Aust. Met. Mag. 48 (1999) 223-230

Numerical prediction model performance summary January to March 1999

W. Skinner and T. Hart
National Meteorological Operations Centre, Bureau of Meteorology, Australia

(Manuscript received June 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.

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

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

ECMWF introduced a new version of the forecast model on 9 March 1999, increasing the number of vertical levels from 31 to 50. The horizontal resolution is unchanged at TL319. The upper level has been raised from 10 to 0.1 hPa with most of the extra resolution in the stratosphere between 150 and 0.1 hPa. Below 150 hPa the distribution of the 50-level version is the same as the 31-level system. The enhanced vertical resolution and raised top level provide significantly better stratospheric analyses and predictions. The change has no significant impact on the characteristics of the ECMWF tropospheric products.

The extra levels will assist in the assimilation of radiances from the TOVS and ATOVS instruments where a forward calculation of the model estimate of the radiance is required. In the 31-level model extrapolation above 10 hPa was required to compute the contribution from the upper stratosphere to the estimated radiances.

ECMWF plan to increase the number of levels to 60 later in the year. In this case the resolution will be enhanced in and just above the planetary boundary layer. It is planned to use the 60-level system in the ECMWF reanalysis project ERA-40, covering 1957 to the present (Untch et al. 1999).

TOVS soundings of temperature and moisture

Acronyms:

TIROS	Television Infrared Observational Satellite
TOVS	TIROS Operational Vertical Sounder
ATOVS	Advanced TOVS
NOAA	National Oceanic and Atmospheric Administration (USA)
MSU	Microwave Sounding Unit
NESDIS	National Environmental Satellite Data and Information Service (USA)

A major problem occurred on 26 February with the MSU instrument on NOAA-11. Availability of TOVS from NOAA-11 ceased from that date, both from NOAA/NESDIS and the local readout. Soundings from NOAA-14 continued apart from a brief scare in early March, but it meant that the coverage of TOVS soundings was halved. Soundings produced from the TOVS instrument on NOAA-15 were successfully introduced into operations on 5 May, but only at a spatial resolution of 500 km. Transmission of the code containing the 120 km resolution soundings has still not resumed. Local processing of NOAA-15 ATOVS is under development.

Given the importance of satellite soundings for prediction in the southern hemisphere, some degradation of NWP performance could be expected from the reduction in the coverage of soundings. Centres which analyse radiances directly could have processed NOAA-15 data during this period. NCEP in the US for instance introduced NOAA-15 radiances on 8 March. Some other centres such as ECMWF and UKMO use NOAA-15 data experimentally but it is unclear whether they were used operationally.

January to March 1999 intercomparisons

Local models: (LAPS, TLAPS, GASP)

GASP continues to score better than either TLAPS or LAPS at both MSLP and 500 hPa (Fig. 1). TLAPS has lower rms-errors than LAPS and lower skill-score at MSLP, but at 500 hPa, LAPS has the better skill-score. This continues a trend established during the October to December quarter when the relative ranking of skill between TLAPS and LAPS at MSLP was reversed and TLAPS outperformed LAPS. The breakdown into individual months (Figs 3 (a), (b)) shows that TLAPS scored lower (i.e. better) than LAPS on two of the three months at MSLP but not on any month at 500 hPa. GASP consistently scored better than the other two models.

The longer term relationship between the models (Fig. 3(e)) confirms these trends. LAPS at MSLP has higher skill-scores than for the same three-month period last year, while TLAPS has scored lower than before. At 500 hPa (Fig. 3(f)) both models had higher than usual scores since November, however LAPS continues to score lower than TLAPS at this level. There was a problem with the soil-moisture fields in LAPS which could account for the decrease in skill at MSLP relative to TLAPS. The overall tendency for higher scores seems to affect all models as persistence skill-scores over the Australian region have tended to be high for about 18 months.

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.

The easterly trough deepening over Western Australia on 4 to 5 January generated some of the highest skill-scores for the month. LAPS and GASP over-developed the trough and TLAPS moved the following anticyclone into the Bight too quickly. The next day, TLAPS also over-developed the leading edge of this easterly trough. Later, on January 30-31, another Western Australian trough deepened more than forecast, and generated moderate errors, particularly by LAPS.

The low development near Hobart on 6 February was easier to forecast at +24 h than it was at +72 h. All three models placed a trough in the area, but all tended to move it too rapidly and underestimated the intensity so that even at the shorter interval, the situation was not very well handled.

The tropical depression in the Coral Sea which moved southwards on 25 February was fairly well forecast at +24 h although owing to its intensity, a small displacement in position caused 7 hPa errors for TLAPS on 24 Feb. All three models moved this low too quickly into the westerlies.

The development of a cut-off over Tasmania on 12 March was underestimated by all the models. The weakening of the cradling high was also a source of error on 14 March.

Fig. 1(a) **Comparison for LAPS/TLAPS/GASP from January to March 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.**

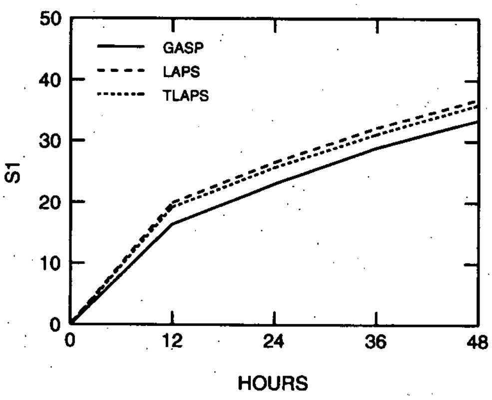


Fig. 1(b) **Comparison for LAPS/TLAPS/GASP from January to March 1999. Root mean square errors of MSLP for combined base-times 0000 UTC / 1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.**

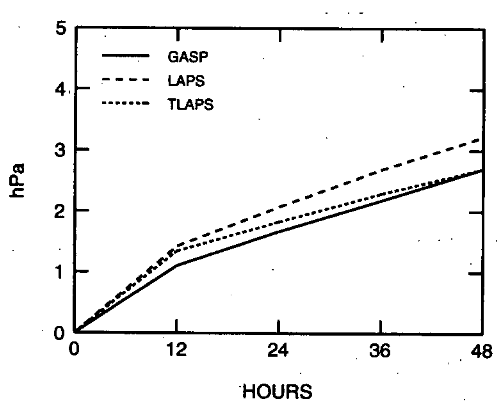


Fig. 1(c) **Comparison for LAPS/TLAPS/GASP from January to March 1999. S1 skill-scores of 500hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals 12, +24, +36, +48 h over the irregular Australian verification grid.**

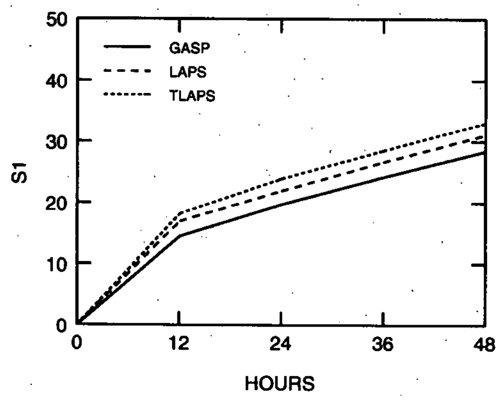
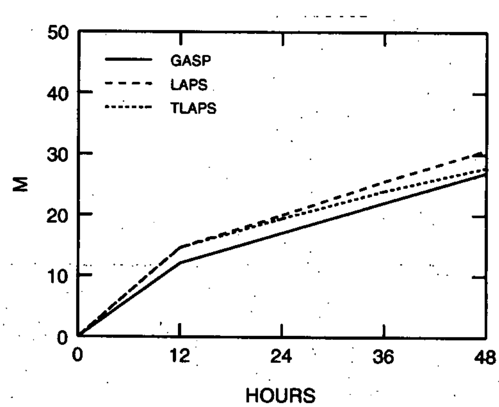


Fig. 1(d) **Comparison for LAPS/TLAPS/GASP from January to March 1999. Root mean square errors of 500 hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +12, +24, +36, +48 h over the irregular Australian verification grid.**



Tropical cyclones *Elaine* and *Vance* caused some forecast errors, as small intense systems sometimes do, when small errors in position can result in large errors in pressure at a particular point. By 23 March, when the remnants of the cyclone had curved to the southeast and were at the head of the Bight, all models had moved the system too slowly and GASP and LAPS had weakened it too quickly.

Yet another Tasman Sea cut-off low developed on 29 March much more deeply than any model prediction and caused some of the worst errors for the month.

Over the 90 situations between January and March, GASP not only had the lowest average score, but was also the most reliable. It had the best (or equal best) score on 57 occasions and the worst on only 13 (57/13). The equivalent scores for TLAPS and LAPS were 27/39 and 16/42. GASP has the advantage over the other models of later data as its cut-off is several hours after the limited area local models. It is not available in time for short-term forecast deadlines.

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

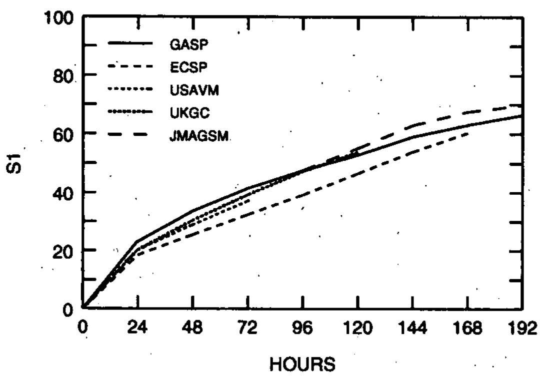


Fig. 2(b) Comparison for GASP/EC/US/UK/JMA from January to March 1999. Root mean square errors of MSLP for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +192 h over the irregular Australian verification grid.

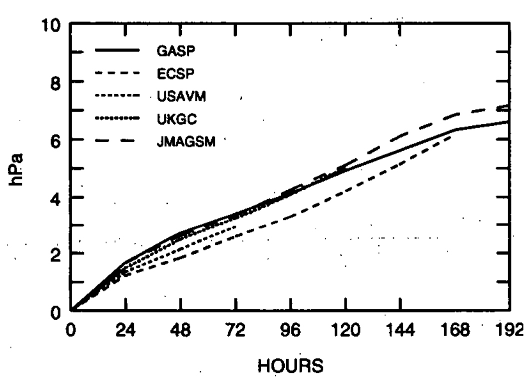


Fig. 2(c) Comparison for GASP/EC/US/UK/JMA from January to March 1999. S1 skill-scores of 500 hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.

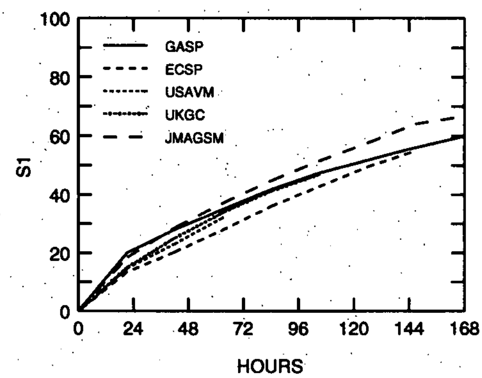
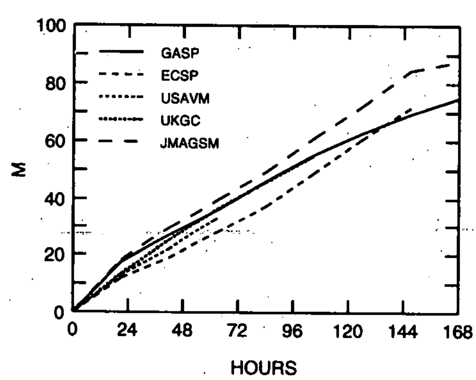


Fig. 2(d) Comparison for GASP/EC/US/UK/JMA from January to March 1999. Root mean square errors of 500 hPa geopotential height (m) for combined base-times 0000 UTC / 1200 UTC and intervals +24 h to +168 h over the irregular Australian verification grid.



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 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 seven and eight-day prognoses. At +72 h, however, (Figs 3(c) (d)) GASP was most successful in January when only ECSP and USAVM out-scored it at MSLP and only ECSP at 500

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

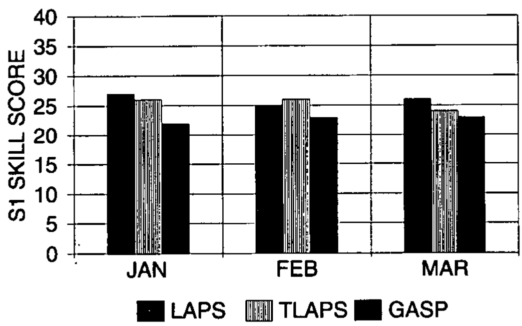


Fig. 3(b) **Monthly S1 skill-scores of 500 hPa geopotential height (m) for LAPS/TLAPS/GASP from January to March 1999 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.**

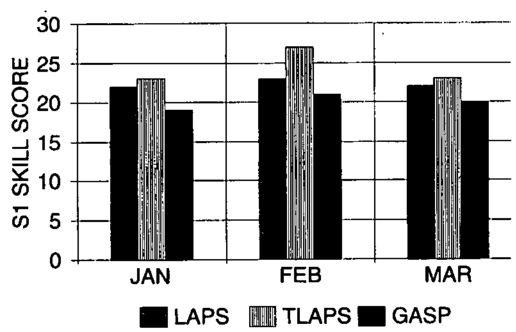


Fig. 3(c) **Monthly S1 skill-scores of MSLP for GASP/EC/UK/US/JMA from January to March 1999 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.**

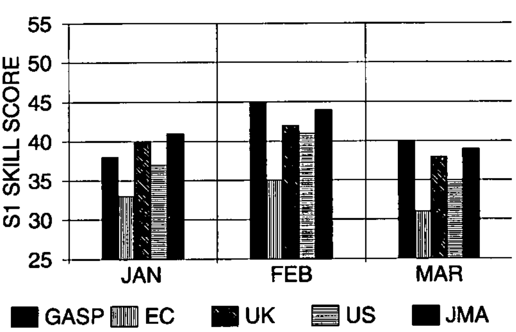


Fig. 3(d) **Monthly S1 skill-scores of 500 hPa geopotential height (m) for GASP/EC/UK/US/JMA from January to March 1999 for base-time 1200 UTC and interval +72 h over the irregular Australian verification grid.**

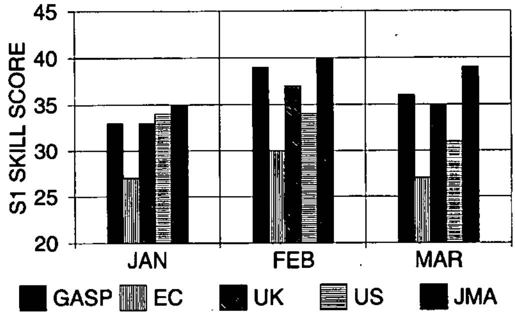


Fig. 3(e) **Monthly S1 skill-scores of MSLP for LAPS/TLAPS/GASP from July 1996 to March 1999 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.**

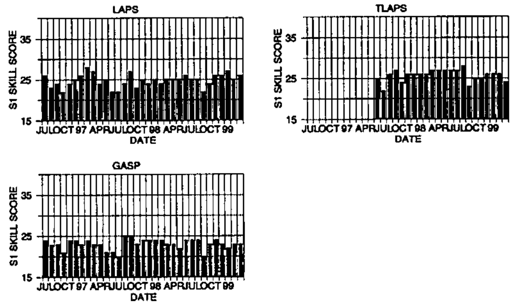


Fig. 3(f) **Monthly S1 skill-scores of 500 hPa geopotential height for LAPS/TLAPS/GASP from July 1996 to March 1999 for base-time 1200 UTC and interval +24 h over the irregular Australian verification grid.**

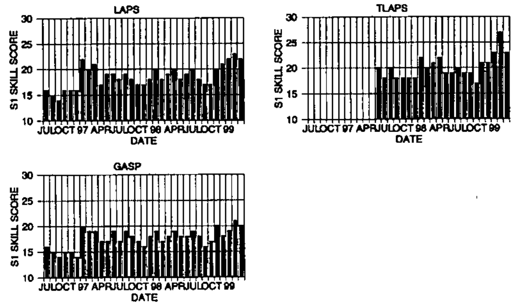
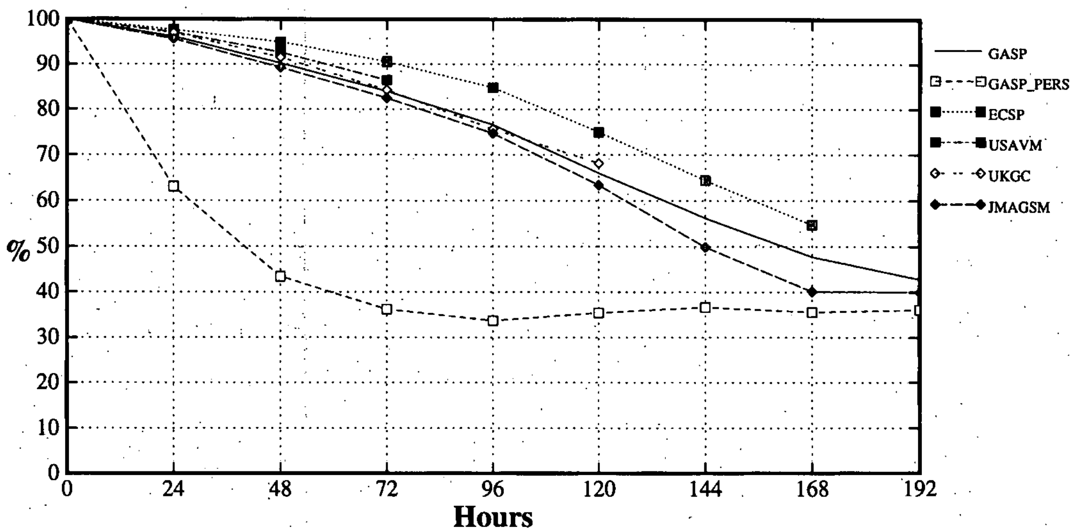


Fig. 4 **Anomaly Correlation of MSLP for GASP/EC/US/UK/JMA from January to March 1999 for combined base-times 0000 UTC / 1200 UTC over the irregular Australian verification grid.**



hPa. The models display similar relative performance when using anomaly correlation (Fig. 4).

If the 60 per cent level is taken as the mark of usefulness, then ECSP provides useful forecasts out to 6+ days, GASP and JMAGSM to 5+ days.

Synoptic intercomparison of global models

This section considers the 1200 UTC 72-hour MSLP predictions from the global systems for the Australian region, broadly defined as between 10° and 40°S from 100° to 140°E and between 10° and 45°S from 140° to 160°E. The focus on MSLP biases the comments towards higher latitudes, and tropical systems are not considered in detail.

January - a month of easterly flow

During January a high pressure ridge was a persistent feature south of the continent, disturbed briefly on only a few occasions by frontal systems. The flow over the mainland was mainly easterly, with transient troughs. Pressure gradients were generally slack and large pressure errors were less frequent than in previous months. The US model, in particular, handled this type of situation well. JMA predictions, especially in the first week, had difficulty in defining the relatively weak features. The occasions which gave the most problem were the cases when troughs in the easterlies developed more than normal such as occurred over Western Australia on 4 and 5 January. The developments of the tropical depressions over the Coral Sea which became tropical cyclones *Pete* and *Olinda* (21-25 January) were not well handled in the 72-hour predictions.

February - two major errors in common

In February there were two cases in which the model predictions had large and surprisingly similar errors. The first was a common failure to predict the rapid development of a mid-latitude low, reminiscent of the predictions at 72 hours in the case of the Sydney-Hobart yacht race on 27 December 1998 discussed in the last summary. On 5 February a deep trough was moving over waters south of the Bight with a weak low to the northeast, over the Tasman Sea (similar to the Sydney-Hobart case). A deep low developed during 6 February from a cut-off at the apex of the trough. All the predictions available showed a mobile trough

Fig. 5(a) ECMWF verifying analysis for 1200 UTC 7 February 1999 at MSLP.

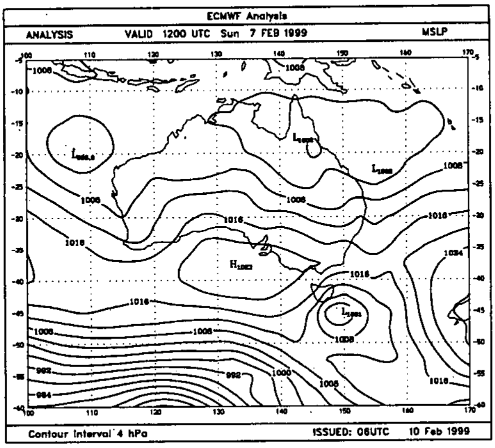


Fig. 5(b) ECMWF +72 h forecast for 1200 UTC 7 February 1999 at MSLP.

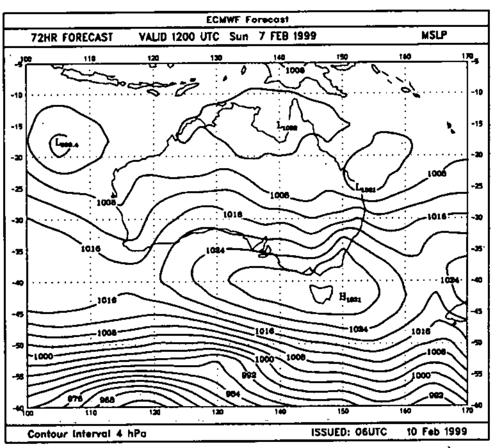
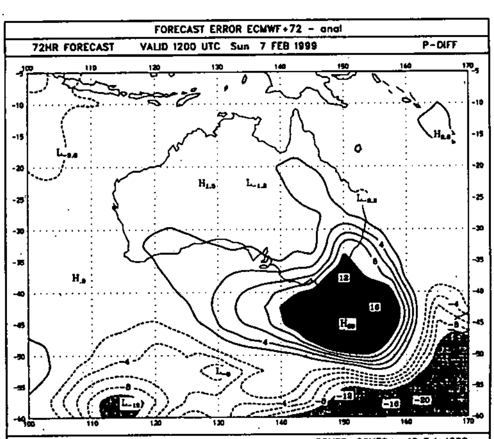


Fig. 5(c) ECMWF +72 forecast/analysis difference for 1200 UTC 7 February 1999 at MSLP.



with strong ridging over Tasmania. In fact, there was rapid development of a low of about 993 hPa centred just south of Tasmania, bringing damaging winds to southern and eastern Tasmania. The low remained near Hobart during 7 February when all the predictions showed a high over the area. Accordingly, the MSLP forecast errors were large, exceeding 20 hPa in all five predictions and with a very similar pattern. Verifying analyses for 1200 UTC 7 February, 72-hour predictions and error patterns are shown for ECMWF (Figs 5 (a), (b), (c)), and GASP (Figs 6(a), (b), (c)).

The 500 hPa long wave pattern for both cases showed a northeastwards tilting trough at eastern Australian longitudes, with a strong ridge over New Zealand. In each case the GASP predictions of the long wave pattern failed to capture the strength of the downstream ridge and the extent of tilt in the trough towards the northern side of the ridge. This event and the case of the Sydney-Hobart race raise the question: what is it about these systems that made them so unpredictable even at the 72-hour range?

A second case of large and similar errors occurred in the 72-hour predictions for 25 February. In this situation a tropical depression which had originated over the Coral Sea had moved southwards over the Tasman Sea. All the models predicted movement to the south-southwest with a strong centre close to the east coast of New South Wales. The UKMO predictions showed a very intense centre of 983 hPa. In fact, the low continued to drift south-southeastwards and weaken. The model behaviour was unusual for cut-off systems such as this, which the models are inclined to move too rapidly eastwards. It was surprising that each model showed a similar, and seemingly uncharacteristic, error pattern.

March - cut-off lows and tropical cyclones cause problems

Cut-off lows over southeastern Australia and the Tasman Sea continued to cause prediction problems on 4 and 12 March, but the predictions for a weak cut-off low over Tasmania on 13 March were all good. Major features of the later period of March were tropical cyclones *Elaine* (16-20 March) and *Vance* (18-24 March) which developed over the North-west Shelf, tracked southeastwards initially and then recurved to the south.

Elaine weakened before crossing the coast but as a rain depression brought heavy rainfall and flood damage to parts of Western Australia. The 72-hour predictions from GASP and ECMWF failed to capture the cyclone to any degree, and their analyses were also poor.

Fig. 6(a) GASP verifying analysis for 1200 UTC 7 February 1999 at MSLP.

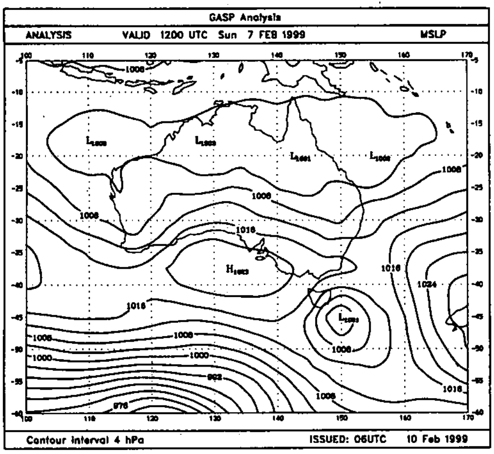


Fig. 6(b) GASP +72 h forecast for 1200 UTC 7 February 1999 at MSLP.

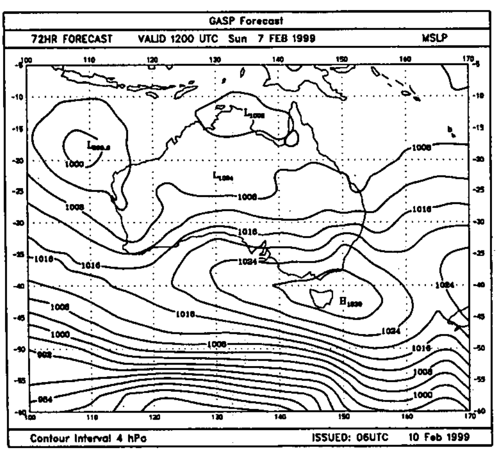
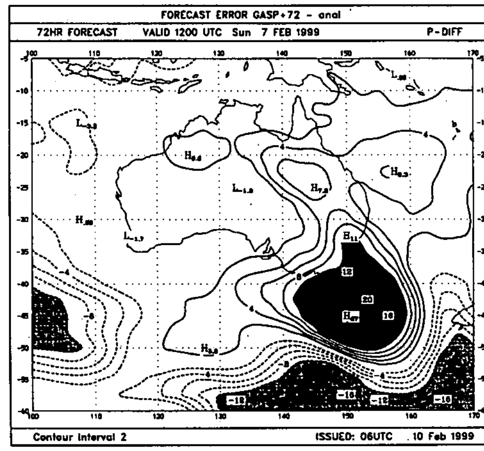


Fig. 6(c) GASP +72 h forecast/analysis difference for 1200 UTC 7 February 1999 at MSLP.



Vance crossed the coast near Exmouth inflicting severe damage and registering the highest wind gust (267 km/h) registered on the Australian mainland. It then weakened but traversed Western Australia and could be traced as it moved across the Bight, ultimately affecting areas of southeastern Australia before merging into the background flow. The 72-hour predictions from GASP, and ECMWF in the early stages were poor. The best predictions throughout came from UKMO, particularly for 22 March, the day of the coastal crossing. The JMA prediction for 21 March was also excellent.

References

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

Untch, A., Simmons, A. et al 1999. Increased stratospheric resolution in the ECMWF forecasting system. *ECMWF Newsletter Number 82* (Winter 1998/99)