog and Fog Collection, St John's, New Foundland, Canada 15-20 Jul.,2001*.

Stern, H. and Dawkins, S. S., 2004: Weather derivatives as a vehicle to realise the skill of seasonal forecasts. *15th Conference on Global Change and Climate Variations & 14th Conference on Applied Climatology, Seattle, Washington, USA 11-15 Jan., 2004.*

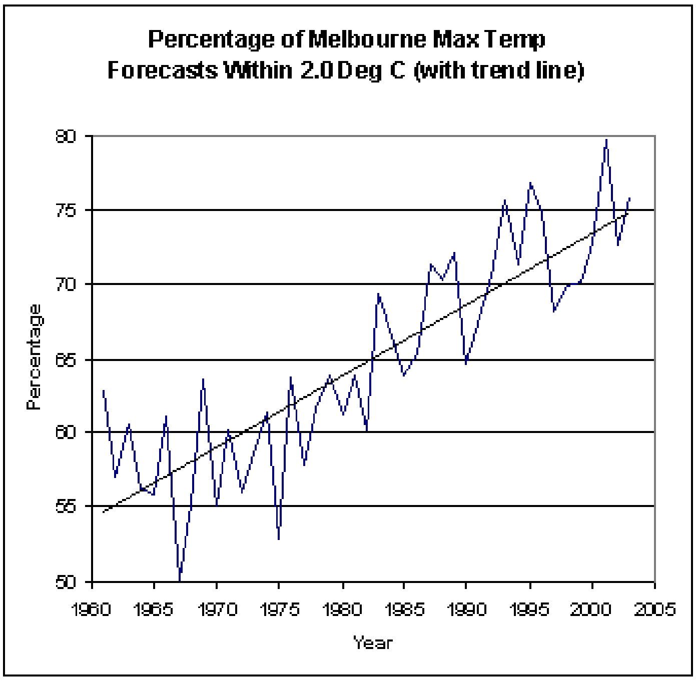

Figure 1(a) Long-term trend in accuracy of Day-1 Melbourne maximum temperature forecasts 1961-2003, as measured by the percentage of forecasts within 2 deg C.

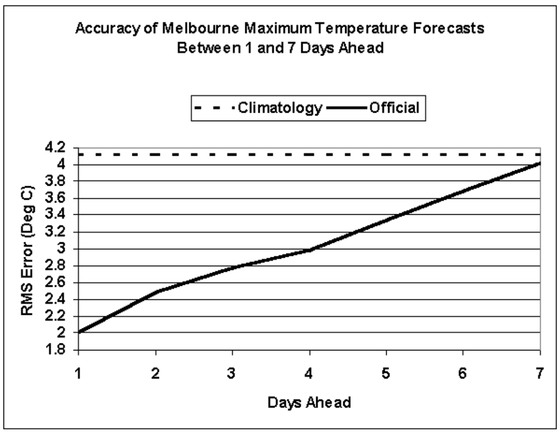

Figure 1(b) Accuracy of day 1 to day 7 Melbourne maximum temperature forecasts 1998-2003, as measured by the Root Mean Square (RMS) error (after Stern, 2004a).

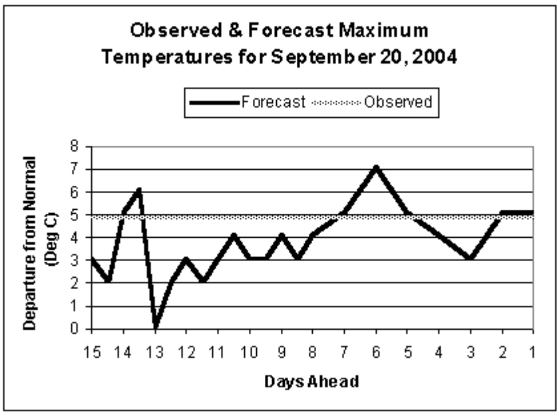

Figure 2 Observed and forecast maximum temperatures for 20 September, 2004 (GFSlr based forecasts made every 12 hours between 15 and 8 days ahead; forecasts between 7 days and 1 day ahead are RFC forecasts).

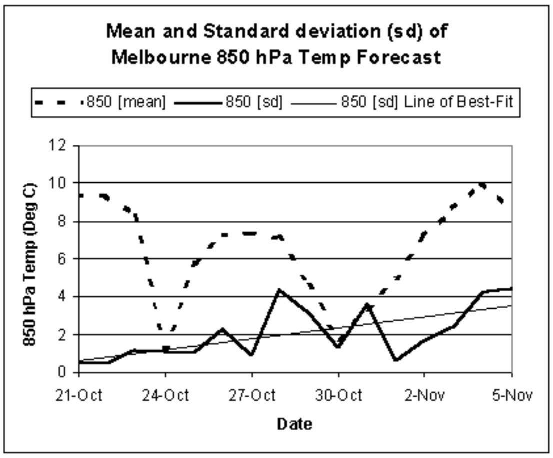

Figure 3(a) Mean and standard deviation (uncertainty) of the Melbourne 850 hPa temperature forecast by the 20 October, 2004, "runs" - 00UTC, 06UTC, 12UTC and 18UTC, and a line of best-fit based on the standard deviation (sd) data.

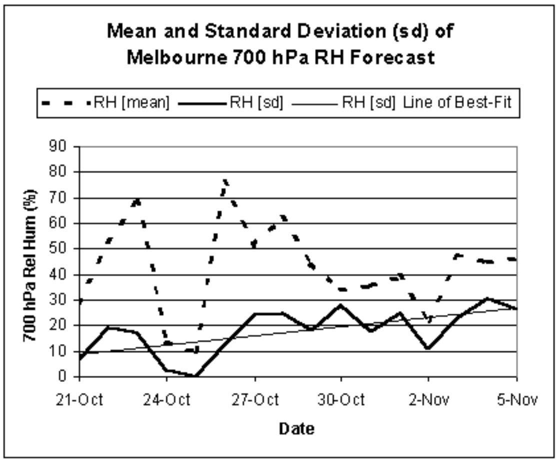

Figure 3(b) Mean and standard deviation (uncertainty) of the Melbourne 700 hPa relative humidity (RH) forecast by the 20 October, 2004, "runs" - 00UTC, 06UTC, 12UTC and 18UTC, and a line of best-fit based on the standard deviation (sd) data.

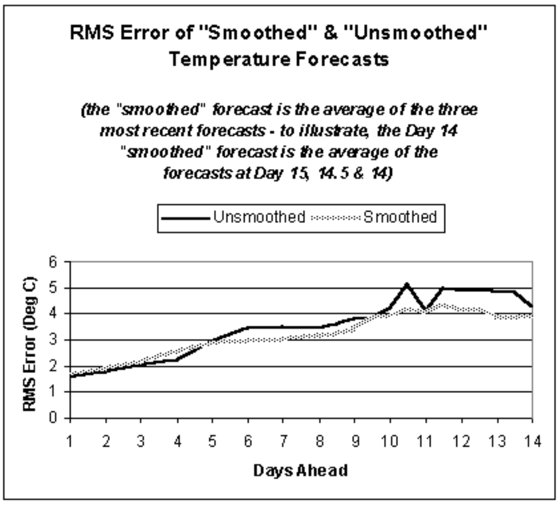

Figure 3(c) RMS Error (deg C) of "smoothed" and "unsmoothed) forecasts of maximum temperature.

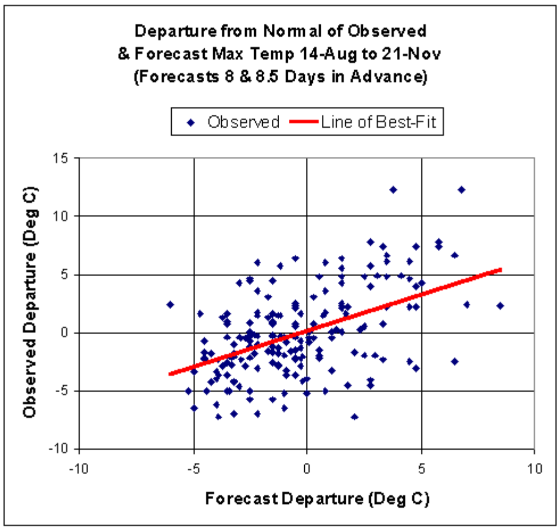

Figure 4 Comparison between the departures from normal of observed and forecast maximum temperatures 8 and 8.5 days in advance. The slope of the regression line may be regarded as a measure of predictability.

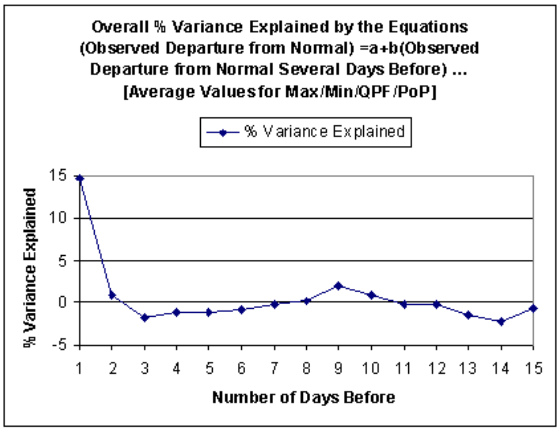

Figure 5 Overall % variance explained by the variables *Max*, *Min*, *QPF* and *PoP,* in the serial relationships defined by Equation (1). Cases of negative values arise from the relationships between the observed departure from normal, and that observed several days before, being negatively correlated.

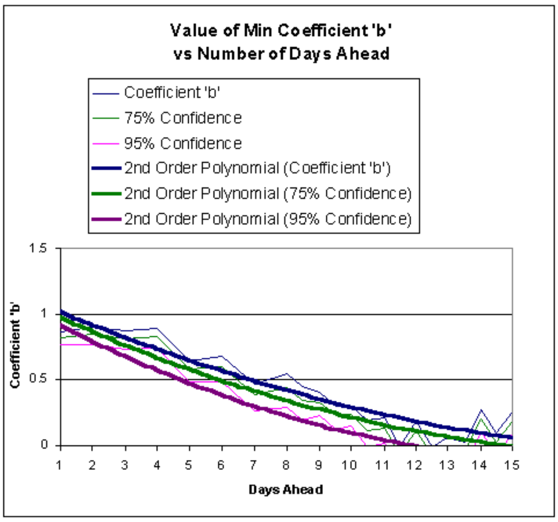

Figure 6 Confidence limits for the regression coefficients 'b' in the *Min* equations. In calculating confidence limits, the number of degrees of freedom is reduced by half. Positive values of 'b' suggest skill.

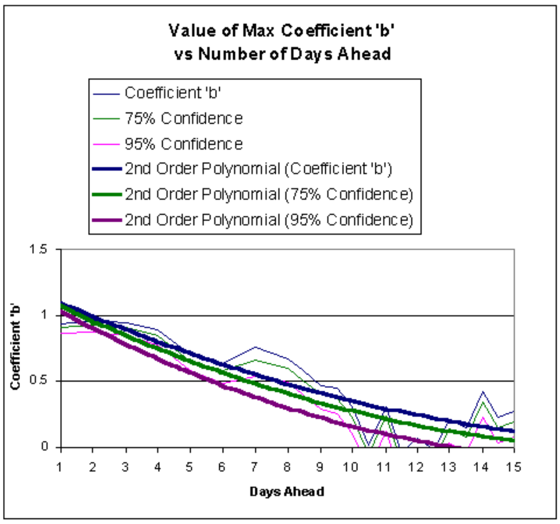

Figure 7 Confidence limits for the regression coefficients 'b' in the *Max* equations. In calculating confidence limits, the number of degrees of freedom is reduced by half. Positive values of 'b' suggest skill.

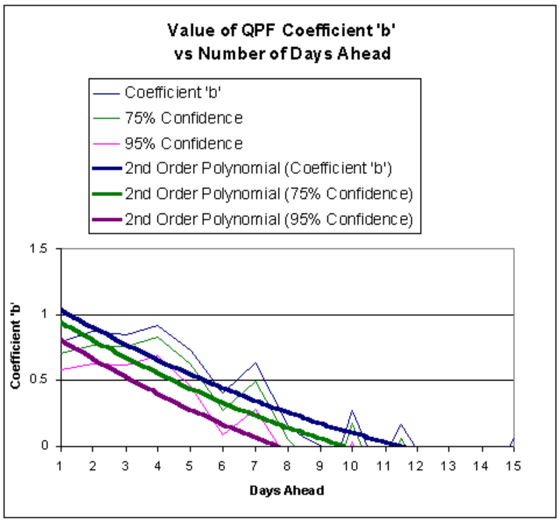

Figure 8 Confidence limits for the regression coefficients 'b' in the *QPF* equations. In calculating confidence limits, the number of degrees of freedom is reduced by half. Positive values of 'b' suggest skill.

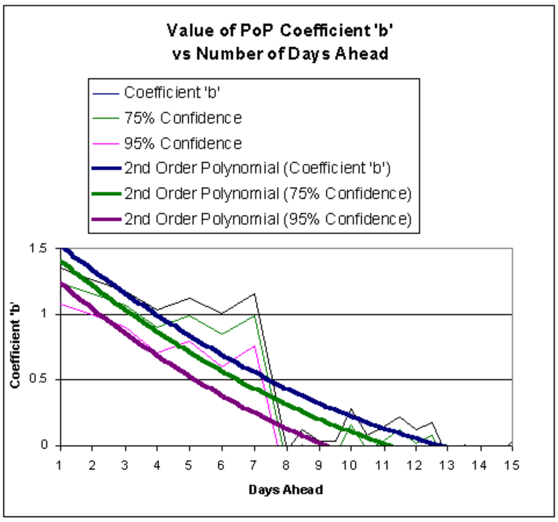

Figure 9 Confidence limits for the regression coefficient 'b' in the *PoP* equations. In calculating confidence limits, the number of degrees of freedom is reduced by half. Positive values of 'b' suggest skill.

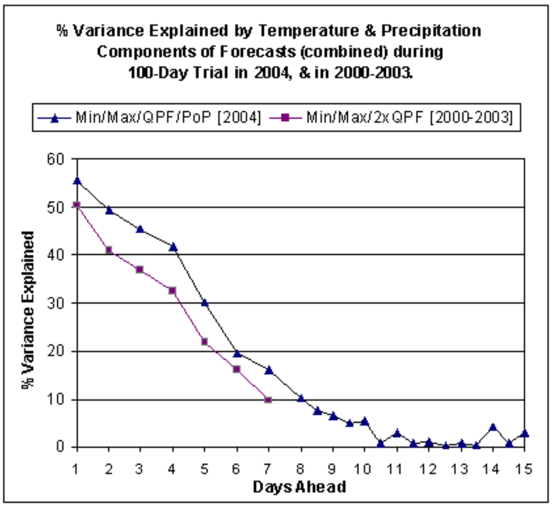

Figure 10 An analysis of the variance explained by

i. The 2000-2003 forecasts (the official forecasts for 1, 2, 3, and 4 days ahead, and the official trial forecasts for 5, 6, and 7 days ahead); and,

ii. The 2004 forecasts (forecasts between 1 and 15 days ahead during the 100-day trial).

For the purpose of this analysis, temperature and precipitation components of the forecasts are combined.

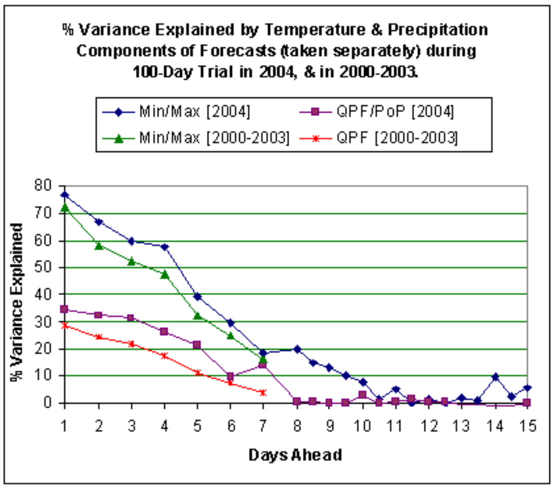

Figure 11 An analysis of the variance explained by

i. The 2000-2003 forecasts (the official forecasts for 1, 2, 3, and 4 days ahead, and the official trial forecasts for 5, 6, and 7 days ahead); and,

ii. The 2004 forecasts (forecasts between 1 and 15 days ahead during the 100-day trial).

For the purpose of this analysis, temperature and precipitation components of the forecasts are taken separately.