Report on the Extended Range Weather Forecast Project : Phase II

A.B. Mullan A.A. Neale
C.G. Revell R.M. Smith

Report on the Extended Range Weather Forecast Project - Phase II

A.B. Mullan, A.E. Neale, C.G. Revell and R.M. Smith

Technical Information Circular 192

UDC 551.509.333 (931)

June 1984

New Zealand Meteorological Service
PO Box 722
WELLINGTON

REPORT ON EXTENDED RANGE WEATHER FORECAST PROJECT - PHASE II

A.B. Mullan, A.A. Neale, C.G. Revell and R.M. Smith

ABSTRACT

The objectives of Phase II of the Extended Range Weather Forecast Project were to (i) evaluate the impact on the Project of forecast charts for the next four days, received from the European Centre for Medium Range Weather Forecasts; (ii) assess the value of extended range forecasts, based on the Extended Range Weather Forecast Project, which were issued in real time to a limited number of users; and (iii) briefly consider the implications of introducing an operational extended range forecast service within the National Weather Forecasting Centre.

The 88 forecasts prepared during Phase II were assessed in the same manner as for Phase I; both subjectively by verifying the broadscale surface flow pattern in the New Zealand region, and objectively by verifying categories of rainfall amount and maximum temperature prepared for six regions of New Zealand. Surface flow pattern prediction during Phase II was significantly more accurate than in Phase I, attributable to a large degree to the availability of European Centre predictions. However, there was no comparable improvement in skill for the categorical forecasts of rainfall and maximum temperature.

Users of extended range forecasts and prognostic MSL charts (issued by the Project to a limited audience) generally considered them to be very useful.

SECTION I : Introduction and method

The Report on the Extended Range Weather Forecast Project - Phase I (Mullan et al., 1983) gives the background to the Project and describes the method employed. While the latter remained largely the same in Phase II, there were a number of differences from Phase I.

(i) Extended range forecasts were made several hours later in the day (between 1515 and 1645 NZST), to allow time for the predictions of the European Centre for Medium Range Weather Forecasts (ECMWF) to be received. Even so, there were a few occasions when these predictions were not received in time to be used.

(ii) The predicted behaviour of the planetary and medium waves (derived from Fourier analyses of hemispheric grid point 500 hPa geopotential height data, supplied by the Southern Hemispheric Analysis Centre in Melbourne), was compared with the forecasts received from the ECMWF. On occasions when these two sources of information indicated different types of situation development, it was necessary to evaluate the most likely course events would take. This evaluation became easier to make with increasing experience. Fourier analyses also provided a backup facility which enabled extended range forecasts to be issued when ECMWF data had not been received.

The data received daily from the ECMWF consisted of a series of messages in a numerical code. These were processed by the N.Z. Meteorological Service's computer, and printed out as a series of charts (for M.S.L. and the 500 hPa level) giving the analysis at 1200 GMT and forecasts for subsequent days out to the fourth day.

In Phase II the 48-hour prognosis (T+48) was obtained by direct interpolation of ECMWF prognoses, and seldom was this modified significantly. The T+96 prognosis was obtained by extrapolation of the ECMWF series of prognoses, with some modification as indicated above.

(iii) A change was made in the timing of the forecast validity periods. In Phase I, there were two periods covering days 1-2 and days 3-5. In Phase II, the periods covered days 2-3 and days 4-5, where the forecast 'day' begins at noon on the date of issue. (Hence the first period for days 2-3 begins verifying about 19 to 21 hours after issue time). This timetable is shown in Fig. 1. Once again, flow pattern forecasts were produced for the mid-points of the two periods, indicated by T+48 and T+96 respectively. For practical verification purposes, data referring to climatological days beginning at 0900 were used.

(iv) A change was made to the categories of maximum temperature forecasts and to their method of verification.

An extreme category (1 = cold, 3 = warm) occurred if the daily maximum temperature of at least three stations in the district deviated from normal by at least 3^{o}C. A given 2-day period verified as an extreme if either day met this criterion, with category 2 being the default or null forecast. Sudden changes from one extreme to the other were treated by verifying each day separately.

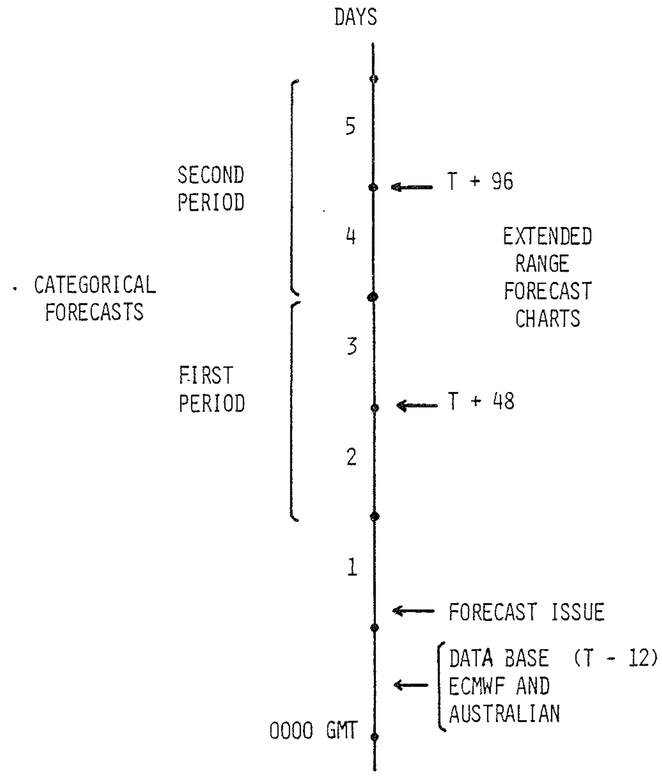

Fig. 1. Timetable of activity during the Extended Range Weather Forecast Project - Phase II. The sequence of 0000 GMT times is shown by dots on the time scale.

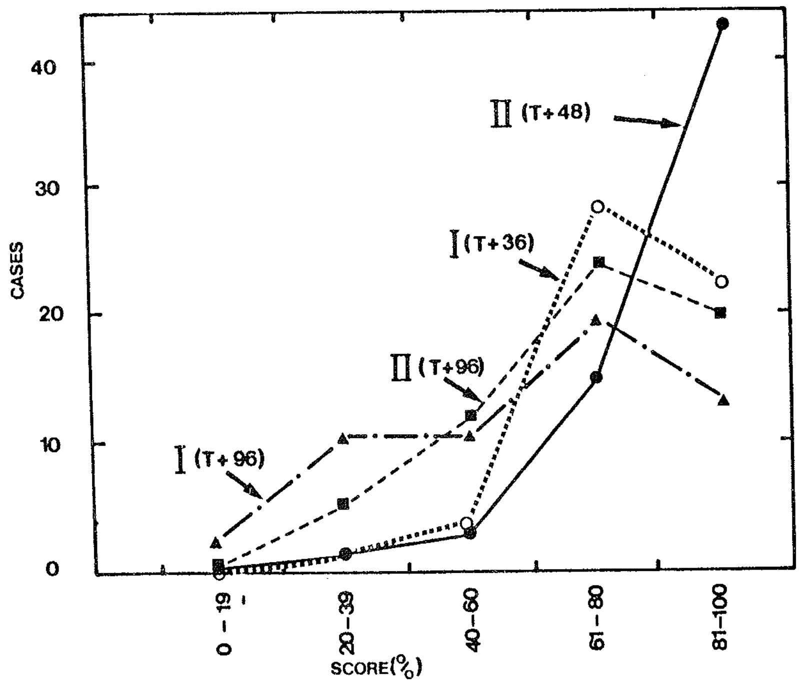

Fig. 2. Frequency distribution of verification scores of flow pattern forecasts in both Phase I and Phase II. Scores are averages of the scores at four sites in New Zealand. The 36-hour forecast in Phase I is indicated by I(T+36), etc.

Table 1: Accuracy ratings of MSL flow pattern prognoses. Separate values for direction and curvature (in addition to the combined value) are given for 'all sites combined' in ERP-I. Ratings for random forecasts are included.

		All sites Dir	All sites Curv.	All sites Comb.	North Cape	Cook Strait	Stewart Island	Chatham Islands	No of prognoses
ERP-II	T+48	-	-	81	82	82	83	76	87
	T+96	-	-	68	69	66	74	64	87
ERP-I	T+84	47	66	53	58	53	51	51	54
Random		36	63	45	-	-	-	-	-

Table 2: Monthly accuracy ratings of MSL flow pattern prognoses. Different people did the evaluations in the two periods February to April and April to August. Reasonable consistency is shown in April when both did the evaluation.

	Month	T+48		T+96	
1982	Feb	75		59	
	Mar	74		62	
	Apr	87	84	61	59
	May		90		78
	Jun		86		71
	Jul		86		75
	Aug		74		70

(v) Operational extended range forecasts for New Zealand, extended range forecast statements giving the reasoning used to develop the extended range forecasts, and surface prognostic charts for T+48 and T+96 (flow pattern only), were given a limited distribution within the Meteorological Service, for use as guidance material. In addition experimental long range forecasts were issued to a weekly radio programme on farming, and to two non-meteorological organisations in southern New Zealand for their use and evaluation.

SECTION II : Results

(I) Verification of flow pattern forecasts

As in Extended Range Weather Forecast Project - I (ERP-I) flow patterns were verified by scoring the forecasts in terms of isobar direction and curvature at four points (North Cape, Cook Strait, Stewart Island, and Chatham Islands). The accuracy rating is the score as a percentage of the maximum possible score. The distribution of scores averaged over the four sites for both Phase I and Phase II is shown in Fig.2. For the shorter range forecasts, almost all the 36-hour forecasts in Phase I [I(T+36)], and almost all the 48-hour forecasts in Phase II [II(T+48)] have scores greater than 60%. Moreover, in Phase II, despite the longer range of the forecast, a greater proportion of forecasts score more than 80% than in Phase I. For the longer range forecasts (96-hours), the two curves are similar, but in Phase II considerably fewer forecasts were in the lower scoring categories.

Table 1 gives the accuracy rating for separate sites (direction and curvature combined) as well as for all sites combined at T+48 and T+96. Results for ERP-I at T+84 and for random forecasts are included for comparison. The method of scoring random forecasts is given by Neale (1982). The scores for T+48 are similar at all sites except Chatham Islands where the accuracy is lower. For T+96 however the score is distinctly higher for Stewart Island (74%) and again lowest (64%) for Chatham Islands. The improvement over ERP-I ranges from 23% for Stewart Island to 11% for North Cape.

No trend appears in scores for individual months (Table 2), the accuracy at T+48 being about 75% in February, March and August and some 10-15% higher in the other months. At T+96 though, the accuracy was significantly better in the months May to August than in February, March and April. The differences cannot be explained on a monthly basis but the scores appear to depend on the type of situation prevailing.

Table 3: Dependence of the accuracy of forecasts at T+96 on those at T+48. Average accuracy scores (%) of flow pattern forecasts over New Zealand at T+96 for given ranges of scores for corresponding T+48 forecasts. Two different scoring methods (a and b) were used in ERP-I

	Accuracy range (%) at T+48		
	0-60	61-80	81-100
Average accuracy (%) at T+96			
ERP-I (a)	57	52	52
(b)	63	62	59
ERP-II	70	67	67

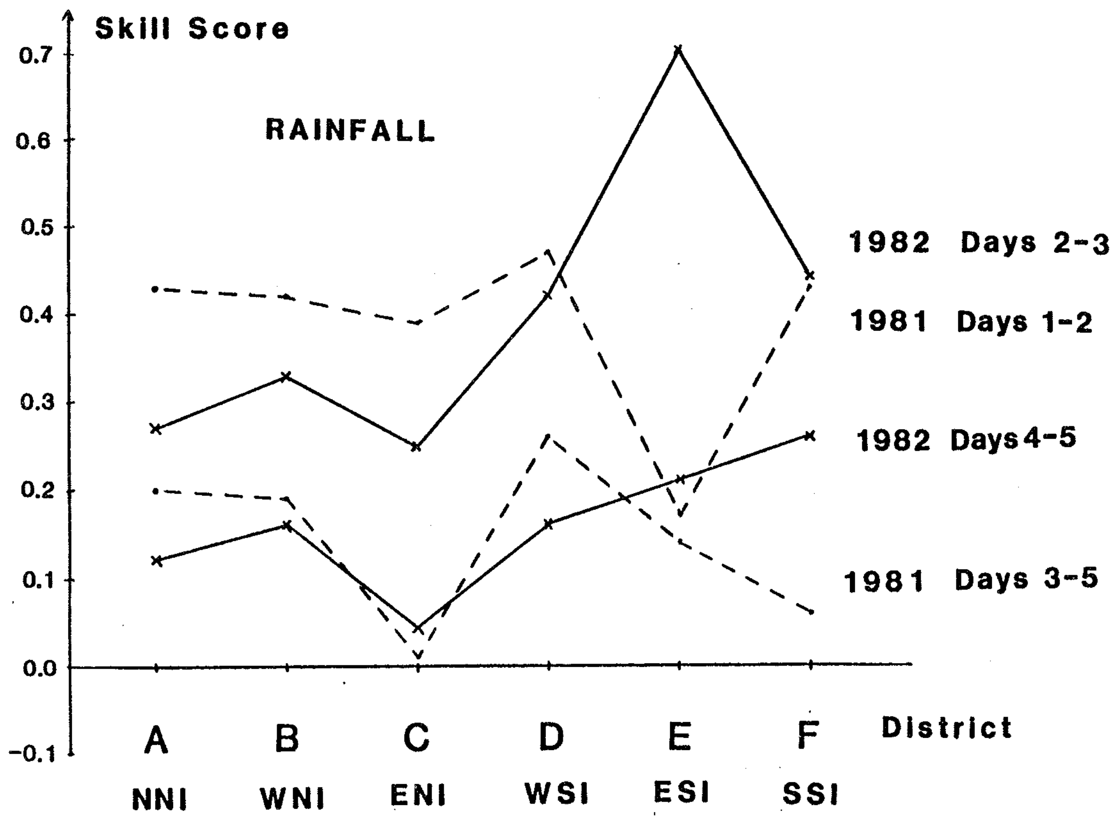

Fig. 3. Rainfall skill scores for the six districts (A to F - see Fig.5). Phase I results (dashed) are for days 1-2, and 3-5; Phase II results (solid) are for days 2-3, and 4-5.

Table 4: Distribution of lag between forecast and occurrence of a major trough or ridge axis near New Zealand.

		Lag (days)												
	Quarter	-2	-1½	-1	-½	0	½	1	1½	2	2½	3	> 3	Average Lag
	First	0		0		3		2		3		1	2	+1.9
	Second	1		4		4		2		1		0	2	+0.5
ERP-I	Third	1		2		5		1		3		1	0	+0.5
	Fourth	1		2		9		1		0		0	0	-0.2
	Month													
	Feb	0	0	4	4	3	1	1	0	0	0	0	0	-0.3
	Mar	0	0	2	0	7	0	0	2	2	0	1	0	+0.6
	Apr	0	0	1	1	1	1	1	1	0	0	1	3	+1.4
ERP-II	May	0	0	2	2	4	3	0	1	0	0	0	1	+0.3
	Jun	0	0	1	1	3	1	1	2	0	0	0	2	+1.0
	Jul	0	0	3	0	4	1	1	1	0	0	0	2	+0.7
	Aug	0	0	2	2	1	3	3	0	1	0	1	0	+0.5

Table 5: Contingency table (a) of forecast (at T+96) and subsequently observed westerly index category (H = high, L = low, ½ = intermediate case).
Comparison table (b) of percentages of each westerly index category forecast correctly, and overall skill score, in ERP-I and ERP-II.

(a)

		Forecast			
		H	½	L	
	H	16	5	1	22
Observed	½	9	36	5	50
	L	2	8	5	15
		27	49	11	87

(b)

	Index Category				Skill
	H	½	L	ALL	S
ERP-I	65%	57%	40%	57%	0.38
ERP-II	73%	72%	33%	65%	0.41

It was found in ERP-I that the accuracy of the forecast flow for the first period did not appear on average to influence the accuracy of that for the extended range forecast period. Table 3, which repeats the ERP-I results, shows that in 1982 also there was no dependence.

Broad-scale aspects of the extended period flow pattern forecasts were evaluated by two schemes in ERP-I. The first gave an indication of timing errors involved in forecasting the eastward progression of a main trough or ridge in the New Zealand area. The result indicated that in the early stage of the project long waves were moved eastward too quickly but later in the project large individual lags became less frequent. A comparison with Phase II is given in Table 4 where the number of forecasts in each period is similar. There is no corresponding trend in the ERP -II results where the pattern of lag is different. The majority of cases are clustered within 1 day of zero lag but there is a group of "outliers". The latter corresponds to situations which simply did not evolve in the way expected. The average lag fluctuates from month to month and it is interesting to note that the maximum occurred in April which was characterised by a lengthy period when the major systems were stationary.

The second scheme involved placing the forecast and verifying flow patterns in one of three westerly index categories determined by the distribution of surface pressure over the north-south length of New Zealand. The categories are :-

(i) High Index (H) - in which pressure is highest over northern New Zealand and decreases southwards;

(ii) Low Index (L) - in which pressure is highest over southern New Zealand and decreases northwards;

(iii) Intermediate Index (½) - in which pressure is highest or lowest over central New Zealand.

Results for the two phases are given in Table 5.

The overall small improvement in ERP-II shown by the skill scores (which are calculated in the same way as for the categorical forecasts) is the result of better predictions of ½ - index situations being almost nullified by worse predictions of low-index situations. This probably reflects the influence of the ECMWF guidance which is known to have an inherent error whereby the forcing of stationary waves is poorly represented.

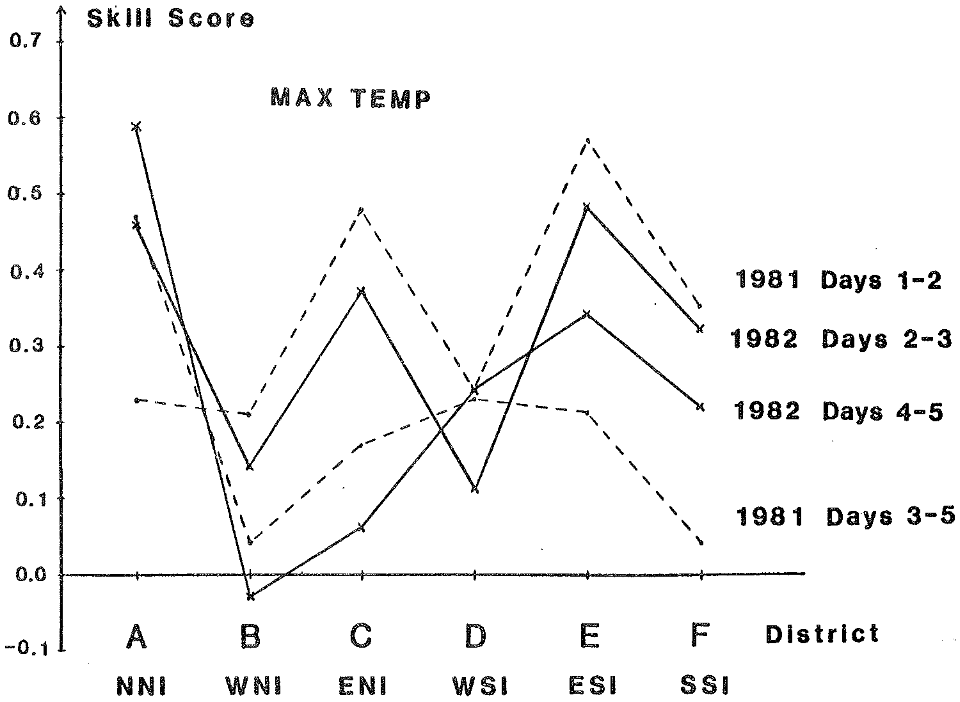

Fig. 4. Maximum temperature skill scores for the six districts A to F. See Fig. 3 for additional explanation.

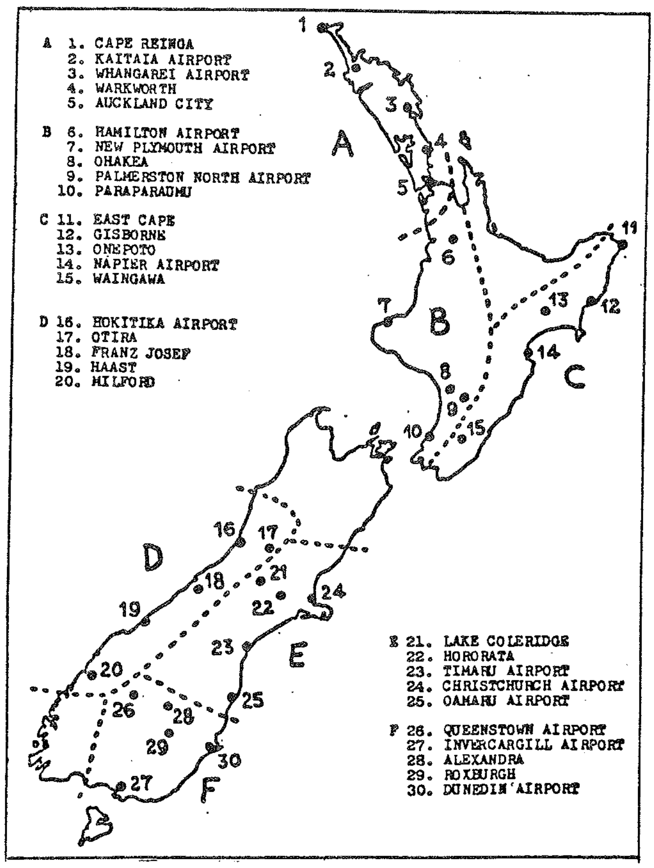

Fig. 5. Dots mark the location of reporting stations used for verifying categorical forecasts for the six regions (5 stations per region). A, B and C are northern, western and eastern North Island; D, E and F are western, eastern and southern South Island.

Table 6: Phase II skill scores (S) for Actual forecasts (A) and Persistence forecasts (P). Also shown is the skill improvement over Persistence (A-P), and the standard deviation (σ) of the Actual skill scores.

Rain: Days 2-3

District

		A	B	C	D	E	F	A-F Average
S:	A	.27	.33	.25	.42	.70	.44	0.400
	P	.02	-.00	-.07	.12	.68	.08	0.137
	A-P	.25	.33	.32	.29	.02	.36	0.263
σ:	A	.15	.13	.15	.10	.09	.23	

Rain: Days 4-5

		A	B	C	D	E	F	A-F Average
S:	A	.12	.16	.04	.16	.21	.26	0.159
	P	-.03	.01	-.04	.05	-.03	-.04	-0.013
	A-P	.15	.15	.08	.10	.24	.30	0.172
σ:	A	.18	.14	.18	.09	.50	.23	

Temp: Days 2-3

		A	B	C	D	E	F	A-F Average
S:	A	.46	.14	.37	.11	.48	.32	0.314
	P	.50	.08	.10	.14	.10	.22	0.189
	A-P	-.04	.06	.26	-.03	.38	.11	0.125
σ:	A	.25	.26	.11	.20	.10	.09	

Temp: Days 4-5

		A	B	C	D	E	F	A-F Average
S:	A	.59	-.03	.06	.24	.34	.22	0.233
	P	.50	-.03	.02	.02	.11	-.08	0.092
	A-P	.09	-.01	.03	.22	.22	.30	0.142
σ:	A	.22	.28	.12	.20	.10	.09	

(II) Verification of categorical forecasts

Table 6 summarises the categorical skill scores for the 88 forecasts made during Phase II (Feb-Aug 1982), and can be compared with Table 6 in the Report on Phase I (Apr-Sep. 1981). The 1981 and 1982 results are also presented graphically in Fig.3 (rainfall skill scores) and Fig. 4 (maximum temperature skill scores). Direct comparison between the two years is made difficult by the change in forecast periods, but there does not appear to be any definite increase in skill in Phase II over Phase I.

The variation in skill between districts follows a similar pattern in both years. There are a few irregularities, such as the high skill achieved for district E rainfall (districts are shown in Fig. 5). This dramatic 'improvement' in Phase II is probably fortuitous. Persistence forecasts score just as well here, due to the continuing drought conditions in the east of the South Island. Concentrating on the rainfall skill scores in the extended range forecast period (days 4-5, we find the scores vary from 0.04 to 0.26, with none more than two standard deviations above zero. In both years the skill is lowest for district C (East of North Island). Examining the forecast/verification contingency tables for bias, it is apparent that category 2 rain (3-15 mm) is overforecast for all North Island districts. This is due to underforecasting of category 1 (up to 3mm) in districts A and B, but to underforecasting of category 3 (over 15 mm) in district C. In the South Island, the main bias is in overforecasting category 3 rain for district F (south of South Island).

For maximum temperature, the predictions are better in the short-term (days 2-3) for eastern districts (C, E). In the extended range forecast period, forecasts are again quite successful for district E, but less so for district C. The worst scores are for district B, with no apparent skill at the extended range in either year. The only marked bias in the temperature predictions occurs for districts E and F where there is a distinct tendency towards overforecasting cold temperatures. The same bias was noted last year in the Report on Phase I.

The temperature skill scores show a number of irregularities such as skill at day 4-5 greater than skill at day 2-3, very high persistence skill, or large standard deviation in the actual scores. These irregularities are confined to districts A, B and D (northern and western districts) where the temperature seldom fluctuates far from normal, and so the great majority of days register as "category 2". Since the extreme categories are so rare, any incorrect forecast

Table 7: Comments on extended range forecasts

Office	Accuracy	Value to staff	Comparison*
Christchurch	Frequently impresssive	Excellent	Quite superior
Hokitika	Fairly good	Very useful	Better
Invercargill	Reasonable	-	Much better
Kaitaia	Reasonable	Adequate	-
Kelburn	Good	Useful	-
Nelson	Generally good	Inadequate	Worse
New Plymouth	Quite good	Very helpful	Better
Ohakea	-	Useful	-
Rotorua	About 75%	Adequate	Better
Wellington Airport	-	Useful	-
Whenuapai	About 80%	Much assistance	-
Wigram	Generally good	Most useful	-

* Comparison between this scheme and the one which it replaced early in 1982.

of extremes results in heavy penalties under the skill score rules. Clearly, it was not a good idea from the scoring point of view to have the same category definitions for all parts of the country. This choice was made for convenience, and because it was thought a $\pm 3^{o}C$ deviation from normal would be felt as a noticeable change by 'the man in the street'.

In eastern and southern districts the daily maxima are more evenly distributed across the three categories, and so the calculated skill scores are a more stable estimate of our skill at predicting deviations from normal. However even for these districts, a close examination of the daily variations shows we probably have greater skill at forecasting trends in temperature than absolute values, or deviations from normal.

(III) Evaluation of extended range forecasts and charts

(a) Extended range forecasts.

Offices which received the extended range forecasts issued each Monday, Wednesday and Friday were asked for their comments on the service. Their replies are given in Table 7.

Summarising these comments, it appears that:

- Accuracy is comparable with that achieved in shorter range forecasts;
- Staff generally find the forecast (very) useful; and
- The scheme is generally better than the one which it replaced.

Suggestions from the twelve offices for improving the scheme included requests that forecasts should be:

- Issued more frequently than Mondays, Wednesdays and Fridays (6 offices), though two offices thought that this should apply over weekends only;
- Issued earlier in the day (4 offices); and
- More detailed (3 offices).

Table 8: Comments on extended range forecast charts.

Office	Helps interpretation of extended range forecast	Helps with answering public enquiries	Suggestions for improving the service
Auckland Airport	Considerable	Very helpful	Delete fronts
Auckland City	Yes	Very helpful	Issue ECMWF charts too
Christchurch	Most valuable	Most valuable	Extend period to 7 days
Kelburn	To a certain degree	Yes	Include T+72
Ohakea	-	Extremely useful	Depict isobars on T+96
Whenuapai	Yes	Much assistance	-
Wigram	Quite satisfactory	Most helpful	-

(b) Extended range forecast charts

Table 8 summarises comments from offices which received prognostic MSL charts for T+48 and T+96 (flow pattern only).

These comments indicate that, in general, the charts are:

- Valuable in helping staff interpret extended range forecasts; and

- Most helpful to staff when they are answering enquiries from the public.

(c) Extended range forecasts issued to other organisations

On an experimental basis, extended range forecasts were issued to two organisations:

(i) The Electricity Division, for use in forward planning of transmission line construction in Otago and Southland; and

(ii) The Agricultural Pests Destruction Council, for use in planning operations associated with combating the spread of rabbits in the Alexandra area.

The Electricity Division estimates "Forecasts are about 75% correct and the outlook is 50%. .....We would appreciate your service being continued."

The Alexandra Pest Destruction Board reported back to the Council, and stated, in part, "This Board found the five-day forecasts .... most beneficial. They had an accuracy rate of approximastely 90 percent."

Also "The only improvement that would be of benefit to us is wind direction and an estimate of velocity."

SECTION III : Conclusions

Analysis of the 88 forecasts prepared during Phase II of the project leads to the following conclusions.

(I) Accuracy of flow pattern forecasts

- Accuracy was significantly greater than during Phase I. This is believed to be due largely to the availability of prognostic charts received from the ECMWF. This increase in accuracy was achieved at the same time that the forecast periods were extended up to one day further into the future than in Phase I.

- Accuracy varied from one month to another, probably being related to the type of synoptic situation.

- In Phase II low index situations were forecast more poorly than in Phase I. This probably reflects the increased dependence in Phase II on the ECMWF prognoses, which are known to have a bias when stationary (long) waves are prominent. The forecast error caused by this bias probably has most effect in low index situations.

(II) Accuracy of categorical forecasts

- Because forecast skill scores are based on rather small samples, it cannot be claimed that extended range forecast skill is statistically significant (two standard deviations or more above zero). However, there is a considerable consistency from one year to the next, particularly in the pattern of variation between districts, and this gives some confidence that the skill scores are meaningful. Therefore some skill is claimed for extended range forecasts for all districts except the east of the North Island in the case of rainfall, and for all districts except the west of the North Island in the case of maximum temperatures.

- The use of ECMWF prognoses in Phase II did not produce an overall improvement in the extended range categorical forecasts. However, it is possible the increased skill shown for districts in the east and south of the South Island (in both rainfall and maximum temperature predictions) is due to the better guidance (from the ECMWF prognoses) for the area south of New Zealand.

- Forecast bias is present for a number of districts, the worst excesses being:

(a) Underestimating the incidence of category 1 rain (3mm or less) in the north and west of the North Island, and of category 3 rain (more than 15 mm) in the east of the North Island; and

(b) Overestimating the incidence of category 3 rain in the south of the South Island, and of category 1 maximum temperature (cold conditions) in the east and south of the South Island.

(III) User reaction to extended range forecasts

Extended range forecasts were issued to a limited number of users. Their reaction to these indicates that

- Extended range forecasts of this type should continue to be issued. This includes extended range charts;

- Requests for more frequent and earlier (in the day) issues will need to be considered; and

- The response from two organisations outside the Meteorological Service indicated that consideration should be given to issuing extended range forecasts to the public*.

---

* From February 1984 extended range forecasts have been issued to the public on a routine basis.

References

Mullan, A.B., Neale A.A. and Revell, C.G., 1983 . Report on the Extended Range Weather Forecast Project - Phase I. N.Z. Met. S. Tech. Inf. Circ. No.189.

Neale, A.A., 1982. Subjective evaluation of MSL prognoses. Research Reports Nov. 1981-Feb. 1982, internal N.Z. Met. S. Report, 69-72.