

Numerical prediction model performance summary October to December 1998

W. Skinner and T. Hart

National Meteorological Operations Centre, Bureau of Meteorology, Australia

(Manuscript received March 1999)

Introduction

This summary continues the series comparing the performances of Numerical Weather Prediction (NWP) models.

Models and methods

A description of the Australian verification methods can be found in a previous article (Skinner 1995).

Models are from the National Meteorological Operations Centre (NMOC) Melbourne and from ECMWF (European Centre for Medium-range Weather Forecasts), NCEP (National Centers for Environmental Prediction) and UKMO (United Kingdom Meteorological Office). Reference is also made to the JMA (Japan Meteorological Agency) model.

The three models considered from NMOC, Melbourne, are: LAPS (Limited Area Prediction System); TLAPS (Tropical Limited Area Prediction System); and GASP (Global Assimilation and Prediction).

Overseas global models included in the comparisons are: ECSP (ECMWF Spectral Assimilation); USAVM (NCEP Washington Spectral model for aviation); UKGC (UK Meteorological Office Grid PE model); and JMAGSM (JMA Global Spectral Model).

Very short summaries of the models can be found in the initial article (Skinner 1995) with references to model updates in subsequent issues.

All results have been calculated within NMOC, Melbourne, where the models were verified against their own analyses. Results are presented for the irregular Australian verification area only.

The statistics are a measure of the skill in forecasting geopotential height at 500 hPa or mean sea-level pressure, MSLP. Other field types are not included in these summaries.

LAPS and TLAPS models are run several hours earlier than GASP and this premature data cut-off, particularly for satellite information, adversely affects their skill compared to GASP.

Note that the Australian region verification grid has southerly points which are outside the TLAPS grid. TLAPS scores are calculated without these points and are therefore not strictly comparable with those from other models.

Notes on NWP systems

GASP

The Global Analysis and Prediction System (GASP) was upgraded at the analysis-prediction cycle based at 0000 UTC on 9 December. The upgrade comprises a three-fold increase in the horizontal resolution and an increase in the number of vertical levels from 19 to 29. The new system has a triangular truncation of TL239 compared with T79 in the replaced system. The TL nomenclature indicates the use of a lower resolution linear grid where required for computing dynamical and physical processes.

Extensive changes have also been made to the formulation of the numerical methods in both the analysis and prediction components. The model changes have significantly reduced the computational time

Corresponding author address: Ms Wilma Skinner, National Meteorological Operations Centre, Bureau of Meteorology, GPO Box 1289K, Melbourne, Vic 3001, Australia

Fig. 1(a) Comparison for LAPS/TLAPS/GASP from October to December 1998. S1 skill-scores of MSLP using combined base-times 0000 UTC / 1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.

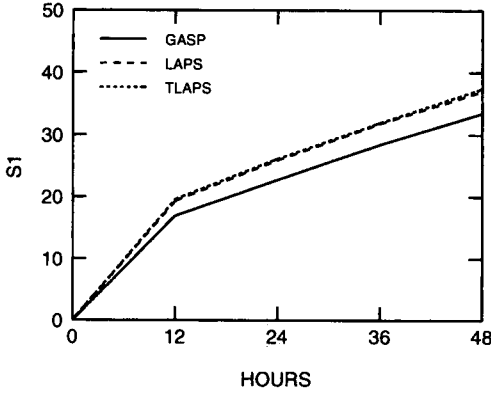


Fig. 1(b) Comparison for LAPS/TLAPS/GASP from October to December 1998. Root mean square errors of MSLP for combined base-times 0000 UTC / 1200 UTC and intervals +12, +24, 36, +48 h over the irregular Australian verification grid.

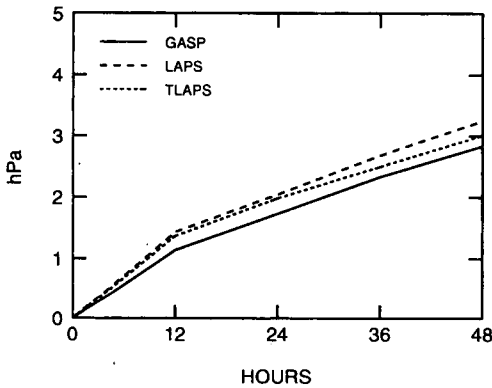


Fig. 1(c) Comparison for LAPS/TLAPS/GASP from October to December 1998. S1 skill-scores of 500 hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +12,+24, +36, +48 h over the irregular Australian verification grid.

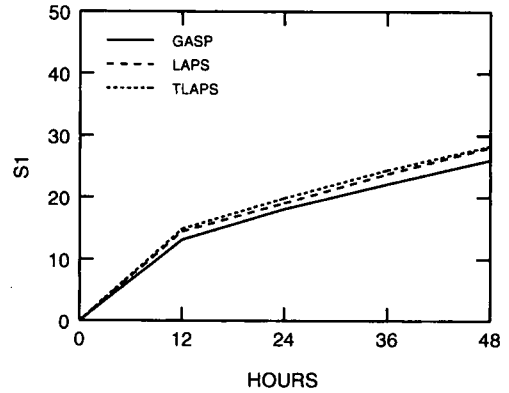
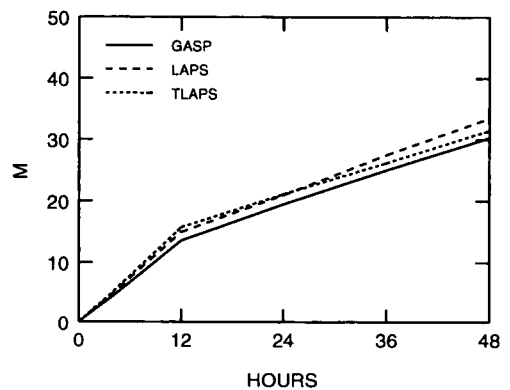


Fig. 1(d) Comparison for LAPS/TLAPS/GASP from October to December 1998. Root mean square errors of 500 hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.



required. The developments in the analysis code have generalised the software to allow its use within the limited area prediction systems and to extend the range of observations the system can handle. Further details, including the impact on prediction performance, are contained in an internal publication (NMOC 1998).

This is the first significant upgrade of an operational system resulting from the installation of the NEC SX-4 supercomputer. Upgrades of the other numerical systems are in train and are expected over the next few months now that the higher resolution boundary conditions have been established.

It is too early to detect any impact of the changes to GASP in the average verification statistics for this quarter. However, individual cases do show greater intensity in analysed and predicted features than was typical of the lower resolution T79 version.

October to December 1998 intercomparisons

Local models: (LAPS, TLAPS, GASP)

As in previous three-monthly periods, GASP continues to produce the best scores at both MSLP and 500 hPa. (Fig. 1, Figs 3(a), 3(b)).

Fig. 2(a) Comparison for GASP/EC/US/UK/JMA from October to December 1998. S1 skill-scores of MSLP for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

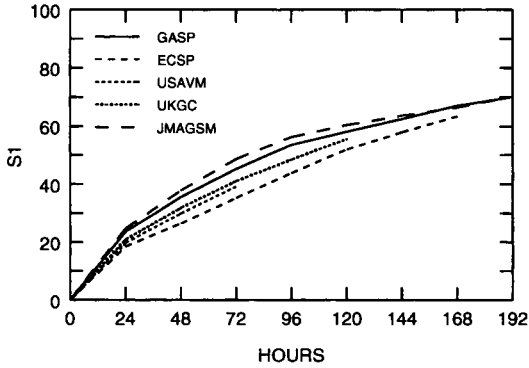
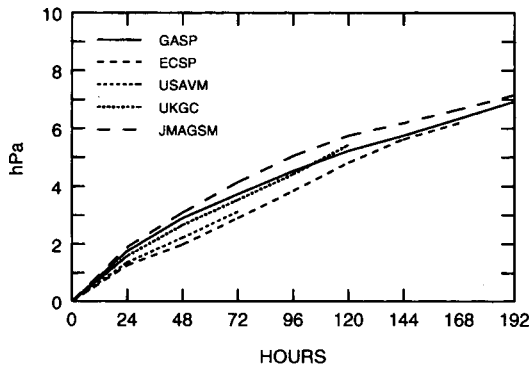


Fig. 2(b) Comparison for GASP/EC/US/UK/JMA from October to December 1998. Root mean square errors of MSLP for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.



LAPS has historically outperformed TLAPS over the Australian region, but during this three-month period, their skill scores at MSLP are indistinguishable (Figs 1 (a), 1(c)) and the rms errors show superior performance by TLAPS (Figs 1 (b), 1(d)). Figure 3(a) indicates that this improvement of TLAPS relative to LAPS occurred during November and December at MSLP.

Figures 3(e) and 3(f) show the time-series of skill-scores from July 1996 and establish the longer term relationship between the models. Figure 3(e) indicates higher scores (lower skill) for LAPS compared to the same three-month period in 1997 and lower scores (higher skill) for TLAPS. This may be due to a sig-

Fig. 2(c) Comparison for GASP/EC/US/UK/JMA from October to December 1998. S1 skill-scores of 500 hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

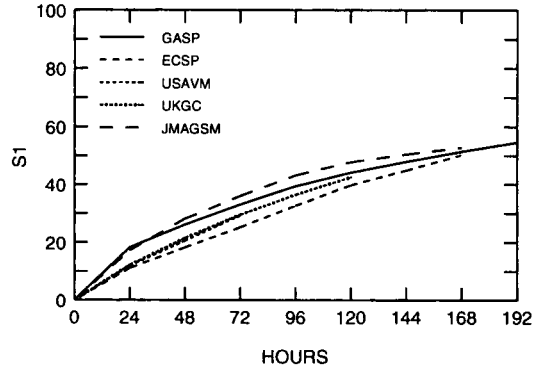
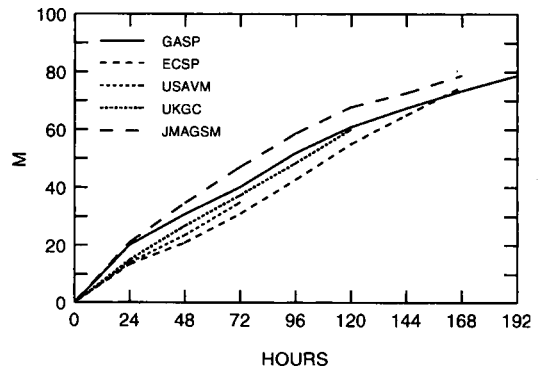


Fig. 2(d) Comparison for GASP/EC/US/UK/JMA from October to December 1998. Root mean square errors of 500 hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.



nificant number of systems occurring in the high latitude band outside the area of TLAPS but included in the verification of LAPS. An error in the LAPS soil moisture fields during this period may also have affected its performance.

Synoptic overview for 24 h MSLP predictions

An examination of the 1100 UTC prognoses for +24 hours at MSLP showed that of the 92 examples between October and December, the number scoring more than 30 skill-score points was 4, 12 and 8 for

Fig. 3(a) Monthly S1 skill-scores of MSLP for LAPS/TLAPS/GASP from October to December 1998 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.

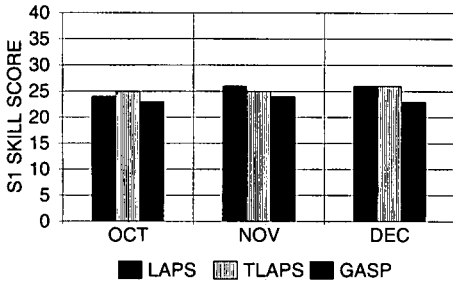


Fig. 3(b) Monthly S1 skill-scores of 500 hPa geopotential height (m) for LAPS/TLAPS/GASP from October to December 1998 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.

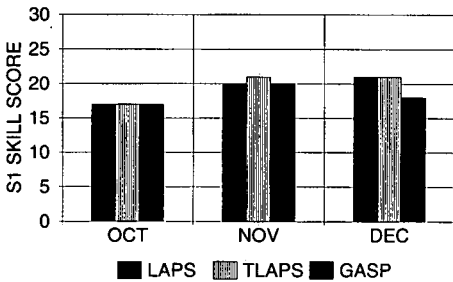


Fig. 3(c) Monthly S1 skill-scores of MSLP for GASP/EC/UK/US/JMA from October to December 1998 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.

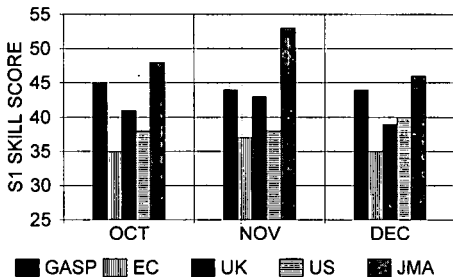


Fig. 3(d) Monthly S1 skill-scores of 500 hPa geopotential height (m) for GASP/EC/UK/US/JMA from October to December 1998 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.

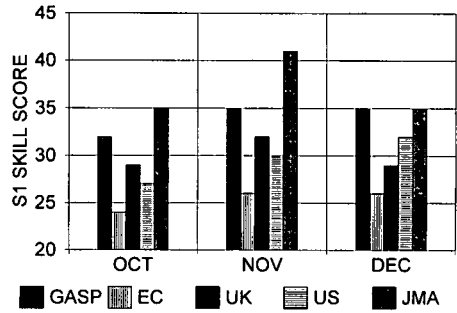


Fig. 3(e) Monthly S1 skill-scores of MSLP for LAPS/TAPS/GASP from July 1996 to December 1998 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.

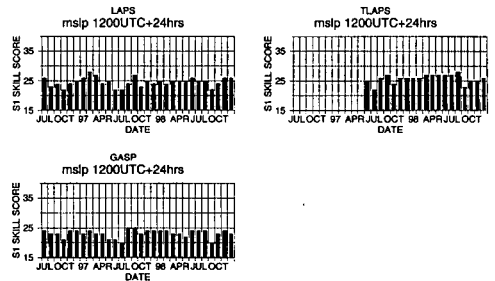


Fig. 3(f) Monthly S1 skill-scores of 500 hPa geopotential height for LAPS/TAPS/GASP from July 1996 to December 1998 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.

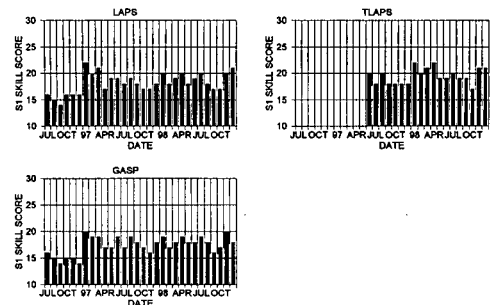


Fig. 3(g) Monthly S1 skill-scores of MSLP for GASP/EC/UK/US from April 1996 to December 1998 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.

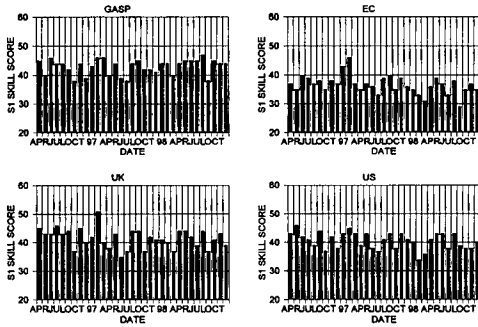
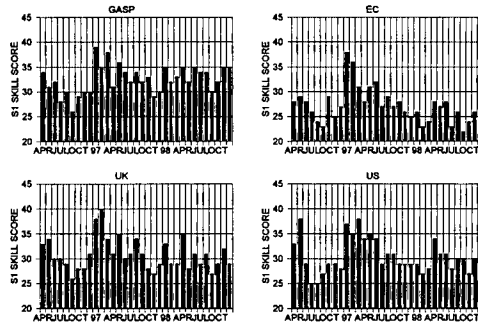


Fig. 3(h) Monthly S1 skill-scores of 500hPa geopotential height for GASP/EC/UK/US from April 1996 to December 1998 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.



GASP, LAPS and TLAPS respectively. The score of 30 was a convenient reference for a ‘poor’ prognosis. On four of these occasions (8, 9, 31 October and 22 November) all three models scored badly simultaneously. All of these were low index situations.

GASP, in addition to having the least number of ‘poor’ scores also provided the best score on 56 of the 92 days and the worst on only 11 days. The scores (Fig. 3(a)) indicate that GASP also had the lowest monthly average for each of the months.

In this period TLAPS outscored LAPS during November and equalled it during December. All models handled zonal synoptic situations well but tended to have more trouble when this flow broke down. The incorrectly forecast detail in cut-off and shearing situations often led to higher than average skill scores.

Tropical cyclones *Billy* and *Thelma* affected the Australian region during early December and caused

some small-scale but high-value errors, particularly for TLAPS and LAPS. GASP analyses and prognoses frequently do not capture the detail of tropical cyclones and so do not generate these errors. The model can score better than the higher resolution models even though it is a less useful indicator of these features.

On 26 and 27 December, a developing deep depression over eastern Bass Strait resulted in storm-force winds and loss of life in the Sydney to Hobart yacht race. The 24-hour forecasts for the 26 December show LAPS and TLAPS underestimated the deepening of this system, but GASP was quite accurate. All three models produced good forecasts for 27 December. Reference is made in a later section to the +72-hour forecasts for the same situation.

Global Models: (GASP, ECSP, UKGC, USAVM, JMAGSM)

The skill-scores and rms errors (Fig. 2) clearly identify ECSP as the best overall of the five models.

USAVM is slightly better than UKGC at MSLP (Figs 2(a), 2(b)) and at 500 hPa for rms errors (Fig. 2(d)) but there is no significant difference between these two for skill-scores at 500 hPa (Fig. 2(c)).

The skill-scores at +72 hours for individual months (Figs 3(c), 3(d)), confirm ECSP as the best performer in each month at both levels. Also USAVM scores lower than UKGC during October and November but higher in December. The plots show GASP mostly scoring lower than JMAGSM, but they were equal in December.

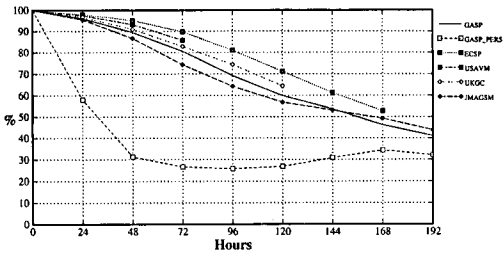
The time-series of these relationships (Figs 3(g), 3(h)) does not include the JMAGSM model, but establishes the relationship between the other models over the longer term.

The Anomaly Correlation (Fig. 4) includes plots of persistence generated from GASP analyses. If the 60 per cent correlation line is taken as a measure of usefulness, all models except JMAGSM are useful at +120 hours and ECSP is still useful at +144 hours. Longer interval forecasts are, on average, below this reference level although on an individual day, the forecast may be good. This repeats the pattern established during the period July to September 1998.

Synoptic intercomparison of global models

This summary looks at the 1200 UTC 72-hour predictions from the global systems for the Australian region, in an attempt to determine the major types of

Fig. 4 Anomaly Correlation of MSLP for GASP/EC/US/UK/JMA from October to December 1998 for combined base-times 0000 UTC / 1200 UTC over the irregular Australian verification grid.



synoptic situations for which there are large errors common to most of the predictions. The main diagnostic tool was the MSLP forecast error chart comparing the prediction with the corresponding analysis for the valid date. Admittedly, this narrow focus on MSLP biases the comments towards higher latitudes. The errors also depend on the details in each system's analysis and prediction, with those systems producing more intense features in either the analysis or prediction being more prone to larger errors than those with smoother features.

The performance of the global NWP systems for the period of the Sydney-Hobart yacht race is of particular interest. Of all the 72-hour predictions during the October-December period the forecasts for 1200 UTC on 27 December were the poorest collective set from the ensemble of models available. The US, UK and GASP predictions failed to show any significant feature in the area of the observed low, and consequently had errors exceeding 25 hPa. The JMA system did predict a low near the observed location, but the central pressure was about 10 hPa too high and the secondary vorticity maximum through Bass Strait was absent. The ECMWF result was best in intensity but the speed of movement was too fast and the predicted low centre was about five degrees southeast of the observed centre. The predictions for 1200 UTC 26 December, the cyclogenesis stage, were also collectively poor, with errors exceeding 10 hPa, except in the ECMWF case.

The performance of the global systems for this situation indicates a general weakness at the 72-hour range in handling such intense, small-scale and rapidly-developing storms. Within the 48-hour forecast range the performance was much better. Several of the models did predict some cyclogenesis at ranges longer than 72 hours, although the intensity and location were poorly represented. This case reinforces common experience that even where the global models do portray cyclogenesis, with some degree of skill,

the details of the system may not be predicted well and the possibility of intense embedded centres should be considered.

As a criterion of a common serious error, cases were considered where at least four of the available predictions had errors exceeding 10 hPa over the 'Australian region', loosely defined as 10°-40°S, 100°-140°E and 10°-45°S, 140°-160°E. Excluding tropical cyclones there were sixteen such cases in the 1200 UTC 72-hour predictions valid during October to December, including 26 and 27 December. The frequency of occurrence of such errors in the individual models ranged from about 35 to 40 per cent for JMA, GASP and UK to 21 per cent for ECMWF.

In the 16 cases of common large prediction problems, the errors were generally associated with a synoptic feature which was common across the predictions. They represented several major types of synoptic situations:

(a) cases of strong mobile high latitude cyclones
These occurred mainly early in October when errors of detail in location or structure within strong gradients were enough to meet the criterion. Another case occurred in the southwest of the region on 29 November.

(b) cut-off lows or easterly dips
These were by far the most common type and occurred right across the region. In addition to the Sydney-Hobart yacht race cases, particular examples were:

- underdevelopment and then too rapid movement of a low over the north Tasman Sea on 10 and 11 October (except for ECMWF), with similar problems for lows further south on 27 October, 29 November and 16 December,
- very variable performance over the Bight on 12 October
- under-development of a cut-off low southwest of Perth on 31 October.

Despite cases of over-development of lows on occasions by individual models, the cases of common large errors were associated with under-development or too rapid movement of low pressure systems.

(c) cases of interaction over the Bight between remnants of a heat low from Western Australia and a higher latitude trough (25 November and 1 December).

(d) tropical cyclones, particularly TC *Thelma* from 7 to 12 December.

References

- Skinner, W. 1995. Numerical prediction model performance summary April to June 1995. *Aust. Met. Mag.*, 44, 309-312.
- NMOC 1998. *Operations Bulletin No.45*, Bureau of Meteorology, GPO Box 1289K, Melbourne, Vic 3001, Australia.