MET O 11 TECHNICAL NOTE NO.166.

Subjective assessments of medium-range numerical forecasts produced by ECMWF - a final report.

by

A.G.Higgins.

Met O 11 (Forecasting Research)
Meteorological Office
London Road
Bracknell
Berkshire
UK.

February 1983.

1. INTRODUCTION

Since September 1979, the Forecasting Research Branch of the Met Office has carried out a subjective assessment of the medium range numerical forecasts produced by ECMWF. Originally, these products were compared with those produced using the Met Office 10-level model, Octagon. After a few months of operational working, a preliminary note by Findlater (1980) suggested there was little significant difference in the number of occasions when each model gave reliable guidance. This view was supported by Jones and Findlater (1980) some months later. Dutton and Hall (1980) noted an improvement in ECMWF performance relative to the Octagon in the period 3 March 1980 to 25 August 1980 compared with the six-month period 3 September 1979 to 25 February 1980. They were uncertain to what extent the improvement may have been due to seasonal influences. Hall (1981) concluded that the improvement noted earlier was maintained throughout the period March 1980 to March 1981, eliminating the doubts concerning seasonal influences. Indeed, marks for the ECMWF model were markedly superior to those of the Octagon during days 3 to 5. The improvement was due to modifications made to the ECMWF model during the period under review; the Octagon on the other hand remained unchanged since development work within the Met Office was concentrated on the new 15-level model.

Continued assessment since March 1981 has led to the indisputable conclusion that on average, the ECMWF forecast product was superior to the Octagon. This fact may be influenced to a small extent by a much later cut-off of data for the ECMWF model (T + 11) compared with the Octagon (T + 3). However, recent experiments at ECMWF (1982) suggest that cut-off time makes little difference to the overall quality of their products.

The next phase of the assessment programme was that of accumulating grid point values of zonal and meridional Indices, 850 mb temperatures and 1000 mb height for the ECMWF model. It had been intended to extend this work to the Octagon. Later however, it was felt that the programming effort was not worthwhile since the model was soon to be replaced. Nevertheless, an analysis of the data collected for the ECMWF model was carried out by Ireland (1981) and Ireland (1982) for the two periods Winter 1980-81 and Summer 1981. He made various comments regarding apparent characteristic errors in the ECMWF model.

The purpose of this report is to complete the assessment of ECMWF products as far as possible up to September 1982. In addition, some remarks are made about the relative performance of the ECMWF model and the new 15-level operational model. It is hoped to show that while some points raised by the earlier reports have been resolved, others still remain. Much of this work has been taken from an analysis of data collected in survey form from the weekly assessment carried out in Met O 11 over the period September 1981 to September 1982.

Some of the points investigated were specifically raised by Ireland; hence the attempt at clarifying certain issues regarding the current characteristics of ECMWF model.

In addition, comparisons between the models are given in Appendix 1, with some data available for the new 15-level model as well as the Octagon. The period involved for these most recent comparisons follows the introduction of a revised marking scheme with effect from late September 1981. Only Days one to four were marked according to the original strict scale; Days 4/5 and Days 5/6/7 were considered as two outlook sequences and were marked with more leniency regarding timing, position etc.

2. ECMWF MODEL CHARACTERISTICS

Points to be considered are as follows:-

i) Representation of the deepening of lows over the E Seaboard/W Atlantic, particularly with respect to timing.

ii) Representation of substantial development occurring at or after D+4.

iii) Handling of the mature depression or depression life cycle.

iv) Depression tracks in association with the jet axis.

v) Secondary low development over Europe and the Mediterranean.

vi) Changes of type/blocking patterns.

vii) The handling of trough disruption and cut-off lows.

viii) Hurricanes Tropical Storms.

i) Deepening of Atlantic Lows

Ireland noted a tendency for Atlantic lows to be deepened too late or too little over the W Atlantic after two days into the forecast period; they were subsequently deepened too much over the E Atlantic. This comment was based on analyses of mean errors in the 1000 mb height field over the Atlantic.

The survey results showed 36 occasions of one or more Atlantic lows developing or deepening during the forecast period. On about half these occasions, the model deepening was assessed as too little in the W Atlantic; on a further third of occasions, deepening was too late. On a small number of occasions, the error combined both these failings.

Considerably fewer depressions were assessed as deepening too much in the E Atlantic to compensate. While Ireland's comment regarding errors in the W Atlantic may be correct, the mean height error in the E Atlantic does not seem to be due solely to overdeepening in that region. It seems more likely that the life cycle is in error and that depressions once developed are slow to fill thus causing the same height error. This is dealt with more fully in section iii).

ii) Development in relation to time into Forecast

In broad terms, cyclonic developments in the Atlantic are reasonably handled by the model with regard to intensity and timing in the early stages of a forecast period. However, odd occasions do occur when a developing wave depression is poorly represented in the observational data. This inevitably results in a weak system being predicted and consequent errors ensue. There is still evidence to suggest that developments occurring at or after D+4 in a forecast will be poorly modelled with respect to intensity, timing or position. Indeed, several occasions have been noted of systems being all but completely missed.

There are a few occasions when second generation development is well modelled after D+4 but this is not the norm. The synoptic details of a forecast thus have a marked drop-off in usefulness after D+4.

iii) Representation of the Mature Depression/Life Cycle

The question of deepening of Atlantic systems has been discussed briefly in section i). In addition to information regarding the amount and timing of developments, a careful watch has been made of depression track. The most significant error was found to be that a rather large number of systems were tracked too far south. (9 out of 36).

This may be associated with either the lack of deepening or the lateness of the model development. In fact, occurrence of incorrect S'ly track was always associated with one or other of these errors, the split being half and half between them.

It would seem therefore that the model life-cycle for Atlantic depressions does not yet accurately represent the real cycle, systems in the model maturing at a different rate from reality. These systems are then erroneously steered too far south with insufficient distortion of the upper flow. The net result of this error is a failure of the systems to turn to the left during the maturing stages of their life cycle. Notes made regarding the steering flow of individual systems suggests that on some occasions the direction of the flow was incorrect while on others, the direction was correct but the intensity was weak. Either way, the same result was observed. An extension of the poor life-cycle theme may also provide the answer to Ireland's findings regarding mean 1000 mb height errors in the E Atlantic/Europe. Investigation into the behaviour of lows in these areas suggests not so much that they are overdeepened but that they are too slow to fill.

Of the 40 occasions considered, 20 showed that the model maintained lows too deep for too long ie they failed to fill. Occasions of overdeepening in the E Atlantic/Europe were found to be 13 occasions out of 41 considered and of these, 10 were subsequently maintained too deep.

It seems likely that errors in the representation of the occluding/maturing stage of depressions are a serious problem particularly when combined with the smaller scale errors which are increasingly evident with time into the forecast period.

iv) Depression track in relation to the Jet axis

Another part of the survey of data was concerned with characteristics of the mark 500 mb jet, Ireland again having raised points as a result of his analyses of zonal and meridional indices.

A subjective measure of small changes in direction and speed of flow is difficult and comments were only made when a large error was evident. This was usually associated with a specific synoptic error and could not be regarded as a model characteristic. The suggestion that the mean jet axis is displaced south was specifically investigated. It was found that on 23 occasions out of 45 considered, the jet was too far south over at least some position of its length and not correspondingly compensated for (within the area of assessment at least) by errors to the north. On eight occasions, the error was firmly identified as being in the base of a trough or troughs ie the jet had penetrated or plunged too far south in the base of the trough.

So while it may appear that the mean position of the jet axis is too far south, this tends to mask the real problem that while the representation of the flow around ridges is consistently good, that around troughs is consistently too far south. This in turn is consistent with problems regarding depression tracks. As shown in the previous section, a significant number of systems follow incorrect tracks to the south of the actual track; here we have identified a possible contributory cause, the tendency to dig troughs too far south.

Since its inception, the ECMWF model has shown a distinct preference for meridionality. This feature regarding troughs is simply a facet of the overall model characteristic for meridionality. However it is also felt that there may be some coupling with the problems described regarding depression life-cycles. With a more realistic representation of the life-cycle of depressions, the inter-related problems may well disappear.

It has been noted that occasions occur when the strongest flow around a trough has been held to the rear of the trough axis. While in reality, the strong flow has propagated around the base of the trough and on to the forward side producing a marked turning mechanism for the associated depression, the ECMWF model has held the flow to the rear with considerably less flow than reality on the forward side. The turning mechanism is not present and the associated depression fails to turn left, maintaining its more easterly track; it also fails to fill in the correct manner. While that may be a gross oversimplification of the situation, the evidence certainly suggests that such an explanation is reasonable. An example of this behaviour is shown in Appendix 2, a forecast run from DO 5 October 1981.

v) Secondary development over Europe

Another reason for large errors in mean 1000 mb heights over Europe suggested by Ireland was the possibility of spurious lows forming over Europe and the Mediterranean. These features were checked in the survey, together with a tendency to overdeepen correctly positioned secondaries. Out of 40 occasions considered, seven were found where spurious lows formed over Europe at some stage in the forecast period. Another nine showed secondaries being too deep though correctly positioned. There was only one occasion of a spurious low in the Mediterranean but another seven when Mediterranean lows were too deep, out of a possible 34 occasions.

Once again, the main error of the model does not appear to be specific overdeepening at the development stage of a (secondary) depression's life-cycle but in its failure to fill systems quickly enough in maturity. The formation of spurious lows may also be considered a problem. The vigour in the model, which produces realistic developments from accurate data, produces spurious developments in the face of small scale errors some time into a forecast period. This kind of error rarely occurs within the first two days of a forecast.

vi) Changes of Type/Blocking Patterns

Correct predictions of a change of type are an important aspect of medium range forecasting. Earlier evidence suggested that no particular model was best at predicting all changes and although the detail was not always correct, useful guidance was frequently given regarding the likelihood of a change of type. The purpose of this aspect of the survey was to ascertain whether or not any improvement had taken place regarding the accuracy of predicted changes of type over Europe. A good forecast was required to produce a reasonable prediction of the change, correct timing and a good representation of the new weather type. Bad forecasts were those failing on one or more of these counts. Of 37 occasions considered, involving changes over UK and Europe, ECMWF gave good advice on 19 occasions. The model was prone to predict changes which did not occur or to overdo changes, sometimes with large scale errors ensuing.

An analysis of data concerned with blocking patterns was difficult, the description of errors being very varied. However, on occasions when a blocking pattern was maintained for a long period of a forecast, the model generally maintained the feature, though not necessarily with correct detail.

The building of blocks was rather poor, there being only four occasions out of 15 assessed as good. The most significant errors were the building of erroneous blocks on three occasions and the non-formation of blocks on three other occasions.

In contrast, the breakdown of blocking patterns was well modelled, with 11 successes out of 15. This combination of results tends to support those regarding changes of types. The model may not predict a change because no block was formed or it may predict an erroneous change associated with an erroneous block.

Thus changes of type from mobility to a blocking situation should perhaps be treated with some caution while a much greater confidence could be given to changes involving the breakdown of a block resulting in a return to mobility.

An example of an incorrect change of type is shown in Appendix 3, a forecast from DO 19 April 1982. Here the model predicted a breakdown of an anticyclonic blocking situation over UK, replacing it with a showery colder NNW'ly regime. This error may be associated with the model tendency to favour meridional developments, built out of over-extended troughs. The model failed to predict the rebuilding of the UK blocking high at D+4; the high persisted for a further week before the change to NNW'lies actually occurred.

This situation is one of the exceptions to the rule suggested above regarding confidence in the breakdown of blocks.

vii) <u>Trough Disruption and Cut-offs</u>

The performance of the model in handling trough disruption, and the cutting-off process, was mixed. While capable of accurately modelling both processes for some troughs, the model made significant errors in one or other process on more occasions than not. The ECMWF model was assessed as producing a good handling of either one or both processes on 19 out of 50 occasions. On nine occasions either the disruption itself or the subsequent cut-off was missed.

There was little evidence of consistency in timing. Earlier, it had been suggested that model predictions of disruption were slow; recent occasions of specific timing errors were split three fast, four slow. Many other errors of position were identified as were occasions when an incorrect portion of a trough disrupted with ensuing errors in evolution.

What can be said in summary is that the model usually gives adequate guidance that a given trough will disrupt but that the detail of the process cannot be taken as completely accurate.

viii) <u>Hurricanes/Tropical Storms</u>

Limitations of the model to predict the development and track of hurricanes or old tropical storms still remain. If such a system is present in the initial data, the model will make a reasonable attempt at prediction.

If however the feature moves into the N Atlantic some time later in the forecast period, the system is likely to be missed completely. The consequences of these poor predictions and their interaction with Atlantic depressions are therefore still a problem.

3. SUMMARY

Since September 1979 when the assessment of ECMWF medium range products began, a considerable improvement in the overall quality of forecast products has been observed. From as early as 1981, it was noted that ECMWF performance was superior to that of the Octagon. Continued improvements to the model formulation etc have left no doubt that reliable forecast products are produced on average out to D+4.

Despite model improvements, characteristic errors are still present. These are not necessarily peculiar to the ECMWF model alone but may well be characteristic of all numerical models of similar grid length, formulation etc. These errors are:-

- i) A reluctance to develop depressions in the W Atlantic sufficiently, with respect to depth and timing of development.
- ii) Poor representation of significant developments after D+4.
- iii) Inaccurate representation of the life-cycle of depressions; notably in errors in track, failure to turn and fill depressions quickly enough.
- iv) In some way coupled with iii) above, the mean jet axis tends to be predicted too far south due to over extension of troughs.
- v) Similar problems to those of iii) in relation to European secondaries.
- vi) A tendency to miss changes of type involving the building of a blocking situation.
- vii) Detail of trough disruption and the subsequent handling of cut-off lows rather poor.
- viii) Inadequate representation of small scale systems.

While the detail mentioned in some of these characteristics is not in itself always vital to the correct interpretation of a particular forecast chart, it becomes increasingly important for the continued correct development within the model over the remaining period of the forecast.

Appendix 1. Model comparison graphs/tables.

2. Example of incorrect track etc. 5 October 1981.

3. Example of incorrect change of type. 19 April 1982.

REFERENCES.

Dutton, M.J.O. and Hall, B.A. 1980. Subjective assessments of medium-range forecasts produced by ECMWF and Met. Office operational numerical models. Met O 11 Tech Note No. 146.

Findlater, J. 1980 A preliminary report on the subjective comparison of medium-range forecasts produced by ECMWF and the Met. Office. Met O 11 Tech Note No. 137.

Hall, B.A. 1981 Subjective assessments of medium-range numerical forecasts produced by ECMWF and the Met. Office. Met O 11 Tech Note No. 147.

Ireland, D. 1981 Some error statistics for ECMWF forecasts up to 7 days ahead (500mb zonal and meridional indices, 850mb temperature and 1000mb height) - Part I : Winter 1980 - 1981. Met O 11 Tech Note No. 151.

Ireland, D. 1982 Some error statistics for ECMWF forecasts up to 7 days ahead (500mb zonal and meridional indices, 850mb temperature and 1000mb height) - Part II : Summer 1981. Met O 11 Tech Note No. 153.

Jones, D.E. and Findlater, J. 1980 Subjective comparisons of medium-range forecasts produced by ECMWF and the UK Met Office. Meeting on synoptic quality of ECMWF forecasts, ECMWF, June 1980.

Woods, J.A. (ECMWF) 1982 Experimental forecasts from 00z data with a 3-hour data cut-off compared with operational ECMWF forecasts. Operations Department Technical Memorandum No. 57, ECMWF.

Appendix 1.

Over the period under consideration, there were 46 assessment runs allowing complete comparison between ECMWF and the OCTAGON. There were 36 runs comparing ECMWF, the OCTAGON and the New 15-Level Model. The latter was the version run on the Cray computer at ECMWF from their analysis; various changes were made to this model during the period, particularly during the early part of the comparison.

Results of the assessments are given below. These are based on the scoring system

D1-4	A+ = 4	D4/5	A = 3
	A- = 3	D5/6/7	B = 1
	B+ = 2		C = -1
	B- = 1		
	C+ = -1		
	C- = -2		

46 Runs	ECMWF	OCTAGON	Maximum
D1-4	1445	843	2944
D4/5	42	-57	552
D5/6/7	-70	-90	552

36 Runs	ECMWF	OCTAGON	NEW	Maximum
D1-4	1064	619	954	2304
D4/5	18	-45	-10	432

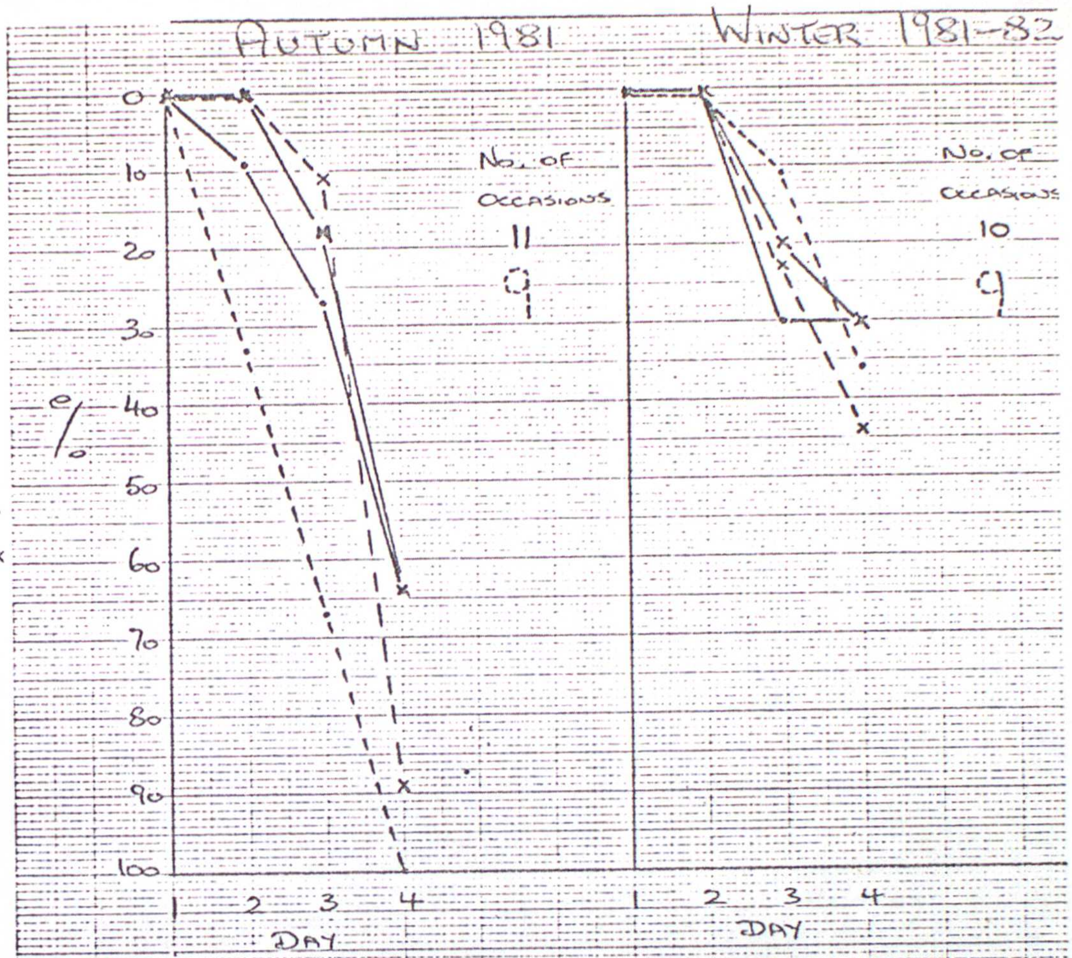
Graphs on the following pages show comparative performance of the models on a seasonal basis. On the whole, ECMWF showed some superiority to the New (15 level) model while both showed considerable superiority to the Octagon, particularly during the Winter period.

Percentage of Occasions of C Markings.

ie Forecast misleading in some important aspect.

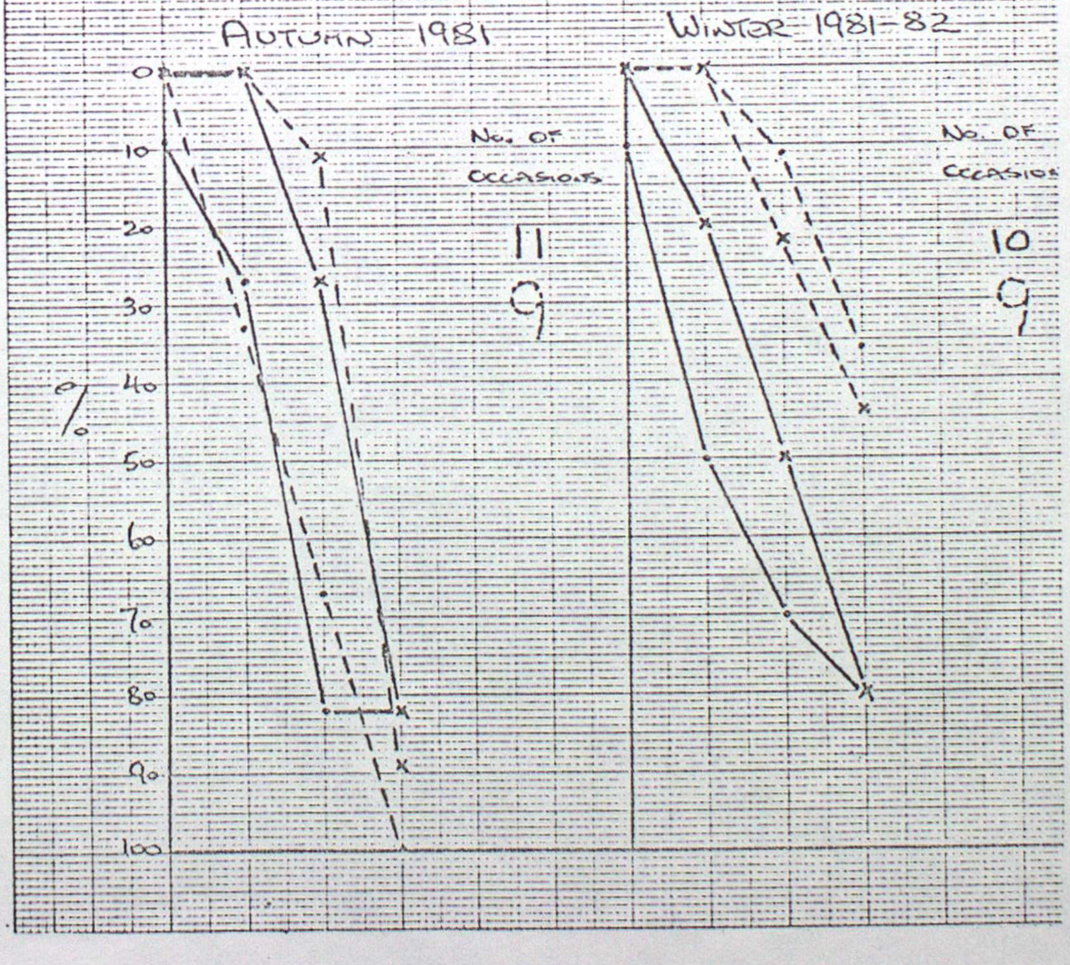
ECMWF
Surface •—•
500mb x—x

NEW
Surface •----
500mb x---x



OCTAGON
Surface •—•
500mb x—x

NEW
Surface •---•
500mb x---x



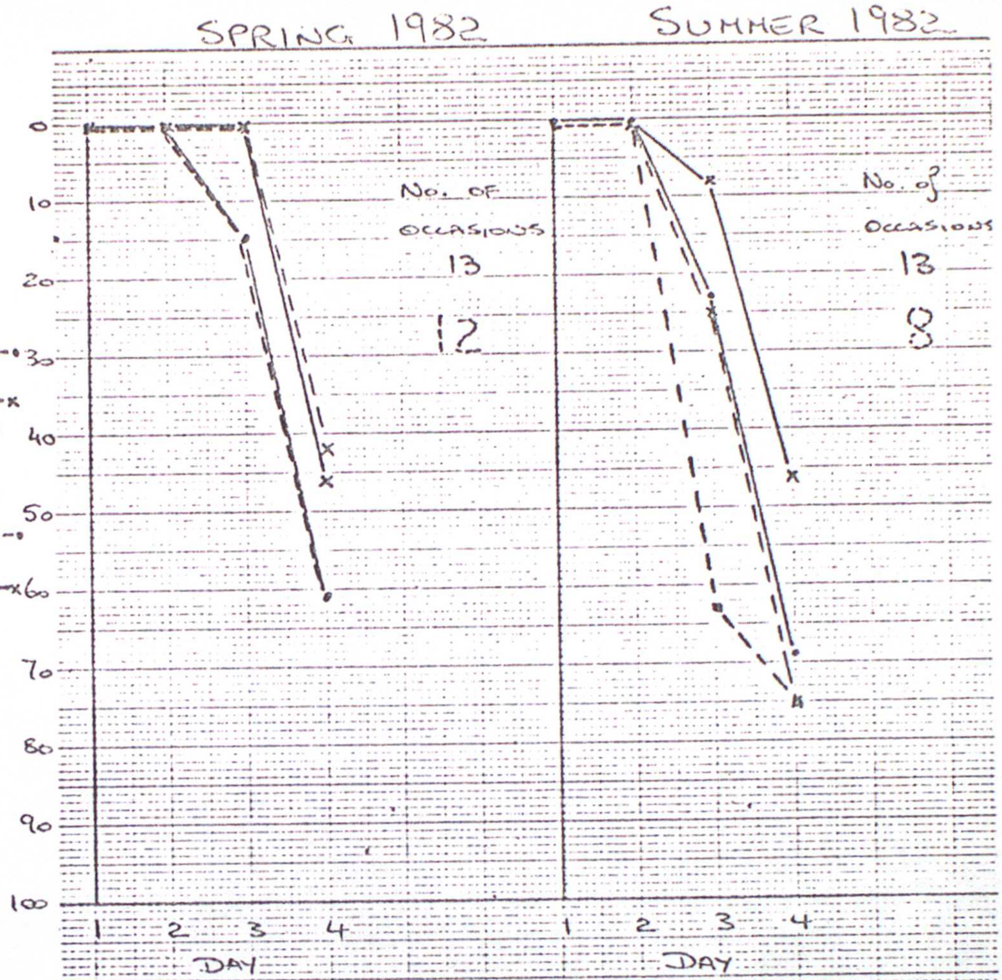
Percentage of Occasions of C Markings.

ie Forecast misleading in some important aspect.

ECMWF
Surface •——•
500mb x——x

NEW
Surface •- - -•
500mb x- - -x

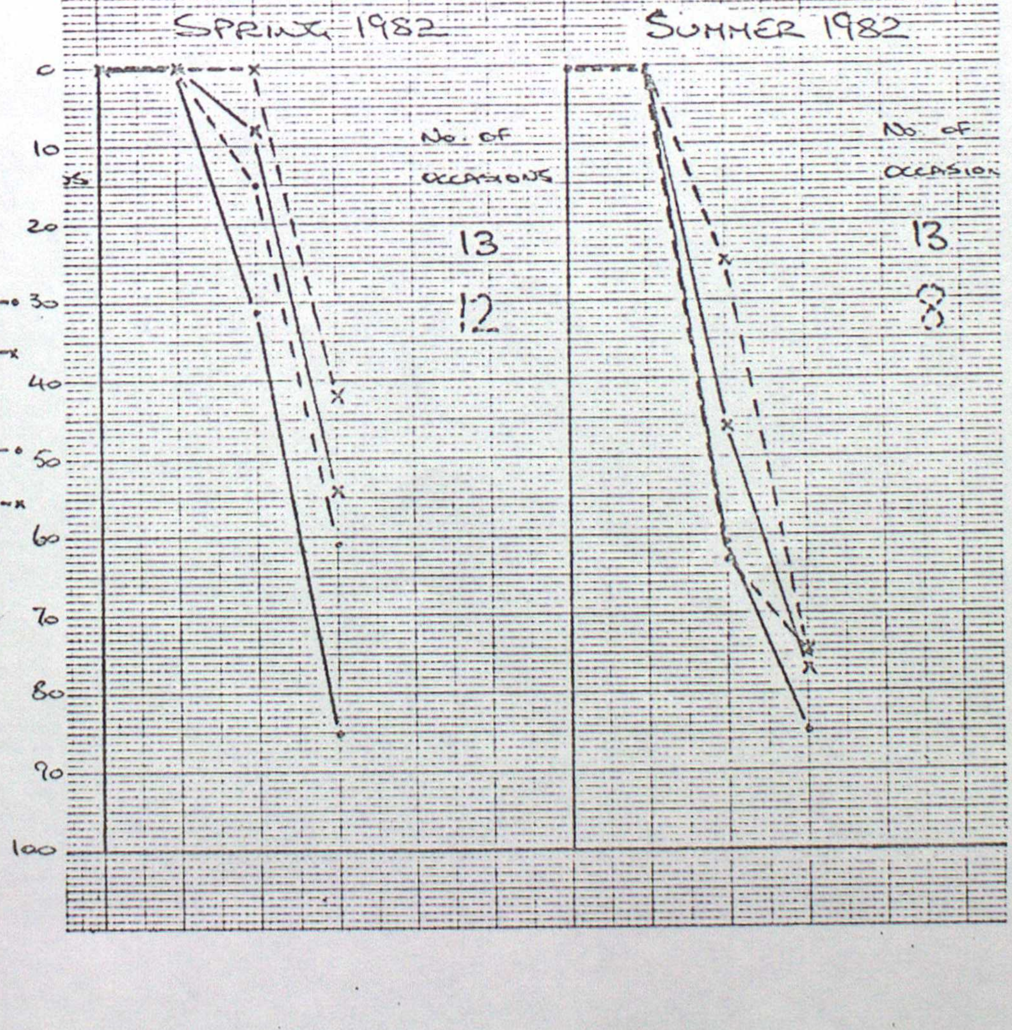
%



OCTAGON
Surface •——•
500mb x——x

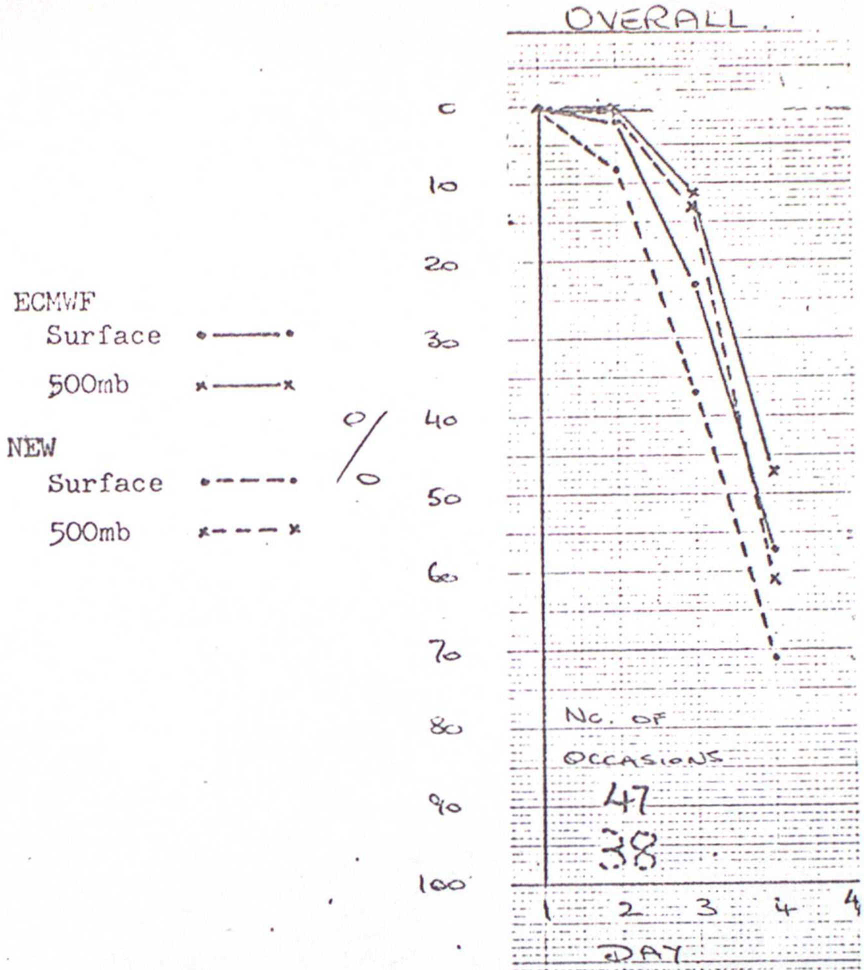
NEW
Surface •- - -•
500mb x- - -x

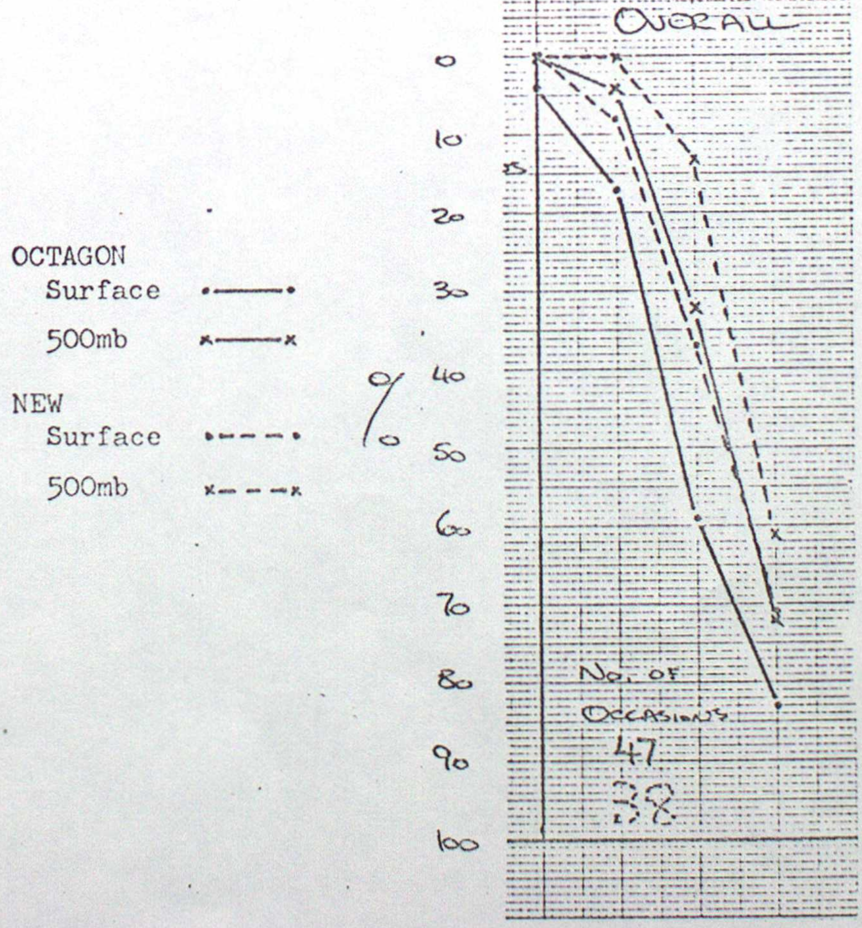
%



Percentage of Occasions of C Markings.

ie Forecast misleading in some important aspect.

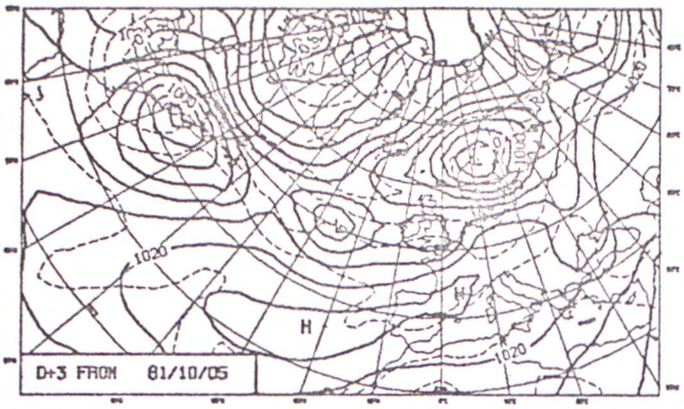


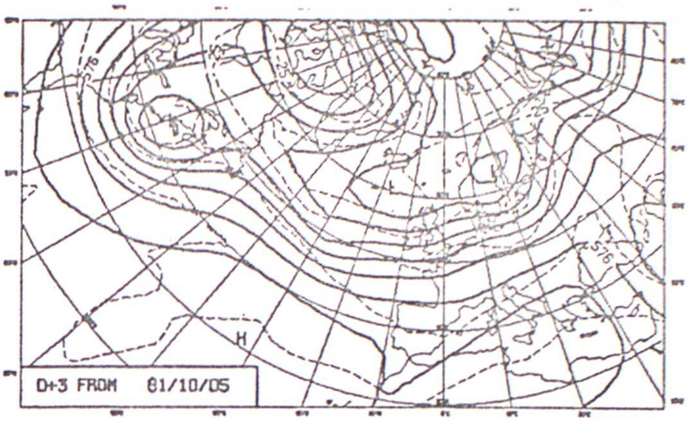


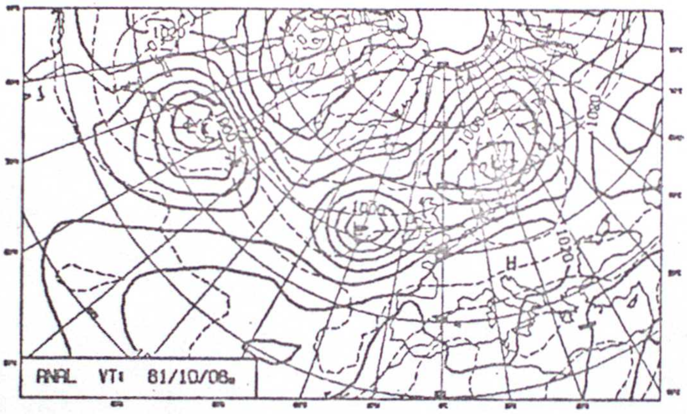
APPENDIX 2.

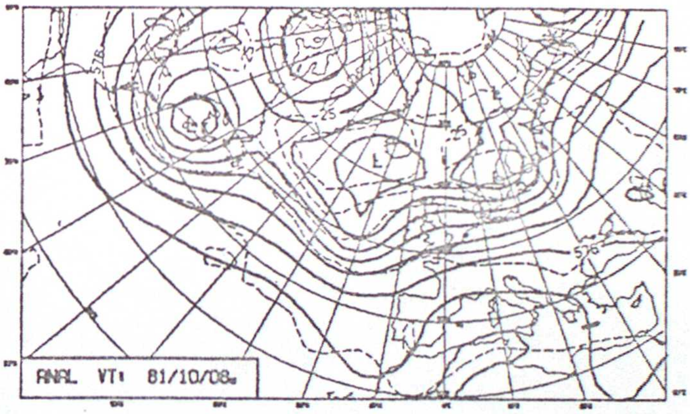
SURFACE (MSL) PRESSURE
(Solid lines, 5mb intervals.)
850MB TEMPERATURES.
(Pecked lines, 5K intervals.)

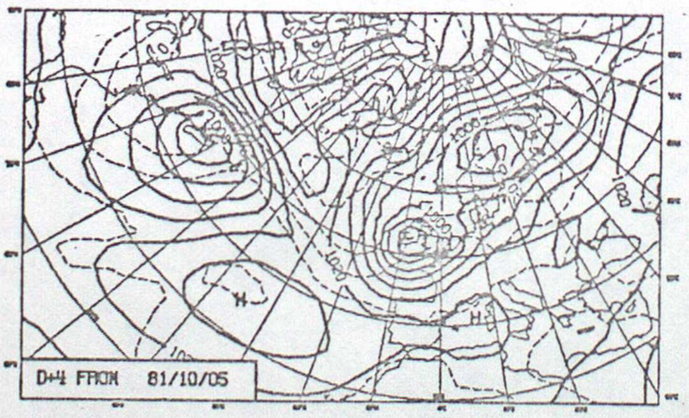
500MB HEIGHTS.
(Solid lines, 8dam intervals.)
500MB TEMPERATURES.
(Pecked lines, 5K intervals.)

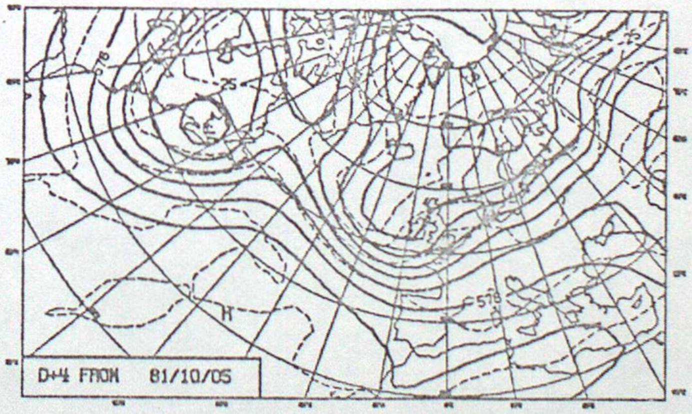


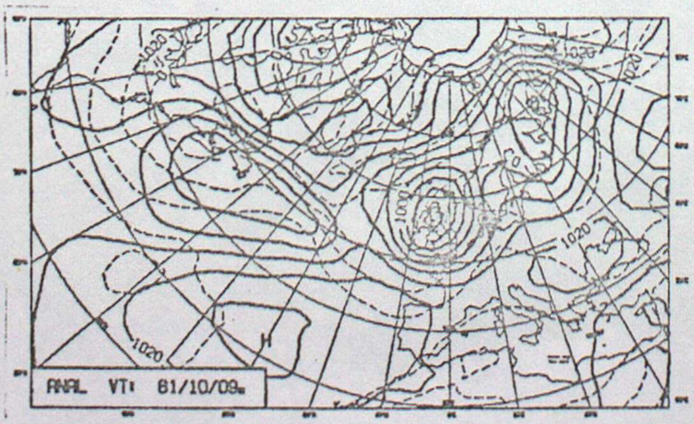


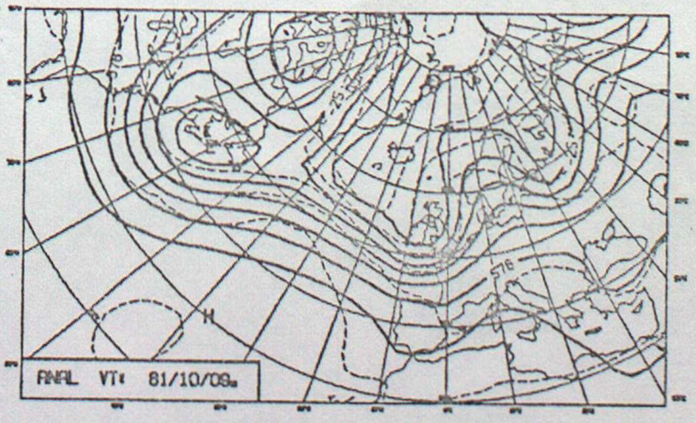


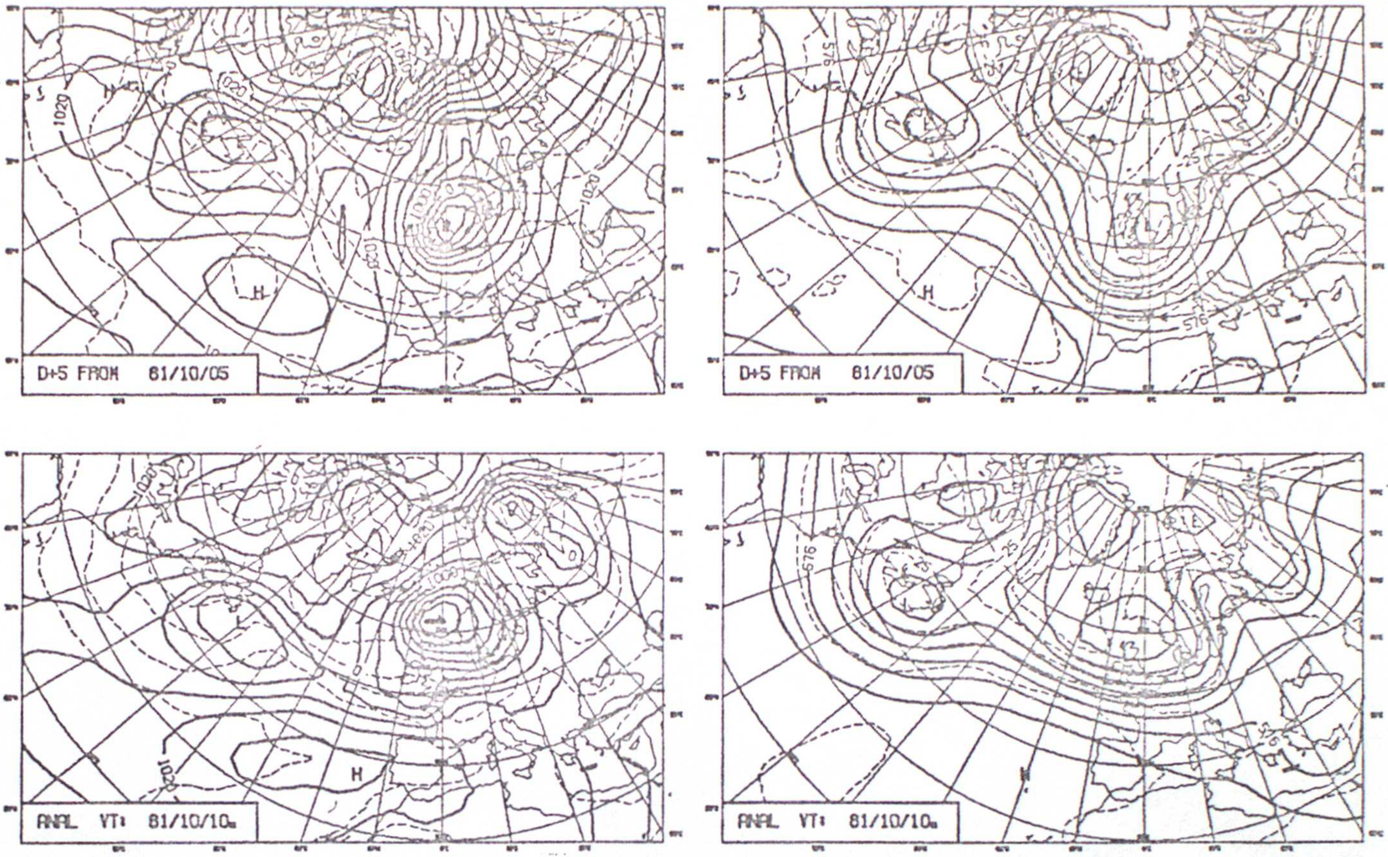












Points to note are:

i. The rather weak representation of the Atlantic wave (ANAL 81/10/08 at 50N 20W) by the D+3 forecast: also the positional error of the same feature at this stage.

ii. There is an overdeepened wave depression over UK with a closed circulation rather than the very minor wave suggested by the ANAL 850 mb temperature fields.

iii. The subsequent lack of turning at D+4, despite the well represented deepening between D+3 and D+4.

iv. The weak 500 mb trough at D+3 around 30W on the forecast cf ANAL 81/10/08 at around 20W.

v. At D+4 the main strength of flow in the 500 mb trough is the WNW flow to the rear whereas on the ANAL 81/10/09, the stronger flow is already into the base of the trough and on the forward side, building the upper ridge over Scandinavia.

vi. The net result - at D+5, a depression much too far south with an associated upper trough erroneously extended into Iberia rather than running East into Europe.

vii. The error in terms of weather for the UK (Europe) is large, the prediction being for a cold showery regime over the whole country rather than being confined to the North of UK with further shallow waves affecting the south.

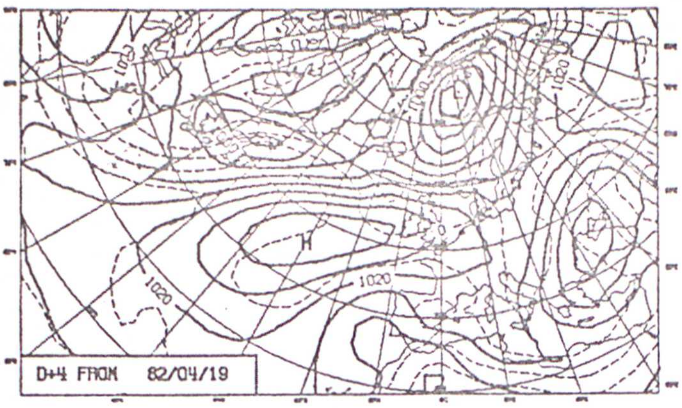
APPENDIX 3.

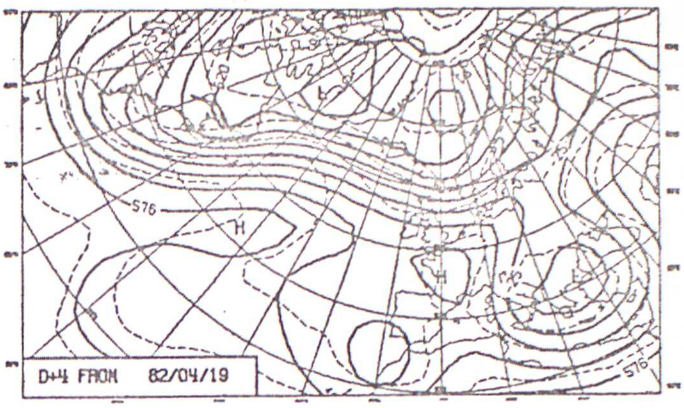
SURFACE (MSL) PRESSURE.
(Solid lines, 5mb intervals.)

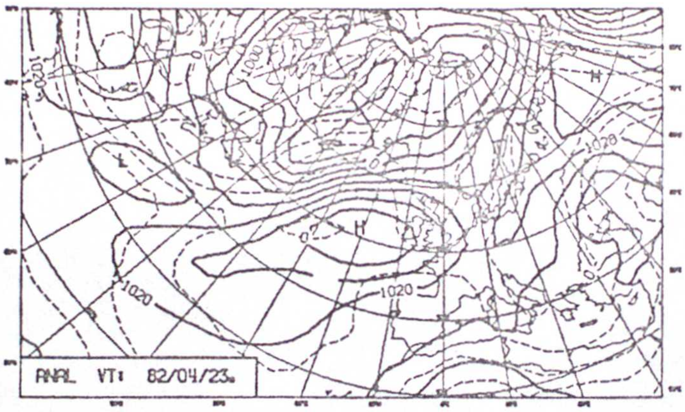
850MB TEMPERATURES.
(Pecked lines, 5K intervals.)

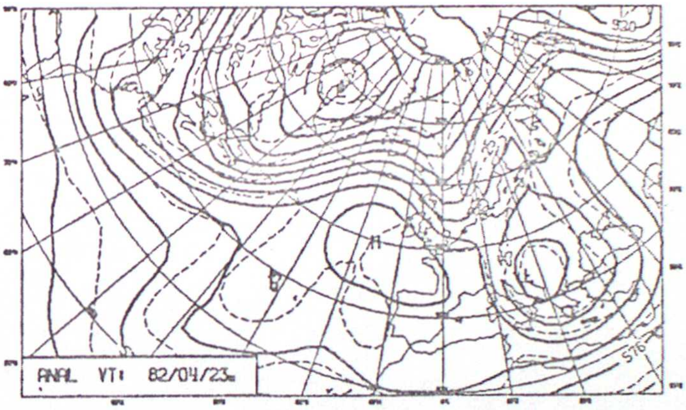
500MB HEIGHTS.
(Solid lines, 8dam intervals.)

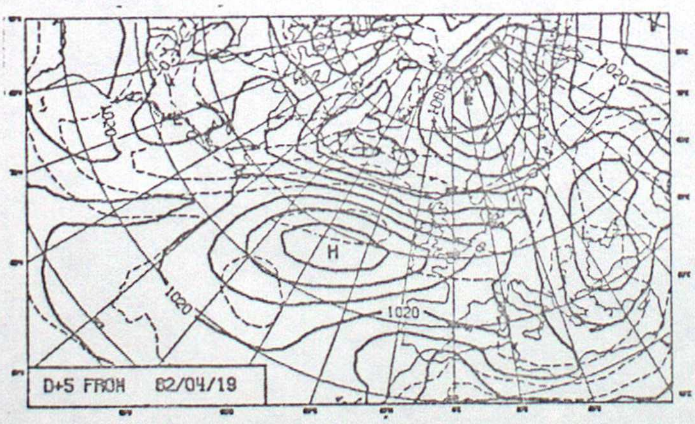
500MB TEMPERATURES.
(Pecked lines, 5K intervals.)

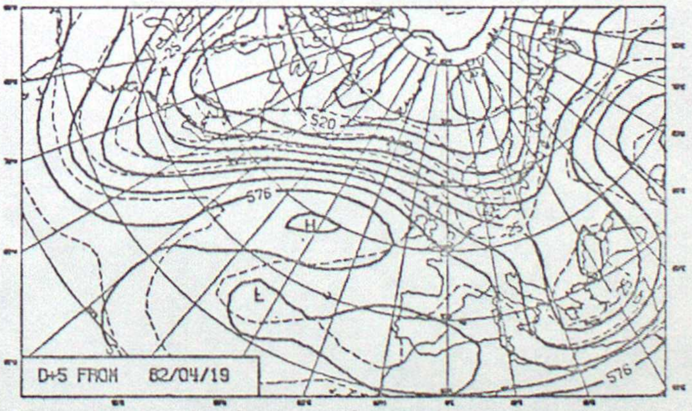


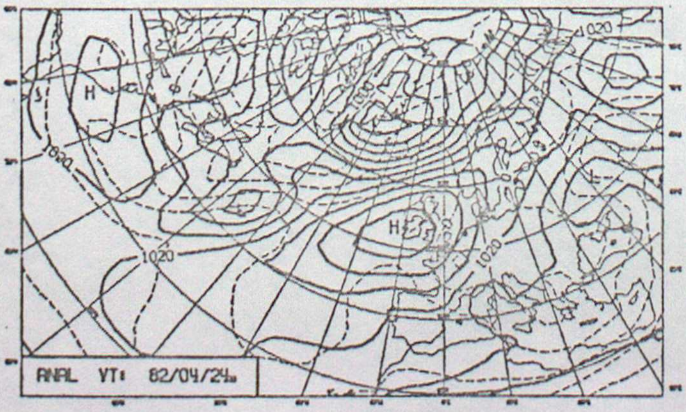


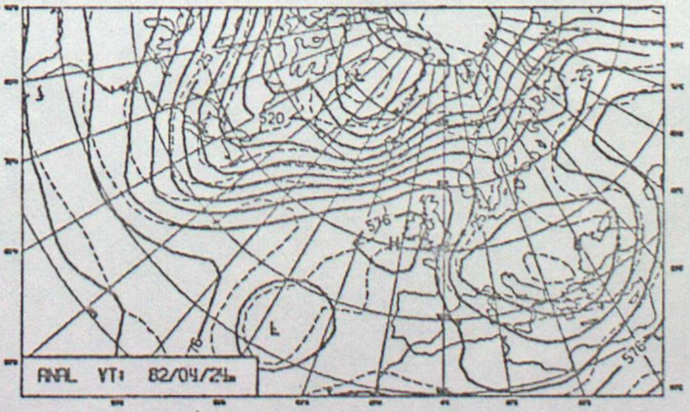


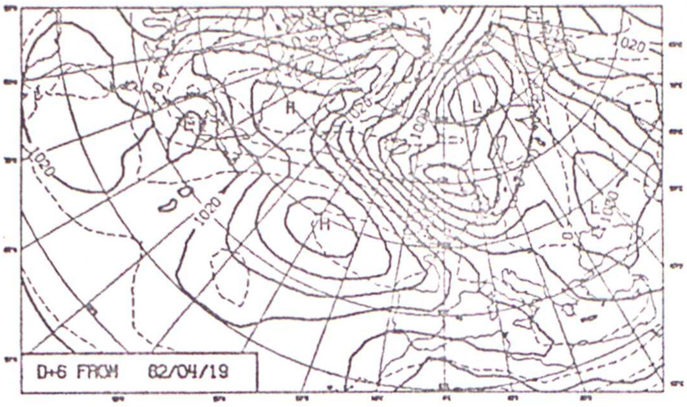


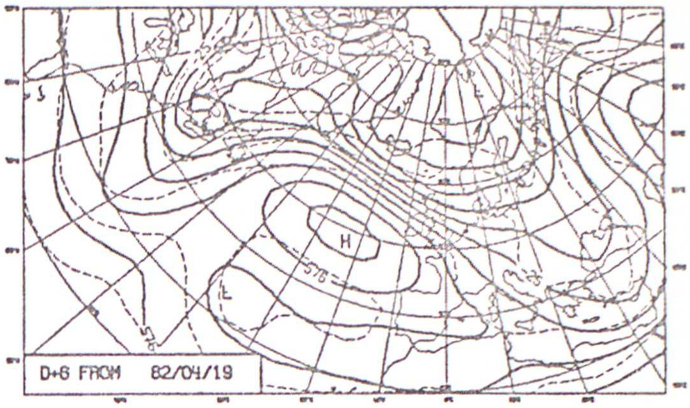


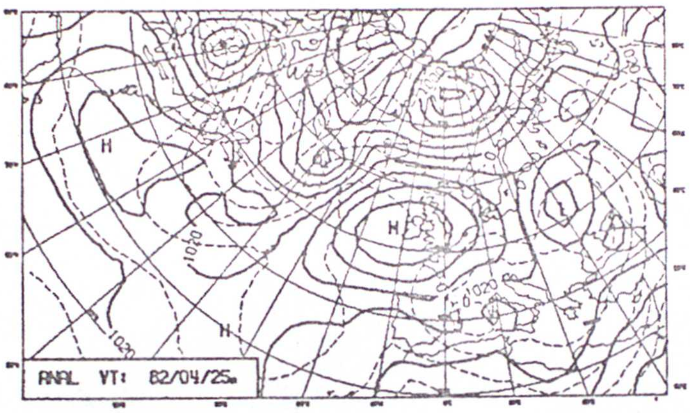


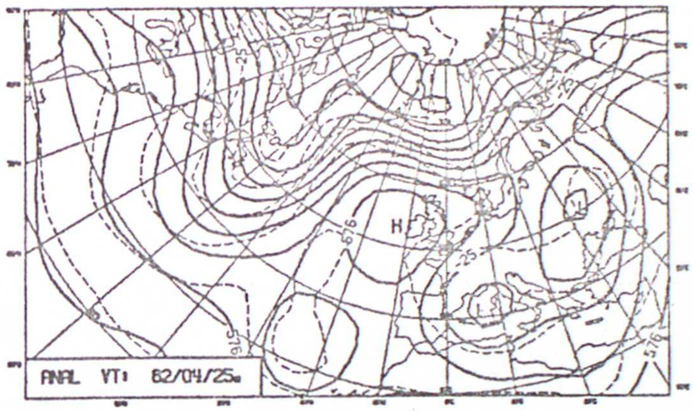












Points to note are:

i. At D+4, the low S of Spitzbergen is much too large and deep; it has not filled correctly and acts as a block to the progress of

ii. the low ANAL (82/04/23) near the southern tip of Greenland. In the forecast, this feature is slow and weak.

iii. At 500 mb D+4, the trough on the Greenwich meridian is too broad with a closed circulation erroneously maintained.

There is also a lack of ridging over Greenland, the blocking high being split.

iv. At D+5, the Spitzbergen low still persists. The Greenland low is still weak, slow and heading in the wrong direction in association with the erroneous 500 mb flow.

v. By D+5, the blocking high has retrogressed to mid-Atlantic.

vi. At D+6, the original Spitzbergen low is still erroneously maintained in the forecast. The Greenland low is now moving ESE into the N Sea rather than NE to Spitzbergen.

vii. The predicted change of type for the UK and near continent is completely wrong; a major contributory factor would seem to be the effect of the erroneously maintained low near Spitzbergen.