

Numerical prediction model performance summary July to September 1996

W. Skinner and T. Hart

National Meteorological Centre, Bureau of Meteorology, Australia

(Manuscript received December 1996)

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 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 as follows: GASP (Skinner and Hart 1995c), TAPS (Skinner and Hart 1995d), ECSP (Skinner and Hart 1995d), USAVM (Skinner and Hart 1995c), UKGC (Skinner and Hart 1995c and 1995d).

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°S 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 were run several hours earlier than GASP and this premature data cut-off, particularly for satellite information, adversely affected their measured skill against GASP.

Note that the Australian region analysis 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

LAPS

From 3 July 1996, a new model, LAPS, became the regional operational model for NMC Melbourne. It is a four-dimensional intermittent data assimilation scheme consisting of an analysis-initialisation-forecast cycle, using the finite difference approach on a latitude/longitude grid and sigma levels for both analysis and forecasts. LAPS uses a multivariate statistical interpolation (MVISI) (Seaman et al. 1995) analysis method. The physics package is now basically a regional grid-point

Fig. 1(a) Comparison for LAPS/TAPS/GASP from July to September 1996. S1 Skill-scores at MSLP for combined base-times 0000 UTC/1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.

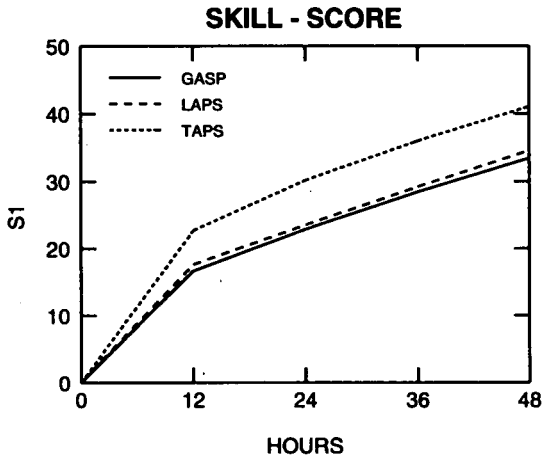


Fig 1(c) Comparison for LAPS/TAPS/GASP from July to September 1996. S1 Skill-scores at 500 hPa for combined base-times 0000 UTC/1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.

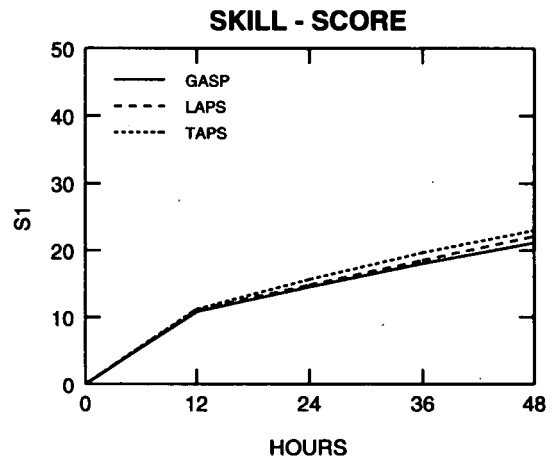


Fig 1(b) Comparison for LAPS/TAPS/GASP from July to September 1996. Root mean square errors at MSLP for combined base-times 0000 UTC/1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.

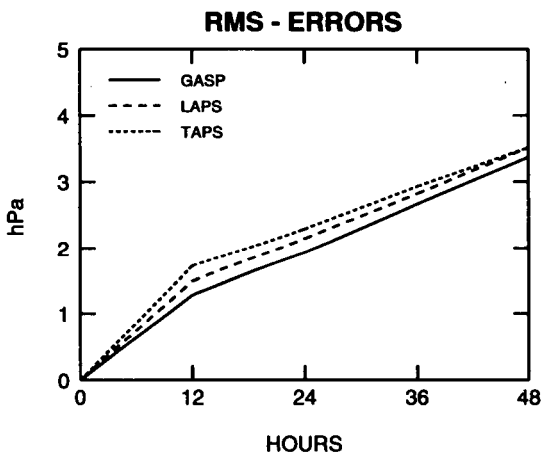
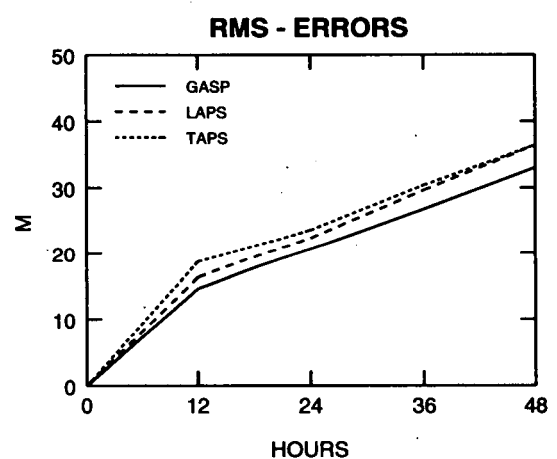


Fig 1(d) Comparison for LAPS/TAPS/GASP from July to September 1996. Root mean square errors at 500 hPa for combined base-times 0000 UTC/1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.



version of that used in GASP. Also, LAPS takes boundary conditions from GASP fields. The new model incorporates GMS moisture data and, when appropriate, tropical cyclone bogus data. The domain now covered is 16.75N - 65.00S and 65.00E - 184.25E. Horizontal resolution is 0.75 degrees. There are 19 vertical levels for both analysis and prognosis. The model is run daily from base times 0000 UTC and 1200 UTC for forecast periods up to +48 h. Although the domain is extended, verification statistics presented here are still for the Australian verification area and only for verification of

geopotential or mean sea-level pressure. A short summary of LAPS can be found in Le Marshall and Mills (1995).

July to September 1996 intercomparisons

Local models: (LAPS, TAPS, GASP)

A marked improvement for the regional operational model scores was apparent for this quarter. The new model showed better scores than TAPS for both MSLP

Fig 2(a) Comparison for GASP/EC/US/UK from July to September 1996. S1 Skill-scores at MSLP for combined base-times 0000 UTC/1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

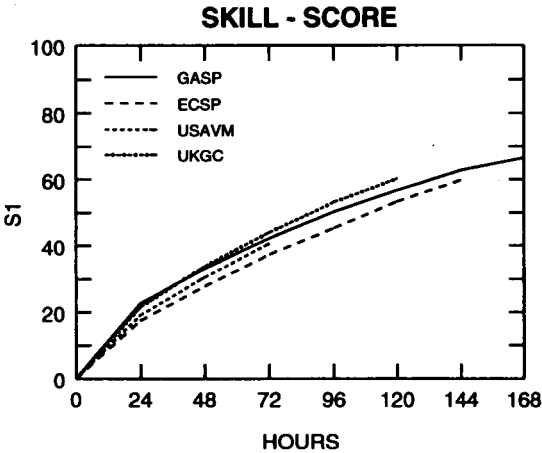


Fig 2(c) Comparison for GASP/EC/US/UK from July to September 1996. S1 Skill-scores at 500 hPa for combined base-times 0000 UTC/1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

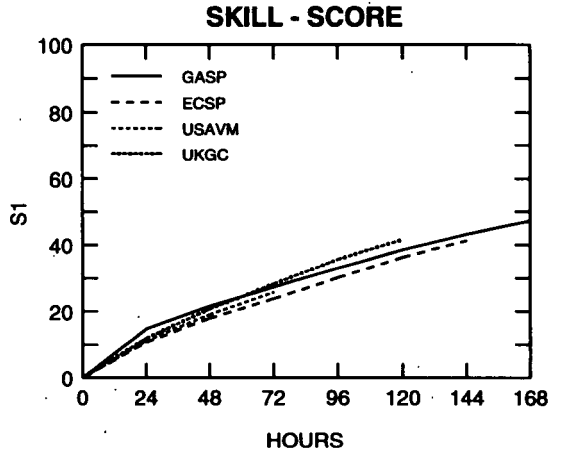


Fig 2(b) Comparison for GASP/EC/US/UK from July to September 1996. Root mean square errors at MSLP for combined base-times 0000 UTC/1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

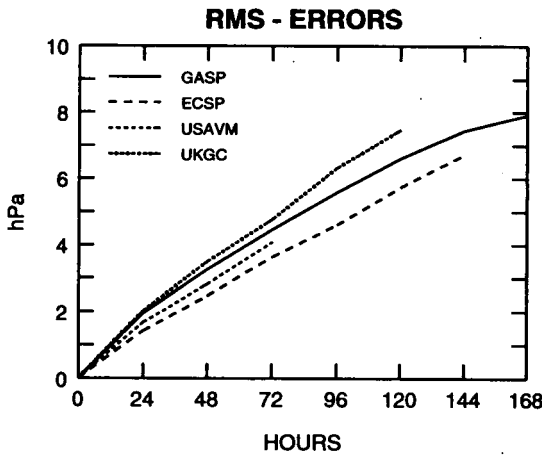
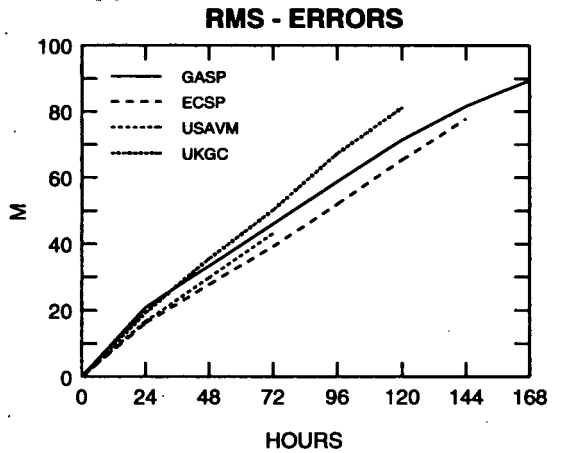


Fig 2(d) Comparison for GASP/EC/US/UK from July to September 1996. Root mean square errors at 500 hPa for combined base-times 0000 UTC/1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.



and 500 hPa and for skill-scores and rms errors (Fig. 1). GASP was still the best model, but its margin of superiority was significantly reduced. It was only 3 h better than LAPS, while still 18 h better than TAPS for MSLP skill-scores (Fig. 1(a)), i.e. at +48 h the GASP skill-score was about the same as for TAPS at +30 h. GASP was run later than LAPS and TAPS and from a better observational base, so the performance of LAPS was impressive. TAPS is handicapped by not having grid-points over the whole verification area. The monthly breakdown of the quarterly statistics (Fig. 3(a,b)) confirms that the operational regional model, LAPS, is now

closer in skill to GASP than to TAPS. This is evident for each separate month, and at both MSLP and 500 hPa although the difference between models is less marked at the upper level. GASP and LAPS both handled low-index situations better than TAPS, particularly in estimating central values for lows. GASP was best at estimating the speed and intensity of troughs in westerly situations.

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

Figure 2 shows ECSP as superior to the other models for each statistic and each level. USAVM, GASP and

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

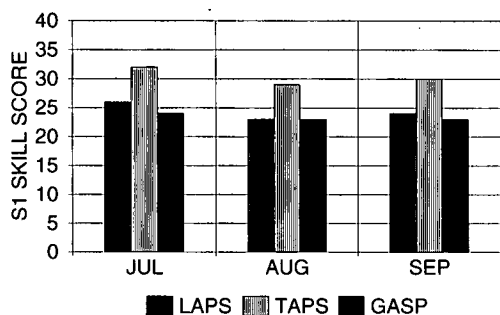


Fig. 3(c) Monthly S1 skill-scores at MSLP for GASP/EC/UK/US from July to September 1996 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.

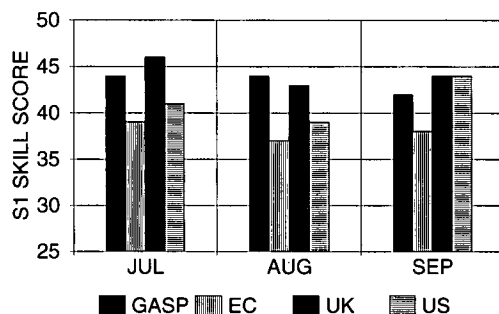


Fig. 3(b) Monthly S1 skill-scores at 500 hPa for LAPS/TAPS/GASP from July to September 1996 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.

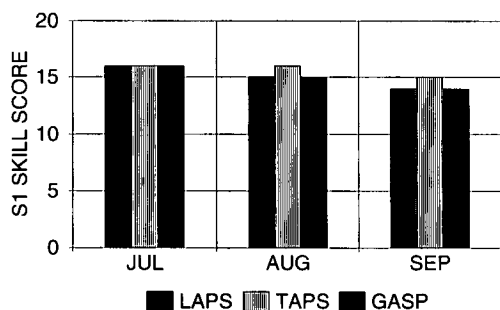
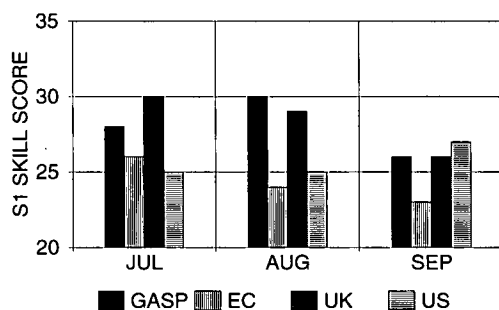


Fig. 3(d) Monthly S1 skill-scores at 500 hPa for GASP/EC/UK/US from July to September 1996 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.



UKGC are next best respectively except that UKGC outscores GASP out to +48 h. (Fig. 2). The monthly values for each model at +72 h (Fig. 3(c, d)) reveal a more volatile relative skill between models with the USAVM at 500 hPa having the best score in July and the worst in September (Fig. 3 (d)). The ECMWF has the best result in every other case.

A comparison of global models at +120 h

The use of +120 h for this synoptic comparison has to be restricted to ECSP, GASP and UKGC as the USAVM model is only issued to +72 h.

The biases of those situations where large errors (more than 20 hPa at MSLP) occurred illustrated the previously noted tendencies for ECMWF to over-intensify systems and for GASP to miss or weaken troughs and lows. In the case of UKGC it seemed to be a case of missing developments of lows combined with a tendency for strong high latitude ridging. The average

skill-score for the five-day predictions is approaching the marginal level. For example, it is about the same level as using persistence for a 24 h forecast. There were some occasions when the predictions had the broadscale pattern completely wrong. One such example was the set of forecasts for 10 September when the real world had lows over the Bight and Tasman and all the models had the lows where the ridges should have been. More generally, the problems were with the representation of the synoptic features within the broader flow. Despite the marginal level of average skill, there were some very impressive predictions including cases of rapid developments, especially from the ECSP system. One notable example was an east coast low which developed on 30 August out of a trough over eastern Australia. The ECSP predictions were clearly the most successful in indicating the development of such an intense system and its subsequent movement. On the other hand, the ECSP predictions continued their previ-

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

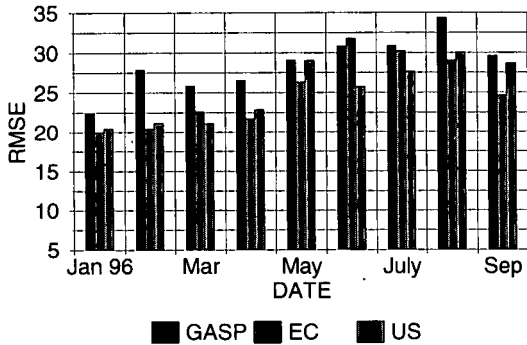


Fig. 4(b) Monthly root mean square errors against radiosondes at 500 hPa for GASP/EC/US from January to September 1996 for base-time 1200 UTC and interval +72 h over the southern hemisphere area 20°S to 90°S.

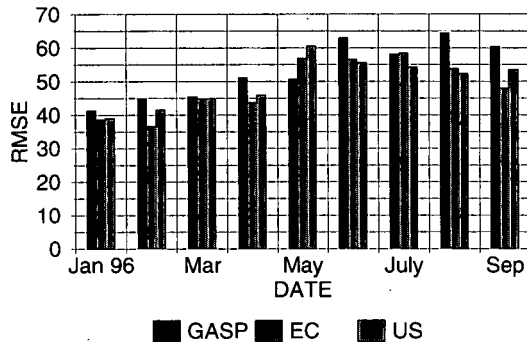
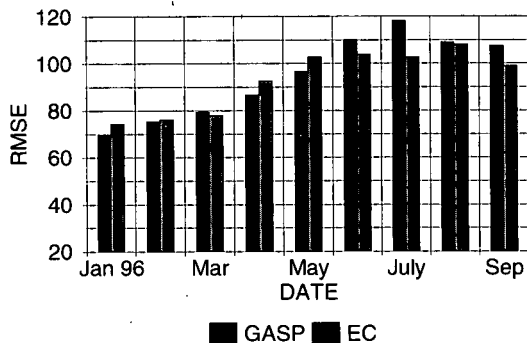


Fig. 4(c) Monthly root mean square errors against radiosondes at 500 hPa for GASP/EC from January to September 1996 for base-time 1200 UTC and interval +168 h over the southern hemisphere area 20°S to 90°S.



ously noted over-active characteristic. Even in the 30 August case, they were a day too early in developing the system, and over-deepened the system by 12 hPa on 31 August. There were also some cases of predicted development which did not occur. Despite this tendency for over-activity, a low over Bass Strait on 29 July was missed almost completely in the five-day prediction as were some cut-off lows in the east Indian Ocean such as on 13 July. In contrast, predictions from GASP and UKGC continued to miss or under-predict developing lows, and to move them too quickly southeastwards into the westerly flow when they did occur.

Figures 4(a, b, c) show the performance of GASP, ECSP and USAVM verified against radiosondes for the southern hemisphere for January to September 1996. The area covered is 20°S to 90°S.

Root mean square errors for 500 hPa at +24 and +72 h are included for all three models, and to +168 h for GASP and ECSP. Errors peak in winter and by September are reducing. GASP was generally not as skillful as the other models although its comparative performance improves for longer intervals. ECSP and USAVM do not seem to have established a pattern of superiority as sometimes one scores better and sometimes the other.

References

Le Marshall, J.F. and Mills, G.A. 1995. Tropical cyclone *Bobby* - a notable example of the impact of local TOVS data. *Aust. Met. Mag.*, 44, 293-7.

Seaman, R., Bourke, W., Steinle, P., Hart, T., Embery, G., Naughton, M. and Rikus, L. 1995. Evolution of the Bureau of Meteorology's global assimilation and prediction system. Part 1: analysis and initialisation. *Aust. Met. Mag.*, 44, 1-18.

Skinner, W. 1995. Numerical prediction model performance summary April to June 1995. *Aust. Met. Mag.*, 44, 309-312.

Skinner, W. and Hart, T. 1996a. Numerical prediction model performance summary July to September 1995. *Aust. Met. Mag.*, 45, 59-62.

Skinner, W. and Hart, T. 1996b. Numerical prediction model performance summary October to December 1995. *Aust. Met. Mag.*, 45, 131-5.

Skinner, W. and Hart, T. 1996c. Numerical prediction model performance summary January to March 1996. *Aust. Met. Mag.*, 45, 189-92.

Skinner, W. and Hart, T. 1996d. Numerical prediction model performance summary April to June 1996. *Aust. Met. Mag.*, 45, 261-5.

