Predicting Regional Forecast Skill Using Single and Ensemble Forecast Techniques

LANCE M. LESLIE AND GREG J. HOLLAND

Bureau of Meteorology Research Centre, Melbourne, Australia

(Manuscript received 4 May 1990, in final form 25 August 1990)

ABSTRACT

The potential for predicting the skill of 36-h forecasts from the Australian region limited area model is investigated using three predictors of model forecast error (MFE) for mean sea level pressure. Two of the predictors utilize single forecasts: one is based on statistical regression of the MFE against the initial analysis and the forecast; and the other uses a measure of the degree of persistence in the forecast. The third predictor utilizes the divergence, or spread, of an ensemble of forecasts from other NWP centers.

Based on a 5-month period of daily 36-h forecasts, correlations were found between the above predictors and the MFE of 0.58, 0.18, and 0.40, respectively. Combining the three predictors in an optimal linear manner increased the correlation to 0.71. Further testing of the combined predictors on a 2-month independent dataset produced a correlation of 0.67. Thus, application of the technique to both dependent and independent datasets explained approximately 50% of the variance in the MFE. This demonstrates that the technique has operational utility for differentiating overall poor and good model forecasts. Using case studies concentrating on southeastern Australia, it is further demonstrated that the predictors can provide excellent differentiation of forecast skill across the forecast domain.

1. Introduction

Numerical weather prediction (NWP) has traditionally focused on reducing model forecast error (MFE). Over the twenty year history of NWP development, great success has been achieved in reducing MFE, with this success coming from a number of directions, including: developments in model numerics, dynamics, and representation of physical processes; increased resolution and emergence of global models made possible by the rapid advances in computing power; enhanced databases arising both from remote sensing and communication advances; and improved methods of analysis and assimilation.

At the same time it has been recognized that there are inherent limits to the predictability of these deterministic models. The predictability limits are imposed by the chaotic nature of the atmospheric system in the sense that future weather states are sensitive to small errors in the initial state. The predictability limit of around two weeks for explicit global weather forecasts suggested by the early work of Lorenz (1969) is still the generally accepted value, although recent work (e.g., Murphy 1990) indicates potential skill for probabilistic forecasts out to a month or so. It also is known that regional forecast skill is sensitive to a range of factors, such as latitude, season, flow regime, scale, hemisphere, relative location of other weather systems, and the influence of lateral boundaries (Anthes 1983).

Until the 1980s forecasts in most centers were restricted to 48 hours, or less, which is well within the predictability limit. Further, the initial state was only poorly known and considerable gains were obtainable by analysis and model refinements. Statistical approaches, such as model output statistics (MOS) (Glahn and Lowry 1972) and statistical–dynamical techniques (Neumann and Pelissier 1981), also were being developed to provide improved forecasts of specific parameters and systems. Thus, despite the acknowledged presence of predictability limits for more than 20 years, only a modest amount of related research has been carried out until the last few years.

Forecasts are now issued to ten days or more and are nearing the accepted predictability limit. Forecast breakdown now is a major issue in medium and extended range model prediction. Even on shorter time scales, there are large day-to-day variations in predictive skill. This variation is illustrated by the time series of 36-h root-mean-square (rms) errors in Fig. 1 for the Australian region data assimilation and prediction system (RASP; Leslie et al. 1985). Most forecast offices also receive model predictions from a variety of NWP centers, and it is recognized that at times there are disparities between forecasts from these different models, while at other times they are in close agreement.

Forecasters therefore are forced to choose between conflicting guidance, with little quantitative information on the actual predictability of the current situation.

Corresponding author address: Dr. Lance M. Leslie, BMRC, GPO Box 1289K, Melbourne Victoria 3001, Australia.

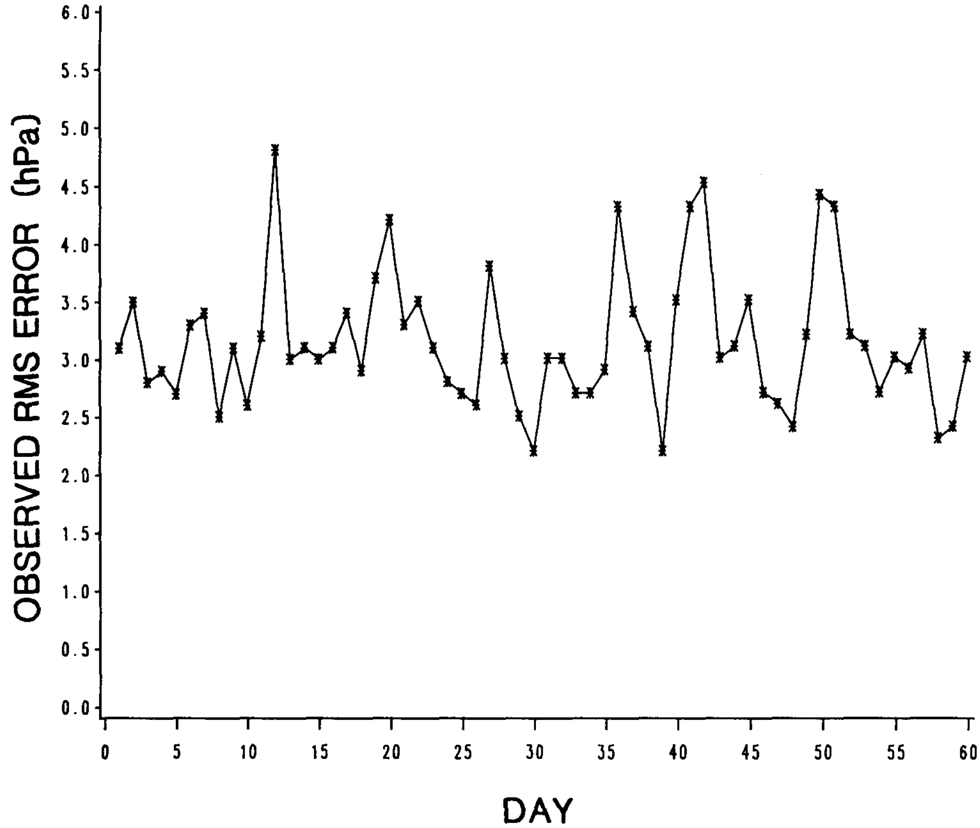

FIG. 1. Time-series of observed rms error (hPa) of the 36-h forecasts of mean sea level pressure over the Australian domain for the two month independent forecast set, December 1989–January 1990.

The motivation for this study has been the recognition that a priori estimates of forecast skill are a necessary adjunct to direct predictions. Other research has been similarly motivated. Techniques based on ensemble forecast approaches have been developed by Leith (1974), Hoffman and Kalnay (1983), and Kalnay and Dalcher (1987). An extensive summary of this work may be found in Murphy (1990). Others have used single forecast model approaches (Leslie et al. 1989; Chen 1989). Overall, the results of these studies were encouraging and indicated that further research into methods of forecasting forecast skill would be beneficial.

In this study, 36-h forecasts from RASP are concentrated on, and both single and multiple forecast model approaches are used to derive techniques for predicting the skill of these model forecasts. This is done by correlating the MFE with three predictors. Two of these predictors are derived from the Australian region model forecasts alone and consist of: the statistical regression scheme developed by Bennett and Leslie (1981) and improved by Glowacki (1988) and the persistence predictor developed by Chen (1989). The third predictor arises from the divergence or "spread" of forecasts from other centers.

The data used and the methodology are described in the following sections. Then the derived relationships between the predictors and the MFE are presented and discussed, both individually and in an optimal combination similar to that proposed by Kalnay and Ham (1989). Finally, an independent dataset is used to verify the performance of the predictors of forecast skill over the entire Australian region NWP domain using a pair of case-study examples and over a subdomain centered on southeastern Australia.

2. Data

The data used in this study are the archived analyses and forecasts from the Australian National Meteorological Centre (Australian NMC) and the archives of the Australian region subgrid of the European Centre for Medium-Range Weather Forecasts (ECMWF), United Kingdom Meteorological Office (UK), and United States National Meteorological Center (US) models that are received at NMC on a daily basis. Although a number of levels are available in the archive, this study has been confined to forecasts of mean sea level pressure (MSLP). The developmental dataset extended over five months from July to November 1989, inclusive. December 1989 and January 1990 were reserved for independent verification of the findings.

3. Methodology

The MFE is defined as the root areal mean-square difference, E, between the model forecast and its verifying analysis:

$$E^2 = \overline{(F^t - A^t)^2} \quad (1)$$

where the overbar denotes an areal average over the Australian NWP model domain (Leslie et al. 1985), and F^t and A^t are the forecast and the verifying analysis fields valid at time t, respectively.

The basic approach here is to correlate the MFE with the range of predictors defined in the following sections to arrive at a prediction of forecast skill. The base forecast is the 36-h RASP prognosis over the Australian region. Both the model domain and the spatial distribution of variance explained for the independent dataset (December 1989–January 1990) are shown in Fig. 2. A meridional gradient is evident, the result of the smaller amplitude of variations in the tropics compared to the higher latitudes. There also is a marked zonal trend, with maximum errors upstream and poleward of the continent at the interface between the regional and global models. These interface errors vary from summer to winter. In summer, they are mainly associated with migrating centers in the subtropical ridge across the western edge of the domain. In winter, the errors tend to arise from poor prediction of the development of baroclinic cyclones and frontal systems at the southern edge of the domain.

a) *Single model predictors*

This predictor initially was developed for the Australian region model by Bennett and Leslie (1981) and subsequently was improved by Glowacki (1988). In this approach the MFE is obtained by multiple linear regression with the detrended initial analysis and model forecast on overlapping subgrids of the Australian domain as predictors. The detrending involves subtracting, at each point of grid, the long term difference between forecast and analysis fields on a seasonal basis. The Glowacki method removes much of the error field shown in Fig. 2. It provides a very effective estimate of model skill, together with a means for correcting the model forecasts, and was implemented operationally in December 1987. Further details, together with the results of a series of predictability experiments, may be found in Leslie et al. (1989). It is noted that the Glowacki scheme actually predicts forecast error but the areal mean-square of this prediction is the estimate of E^2 for this technique.

Chen (1989) suggested that a very simple a priori estimate of the MFE could be obtained using the persistence of the forecasts in time as a predictor. The persistence predictor, P, is defined by:

$$P^2 = \overline{(F^{36} - F^{24})^2} \quad (2)$$

where F^{24} and F^{36} are the 24- and 36-h forecasts from the single model forecast run. The idea is to use the degree of development and changes in a single model run as an indicator of sensitivity to initial conditions, and thus as a predictor of the forecast error.

b) *Ensemble forecast methods*

The Australian Bureau of Meteorology's NMC receives operational forecasts from five local and international NWP forecast centers: the Bureau's global and regional models, the ECMWF model, the UK model, and the US model. These forecasts are used to derive a predictor based on the divergence, D, defined by:

$$D^2 = \frac{1}{(M-1)} \sum_{m>1} \overline{(F_1{}^t - F_m{}^t)^2} \quad (3)$$

where M is the total number of ensemble members, m, and the other terms are as in Eq. (1). The base ensemble member, $m = 1$, is the Australian regional model, RASP. No attempt has been made to investigate the use of a different model as the base member.

The Australian region model predicts to 36 h and has a short data cutoff time of 2.5 h after observation time. This is required from considerations of timeliness of the forecasts. Since all the other models described above are not received earlier than 12 h after observation time, 48-h forecasts from these models were combined with the 36-h forecast from the Australian model in deriving the forecast divergence described in using Eq. (3).

It is noted in passing that an alternative ensemble forecasting approach has been proposed by Hoffman and Kalnay (1983). This is the so-called lagged-average-forecast (LAF) method using a single model initialized at different times in preference to multiple model forecasts. Unfortunately, this method cannot be used as an additional predictor of MFE in the present study, as 36 h is the limit of the validity of the model. The technique is being tested in BMRC for application to 12- and 24-h forecasts and will be reported elsewhere.

4. Domain mean predictions

The regression equations between the MFE and the statistical regression method (S), divergence (D), and persistence (P) predictors were developed on the five month dependent dataset, with the results shown in Table 1. Statistical regression produced the largest reduction in error variance (34%) followed by the divergence of forecast models from the different forecast centers (16%). Persistence of the model forecast had only a marginal contribution, with less than 3% of the

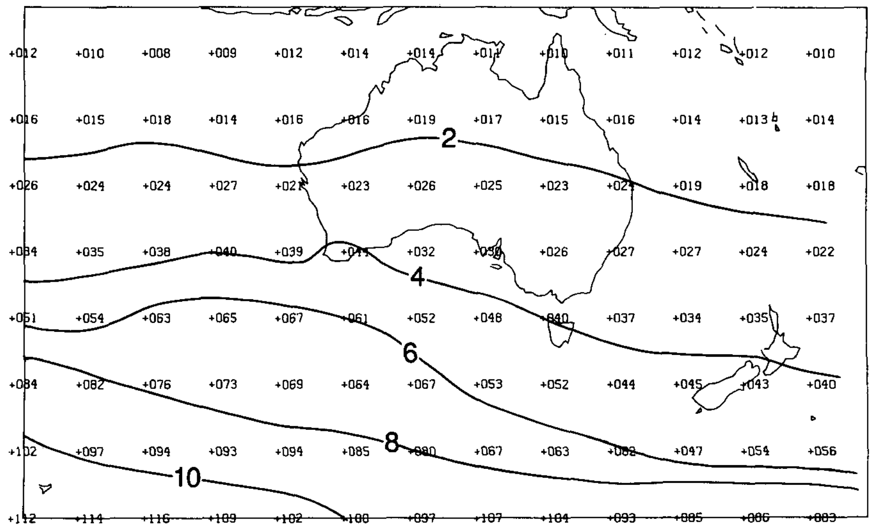

FIG. 2. Spatial distribution of rms errors averaged over the 2-month independent forecast set, December 1989–January 1990.

variance explained. As is shown in the second column of Table 1, similar results were obtained by application of the predictors to the two month independent dataset.

Combining all three predictors in an optimal linear manner explained around 50% of the error variance in both the dependent and independent forecasts. This is a remarkably high result from such a simple approach and implies that substantial operational information can be provided on the quality of the forecast products at the time of issue.

It is necessarily the case that optimal combinations of predictors explain more of the total variance than does any single predictor alone (e.g., Fraedrich and Leslie 1987; Leslie and Fraedrich 1990). However, the markedly increased variance explained by the combined predictor indicates that each of the individual predictors is largely independent of the other and that each contributes to a different part of the error variance. This is confirmed by correlating the three predictors and finding coefficients varying from 0.19 to 0.23. This finding is also supported by the following heuristic reasoning based on the different approaches adopted by each predictor.

TABLE 1. Correlation coefficient and percentage of explained model error variance (in parentheses) using three predictors individually and in optimal combination.

Techniques	Dataset: Dependent	Dataset: Independent
Statistical regression (*S*)	0.58 (34)	0.56 (31)
Divergence (*D*)	0.40 (16)	0.42 (17)
Persistence (*P*)	0.18 (03)	0.23 (05)
Combined S + D + P	0.71 (51)	0.67 (48)

The statistical regression scheme removes the spatially varying model bias. For example, Fig. 3 shows the mean forecast errors over the Australian domain for both the dependent (Fig. 3a) and independent (Fig. 3b) forecasts. Although there are variations between the two (largely due to moving from winter to sum-

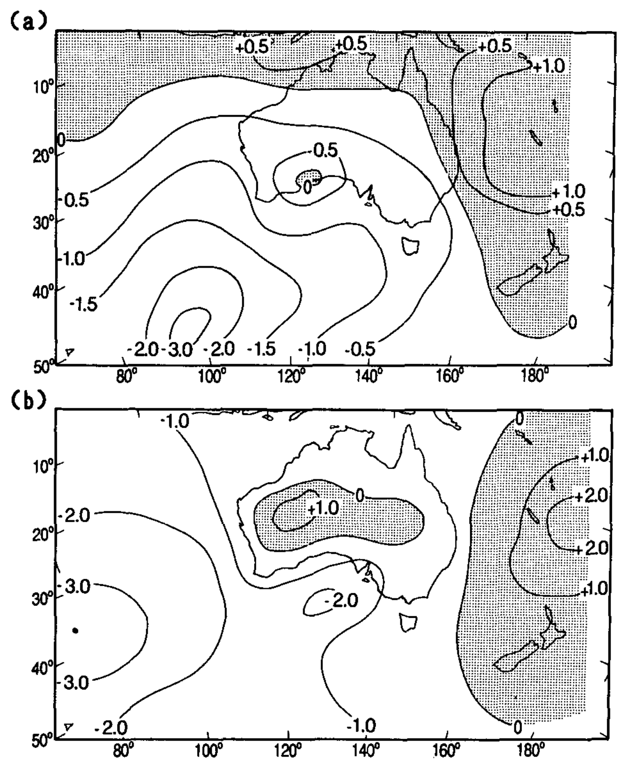

FIG. 3. Spatial distribution of the mean forecast errors (10^{-1} hPa) for: (a) the dependent forecast set and (b) the independent forecast set.

mer), a consistent basic pattern can be seen. Large negative forecast errors over the southwestern part of the domain change to smaller positive errors in the northeast. The across domain variation is related to problems with the nesting RASP to forecasts from the global model that are 12 h old and have poor resolution. Deficiencies in the physical parameterisation of the continental heat flux in summer also are apparent in the underprediction of the continental heat low in summer (Fig. 3b). The statistical regression technique effectively predicts much of this bias pattern, which contributes heavily to the error variance explained by this technique.

Thus, the divergence and persistence predictors are based on similar physical processes and would be expected to be partially related. The relatively low observed correlation indicates that the sensitivity of the atmosphere to initial conditions changes sharply on a day-to-day basis. This is supported by the rapid changes in model error shown in Fig. 1a and emphasizes that there is little persistence in model error characteristics from one day to the next.

The transients that effect the divergence predictor would be expected to have little effect on the mean statistical regression procedure, and vice versa. Thus, it is to be expected that the statistical and the sensitivity predictors will be largely independent of each other, as is confirmed by the weak correlations between these predictors.

5. Variations in spatial predictability

Although the domain mean results are encouraging, the spatial characteristics of the errors are of far more

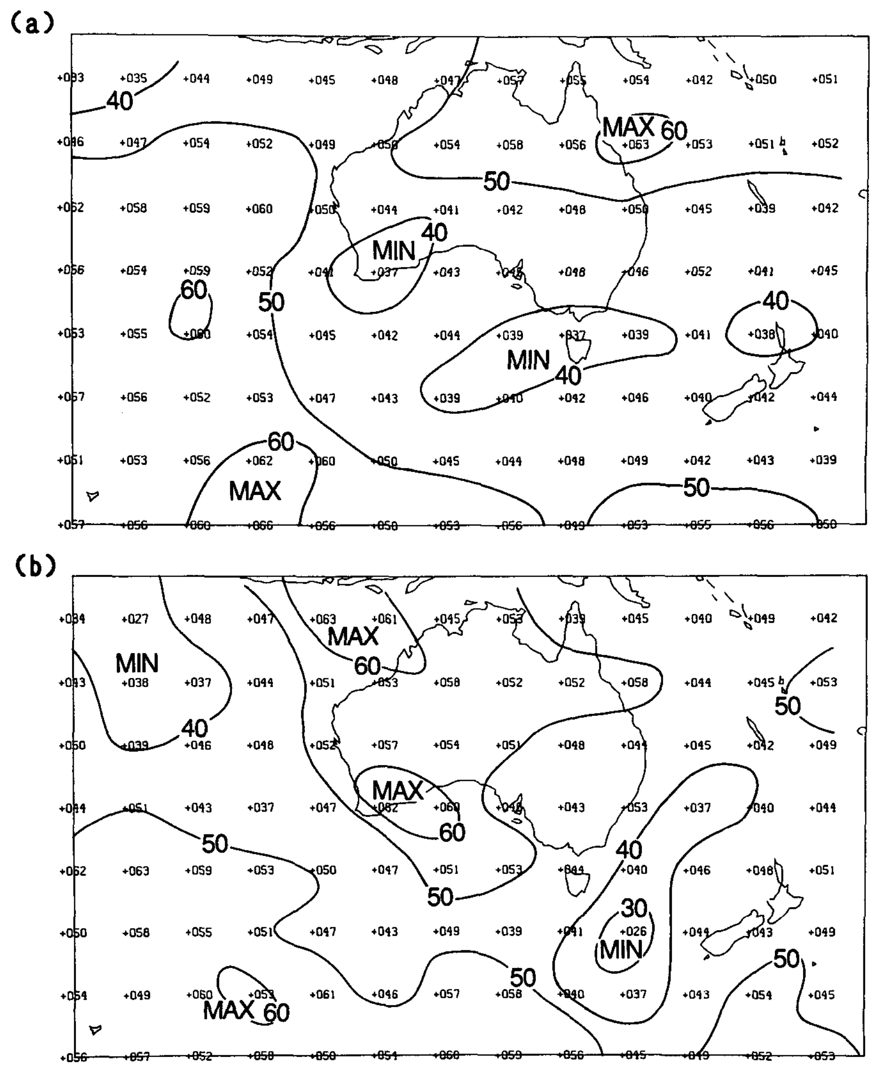

FIG. 4. Spatial distribution of percentage of forecast error explained by the combined predictor for: (a) the dependent forecast set and (b) the independent forecast set.

interest to forecasters who provide forecasts at specific localities. Therefore, examined next is the spatial variability of the error prediction using time-mean charts and individual case studies. These are obtained directly from the results of the regional studies applied to individual grid points.

a. Time-mean spatial variability

The spatial variation of predictive skill, averaged over the dependent and independent time periods, is shown in Fig. 4, which contains analyses of the mean percentage error explained for the dependent (winter) forecast set (Fig. 4a) and the independent (summer) forecast set (Fig. 4b). Very strong predictive capacity is indicated over the western half of the continent in summer and over the tropics in both seasons. But the skill is markedly weaker in the southeastern part of the domain, especially in the independent forecast set.

The effects of the statistical regression predictor are evident in the high variance explained in the southwestern corner of the domain and over the continent in summer (compare the patterns in Fig. 4 with those in Fig. 3). Consider that time-transient effects are responsible for the relatively low forecast skill in the southeastern part of the continent. For example, the band of low skill across Tasmania in winter seems to be associated with the development of frontal systems

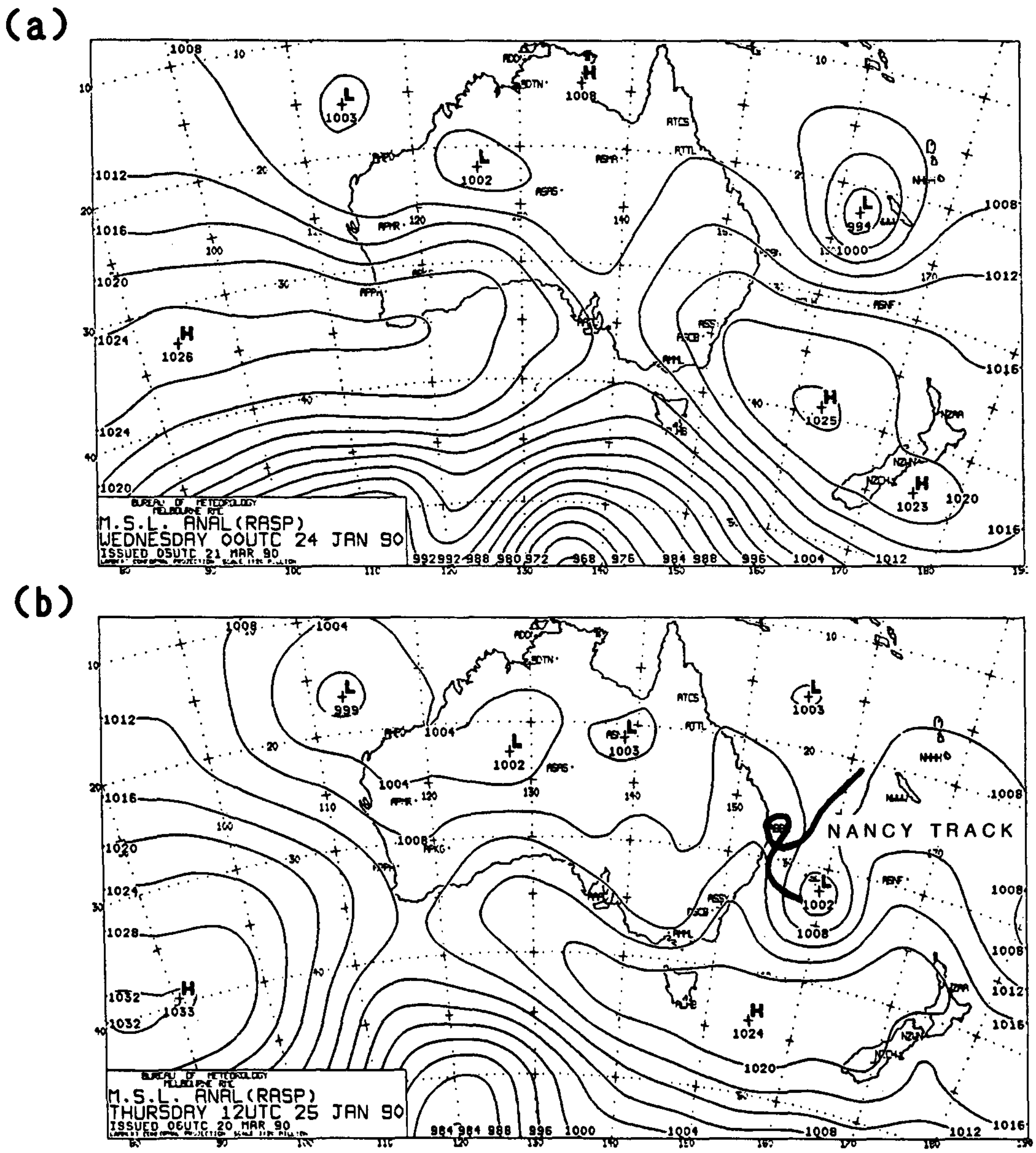

FIG. 5. Australian region analyses of MSLP (hPa) for: (a) 0000 UTC 24 January and (b) 1200 UTC 25 January 1990. Also shown on (b) is the 36-h track of Tropical Cyclone Nancy.

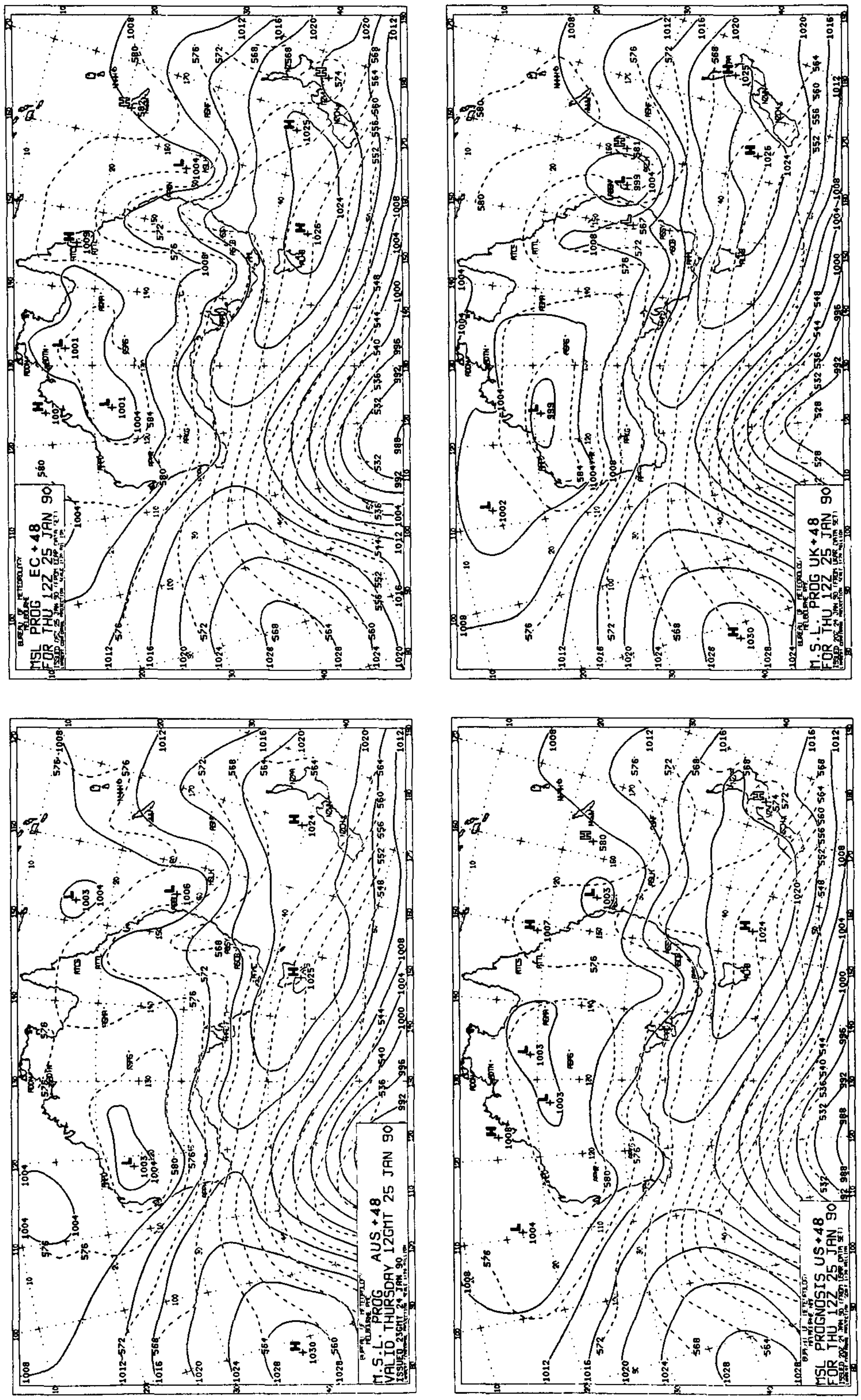

FIG. 6. 48-h forecasts from the ECMWF, UK, NMC Washington, and HASP models valid at 1200 UTC 25 January 1990.

in this region. Section 5 will return to a discussion of the predictive capacity in this region.

Thus, the statistical regression technique provides an overall correction to the prognosis fields. Since the localized variations in space and time are due largely to transients, the divergence predictor is the most important on a case-by-case basis, as is demonstrated in the following two case studies.

b. Case study 1: predicting a good forecast

One of the most difficult forecast situations off the eastern coast of Australia is the poleward movement and transition of a tropical cyclone into a midlatitude depression. An example is shown in Fig. 5 for severe Tropical Cyclone Nancy in January 1990. AT 0000 UTC 24 January (Fig. 5a) Nancy was an intense tropical cyclone near 22°S, 162°E. In the next 36 h Nancy moved close to the Australian coast, transformed into an east coast cyclone (Holland et al. 1987), and moved back out to sea, as is shown by the track in Fig. 5b. The situation at 1200 UTC 30 January is shown in Fig. 5b, with the transformed Nancy near 32°S, 160°E. Other notable features are the southeastward movement of a midlatitude frontal system from near 140°E out of the domain, the development of a frontal system at 120°E, and the development of a trough in the easterly flow over southeastern Australia.

As shown in Fig. 6, the 48-h forecasts from the four large-scale models received in the Australian NMC were quite consistent. The location of the extratropical system that was Tropical Cyclone Nancy was well predicted by all models, though there was a tendency for the movement to be too slow. Further, all of the synoptic features noted in the previous paragraph were predicted accurately by each model.

A similarly consistent forecast was made with the RASP model, as is shown in Fig. 7a. The predicted domain mean rms error for this forecast using the combined predictor set was 2.5 hPa, which is a very low value (see Fig. 1). Close agreement was found with the actual forecast error of 2.7 hPa.

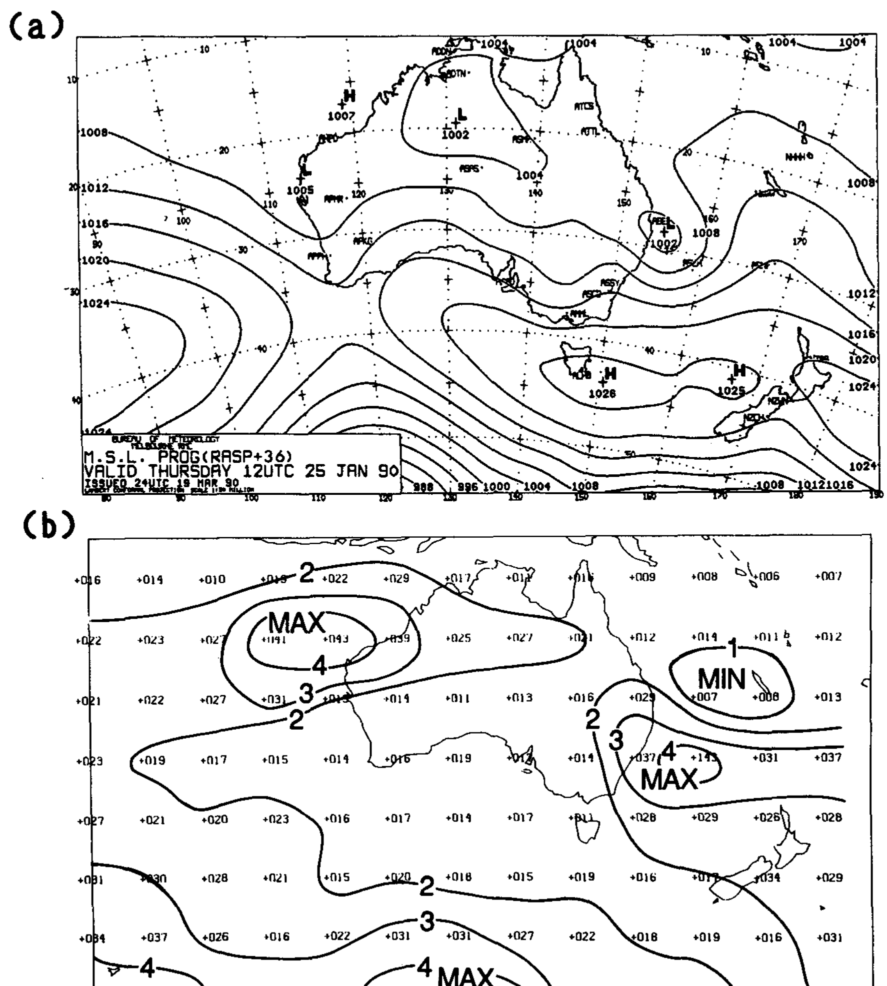

FIG. 7. (a) RASP 36-h MSLP forecast valid at 1200 UTC on 25 January 1990 and (b) the predicted spatial distribution of rms errors in this forecast.

Of considerable interest is the ability of the predictor to differentiate regions of relatively poor and relatively good forecast quality. As may be seen in Fig. 7b, the largest predicted forecast errors were in the regions of developing and moving systems off the Australian west coast and in the south near 120°E, and of Tropical Cyclone Nancy. Note, however, that the forecasts of the rapidly moving and decaying frontal system along 140°E in Fig. 5a were predicted to be very accurate. Comparing the RASP forecast and the verifying analysis in Figs. 7a and 5b shows that this predicted event did indeed occur.

c. Case study 2: predicting a poor forecast

To indicate the ability of the combined predictors to predict a poor NWP forecast, the situation of 6 and 7 August 1989 is shown in Fig. 8. At 0000 UTC 6 August (Fig. 8a), a strong frontal system was approaching western Australia and a large cyclone was moving away from the east Australian coast. Within the next 36 h (Fig. 8b), a small cutoff low had formed from a wave on the frontal system. The east coast cyclone had moved southeastward, weakened, and been replaced by new development off the Australian coast.

Each of the large-scale models (not shown) provided conflicting forecasts of the development of both of these systems. As a result, the combined prediction was for a very poor forecast with a domain average rms of 4.6 hPa. The spatial distribution shown in Fig. 9a further indicated that the forecast would be excellent off western Australia, over the central continent, and in the vicinity of the high pressure belt across Tasmania. Very poor forecasts were predicted in the vicinity both of the frontal system south of western Australia and of the east coast cyclone.

The quality of the 36-h RASP forecast, shown in Fig. 9b, was in close agreement with these predictions.

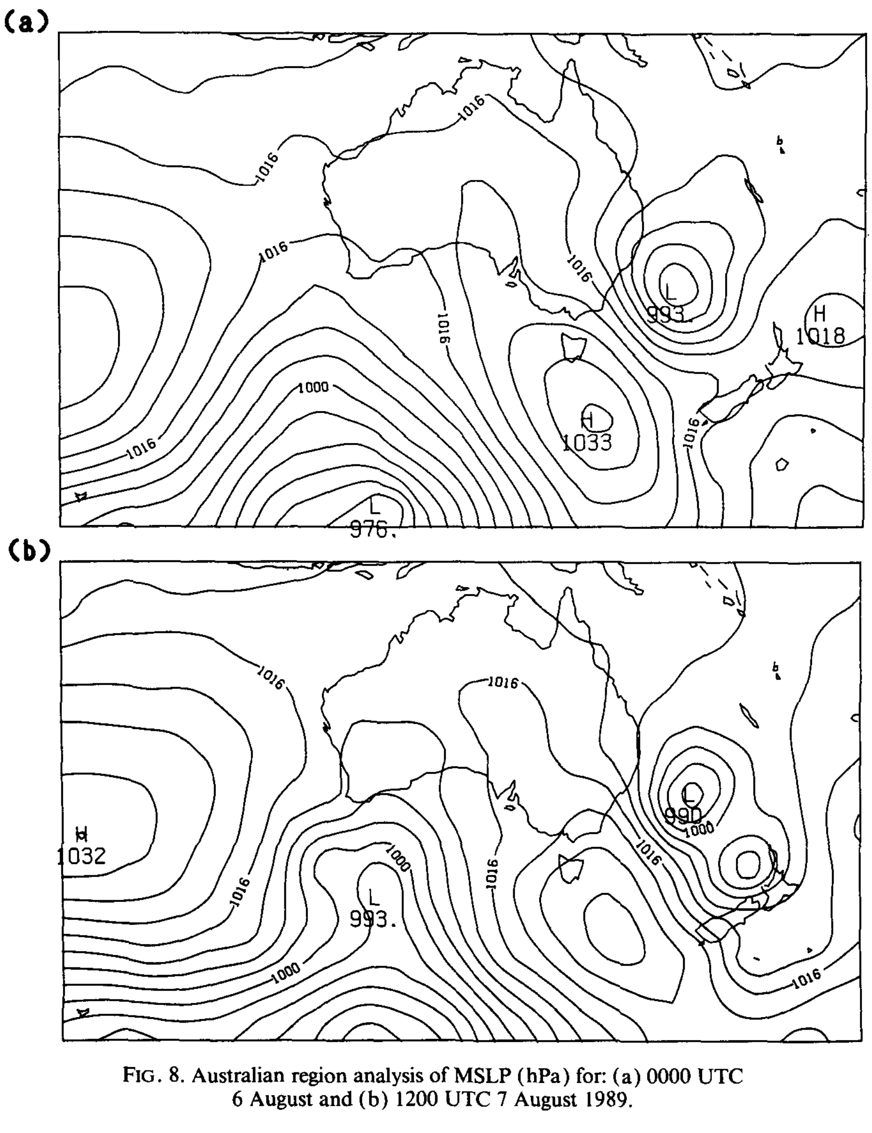

FIG. 8. Australian region analysis of MSLP (hPa) for: (a) 0000 UTC 6 August and (b) 1200 UTC 7 August 1989.

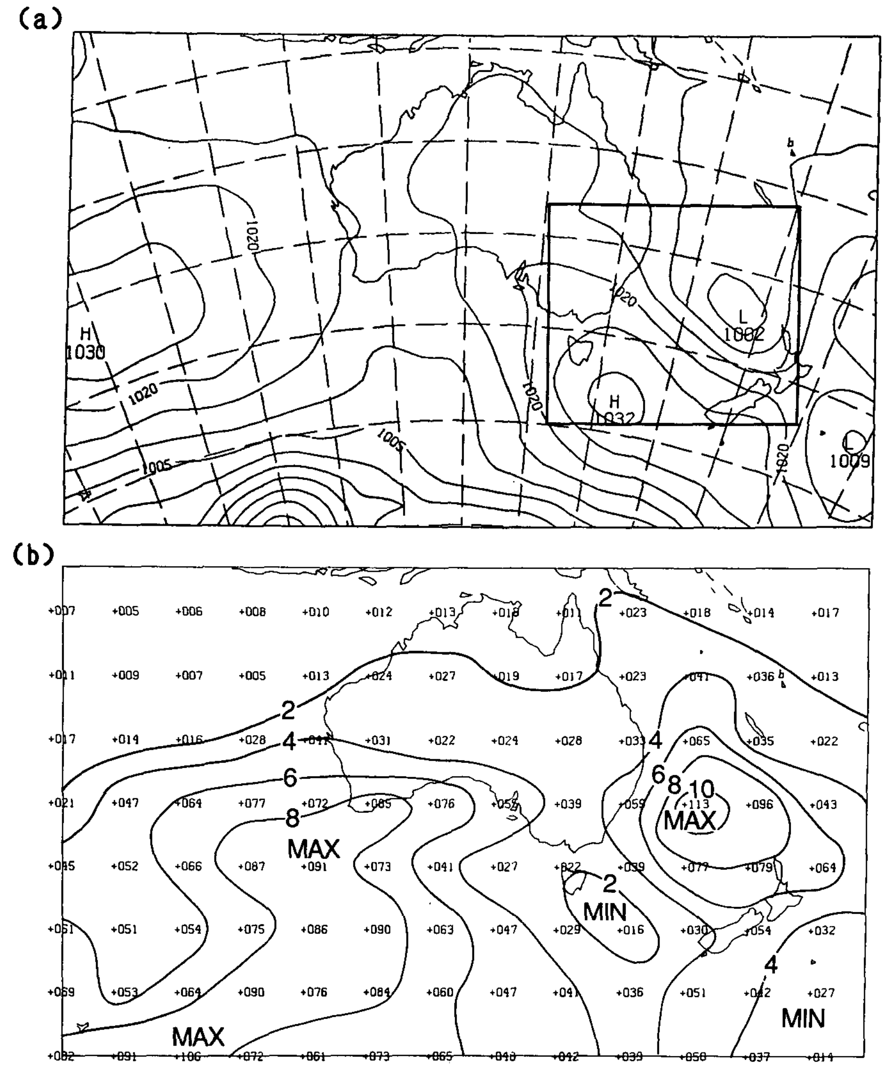

FIG. 9. (a) Predicted spatial distribution of forecast errors and (b) the RASP 36-h forecast for 1200 UTC 7 August 1989. The box over southeastern Australia and New Zealand in (b) indicates the domain used for testing our capacity to predict model skill over a limited region.

The domain mean rms error was 4.9 hPa, and the model missed both the redevelopment of the east coast cyclone and the wave development on the frontal system.

6. Regional predictions of forecast skill

The case studies presented in sections 4b and 4c indicate that the predictive equations have some capacity both to predict the domain mean model skill and to indicate those parts of the domain that will be poorly forecast.

This regional predictive capacity was tested by the following method. First, the region shown in Fig. 9b covering southeastern Australia and New Zealand was chosen; this region covers the major population centers for Australia. It also is the area of least variance explained by the predictive equations (see Figs. 4a,b) and thus provides the severest possible test of the system. Next, the dependent dataset was used to separate the forecasts into three roughly equal categories of high, medium, and low quality based on domain rms errors (high for MFE < 2.5 hPa, medium for $2.5 <$ MFE < 3.5 hPa, and low for MFE > 3.5 hPa). Then the category for each forecast in the independent sample was predicted, with the results shown in Table 2.

Following Panofsky and Brier (1958), the overall skill of these forecasts can be estimated from

$$S = \frac{C - X}{T - X}$$

$$X = \frac{1}{T} \sum_{i,j} O_i P_j \qquad (4)$$

where C is the number of correct forecasts, $T = 60$ is the total number of cases, and O_i and P_j are the observed row i and predicted column j, respectively. Here X is simply an estimate of the number of correct forecasts that would be made by chance, and $S = 0$ for no skill and 1 for perfect forecasts. Substituting the relevant numbers from Table 2 into Eq. (4) gives $S = 0.28$ for the categorical forecasts, which is regarded as a good level of skill (Panofsky and Brier 1958).

The results in Table 2 are therefore very encouraging. The predicted category was verified more frequently than any other category, as is shown by the good skill score that was achieved. Perhaps more importantly, major prediction errors across two categories were quite rare. For example, of the 20 predictions of high forecast quality, 11 verified and only 1 turned out to be a poor forecast. Conversely, of the observed 22 high quality forecasts, only 2 were predicted to be in the low category.

7. Conclusions

The capacity for predicting the skill of NWP model forecasts on a daily and regional basis has been examined. Three predictive methods were tested: 1) a statistical regression approach that removes the overall spatial bias in the forecasts; 2) a persistence parameter that measures the time consistency of the Australian region model; and 3) a divergence parameter that measures the sensitivity to initial conditions from the consistency of 2 Australian and 3 international NWP forecasts received at the Australian NMC.

Using both dependent and independent forecast sets, the statistical regression technique explained the highest proportion of the domain-mean forecast error variance at around 30%–35%; the divergence parameter was next with 15%–20%; and the persistence parameter was of only marginal use at less than 5% variance explained. Combining all three techniques increased the explained variance of forecast error to around 50%. As far as is known, this is highest reported skill for techniques of this type.

The capacity of the predictive system to indicate the spatial variability of skill in a single forecast has been demonstrated using two case studies and series of categorical forecasts over a limited domain. In the case studies of a poor and of a good numerical forecast, the system accurately predicted both the domain average skill and those parts of the domain where the forecasts were skillful or in error. A series of categorical forecasts of high, medium, and low forecast skill over southeastern Australia for the independent forecast set confirmed the case study findings. Based on 60 forecasts, the contingency skill score was a respectable 0.28, and major prediction errors were very rare.

These encouraging results indicate that the methods adopted here have immediate potential for operational application. An operational trial will be introduced in the near future using a combination of Australian region analyses of forecast skill and categorical forecasts for specific regions.

TABLE 2. Contingency table for the categorical predictors of high, medium, and low forecast confidence for southeastern Australia in summer.

Observed forecast skill	Predicted forecast confidence			
	High	Medium	Low	Total
High	11	9	2	22
Medium	8	13	6	27
Low	1	2	8	11
Total	20	24	16	60

Acknowledgments. We are indebted to David Pike for preparation of the datasets and help with drafting of figures. The investigation was partially supported by the U.S. Office of Naval Research under Grant N-000014-89-J-1737.

REFERENCES

Anthes, R. A., 1983: A review of regional models of the atmosphere in middle latitudes. *Mon. Wea. Rev.*, **111**, 1306–1355.
Bennett, A. F., and L. M. Leslie, 1981: Statistical correction of the Australian region primitive equation model. *Mon. Wea. Rev.*, **109**, 453–462.
Chen, W. Y., 1989: Another approach to forecasting forecast skill. *Mon. Wea. Rev.*, **117**, 427–435.
Fraedrich, K., and L. M. Leslie, 1987: Combining predictive schemes in short-term forecasting. *Mon. Wea. Rev.*, **115**, 1640–1644.
Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.
Glowacki, T., 1988: Statistical corrections to dynamical model predictions. *Mon. Wea. Rev.*, **116**, 2614–2627.
Hoffman, R. N., and E. Kalnay, 1983: Lagged-average forecasting. *Tellus*, **35A**, 100–118.
Holland, G. J., A. H. Lynch and L. M. Leslie, 1987: Australian east coast cyclones. Part I: Synoptic overview and case study. *Mon. Wea. Rev.*, **115**, 3024–3036.
Kalnay, E., and A. Dalcher, 1987: Forecasting forecast skill. *Mon. Wea. Rev.*, **115**, 349–356.
——, and M. Ham, 1989: Forecasting forecast skill in the Southern Hemisphere. *Extended Abstracts, Third International Conference on Southern Hemisphere Meteorology and Oceanography*, Buenos Aires, American Meteorological Society, 24–27.
Leith, C., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418.
Leslie, L. M., and K. Fraedrich, 1990: Reduction of tropical cyclone position errors using an optimal combination of independent forecasts. *Wea. and Forecasting*, **5**, 158–161.
——, G. A. Mills, L. W. Logan, D. J. Gauntlett, G. A. Kelly, M. J. Manton, J. L. McGregor and J. M. Sardie, 1985: A high resolution primitive equations model for operations and research. *Aust. Met. Mag.*, **33**, 11–35.
——, K. Fraedrich and T. J. Glowacki, 1989: Forecasting the skill of a regional numerical weather prediction model. *Mon. Wea. Rev.*, **117**, 550–557.
Lorenz, E. N., 1969: Three approaches to atmospheric predictability. *Bull. Amer. Met. Soc.*, **50**, 345–349.
Murphy, J. M., 1990: Assessment of the practical utility of extended range ensemble forecasts. *Quart. J. R. Met. Soc.*, **116**, 89–125.
Neumann, C. J., and J. M. Pelissier, 1981: Models for the prediction of tropical cyclone motion over the North Atlantic: An operational evaluation. *Mon. Wea. Rev.*, **109**, 522–538.
Panofsky, H. A., and G. W. Brier, 1958: *Some Applications of Statistics to Meteorology*. Pennsylvania State University, 224 pp.