

Numerical prediction model performance summary January to March 1997

W. Skinner and T. Hart

National Meteorological Centre, Bureau of Meteorology, Australia

(Manuscript received June 1997)

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 both the National Meteorological Centre (NMC) Melbourne and from ECMWF (European Centre for Medium-range Weather Forecasts), NCEP (National Centers for Environmental Prediction) and UK (United Kingdom Meteorological Office).

The three models considered from NMC Melbourne are: LAPS (Limited Area Prediction System), TAPS (Tropical Analysis and 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) and UKGC (United Kingdom Meteorological Office Grid PE model).

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

In general, results have been calculated within NMC Melbourne, where the models were verified against their own analyses for the irregular Australian verification area only. The above is true for Figs 1, 2 and 3 but Fig. 4 shows verification data supplied by ECMWF and NCEP for models verified against radiosondes in the southern hemisphere.

These statistics are in accordance with the recommendations of the World Meteorological Organization's

Commission for Basic Systems. In this context the southern hemisphere is 20° to 90° South and models are verified against a list of 68 stations.

All statistics are a measure of the skill in forecasting geopotential height or MSLP. Other field types are not included in these summaries. LAPS and TAPS 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 TAPS grid. TAPS scores are calculated without these points and are therefore not strictly comparable with those from other models.

Notes on NWP systems

UKGC

From 5 November 1996, the UKGC global model included convective momentum transport, gravity wave drag, fourth-order advection and assimilation of humidities derived from satellite radiances. In parallel testing this package of changes reduced the root mean square vector wind error by up to 10 per cent in the tropics. MSLP forecasts at longer time intervals also improved with the change to advection. The assimilation of satellite radiances improved tropical forecasts. (UK Met. Office 1996).

Temporary loss of 250 km TOVS temperature soundings

The binary reports containing the TOVS vertical soundings of temperature and moisture at 250 km horizontal resolution ceased on 19 March. These reports had been transmitted from Washington and used in the Bureau's

Corresponding author address: Ms W. Skinner, National Meteorological Centre, GPO Box 1289K, Melbourne, Vic. 3001, Australia.

Fig. 1(a) Comparison for LAPS/TAPS/GASP from January to March 1997. 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.

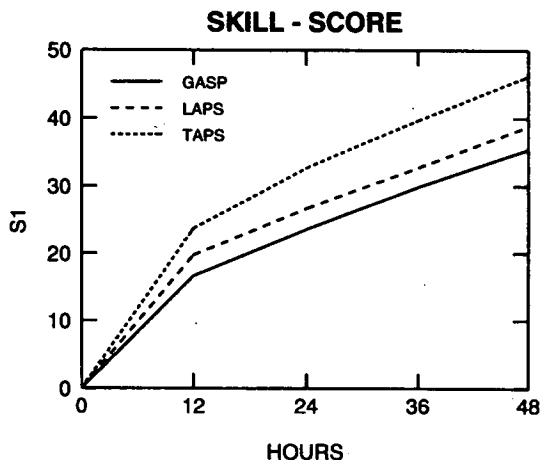
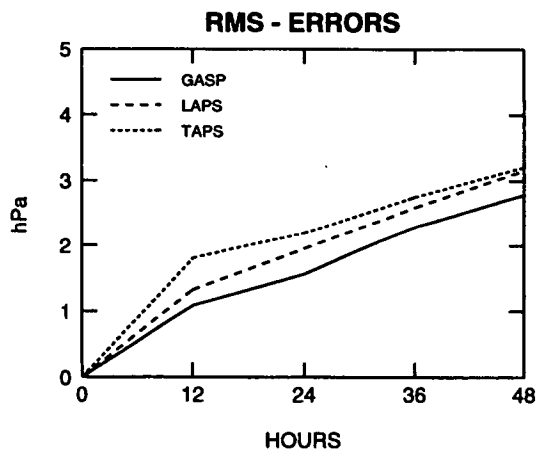


Fig. 1(b) Comparison for LAPS/TAPS/GASP from January to March 1997. Root mean square errors of MSLP for combined base-times 0000 UTC/1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.



numerical analyses where they had been shown to have a positive impact over the use of the 500 km SATEM messages alone. An alternative source of the messages was available from NOAA/NESDIS and has been included in the operational models since 2 April 1997.

LAPS and TAPS also have available the locally received and processed full resolution TOVS for the Australian region, so the greatest impact from the loss of the reports was expected in GASP.

Fig. 1(c) Comparison for LAPS/TAPS/GASP from January to March 1997. S1 skill scores of 500 hPa geopotential height for combined base-times 0000 UTC/1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.

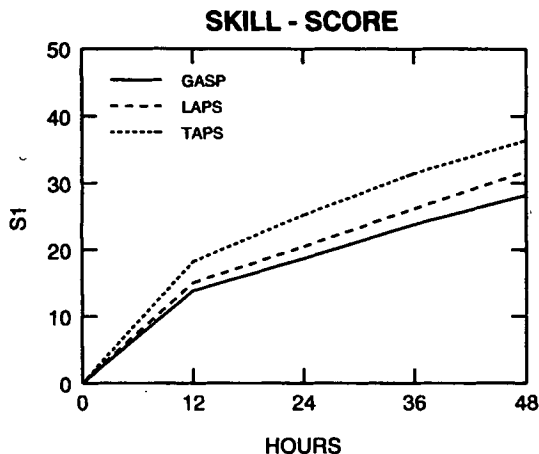
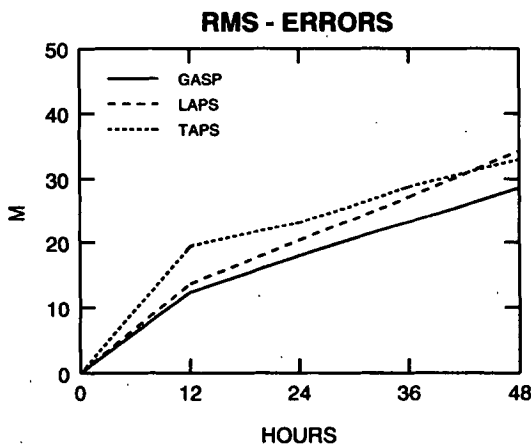


Fig. 1(d) Comparison for LAPS/TAPS/GASP from January to March 1997. Root mean square errors of 500 hPa geopotential height (m) for combined base-times 0000 UTC/1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.



January to March 1997 intercomparisons

Local models: (LAPS, TAPS, GASP)

The local models performed in order GASP, LAPS and TAPS. This order is maintained for skill scores and root mean square (rms) errors and at MSLP and 500 hPa. Both the three-monthly averages (Fig. 1) and the individual monthly scores (Figs 3(a), 3(b)) show this per-

Fig. 2(a) Comparison for GASP/EC/US/UK from January to March 1997. 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.

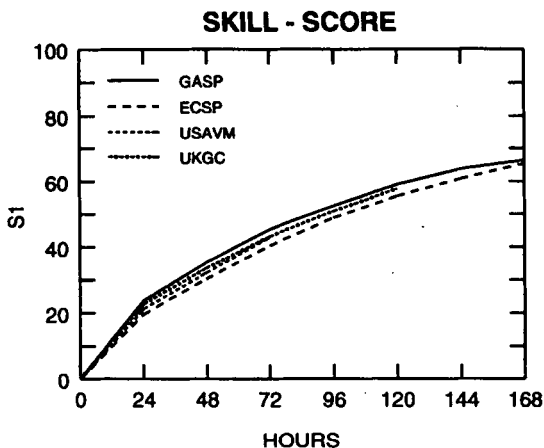


Fig. 2(c) Comparison for GASP/EC/US/UK from January to March 1997. S1 skill scores of 500 hPa geopotential height for combined base-times 0000 UTC/1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

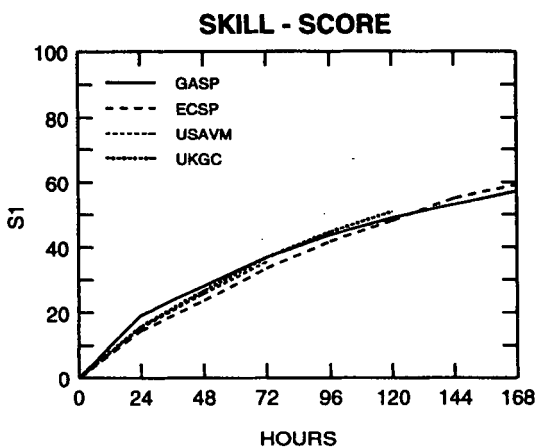


Fig. 2(b) Comparison for GASP/EC/US/UK from January to March 1997. Root mean square errors of MSLP for combined base-times 0000 UTC/1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

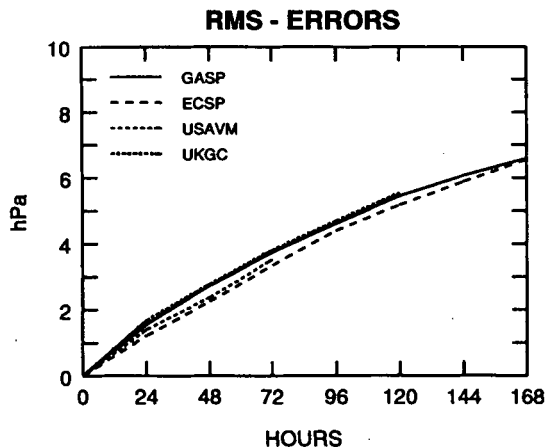
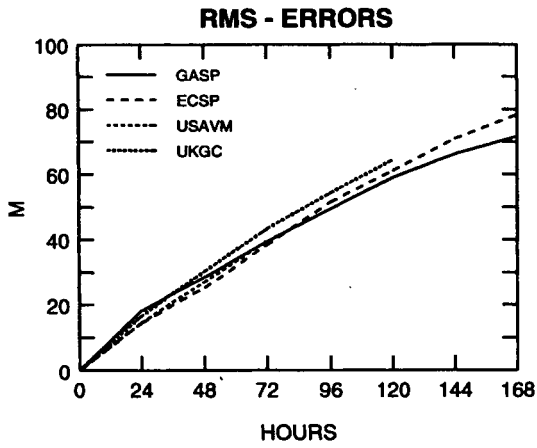


Fig. 2(d) Comparison for GASP/EC/US/UK from January to March 1997. Root mean square errors of 500 hPa geopotential height (m) for combined base-times 0000 UTC/1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.



formance order. The margin of improvement of GASP over LAPS has increased in this period. All the models had higher skill scores at both levels for January to March than were seen for October to December 1996 although GASP showed the lowest increase, particularly at MSLP.

Examination of the daily MSLP synoptic charts for the period shows that significant differences between GASP and LAPS often occurred when tropical cyclones

or low index situations were present. GASP was more skillful with slow-moving anticyclones and cut-off lows and was only rarely out-scored by LAPS.

TAPS and LAPS can be rated as performing quite poorly when tropical cyclones are present. Tropical cyclones caused least damage to GASP's verification because this model makes no attempt to force in these small-scale intense systems. Hence both starting and verifying analyses have very smoothed representations

Fig. 3(a) Monthly S1 skill scores of MSLP for LAPS/TAPS/GASP from January to March 1997 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.

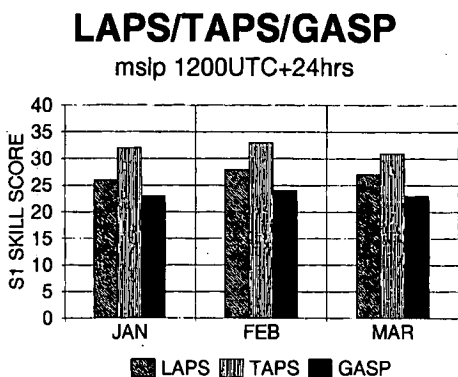
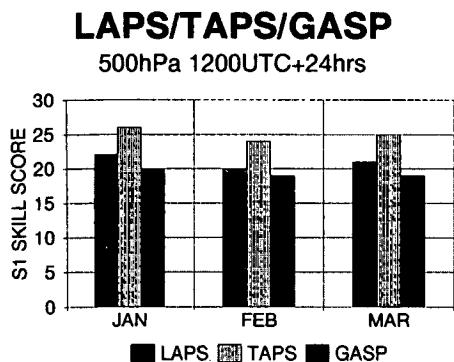


Fig. 3(b) Monthly S1 skill scores of 500 hPa geopotential height for LAPS/TAPS/GASP from January to March 1997 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.



of the cyclone if indeed it is captured at all. LAPS and TAPS on the other hand have a process for forcing the analysis to try to represent the cyclones. Both position and central pressure are difficult to capture accurately at the resolution of these models and small errors in the predicted location of a small intense system can result in large errors in gradient and can impact adversely on the skill of the prediction systems. As seven tropical cyclones (*Rachel, Drena, Pancho, Gillian, Harold, Ita and Justin*) occurred in the Australian region during these three months, there is likely to have been a significant impact on the statistics for LAPS and TAPS. The Australian verification region however does not extend north of 15°S and hence misses many tropical features.

Fig. 3(c) Monthly S1 skill scores of MSLP for GASP/EC/UK/US from January to March 1997 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.

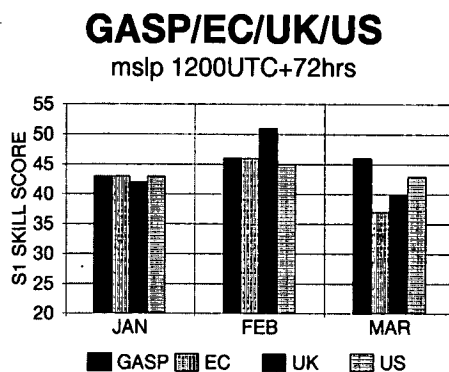
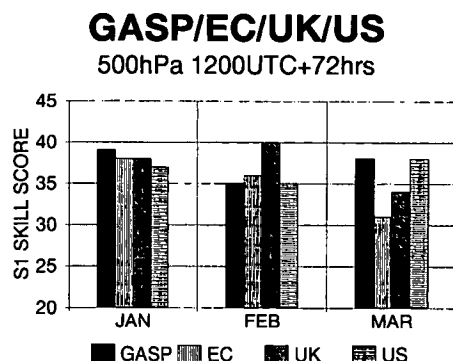


Fig. 3(d) Monthly S1 skill scores of 500 hPa geopotential height for GASP/EC/UK/US from January to March 1997 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.



Global models: (GASP, ECSP, UKGC, USAVM)

Figure 2 shows that ECSP is still the best performed of the global models that are available in NMC, particularly for MSLP. UKGC skill scores outperform GASP at MSLP and to +72 h at 500 hPa. Root mean square errors show greater skill for GASP from +48 h at 500 hPa. USAVM is superior to both UKGC and GASP but is available only to +72 h.

Scores for individual months (Figs 3(c), 3(d)) show more variability. The relatively poor performance of GASP in March may be attributed to the loss of the higher resolution TOVS data for the last fortnight.

A comparison of global models at +120 h

The use of +120 h for this synoptic comparison has to be

Fig. 4(a) Monthly root mean square errors of 500 hPa geopotential height (m) against radiosondes for GASP/EC/US from January 1996 to March 1997 for base-time 1200 UTC and interval +24 h over the southern hemisphere area 20°S to 90°S.

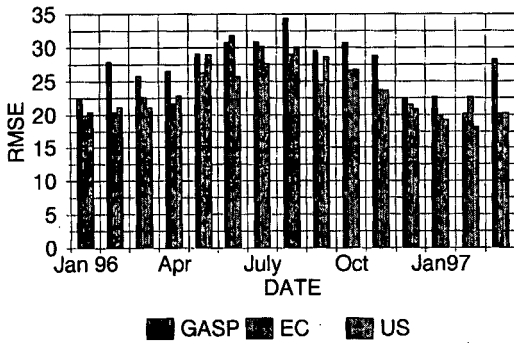


Fig. 4(b) Monthly root mean square errors of 500 hPa geopotential height (m) against radiosondes for GASP/EC/US from January 1996 to March 1997 for base-time 1200 UTC and interval +72 h over the southern hemisphere area 20°S to 90°S.

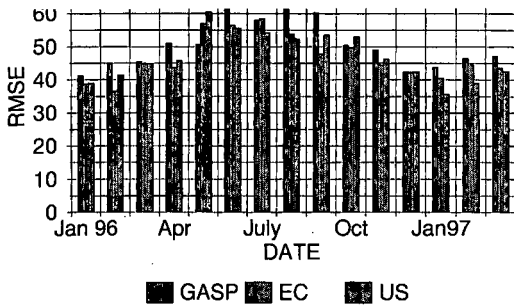
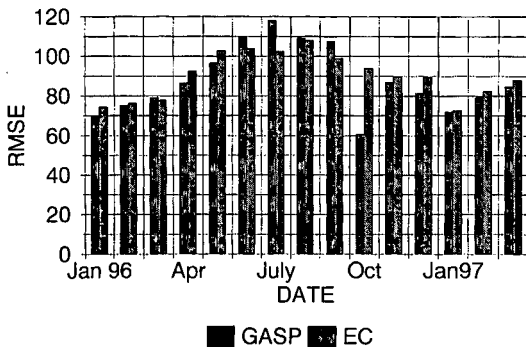


Fig. 4(c) Monthly root mean square errors of 500 hPa geopotential height (m) against radiosondes for GASP/EC/US from January 1996 to March 1997 for base-time 1200 UTC and interval +168 h over the southern hemisphere area 20°S to 90°S.



restricted to GASP, ECSP and UKGC as the USAVM model is only issued to +72 h.

Errors exceeding 10 hPa in magnitude occurred on about 90 per cent of occasions for each of the three models considered. This is slightly lower than the winter and spring periods when errors exceeding 10 hPa were present in virtually every case. The frequency of errors exceeding 20 hPa was also significantly lower in this predominantly summer period, ranging from 18 per cent for UKGC to 24 per cent for GASP. The performance of UKGC on this measure is in contrast to the previous two periods, when the UKGC had the highest percentage of such errors. Again UKGC and GASP showed a preponderance of positive errors while the ECMWF errors were quite evenly distributed in sign. The large errors were located predominantly at high latitudes in the strong westerly regime or were associated with tropical systems. The high detail in the ECMWF analyses and prognoses could lead to large errors resulting from relatively small location errors, so errors due to mislocations of intense tropical features alone were not included in the statistics.

On several occasions relatively weak low pressure systems developed in the eastern Indian Ocean region and became cut-off before drifting slowly eastwards. Although the initial developments were generally captured by the models (except on occasions for GASP) all the models tended to perform poorly in the later stages by weakening the system and moving it much too rapidly southeastwards into the westerly flow. One such case occurred during the period 18-22 January. On 18 January a secondary low developed at the apex of a broad trough in the southwest of the study region. By 19 January a closed low of 990 hPa was analysed at 40°S 116°E. The low then moved slowly eastwards and intensified near Tasmania on 22 January. The GASP predictions failed to capture the initial development. The UKGC and ECMWF predictions showed the development but located the centre about 10 degrees too far to the southeast. On 22 January all models showed only a weak trough over Tasmania and all had errors of about +20 hPa for that region. Similar cases occurred for 25-28 February and 15-20 March.

There were also several cases where developments of low pressure systems were missed by some of the models. Particular examples were developments in the east Australian or Tasman Sea region on 7 January, 3 February and 9 February. In each of these cases the GASP and UKGC predictions failed to realise these developments while the ECMWF forecasts did, although in the 3 February case the predicted centre was located about 600 km too far west. This feature of the ECMWF predictions for 3 February is again

indicative of a slight tendency for overprediction of cyclonic activity. Another notable case was for 28 February when the ECMWF prediction showed a low of 985 hPa south of Western Australia at 48°S 118°E. The verifying analysis showed a ridge in the area resulting in an error of -36 hPa near the forecast centre of the low.

GASP's tendency to overdevelop the heat troughs over the continent persisted, particularly over Western Australia. This problem also occurred occasionally in UKGC but not as consistently as in GASP. On 1 March the GASP forecast MSLP over southern Western Australia was up to 13 hPa below the analysed value.

The five-day predictions, especially those of ECMWF, again provided surprisingly useful guidance on the development and subsequent movement of tropical lows although there were also spurious predictions of cyclogenesis such as 28-29 January and 23-24 February. The ECMWF predictions in early January for tropical cyclone *Rachel* in the northwest and *Drena* over the Coral Sea were quite good in the early stages although falling away in the later stages. The retention of *Justin* by ECMWF for long periods was also noteworthy although the predicted locations may have been in error.

Comparison of models against radiosondes

Figure 4 shows rms errors from a comparison of model forecasts against a list of 68 WMO-nominated radiosonde stations. The level is 500 hPa and intervals +24 h, +72 h and +168 h. Only GASP and ECSP can be shown at +168 h but USAVM is available up to +72 h. The verification area is from 20°S to 90°S and 0° to 360°E. The southern hemisphere area is used here, rather than the Australian region only, as there are not many radiosondes in the Australian region at 1200UTC.

GASP's performance increases with the longer interval and by +168 h is slightly ahead of ECSP in the warmer months. At +24 h it is consistently the worst performer with the +24 h March result being particularly poor. This should improve with the restoration of full TOVS coverage to the model from early in April 1997.

References

- Skinner, W. 1995. Numerical prediction model performance summary April to June 1995. *Aust. Met. Mag.*, 44, 309-12.
- UK Met. Office 1996. nwp Gazette. *A quarterly Newsletter about Numerical Weather Prediction from the UK Met. Office*, Fourth Quarter 1996. Vol. 3 No 1, 14.