

# Numerical prediction model performance summary April to June 1996

W. Skinner and T. Hart

National Meteorological Centre, Bureau of Meteorology, Australia

(Manuscript received September 1996)

## Introduction

This summary continues the series comparing the performances of numerical weather prediction (NWP) models for a three-monthly period.

## Models and methods

Explanations of the National Meteorological Centre (NMC) Melbourne NWP models and global models from other operational centres, together with Australian verification methods, can be found in a previous article (Skinner 1995). The three models considered from NMC Melbourne are: RASP (Regional Assimilation and Prognosis); TAPS (Tropical Analysis and Prediction System); and GASP (Global Assimilation and Prediction). Overseas global models included in the comparisons are: ECSP (European Centre for Medium-range Weather Forecasts (ECMWF) Spectral Assimilation); USAVM (National Centers for Environmental Prediction (NCEP), Washington, Spectral model for aviation); and UKGC (UK Meteorological Office Grid PE model).

Verification entities have been calculated within NMC Melbourne, and models were verified against their own analyses. Quoted results apply to the irregular Australian verification area only. The above is true for Figs 1, 2 and 3.

Figure 4 includes verification data supplied by ECMWF and UK Meteorological Office. These data show models verified against radiosondes for 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 degrees to 90 degrees south. Models are verified against a list of 68 stations.

RASP 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

### RASP

This issue will be the last to contain statistics for the local model RASP. It was replaced on 3 July 1996 with a new model LAPS (Limited Area Prediction System). LAPS will be described briefly in the next article when the first set of comparative statistics will also be available.

## April to June 1996 intercomparisons

### Local models: (RASP, TAPS, GASP)

The clearest feature of the April to April graphs is the superiority of GASP to both RASP and TAPS. Both at MSLP and 500 hPa and for skill scores and rms errors, GASP outperformed the other local models (Fig. 1, Fig. 3(a,b)). At best (Fig. 1(a)) GASP was 18 hours better, i.e. at +36 h GASP scores were similar to the other models' scores at +18 h. It should be remembered, however, that GASP is run later than the others and with a better observational base.

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

Fig. 1(a) Comparison for RASP/TAPS/GASP from April to June 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.

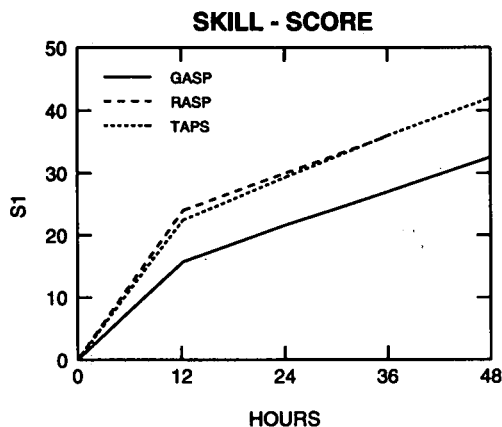
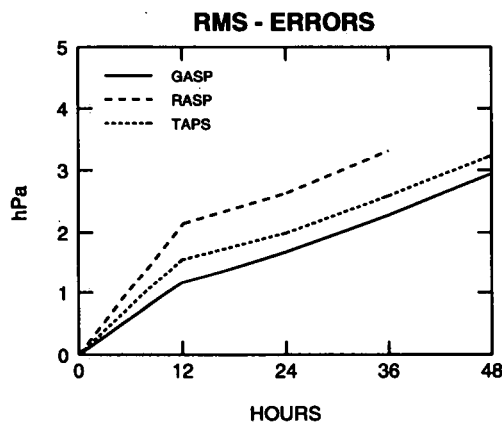


Fig. 1(b) Comparison for RASP/TAPS/GASP from April to June 1996. Root mean square errors at MSLP for combined base-times 0000 UTC/1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.



The differences between RASP and TAPS are slight, except for the rms error at MSLP (Fig. 1(b)). Here TAPS scored significantly better than RASP. TAPS quite often failed to accurately capture the movement of mid-latitude troughs and ridges, tending to underestimate amplifying westerly troughs and overestimating the depth of easterly troughs. TAPS, though, does accurately describe tropical lows in its analyses whereas GASP can have them poorly defined with high central values and weak gradients on both forecast and verify-

Fig. 1(c) Comparison for RASP/TAPS/GASP from April to June 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.

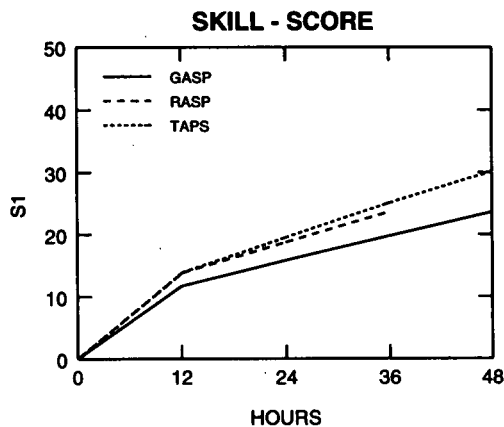
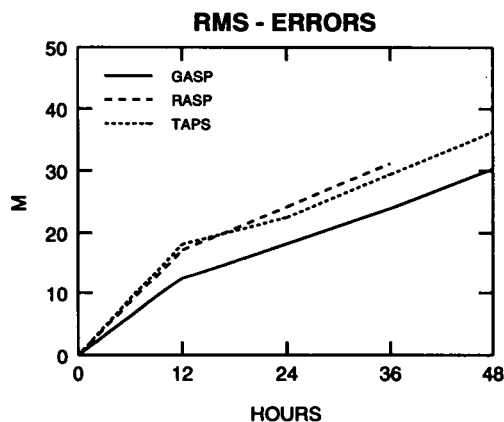


Fig. 1(d) Comparison for RASP/TAPS/GASP from April to June 1996. Root mean square errors at 500 hPa for combined base-times 0000 UTC/1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.



ing analysis. A small error in position for the TAPS forecast can result in a large statistical error and an unflattering comparison with the GASP score, for what is actually a more useful forecast.

#### Global models: (GASP, ECSP, UKGM, USAVM)

Figure 2 shows ECSP retained its usual superiority. Three-monthly MSLP skill scores for the other three models are almost indistinguishable (Fig. 2(a)). The

Fig. 2(a) Comparison for GASP/EC/US/UK from April to June 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.

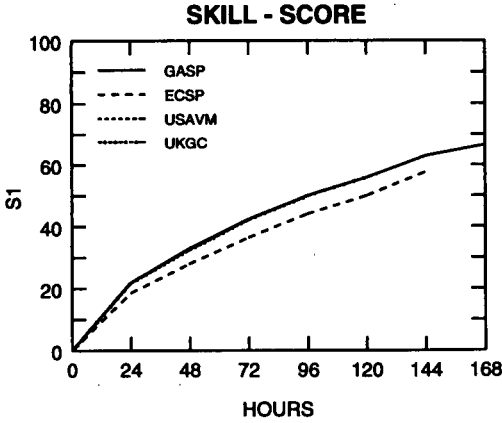


Fig. 2(b) Comparison for GASP/EC/US/UK from April to June 1996. Root mean square errors at MSLP for combined base-times 0000 UTC/1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

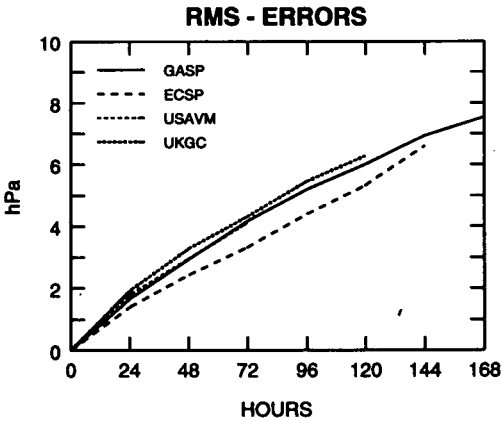


Fig. 2(c) Comparison for GASP/EC/US/UK from April to June 1996. S1 skill-scores at 500hPa for combined base-times 0000 UTC/1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

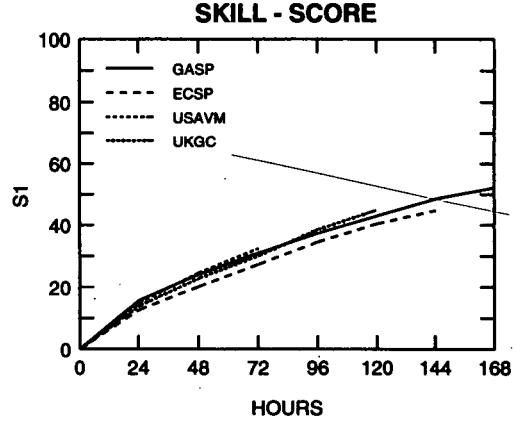
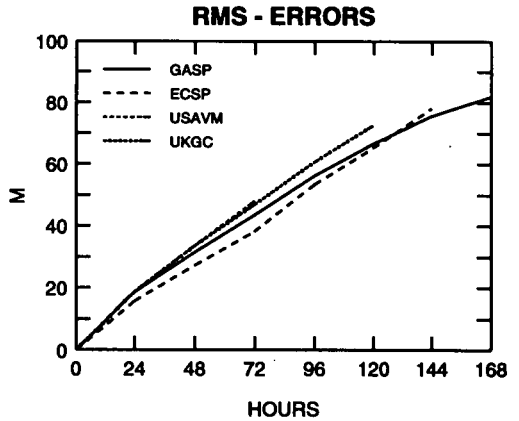


Fig. 2(d) Comparison for GASP/EC/US/UK from April to June 1996. Root mean square errors at 500 hPa for combined base-times 0000 UTC/1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.



monthly breakdown of skill scores (Fig. 3(c)) also captures the skill of ECSP but reveals that the similarity of the three-monthly scores for the other models masks considerable variation from month to month.

Three-monthly 500 hPa skill scores reveal that the UKGC model has an edge over the USAVM model with GASP worse than UK to +72 h and better at longer intervals (Fig. 2(c)). 500 hPa individual monthly results (Fig. 3(d)) again show ECSP being consistently good and the other models sharing second place.

Root mean square errors at MSLP (Fig. 2(b)) show

that GASP and USAVM are similar, with UKGC performing slightly worse. At 500 hPa (Fig. 2(d)), GASP has a slight edge over UKGC and USAVM.

**A comparison of global models at +72 h**

An increasing frequency of large errors during the period reflects the seasonal strengthening of the westerly flow and was evident in all models. All models had difficulty with the situation on 6 May when the remains of tropical cyclone *Jenna* developed into a cut-off low near 35°S 95°E and subsequently moved to the southeast.

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

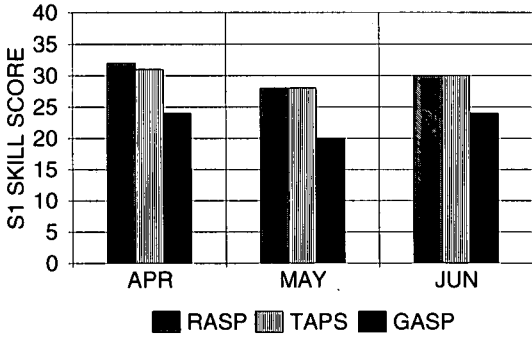


Fig. 3(c) Monthly S1 skill-scores at MSLP for GASP/EC/UK/US from April to June 1996 for base-time 1200UTC and interval +72 h over the irregular Australian verification grid.

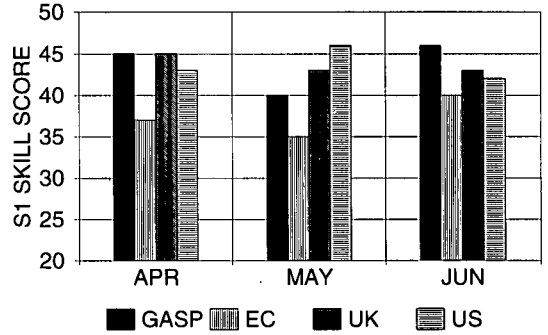


Fig. 3(b) Monthly S1 skill-scores at 500 hPa for RASP/TAPS/GASP from April to June 1996 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.

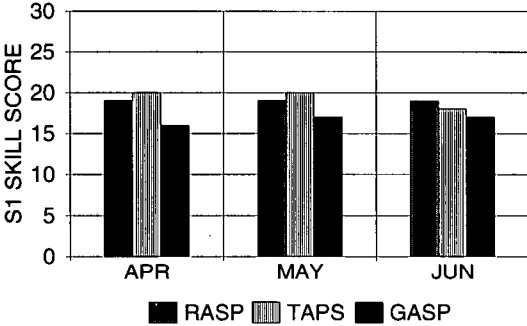
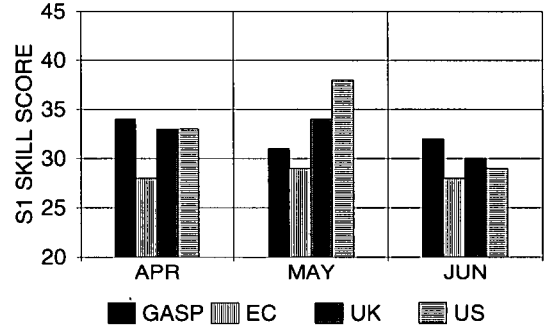


Fig. 3(d) Monthly S1 skill-scores at 500 hPa for GASP/EC/UK/US from April to June 1996 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.



They initially developed the low too far to the southeast in the higher latitude westerly flow and took some time to catch up with the situation.

The tendency to move cut-off systems too fast and to link them with the westerlies too early is still present, particularly for GASP and to a lesser extent UKGC. GASP also is more prone than the other models to miss the initiation or underpredict the intensity of such systems. Once the models have captured a system, the ECSP and USAVM generally perform best in retaining it, although the ECSP tends to retain a cut-off low too long.

The performance in tropical easterly flows was variable with examples of both overdevelopment and of

missed development. The USAVM predictions had a tendency to drag easterly dips too far westwards over eastern Australia. GASP's proclivity last year to overdevelop tropical systems appears to have been remedied by the change in convection parametrisation.

Consistent pressure bias in high pressure systems was not apparent. In this respect GASP seems to have improved, as previously it underestimated central pressures.

A feature of note was the excellent +72 h predictions from the ECSP system for tropical cyclone *Olivia* from 5 to 11 April. The +72 h predictions picked the initial development of a depression and provided reasonable guidance on the track of the cyclone, although the accuracy of the forecast speed of movement was variable.

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

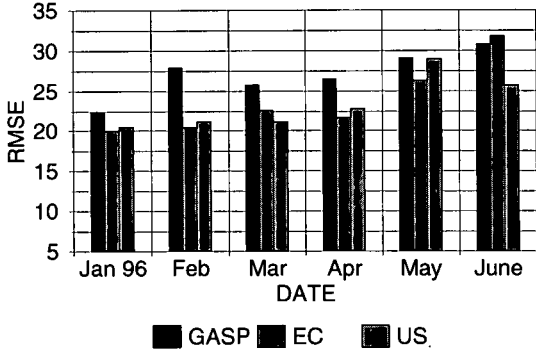


Fig. 4(c) Monthly root mean square errors against radiosondes at 500 hPa for GASP/EC from January to June 1996 for base-time 1200 UTC and interval +168 h over the southern hemisphere.

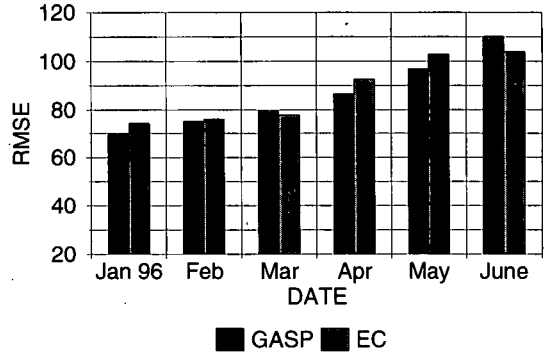
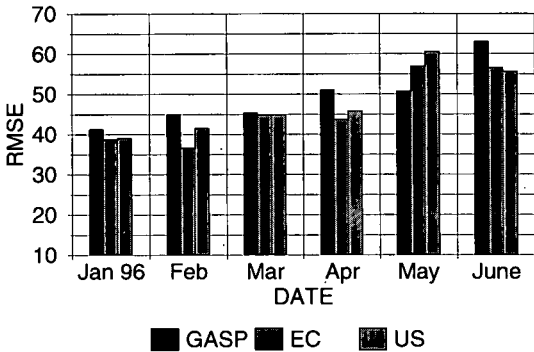


Fig. 4(b) Monthly root mean square errors against radiosondes at 500 hPa for GASP/EC/US from January to June 1996 for base-time 1200 UTC and interval +72 h over the southern hemisphere.



However, the predictions of the coastal crossing and subsequent southeastwards movement across Western Australia was very impressive. The performance of other models was poor in comparison.

Figure 4(a, b, c) show the performance of GASP, ECSP and USAVM verified against radiosondes for the southern hemisphere.

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. The increasing error from summer to winter is evident for all models at all periods. GASP was generally not as skilful as the other models although the errors vary considerably.

## References

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

