Aust. Met. Mag. 48 (1999) 295-301

Numerical prediction model performance summary April to June 1999

W. Skinner and D. Thomas
National Meteorological Operations Centre, Bureau of Meteorology, Australia

(Manuscript received October 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, USA) 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

ECSP

On 5 May, model changes were made such that raw TOV/ATOVS data from NOAA-14 and NOAA-15 are processed and assimilated (McNally et al. 1999). This has required a significant revision of the errors assigned to radiance observations in the analysis.

At the same time a re-coded sea and lakes temperature prescription software and a modified oceanic waves code was introduced. The improved forecasts skills are largest in the southern hemisphere and lower stratosphere (Lalaurette 1999).

GASP

On 17 June a system was implemented to force the GASP analysis system to accept bogus observations (PAOBS) when supported by at least one nearby

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

observation. Previously, these were likely to be rejected as gross errors in areas of high gradient such as in tropical cyclones. It is hoped that this will result in better representation of these systems at MSLP and help resolve situations such as tropical cyclone *Gwenda* and the Tasman Sea low (24 May) – both mentioned later in this article.

April to June 1999 intercomparisons

Local models: (LAPS, TLAPS, GASP)

GASP continued to score better than either TLAPS or LAPS at both MSLP and 500 hPa (Fig. 1). TLAPS had lower rms-errors than LAPS at MSLP and for +36/+48 hours at 500 hPa but LAPS had the better skill-scores. The breakdown into individual months (Figs 3 (a), (b)) showed that LAPS scored lower (i.e. better) or equal to TLAPS on every occasion. GASP consistently scored better than the other two models.

The longer term relationship between the models (Figs 3(e), 3(f)) confirmed these trends. LAPS at MSLP (Fig. 3(e)) has generally lower skill-scores than TLAPS and while the previous six months (Oct to Mar) reversed this, the tendency is now evident again. Also, all models showed lower scores than for the same three-month period last year.

At 500 hPa (Fig. 3(f)) there were higher than usual scores from November to April. In May and June however, scores decreased, and returned to more average values. LAPS continued to score lower than TLAPS at this level.

Synoptic overview for 24 h MSLP predictions

A later section, 'Synoptic intercomparison of global models', gives general descriptions of the synoptic situations which caused forecasting difficulty for global models at +72 h and it is not surprising that the same situations tended to also generate forecast errors at +24 h. This section looks at the behaviour of local models at the shorter interval.

Although only one tropical cyclone occurred in the Australian region during this period, *Gwenda* was classified as category 5 and had an effect on the model forecasts. None of the three models captured the central pressure, which dropped as low as 915 hPa on 6/7 April. GASP analysis showed a diffuse low pressure system with central pressure of 1004 hPa while LAPS and TLAPS analysed the central value at 997 hPa. However, allowing that the capture of detail of the system was poor, the 24 h forecasts did show the correct movement of the cyclone – first southwest, then southward and weakening. Such systems highlight one of the limitations of skill scores as a measure of forecast accuracy. A small error in position of an intense system can cause poor scores for a forecast otherwise providing quite good guidance, whereas a model with a low resolution, that fails to capture the storm at all, can score well.

On 5–6 April, the intensity of a trough amplifying into the Tasman Sea was underestimated by all three models. Also, the subsequent southeastward movement of the southernmost part of the same trough was too slow.

There were two significant Tasman Sea low pressure systems in late April. The first developed as a cut-off low on 22 April. The +24 h forecasts by LAPS and GASP were good while TLAPS moved it a little too fast. The next four days showed a slow eastwards movement which all models caught quite well, but the cradling high and upstream systems caused problems. By the 25 April, LAPS and GASP had largely underestimated the strength of a southerly trough near 140°E and both had major (18 hPa) errors in the southern ocean. The TLAPS grid did not cover this area, but errors similar to the other two models were evident on the northern edge of the trough.

The second cut-off developed on 27 April and was well forecast at +24 h by all the models.

On 5 May the easterly trough over Western Australia deepened and moved eastwards towards the head of the Bight. It then merged with a southern trough amplifying towards the northeast. Another deepening of the West Australian trough on 7 May occurred but this low moved southwest initially. These western lows were not well handled. TLAPS in particular had large errors in position and correspondingly high skill scores.

TLAPS also developed a low in an easterly trough on 9, 10 and 11 May more quickly than in fact occurred. GASP also showed this tendency.

By 12 May, all models scored badly with TLAPS the worst with 10 hPa errors on 12 May while LAPS and GASP had 8 and 7 hPa errors respectively in the Tasman. At the same time, the Western Australian trough was again a source of error. This deepened on 12 May and developed complexity. It subsequently moved southeastwards on May 16–17. Another trough downstream was amplifying into the Tasman Sea and all the models underestimated its strength.

A low deepening in the Coral Sea was fairly well forecast at +24 h to 24 May. By this time a low in the eastern Bight had also deepened and LAPS was underestimating the central values of both systems. The following day the Bight low weakened and the Tasman low moved southwards as forecast.

Fig. 1(a) Comparison for LAPS/TLAPS/GASP from April to June 1999. 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.

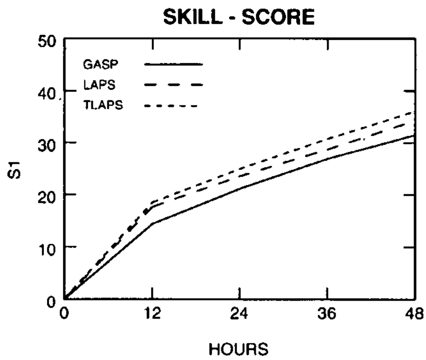


Fig. 1(c) Comparison for LAPS/TLAPS/GASP from April to June 1999. S1 Skill-scores of 500 hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.

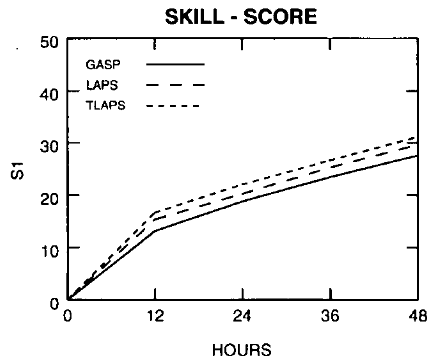


Fig. 1(b) Comparison for LAPS/TLAPS/GASP from April to June 1999. Root mean square errors of MSLP for combined base-times 0000 UTC / 1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.

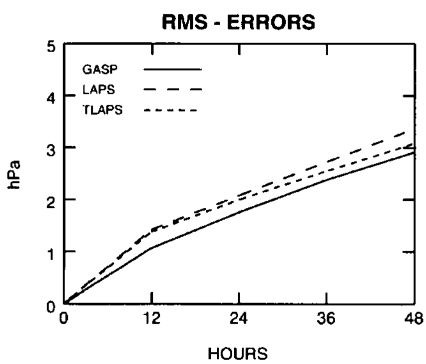
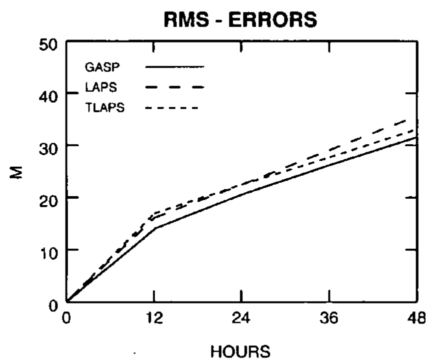


Fig. 1(d) Comparison for LAPS/TLAPS/GASP from April to June 1999. Root mean square errors of 500 hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.



Another low developed off Western Australia on 26 May and moved rapidly to the southeast. All models forecast this development and movement but details of pressure and position varied. The largest error was the LAPS forecast for 27 May when the pressure south of Western Australia was 12 hPa too high.

Another deepening low in the Bight on 4 June was moved too quickly to the southeast by LAPS and TLAPS.

The situation of 10 and 11 June mentioned in the +72 h intercomparison was fairly well handled at +24 h. However the pattern was both meridional and mobile and by 13 June all models were in error and tended to underestimate the speed of movement of the systems.

The three monthly period April to June provided 91 situations. GASP not only had the lowest average score, but was also the most reliable. It had the best (or equal best) score on 62 occasions and the worst on only 11 (62/11). The equivalent scores for TLAPS and LAPS were 19/51 and 23/44. GASP has the advantage over the other models of later data as its cut-off is several hours after cut-off for the limited area local models. GASP is not available in time for short-term forecast deadlines.

Fig. 2(a) Comparison for GASP/EC/US/UK/JMA from April to June 1999. S1 Skill-scores of MSLP for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +192 h over the irregular Australian verification grid.

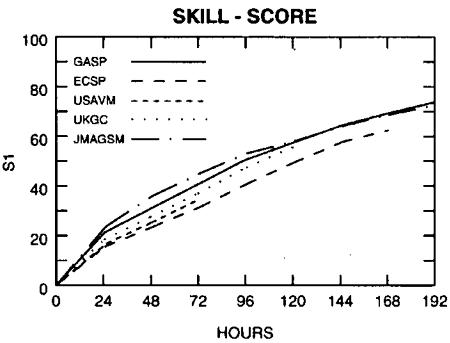


Fig. 2(c) Comparison for GASP/EC/US/UK/JMA from April to June 1999. S1 Skill-scores of 500 hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

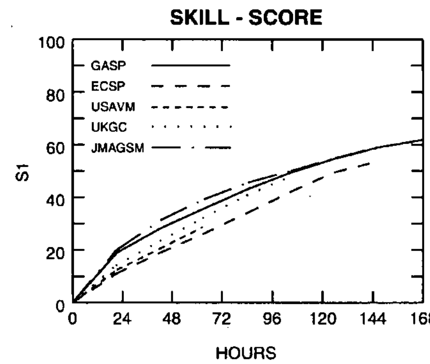


Fig. 2(b) Comparison for GASP/EC/US/UK/JMA from April to June 1999. Root mean square errors of MSLP for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +192 h over the irregular Australian verification grid.

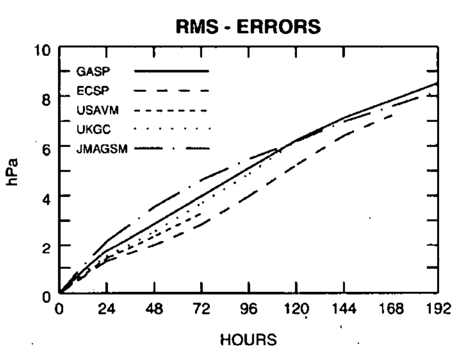
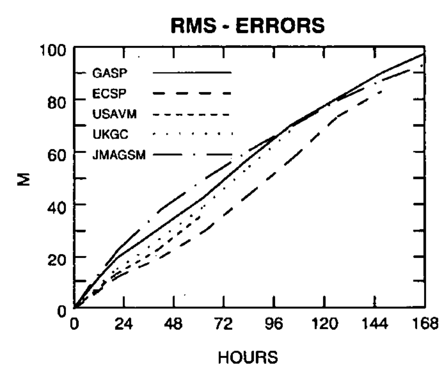


Fig. 2(d) Comparison for GASP/EC/US/UK/JMA from April to June 1999. Root mean square errors of 500hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.



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

The comparison of all models at all forecast intervals is shown in Fig. 2 and indicates that ECSP still substantially outscores the other models. This superiority is established at both MSLP and at 500 hPa and for both skill-score and rms-error. USAVM out to +72 h is next best followed by UKGC. Again, this relativity applies to both levels and both verification types. GASP performance tends to decrease less than the other models as the interval increases and its performance compares favourably with other models for 7 and 8 day prognoses. At +72 h however, (Figs 3(c), (d)) GASP was most successful in June when only ECSP and USAVM out-scored it. In both April and May, GASP was outscored by all models except JMA. The models display similar relative performance when using anomaly correlations (Fig. 4).

If the 60 per cent level is taken as the mark of usefulness, then ECSP and UKGC provide good forecasts out to five days, GASP to 4+ days and JMAGSM to four days. This is approximately one day less skilful than during the January to March quarter.

Fig. 3(a) **Monthly S1 Skill-scores of MSLP for LAPS/TLAPS/GASP from April to June 1999 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.**

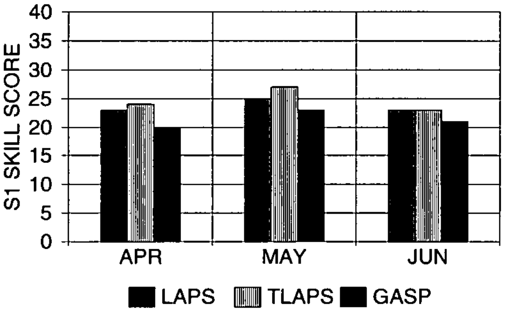


Fig. 3(b) **Monthly S1 Skill-scores of 500 hPa geopotential height (m) for LAPS/TLAPS/GASP from April to June 1999 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.**

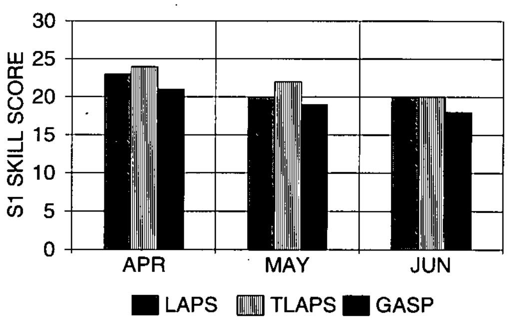


Fig. 3(c) **Monthly S1 Skill-scores of MSLP for GASP/EC/UK/US/JMA from April to June 1999 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.**

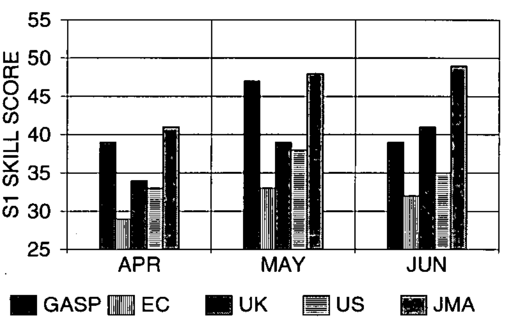


Fig. 3(d) **Monthly S1 Skill-scores of 500 hPa geopotential height (m) for GASP/EC/UK/US/JMA from April to June 1999 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.**

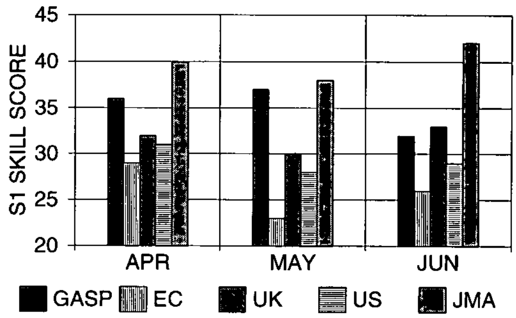


Fig. 3(e) **Monthly S1 Skill-scores of MSLP for LAPS/TLAPS/GASP from July 1996 to June 1999 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.**

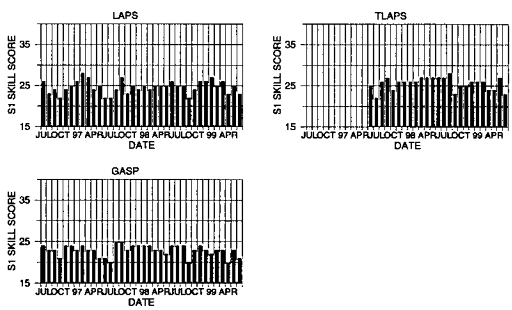


Fig. 3(f) **Monthly S1 Skill-scores of 500 hPa geopotential height for LAPS/TLAPS/GASP from July 1996 to June 1999 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.**

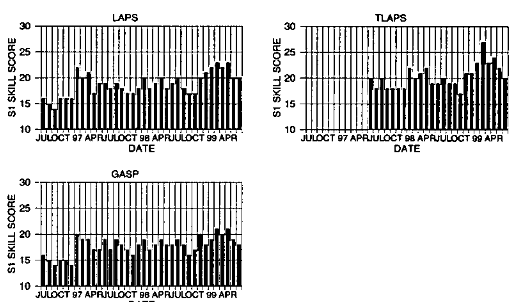


Fig. 3(g) **Monthly S1 Skill-scores of MSLP for GASP/EC/UK/US from July 1996 to June 1999 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.**

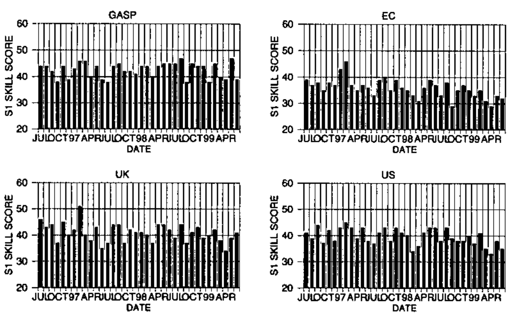


Fig. 3(h) **Monthly S1 Skill-scores of 500 hPa geopotential height for GASP/EC/UK/US from July 1996 to June 1999 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.**

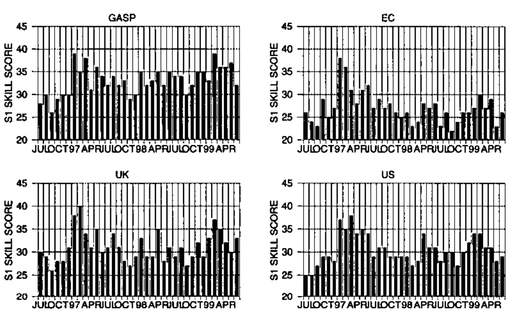
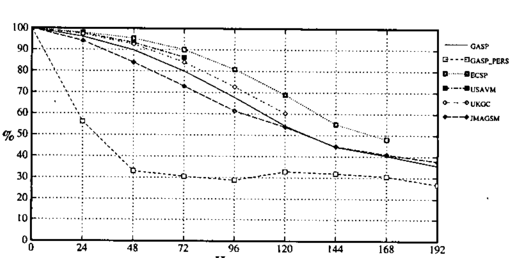


Fig. 4 **Anomaly Correlation of MSLP for GASP/EC/US/UK/JMA from April to June 1999 for combined base-times 0000 UTC / 1200 UTC over the irregular Australian verification grid.**



Synoptic intercomparison of global models

This section considers the 1200 UTC 72-hour Mean Sea Level pressure predictions from the global systems for the Australian Region, broadly defined as between 10° and 45°S and between 100° to 160°E.

April

During 5 to 7 April, a cold front moved across southeast Australia, formed a weak low in the Tasman Sea, then moved off quickly to the southeast. Models had trouble with the speed of the front and over-forecast the low's intensity with errors on 7 April of the order of 15 hPa in the Tasman Sea. By the third day, the EC 72 hour prognosis predicted the situation well, but other models still had trouble.

A rapid development of a low on a cold front moving into the Tasman Sea occurred during the 26 to 28 April. The UK 72 hour prognosis for 26 April had predicted the location and intensity of the front well compared to the other models which had pressures between 5 and 10 hPa too high east of Bass Strait. The 72 hour forecasts for 27 April had pressures in the western Tasman Sea about 10 to15 hPa too high as the low had stayed further west than forecast. The EC and UK prognoses for 28 April had caught on to the slow movement and errors were reduced to less than 4 hPa. The JMA, GASP and US prognoses still had errors of around 5 to 10 hPa.

During 4 and 5 April, a low moved west across the Timor Sea and developed into tropical cyclone *Gwenda* to the northwest of Broome on 6 April. The cyclone then turned south and crossed the coast near Port Hedland on 8 April. Tropical cyclone *Gwenda* was handled quite differently by the various models. In the initial stages, the low was under-forecast by the EC, GASP and US models but well forecast by the JMA. On the second day, the low was better represented by the models, although GASP failed to predict any significant low where the cyclone should have been, and the subsequent GASP analyses had no significant circulation where the cyclone was. All other models had some representation of the cyclone in their analyses, although of varying intensities. On the final day of the cyclone, the EC, US and JMA models 72 hour forecast had predicted the southerly movement but were still too far west and too slow. The UK model had missed the recurvature, predicting the system would move slowly west southwest.

May

A low developed in an easterly trough over the Coral Sea during the period 9 to 12 May. Model 72 hour forecasts had been predicting this low for the day before it formed, and GASP had the low forming two days too early. Once the low did form on 11 May, it deepened and moved south more quickly than models had suggested so that pressures east of Gabo Island on 12 May were between 12 and 22 hPa too high on the 72 hour prognoses. Interestingly, the period 12 to 14 May, also contained a complex interaction between a low in an easterly dip off southwest Western Australia and a frontal system. The complexity of the situation on 13 May led to very large positive errors (~20/30 hPa) over the Southern Ocean in all models and negative errors (~5/10 hPa) to the southwest and southeast of the continent on most models. 72 hour forecasts for 14 May were significantly better than the previous day, however, errors of the order of 10 hPa were still evident on all models apart from the EC.

The period 19 to 25 May saw a low move southwest from near the Solomon Islands and into the Tasman Sea, passing over Lord Howe Island. The GASP, JMA and UK models 72 hour forecasts for 20 May over-forecast the intensification and southwest movement when compared to their verifying analyses. This also happened for the +72 hour forecasts for 21 May. However, the variation between the verifying analyses of each model was becoming significant on 21 May with 5 hPa separating GASP and the EC. This trend continued on 22 May with the EC and UK showing the Coral Sea low about 5 hPa deeper than the GASP, JMA or US analyses. By 24 May, the models were up to 8 hPa different in their analysis of the low despite the presence of Lord Howe Island observations confirming a more intense system.

During the event of the Coral Sea/Tasman Sea low, an intense cut-off low moved southeast out of the Bight but weakened as it moved east of 150°E on 25 May. This aspect was well forecast by the models. In contrast, a low developing west of southwest Western Australia, was being treated quite differently by each model, producing positive and negative errors of the order of 15 to 30 hPa over waters south of the continent during the period 25 to 27 May. By 28 May the cut-off had moved across the Bight as suggested by all models, except the JMA which had kept the low near Perth. The US and EC models handled the movement of the cut-off best, with errors of the order of 5 to 10 hPa over the Bight and Tasmania compared to the 15 to 20 hPa errors evident with the UK, GASP and JMA models.

June

June was a quieter month in that the models seemed to capture the synoptic evolutions quite well. Still, there was a period between 10 to 17 June associated with a series of fronts which caused significant errors in the 72 hour forecast pressure fields. One front was associated with the development of an east coast low. Other fronts extended into central Australia.

The first problem period was 10 to 11 June, when a cold front moved across southeastern Australia before developing into a low in the Tasman Sea. All models under-forecast the intensity of the Tasman Sea low, which, combined with position errors, produced pressure errors of about 10 hPa. Similarly, all models except the JMA, under-estimated the intensity of the ridging across southeastern Australia and Tasmania immediately behind the front with pressures being 5 to 10 hPa too low west of Tasmania.

Following this system, the pattern became meridional over Australia and a large cold front moved across central Australia during 11 to 13 June. The strong dynamics of the situation led to some large errors in the 72 hour pressure forecasts on 12 June, with forecasts which were 10 to 15 hPa too low east of Tasmania, about 15 hPa too high west of Tasmania and up to 17 hPa too low in the Great Australian Bight. The EC, US and GASP had the situation well captured by the +72 hour forecasts valid for 13 June with the EC showing pressures over Australia within 1 or 2 hPa of the analysis. The GASP too had fairly small errors except over the southeast near Gabo Island where the forecast pressure was about 10 hPa too high. The JMA and UK had larger errors (~4 hPa) over Australia.

As the front moved into the Tasman Sea on 14 June, the JMA and UK models tried to move or weaken the system too quickly and develop ridging around its southwest flank. This resulted in the +72 hour forecast of pressures for the 14 June over Tasmania being about 20 hPa too high on the JMA and 16 hPa too high on the UK model. GASP in contrast, held the system too far to the west which resulted in negative pressure errors of about 10 hPa over Tasmania.

References

Lalaurette, F. 1999. Changes to the operational forecasting system. *ECMWF Newsletter Number 83* (Spring 1999).

McNally, A.P., Andersson, E., Kelly, G. and Saunders, R.W. 1999. The use of raw TOVS/ATOVS radiances in the ECMWF 4D-Var assimilation system. *ECMWF Newsletter Number 83* (Spring 1999).

Skinner, W. 1995. Numerical prediction model performance summary April to June 1995. *Aust. Met. Mag., 44,* 309–12.