Extended-range predictions with ECMWF models: Interannual variability in operational model integrations

By T. N. PALMER, Č. BRANKOVIĆ, F. MOLTENI and S. TIBALDI*

European Centre for Medium Range Weather Forecasts, Reading

(Received 2 June 1989, revised 24 January 1990)

SUMMARY

A set of 30-day integrations made with the operational ECMWF model over a period of three and a half years (from 1985 to 1988) is studied. The impact of model reformulations during this period on the climate drift of the model is assessed, and the level of extended-range forecast skill achieved by the operational model is studied.

Diagnostics of zonal average 30-day mean wind and temperature error show a systematic reduction in the extratropics over three years. It is argued that these improvements are consistent with the changes in model formulation. Eddy fluxes of heat and momentum, and levels of eddy kinetic energy are similarly improved. We also show that the ability to simulate blocking activity in the extended range has improved substantially over the three winters.

From maps showing the ratio of systematic to total error it is concluded that, by the final winter, the middle-latitude extended-range error is almost entirely associated with random errors. This is supported by the growth of the dispersion, or spread, between forecasts initialized 24 hours apart. In particular, for the extended winter 1987/88, it is shown that the asymptotic level of (internal) spread is comparable with the asymptotic level of (external) skill. By contrast, in the tropics, the total error is dominated by the systematic error for all years.

Despite improvements to the model, extended-range forecast skill is modest. Skill has been measured in a spatially-filtered three-dimensional phase space, spanned by rotated EOF (empirical orthogonal function) coefficients of 500 mb height, which contain the principal weather regimes in the atmosphere. On average the skill scores are consistent with those of other major centres studying dynamical extended-range prediction. In particular there is evidence of skill to day 15 and possibly to day 20 in middle latitudes. Beyond day 20 skill is marginal. We demonstrate that in the winter sample of operational-model forecasts, skill is correlated with the observed value of the first rotated EOF, which can be described as a global version of the Pacific/North American mode.

By contrast, the ability to predict monthly mean fields in the tropics depends primarily on a knowledge of the boundary forcing, including sea surface temperature. It is shown that some aspects of the interannual variations in the summer monsoon circulation can be predicted, in particular the tropical easterly jet anomalies, African rainfall, and, to a much lesser extent, Indian monsoon rainfall.

1. INTRODUCTION

Extended-range prediction of large-scale atmospheric flow beyond the limit of instantaneous deterministic predictability is a topic of considerable interest at present. A physical basis for the prediction of monthly means has been discussed by Shukla (1981), and an example of skilful dynamical prediction for 30 days has been presented by Miyakoda et al. (1983). Furthermore, extended-range integrations give valuable insight into the climate drift of a numerical weather prediction (NWP) model. A study of climate drift is important to evaluate changes in the operational model's formulation. For example, the amplitude of medium-range systematic errors in most state-of-the-art NWP models is comparable with the short-range errors of models used less than a decade ago. As a result of these developments, the European Centre for Medium Range Weather Forecasts (ECMWF) embarked on a programme of 30-day extended-range integrations. This is the first in a series of three papers which present results from this programme. The project is part of an international attempt to assess the level of extended-range skill achievable using current NWP models. Some early results from other centres are summarized in a recent workshop proceedings (ECMWF 1988), and we shall compare our results with some of those quoted in this reference.

* Present affiliation: Department of Physics, University of Bologna, Via Irnerio 46, 40126 Bologna, Italy.
The programme commenced in April 1985, and near the middle of every month the operational 10-day integration of the (T106) model was continued out to 30 days, from two consecutive 24-hour analyses. The data-set forms the basis of the study in this paper. From a deterministic point of view, these forecasts can be viewed as having been made, at the time of integration, with the most skilful NWP system available at ECMWF. Our principal objectives are: to discuss the possible impact of model reformulations during the period of the experimental programme on the asymptotic climate drift of the operational model, to assess the average level of extended-range forecast skill achieved by the operational model, and to study interannual variations in the skill of extended-range forecast fields in both the tropics and extratropics.

In addition, during the first year of the programme, each T106 integration was repeated at T63, T42 and T21 resolution. This multi-resolution data-set is used in the companion paper by Tibaldi et al. (1990), hereafter referred to as T, to assess the influence of horizontal model resolution on climate drift and extended-range skill. Finally, approximately every three months, the T63 version of the operational model was integrated for 30 days from nine consecutive 6-hour analyses. The purpose of this was to assess the potential for predicting forecast skill, and to develop techniques for probabilistic forecasting. These T63 forecasts comprised the time-lagged ensembles whose properties are described in Branković et al. (1990) hereafter referred to as B.

In section 2 we list the forecast initial dates, and describe changes to the operational model over the period of the experimental programme. Indeed, since the first year of extended-range integrations there have been some substantial reformulations of the model, and one would expect these changes to lead to a significant impact on the model climate drift. These are documented in section 3.

It might be expected that a decrease in model systematic bias from one year to the next would be associated with an increase in forecast skill. In order to assess this, we have calculated the extent to which systematic error contributes to the total error, for each ‘extended’ winter season of October to March. The impact of the model improvements in reducing this ratio in regions of jet maxima can be clearly seen. Throughout much of the middle latitudes, however, the non-systematic component of error is dominant. This is discussed further in section 4 where we show northern hemisphere skill scores for the three individual extended winter periods. Comparing these scores with the growth of the spread between the individual forecasts initialized one day apart (internal error growth), we assess whether interannual variability in forecast skill is associated with variation in predictability, or with changes in model formulation.

Since the goal of extended-range prediction is to forecast changes in weather regime, we study the skill of predictions of 5-day mean 500 mb height projected onto a three-dimensional space of rotated empirical orthogonal functions (EOFs). According to Molteni et al. (1990), this three-dimensional space contains six principal local density maxima of atmospheric states, which can be interpreted as weather regimes. Skill scores and phase-space trajectories are shown for these truncated fields.

Section 5 of the paper is devoted to a discussion of prediction of monthly mean wind and rainfall in the tropics during the northern summer. As discussed by Charney and Shukla (1981), internal instabilities of the flow on synoptic scales account for most of the interannual variability of monthly-mean quantities at mid-latitudes, but cannot explain the observed variability in low latitudes. The latter, they suggested, is due partly to fluctuations in boundary parameters such as sea surface temperature (s.s.t.). Since these parameters vary on much longer timescales than the synoptic-scale flow, they should be predictable for longer periods of time. We show that interannual variability in the monthly-mean strength and position of the tropical easterly jet, which appears to be
strongly associated with monsoon variability, can be predicted. Also, simulated rainfall anomalies over the African inter-tropical convergence zone (ITCZ) from the four summer seasons do appear to be realistic.

Moreover, it is shown that forecasts of monthly means from integrations initialized one day apart are generally strongly consistent in the tropics. This is in contrast to the predictions of extratropical monthly mean rainfall in the summer extratropics (see, for example, Palmer and Branković 1989), lending support to Charney and Shukla’s hypothesis. Concluding remarks are given in section 6.

2. THE DATA BASE AND FORECAST MODEL CHANGES

As mentioned in the introduction, this paper deals with results from extended-range integrations of the operational ECMWF model. The integrations were run twice a month, around the middle of the month, from consecutive 12 GMT analyses. The initial dates of these integrations are given in Table 1.

<table>
<thead>
<tr>
<th>TABLE 1. INITIAL DATES OF OPERATIONAL MODEL INTEGRATIONS USED IN THE PRESENT STUDY.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
</tr>
<tr>
<td>January</td>
</tr>
<tr>
<td>February</td>
</tr>
<tr>
<td>March</td>
</tr>
<tr>
<td>April</td>
</tr>
<tr>
<td>May</td>
</tr>
<tr>
<td>June</td>
</tr>
<tr>
<td>July</td>
</tr>
<tr>
<td>August</td>
</tr>
<tr>
<td>September</td>
</tr>
<tr>
<td>October</td>
</tr>
<tr>
<td>November</td>
</tr>
<tr>
<td>December</td>
</tr>
</tbody>
</table>

For all the integrations, the spectral horizontal resolution of the model remained fixed at T106. However, during the period of the extended-range programme, the vertical resolution of the model changed, two existing physical parametrizations were revised, and one new physical parametrization was introduced. All the model revisions are listed in Table 2.

In May 1986, the 16-level version of the model was replaced by a 19-level version. This affected the resolution in the stratosphere, but below 200 mb there was little change. The principal motivation for the change was to provide better analyses in the stratosphere at standard pressure levels. The impact of this change in vertical resolution has been discussed by Simmons et al. (1989). In particular, assessment of a series of 10-day forecasts appears to indicate that the increased vertical resolution had positive impact on the Atlantic jet, the jet split, and associated downstream ridge and trough development.

<table>
<thead>
<tr>
<th>TABLE 2. MODEL REVISIONS OVER THE PERIOD OF THE EXTENDED-RANGE FORECAST PROGRAMME.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
</tr>
<tr>
<td>13 May 1986</td>
</tr>
<tr>
<td>15 July 1986</td>
</tr>
<tr>
<td>9 September 1986</td>
</tr>
<tr>
<td>7 April 1987</td>
</tr>
<tr>
<td>5 January 1988</td>
</tr>
</tbody>
</table>
in the troposphere. In Simmons (1987), it is shown that the overall strength of mid-
latitude tropospheric westerlies has been somewhat reduced in the 19-level version of
the model.

In July 1986 the parametrization of orographic gravity-wave drag (GWD) was
introduced, in order to alleviate the tendency toward excessive surface westerly flow in
the northern winter. The impact of this scheme on the ECMWF model has been discussed
in detail by Miller et al. (1989). In addition to a reduction in zonal mean wind, the GWD
scheme has a profound impact on the large- and synoptic-scale eddy heat and momentum
fluxes. Such diagnostics will be shown in the present paper, where we shall make reference
to the appropriate figures in Miller et al. (1989).

In September 1986 some changes to the analysis system were made. These changes
were primarily of a technical nature, and practically all characteristics of the previous
analysis scheme were retained. However, humidity analysis has been changed to three-
dimensional univariate optimal interpolation. It was found that such a new analysis fits
observational data better than the previous successive correction method, and makes the
analysed fields less moist.

In April 1987 the parametrization of land surface processes was revised (Blondin
and Böttger 1987). These revisions impact principally on surface heat fluxes over the
continents, and have larger impact in the summer. Blondin and Böttger found, for
example, an overall reduction of about 20% in convective precipitation at the ground.
Since the decrease in rain was balanced by reduction in local surface fluxes, rather than
reduction in moisture flux convergence, the revision has little impact on mean wind fields
in general, and no significant impact on the tropical easterly jet in particular.

Prior to the introduction of the GWD parametrization, subgrid-scale momentum
transports in the free atmosphere were represented by a vertical diffusion scheme.
However, the values of diffusion coefficients used prior to January 1988 caused free
atmospheric dissipation and associated vertical smoothing of jets to be excessive. A
revised scheme was introduced in January 1988 in which the vertical diffusion term no
longer operated above a diagnosed boundary-layer top, except when static instability
was generated (Miller 1988). Tests indicated that this revision substantially reduced
errors in the level of eddy kinetic energy.

Since the change to the vertical diffusion parametrization was made during the
extended winter period (October to March; hereafter OM), we re-ran the three pairs of
forecasts from October, November and December 1987, with the revised vertical diffusion
scheme (though we did not re-assimilate the initial data). The results presented below
will therefore be such that within each of the three OM periods, there have been no
changes to the forecast model. Some aspects of the influence of these changes on medium-
range forecast skill were discussed by Arpe (1987).

3. INTERANNUAL VARIABILITY IN SYSTEMATIC ERROR

(a) Zonal average diagnostics

A number of zonal-mean diagnostics of heat and momentum are shown below for the
three OM T106 forecasts. Similar diagnostics are given in the companion paper T on
the impact of horizontal resolution on systematic error.

Figure 1, left, shows the 30-day average OM systematic error in zonal-mean zonal
wind. For winter 1985/86 (Fig. 1 (a)) we find excessive westerlies in the extratropics,
particularly in the northern hemisphere (NH), extending from the ground to the top of
the model. In the northern tropical troposphere, there are substantial easterly errors,
becoming westerly in the tropical stratosphere.

In the second winter (Fig. 1 (b)), the tropical errors are largely unchanged, but there
is significant reduction in extratropical westerly errors, particularly in the NH stratosphere and upper troposphere. This is clearly consistent with the model changes introduced between these two winter seasons (see Table 2). In particular, the impact of GWD would certainly be in the sense observed between Figs. 1(a, b) (compare, for example, with Figs. 1(c) and 22 of Miller et al. 1989). This has been confirmed by Branković (1989) who compared OM 1986/87 extended-range integrations of the ECMWF T63 model which included 19 levels but no GWD, with T63 integrations which included both 19 levels and GWD. In the ‘no GWD’ runs, the mean NH lower stratosphere westerly wind error exceeds 8 m s⁻¹ at about 45°N, having almost identical structure to that shown in Fig. 1(a). This would confirm that the reduction in westerly bias between 1985/86 and 1986/87 (see Fig. 1) comes primarily from the GWD parametrization.

In the third winter (Fig. 1(c)), there is some further reduction in NH extratropical westerly error, though the tropical easterly error appears to have increased in the upper troposphere.

It could be argued that the apparent reduction in systematic error over the three winters merely reflects interannual variability in the real atmospheric circulation. For example, if the extratropical zonal-mean flow in the second winter was significantly stronger than in the first winter, the model may show a reduction in systematic error in the absence of any change in model formulation. Figures 1(d–f) show the verifying analysis anomalies relative to the 3-year mean. (Note the reduced contour interval relative to Fig. 1(a–c).) It is clear that, in general, the reduction in systematic error is genuine, except perhaps in the equatorial stratosphere, where the analysis shows strong
westerly anomalies for the third year, presumably associated with the quasi-biennial oscillation.

It is interesting to examine the evolution of these systematic errors as a function of forecast time, in order to assess the extent to which the drift in climate of the model can be inferred purely from systematic errors within the medium-range period. Figure 2 shows the zonal-mean zonal wind error averaged from 300–30 mb for the three CM periods. The errors are growing strongly at the end of the medium range (denoted here by the vertical lines drawn at day 10). Moreover, the evolution of systematic error in the first 10 days can give a misleading impression of the overall longer term climatic drift. For example, the reduction in the 30-day mean NH extratropical zonal-mean wind error between the first, second and third winter (Figs. 1(a–c)) is not apparent in the first 10 days of integration. Indeed, in the first 10 days during the third winter, a 4 m s\(^{-1}\) error is evident by about day 5, earlier in the forecast period than in the previous two winters. However, these errors barely exceed 4 m s\(^{-1}\) throughout the 30-day period, whilst in the second winter they exceed 6 m s\(^{-1}\) by about day 18, and in the first winter they exceed 8 m s\(^{-1}\) by about day 16, reaching 10 m s\(^{-1}\) towards the end of the forecast period. As commented above, 30-day mean tropical easterly errors appear to have increased in the third winter; this is quite clear from Fig. 2, where, for the third winter, there are values in excess of 8 m s\(^{-1}\) in the last 5 days of integration. Indeed it is unclear from this figure whether the climatic drift of the model has reached its asymptotic level at day 30. We believe these remarks highlight the importance of extended-range integrations in the assessment of possible model reformulations, even within a purely medium-range NWP environment.

Diagrams of zonal-mean temperature error are shown in Figs. 3 (a–c). The NH high-latitude stratospheric cooling in the first extended winter is reduced in the second. In the SH (southern hemisphere) there is a cold bias in all three years. The overall warming in the tropical stratosphere persists throughout the period; however, tropospheric warming in the subtropics has been almost halved.

Figure 4 shows 30-day mean vertical integrals of momentum flux, decomposed into zonal wavenumbers 1–3 (left) and wavenumbers 4–9 (right). In the verifying analysis (top panels) the NH high-latitudes equatorward momentum-flux due to long waves was strongest in the first extended winter, while in the subtropics the amplitude of long waves is smallest in the second winter (Fig. 4(a)). In the NH low latitudes there is a marked increase in analysed synoptic-scale poleward momentum flux in the second and third winters relative to the first winter, 1985/86 (Fig. 4(b)). Between the first and second winters, after the introduction of orographic GWD, the maximum in long-wave momentum flux error has shifted from about 60°N to about 50°N (Fig. 4(c)), whilst in Fig. 4(d), a clear decrease in the synoptic-scale momentum flux errors can be seen. As discussed in T, the fact that in the first winter the model simulated insufficient long-wave equatorward momentum flux in high latitudes, and excessive synoptic-scale momentum flux in lower latitudes, may result from inadequate orographic forcing in the ECMWF model. The maximum error in the third winter is positioned close to the latitude of maximum momentum-flux gradient in the analysis, and therefore corresponds to an error in the latitudinal positioning of maximum gradient. This was not anticipated on the basis of experiments with the GWD parametrization alone (see Miller et al. 1989, Fig. 18). In T it is also argued that the excessive long-wave poleward momentum fluxes in the stratosphere centred at about 45°N are the result of spurious reflection of Rossby-wave activity from the top of the model. Consistent with this speculation, one would expect a decrease in the error with the increase in vertical resolution. Such a decrease is indeed observed between the first and second winters (Fig. 4(e)),
Figure 2. Latitude-time diagrams of zonal-mean wind error averaged between 300–30 mb for October to March forecasts: (a) 1985/86; (b) 1986/87; (c) 1987/88. Contours every 2 m s$^{-1}$. 
As shown in Fig. 5, the underestimation in stratospheric poleward heat fluxes is substantially reduced between the first and second winters, consistent with expectations on the basis of GWD experiments (see Fig. 18 of Miller et al. 1989). On the other hand, near the tropopause, there is an increase in (synoptic-wave) heat-flux error in the NH in the second winter, but this is reduced in the third winter consistent with the effective removal of free-atmospheric vertical diffusion.

One of the principal improvements to the drift in climate of the model in the third winter is the reduction in the error in eddy kinetic energy. This is shown in Fig. 6. In the second winter, it appears that the introduction of 19 levels and GWD had some impact in the stratosphere, but little in the troposphere. For the third winter, reduction of the error in eddy kinetic energy throughout the troposphere is evident.

In zonally-averaged zonal wind, by contrast with the OM period, there is no significant reduction in the error amplitude during the four April to September (AS) seasons 1985–1988 (not shown). The differences between the extended summers are mainly in the error pattern and we believe they mostly reflect the influence of interannual natural variation. The largest errors in the zonal wind are found in the SH stratosphere at about 30°S in the region of strong vertical wind shear above the jet core. The error maximum here demonstrates the model’s inability to capture this shear properly. The error in zonal mean temperature remains largely unchanged in most of the model’s atmosphere during the four AS seasons.

(b) Simulation of blocking activity

As documented by Tibaldi and Molteni (1990), prediction of blocking activity in the ECMWF operational medium-range forecasts is poor. In this section we address the
Figure 4. 30-day mean vertically integrated momentum flux (m²s⁻²). Mean of October to March forecasts. Left: contribution from zonal waves 1–3. Right: contribution from zonal waves 4–9; (a), (b) verifying analysis 1000–25 mb; (c), (d) forecast error 1000–25 mb; (e) forecast error 70–25 mb.
Figure 5. 30-day-mean zonal-mean cross-sections of horizontal heat flux error. Mean of October to March forecasts. (a) 1985/86. (b) 1986/87. (c) 1987/88. Contours every 2 K m s$^{-1}$. 
Figure 6. 30-day-mean zonal-mean cross-sections of error in eddy kinetic energy. Mean of October to March forecasts: (a) 1985/86; (b) 1986/87; (c) 1987/88. Contours every 2 kJ m$^{-2}$. 
question whether simulation of blocking activity in the extended range has improved over the three winters considered.

A blocking frequency index is defined on a five-day-mean 500 mb height field using Tibaldi and Molteni’s prescription. The index measures the frequency of easterly flow at some longitude between 40 and 60°N, provided that, in addition, a moderate or strong westerly flow also exists north of 60°N. In Fig. 7, the dashed lines show analyzed blocking frequency, the solid lines show simulated blocking frequency between days 11–30 of the integrations. The results are related to November, December, January and February integrations only.

For the winter 1985/86, it can be seen that the model fails to reproduce the observed blocking frequency at all longitudes. For 1986/87 there is some small improvement. However, in 1987/88 simulated blocking frequency is dramatically improved and, for some longitudes, it exceeds the analysed frequency.

Again, this indicates that changes in model formulation over the period of extended-range forecast programme have improved the model’s ability to simulate atmospheric variability.

(c) Geographical distribution of extratropical systematic error and its contribution to total error

Figure 8 shows the average 30-day-mean 500 mb height-error maps for forecasts in the NH and SH for the three OM seasons. It can be seen that the negative height errors over the north-east Pacific and north-east Atlantic are reduced in the second winter, particularly over the Atlantic (cf. Fig. 8(a, c)). There is, however, an increase in geopotential-height error over the central Pacific between the first and second winter. (This may be associated with atmospheric interannual variability; the 1986/87 El Niño event was at its peak during this winter period.) During the third winter period (Fig. 8(e)) there is an overall reduction in height error; the only significant errors are over the Pacific and east Asian seaboard, though smaller in comparison with the two previous winters. Whilst the error pattern over the northern Pacific and northern Atlantic shown in Fig. 8(a) is typical of the ‘zonalization’ in earlier versions of the ECMWF model (see for example Wallace et al. 1983), from 1986 this error has not only been reduced, but its pattern has also changed.

In the SH for the first year (Fig. 8(b)), there is a band of positive height error in mid-latitudes exceeding 4 dam. There is a small region at about 60°S near the date-line where negative height errors exceed 5 dam. The character of these errors has been changed little in the second year (Fig. 8(d)). The largest change comes in the third year (Fig. 8(f)), when, except for some rather isolated regions, errors are everywhere less than 4 dam.

Another way of assessing the impact of model reformulations on model skill is shown in Figs. 9 and 10. These figures show for each winter period, the fraction of total daily mean-square error in 500 mb height contributed by the systematic error. This fraction is computed for days 21–30 of each forecast; over this time, daily forecast errors are, to a good approximation, statistically stationary. The ratio shown in Figs. 9 and 10 can be written as

\[ \left\{ \frac{1}{120} \sum_{j=1}^{12} \sum_{i=21}^{30} E_q^2(x, y) \right\}^{\frac{1}{2}} \left/ \left\{ \frac{1}{120} \sum_{j=1}^{12} \sum_{i=21}^{30} E_j^2(x, y) \right\} \right. \]

where \( E_q(x, y) \) is the 500 mb height error. The index \( i \) runs over forecast days (i.e. from 21–30), and the index \( j \) runs over forecast cases (i.e. from 1–12 for the set of OM forecasts).
30-day mean error
30-day mean October-March

1985/86

1986/87

1987/88

Figure 8. 30-day mean error of 500 mb geopotential height. Mean of October to March forecasts. Left: northern hemisphere. Right: southern hemisphere; (a), (b) 1985/86; (c), (d) 1986/87; (e), (f) 1987/88. Contours every 4 dam.
\( Z \) 500 error ratio

Days 21-30 mean; October-March

NH  

SH  

1985/86  

1986/87  

1987/88  

Figure 9. Ratio of systematic to total daily mean-square error of 500 mb geopotential height in extratropics. Computed using days 21-30 of October to March forecasts. Left: northern hemisphere. Right: southern hemisphere; (a), (b) 1985/86; (c), (d) 1986/87; (e), (f) 1987/88. Contours every 15\%.
For the first NH winter period (Fig. 9(a)), it can be seen that the contribution of the systematic error to the total error is considerable, particularly in the jet regions (see also Hollingsworth et al. 1987). Over the east Asian seaboard the systematic error contributes up to 70% of the total daily mean-square error (over 80% in terms of the ratio of r.m.s. errors). Across the Eurasian land mass, values exceed 30% everywhere in jet regions. In high latitudes values are generally smaller, though they reach almost 60% over the Bering Strait.

These values are generally reduced in the second NH winter, Fig. 9(c) (except over the subtropical central north Pacific, where as shown above, the 30-day mean systematic error increased). On the other hand, the character of the error ratio remains the same as in the first winter. There is a considerable overall reduction, however, in the third winter (Fig. 9(e)) with the only significant area of high ratio in the extratropics lying over the east Asian seaboard and the subtropical east Pacific. This major reduction is most likely due to the impact of the reduction in free-atmospheric vertical diffusion.

A similar reduction in the relative impact of the systematic error of the model on the total daily mean-square error, can be seen in the southern hemisphere (Fig. 9(b, d, f)), with the only significant area of high ratio in the extratropics, in 1987/88, lying over Australia.

In the tropics, on the other hand, systematic error is a substantial fraction of total error over much of the area during all three winters (see Fig. 10). There is no apparent reduction in the ratio over the three years. For all years, there are maxima over the eastern Pacific/central American region, and over or near Africa. For 1986/87 and 1987/88, the 60% contour (almost 80% in terms of r.m.s.) covers most of the tropics. It should be noted, however, that the model changes introduced over the period of three years were aimed mainly at reducing model systematic deficiencies in mid-latitudes.
4. **Evolution of model skill over three extended winters**

(a) *Interannual variability in extended-range skill*

In earlier sections we have discussed the apparent improvement in systematic error since the beginning of the extended-range forecast programme. In this subsection, we pose the question: have these model improvements led to any change in extended-range skill?

In Fig. 11 (a) we show the conventional NH anomaly correlation coefficient (ACC) of 500 mb height, averaged over the 12 forecasts for each season, for 10-day and monthly-mean fields, for each of the three winters. It can be seen that, despite the improvements to the model climate drift, there is no clear monotonic improvement in skill over the three years. Indeed, over the 30-day period, one cannot clearly distinguish between the first and third winters in terms of ACC.

Figure 11 (b) shows the intrinsic error growth over the three OM periods in terms of the ACC of 500 mb height within each of the twin forecasts. It is interesting to note that growth and apparent asymptotic level of spread in the third winter are noticeably larger than the values in the first two winters. This is also observed in terms of r.m.s. spread (not shown), which for the third winter approaches asymptotically close to the level of persistence at days 16–25. This, in turn, gives the upper limit for potential predictive skill for 10-day mean fields.

It would appear therefore from Fig. 11 that the recent improvements in model formulation have not had a strong impact on time-mean forecast skill throughout the 30-day range. This may indicate that the impact of intrinsic atmospheric interannual variability was dominant (see also below). Consistent with results above, the reduction in free atmospheric vertical diffusion appears to have led to more realistic internal error growth statistics.

(b) *Forecast skill of large-scale weather regime transition*

We now address a crucial question in the assessment of the extended-range programme: is there any indication, in results over the last three winter periods, of skill in the prediction of weather regime transitions beyond the medium range?

In a recent paper (Molteni et al. 1990), regimes of the northern extratropical circulation in winter were identified as local density maxima, or clusters, of atmospheric states in a five-dimensional phase-space generated by the leading EOFs of 500 mb eddy (i.e. deviation from zonal-mean) geopotential-height fields. Six principal clusters were found. Further analysis revealed that a three-dimensional subspace contained these six clusters. The three rotated EOFs of 500 mb height which define orthogonal axes for this subspace are shown in Fig. 12. Whilst the clusters, and not the rotated EOFs, represent the actual circulation regimes, it is worth commenting on the patterns associated with these EOFs. Rotated EOF 1 is very similar in structure to the Pacific/North American pattern identified by Wallace and Gutzler (1981), except that the wave pattern over the Atlantic and Eurasia has a more substantial amplitude. Rotated EOF 2 has a clear wavenumber 3 pattern, with ridges in areas of strong low-frequency and blocking variability (see for example Dole 1986). Rotated EOF 3 has strong amplitude in the Euro/Atlantic sector where it resembles the Eurasian teleconnection pattern described by Wallace and Gutzler (1981).

Since the principal weather regimes of the northern winter lie in the space spanned by these three rotated EOFs, and since the principal objective of extended-range forecasting is the prediction of weather regime (rather than instantaneous weather), we study in detail the verification of the wintertime extended-range forecasts (November to
Figure 11. 10-day and 30-day mean northern hemisphere 500 mb height anomaly correlation coefficients for: (a) model skill; (b) spread (between adjacent forecasts for the October to March period). Dashed: 1985/86. Dotted: 1986/87. Solid: 1987/88.

February inclusive) by projecting each forecast and verifying the 500 mb height field onto these three rotated EOFs.

Results of ACC skill scores of 5-day mean height fields, projected onto these three EOFs are summarized in Table 3(a), which shows the average ACC (over all forecasts in each winter) for the three individual winters (1985/86, 1986/87 and 1987/88). In the
first five days, the height pattern is predicted with ACC > 0.9 for all winters. For days 6–10, there is already considerable interannual variability, with the results for the first winter having the highest skill score, and those for the second winter having the least. In this sense, these filtered ACC scores are consistent with the conventional ACC estimates of skill for the first 10-day mean, shown in Fig. 11. Beyond day 10 the ACC for winter 1 remains positive throughout the forecast period. For winters 2 and 3, on the other hand, the ACCs for days 16–20 are negative. For days 21–30, there is an apparent 'return of skill' during the second and third year, though it is difficult to attach any particular significance to this. Differences between Table 3 and Fig. 11 result from different sampling (only November to February in Table 3), and because data in Table 3 were derived from filtered eddy fields only.

### TABLE 3. (a) 5-DAY MEAN ANOMALY CORRELATION COEFFICIENT FOR 500 mb HEIGHT FIELDS PROJECTED ONTO THE ROTATED EOFs SHOWN IN FIG. 12. (b) 10-DAY MEAN ANOMALY CORRELATION COEFFICIENT FOR 500 mb HEIGHT FIELDS PROJECTED ONTO THE ROTATED EOFs SHOWN IN FIG. 12.

#### (a)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>85/86</td>
<td>0.91</td>
<td>0.82</td>
<td>0.26</td>
<td>0.04</td>
<td>0.05</td>
<td>0.08</td>
</tr>
<tr>
<td>86/87</td>
<td>0.94</td>
<td>0.32</td>
<td>0.48</td>
<td>−0.15</td>
<td>0.17</td>
<td>−0.04</td>
</tr>
<tr>
<td>87/88</td>
<td>0.97</td>
<td>0.61</td>
<td>0.33</td>
<td>−0.22</td>
<td>0.29</td>
<td>−0.10</td>
</tr>
</tbody>
</table>

#### (b)

<table>
<thead>
<tr>
<th>ACC Winter</th>
<th>1–10</th>
<th>11–20</th>
<th>21–30</th>
</tr>
</thead>
<tbody>
<tr>
<td>85/86</td>
<td>0.90</td>
<td>0.32</td>
<td>0.14</td>
</tr>
<tr>
<td>86/87</td>
<td>0.44</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>87/88</td>
<td>0.81</td>
<td>−0.05</td>
<td>−0.12</td>
</tr>
</tbody>
</table>

The ACC skill scores of 10-day mean height fields are summarized in Table 3(b). As above, it can be seen that forecasts have the highest skill score for the first winter, where ACC > 0 for all three 10-day means. Taking the three winters together, it would appear that skill is lost between days 10 and 20, though over a significant fraction of forecasts, skill is maintained to the end of this period.

We compare these figures with skill scores from other dynamical extended-range forecasts (DERF) programmes. For example, for the NMC (National Meteorological Center, U.S.A.), programme 108 30-day forecasts were made covering the period 14 December 1986 to 1 April 1987 which gave average ACC scores from the NMC DERF programme for days 1–10, 11–20 and 21–30 of 0.72, 0.18 and 0.12 respectively (Tracton et al. 1989). For days 1–10 and 11–20, these values lie between the ECMWF three-winter mean and the values for the winter 1986/87 (see Table 3(b)). In this sense the two sets of scores are quite consistent. For days 21–30, the NMC scores are a little higher than those for ECMWF, both in terms of the three-winter mean and the winter 1986/87. However, it is unlikely that they are significantly higher, since the overall level of skill is poor at this range. These results are also broadly consistent with those from the Meteorological Office extended-range forecast programme. Murphy (1990), for example, reports that forecast skill was consistently positive at days 6–15, but very small on average beyond 20 days.

In order to give a visual impression of the overall level of skill of forecasts of the large-scale flow up to day 15, we show in Fig. 13 trajectories of the winter integrations and verifying analyses for the three years, in the space spanned by the first two rotated
EOFs. (For reasons of space we do not show trajectories projected onto the third EOF, though the general conclusions made here are valid for the trajectories in the third dimension.)

Each trajectory in Fig. 13 consists of two line segments. The line segments join values for the mean of days 1–5, 6–10 and 11–15. The two forecast trajectories have closed arrowheads labelled ‘1’ and ‘2’ and the verifying analysis trajectory has an open arrowhead labelled ‘0’. The arrowheads lie on values for days 11–15. The left-hand column shows forecasts from November 1985 to February 1986, the middle column from November 1986 to February 1987, and the right-hand column from November 1987 to February 1988.

It is clear from Fig. 13 that predictive skill is, on average, marginal at days 11–15. Nevertheless, there are clearly some cases when both of the twin forecasts were skilful at this range (e.g. the ‘close to perfect’ predictions for February 1986; this case is discussed in more detail in B). There are others (e.g. January 1988) where one of the twin forecasts
Figure 13. Phase-space trajectories of twin forecasts (numbered '1' and '2'), and verifying analysis (numbered zero) from the wintertime period (November to February inclusive). Line segments join values for the 5-day mean fields 1–5, 6–10, and 11–15. Left-hand column: 1985/86. Middle column: 1986/87. Right-hand column: 1987/88.
evolved correctly, whilst the other was erroneous. There are cases (e.g. November 1985) where both forecasts, although marginal, give an approximately correct sense of direction of the phase-space evolution of the observed flow. Finally, there are cases (e.g. December 1986) when both integrations, although consistent in their prediction of flow evolution beyond day 10, are incorrect.

The visual impression of these phase-space trajectories confirms the results of the objective skill scores. However, it is important to judge the viability of extended-range forecasting not only in terms of the mean level of skill, but also by the ability to identify a priori, those forecasts with above average levels of skill. This topic is discussed in detail in B which addresses techniques of time-lagged ensemble forecasting.

We conclude this section by discussing the possible relationship between forecast skill and transitions between particular weather regimes. As discussed in Palmer (1988), there appears to be a clear dependence of extended-range predictive skill on the sign of the Pacific/North American index of either the observed or predicted flow. This relationship has been confirmed by O'Lenic and Livezey (1989) in the medium range, and by Tracton et al. (1989) in the extended range. It has been studied in more detail by Molteni and Tibaldi (1990) in terms of the rotated EOFs used in this section. In particular, Molteni and Tibaldi found that forecast skill scores in the medium range had a clear bimodal distribution when the rotated EOF 1 had a negative coefficient. They related this to the existence of two major atmospheric clusters found by Molteni et al. (1990), in the region of phase-space where the rotated EOF 1 coefficient was negative. In Fig. 14 we show scatter diagrams of the ACC skill score for days 11–20 plotted against the verifying rotated EOF 1 coefficient on days 11–20. There is a correlation of 44% between these two sets of data, with the most skilful forecasts being associated with transitions to, or persistence of, a positive EOF 1 coefficient. Since positive EOF 1 has positive PNA

![Figure 14](image-url)

Figure 14. Scatter diagram showing the correlation between the anomaly correlation coefficient (ACC) skill score and the observed value of rotated EOF 1 at days 11–20. Triangle markers denote forecasts from the winter 1985/86; circle markers denote forecasts from the winter 1986/87; square markers denote forecasts from the winter 1987/88.
index, these results are consistent with those of studies mentioned above. For EOF 2 and EOF 3, the correlation is much smaller.

5. PREDICTION OF TROPICAL MONTHLY MEAN WIND AND RAINFALL

As discussed in the introduction, internal instabilities in the extratropical wintertime flow appear to be able to account for much of its observed variability up to timescales of a month (e.g. Palmer 1987). Hence the variation of boundary forcing on the extratropical flow is not a dominant source of observed extratropical wintertime low-frequency variability. As a result, monthly timescale prediction in the extratropical wintertime appears to be principally an initial value problem.

On the other hand, as suggested by Charney and Shukla (1981) and demonstrated in a series of GCM (general circulation model) experiments by Manabe and Hahn (1981) and Lau (1981), the internal dynamics of the tropical atmosphere can only explain a fraction of monthly timescale interannual variability of the atmospheric flow. Indeed in our discussion in section 3, we noted that aspects of the model tropical systematic error were correlated with anomalies in the boundary forcing. Hence, unless we can predict this boundary forcing well, we cannot make good extended-range predictions of the tropical atmosphere.

In common with most GCMs, some of the land surface processes in the ECMWF model (e.g. associated with variability in surface temperature, soil moisture) are inter-active. However, over the ocean, surface temperatures throughout the integration are held constant at their initial values. Such noninteractive ocean boundary-conditions may be sufficient for monthly prediction, but are probably insufficient for seasonal timescales. Whilst seasonal prediction with coupled ocean–atmosphere models represents a major goal of the WMO World Climate Research Programme, it is important to ascertain first that accurate monthly prediction in the tropics with a high-resolution NWP model is feasible with only partially interactive boundary conditions. A successful outcome may provide some stimulus for the development of operational NWP models with fully interactive boundary forcing. Moreover, as suggested by Shukla et al. (1988), monthly mean fields during severe drought or flood monsoon years are often broadly representative of the season as a whole.

It can also be mentioned that an ability to simulate accurately tropical low-frequency variability has a profound influence on extratropical extended-range forecast skill. This has been shown in a companion paper (Ferranti et al. 1990) in which a number of extended-range integrations were re-run with tropical winds and temperatures relaxed towards the verifying analysis. The results of Ferranti et al. therefore provide additional motivation for the analysis in this section.

In subsections (a) and (b) we study the prediction of interannual variation in wind fields over the African/Asian monsoon region during the summer periods from 1985 to 1988, and focus on India and the Sahel to study the regional prediction of monthly mean rainfall. Again we must stress that these results should be viewed within the operational framework in which the integrations were made. In particular, since the model was not held fixed during the period of the experimentation, it is necessary to consider the influence of model reformulations on simulated interannual variation.

As discussed in section 2, changes in April 1987 to land-surface processes reduced convective precipitation over the continents, though they had little effect on the mean wind field. Hence, in the absence of observed interannual variability, forecast rainfall values for summers 1987 and 1988 will be underestimated when compared with values for 1985 and 1986, though wind values should be roughly comparable.
(a) *Summer monsoon wind*

Figure 15 shows the verifying analysis of 30-day-mean 200 mb winds over the eastern tropical hemisphere, for the four individual mid June to mid July periods from 1985 to 1988. Diagnostic studies (Newell and Kidson 1984; Lu and Ding 1989, for example), have suggested a strong association between the strength of the tropical easterly jet and the intensity of seasonal monsoon rains over both Africa and India.

Interannual variability in the strength and position of the tropical easterly jet is apparent. Values greater than 20 m s\(^{-1}\) are shown stippled. During 1985 (Fig. 15(a)), the analysed 20 m s\(^{-1}\) isotach extended from Indonesia to east Africa, with the 10 m s\(^{-1}\) isotach extending across most of the eastern hemisphere. During 1986 (Fig. 15(b)), the extent of the 20 m s\(^{-1}\) isotach is clearly reduced, and over western Africa, the 10 m s\(^{-1}\) line is broken. The analysis for 1987 (Fig. 15(c)) shows a further weakening of the tropical easterly jet. The 20 m s\(^{-1}\) isotach is barely evident, and the 10 m s\(^{-1}\) contour over Africa is well broken. During the 1988 period (Fig. 15(d)), the strength of the easterly jet increased significantly compared with 1987 values. It is worth noting that in 1988 the position of the jet core is shifted further west than in any of the earlier years, and that over the four-year period it is strongest over Africa. Interannual variability in the tropical easterly jet for other summer months is qualitatively consistent with these results (not shown).

Figure 16 shows the forecast monthly-mean 200 mb winds for the four years. For each year, we show results from only the first forecast from the consecutive initial conditions. However, as will be shown below, tropical predictions from the twin forecasts are very similar. During 1985, the forecast 20 m s\(^{-1}\) isotach extended from about 90°E to 15°E. During the 1986 period, the stippled region where winds exceed 20 m s\(^{-1}\) is reduced, extending now from the tip of India to central Africa. For 1987 it can be seen that there is only a very small region of 20 m s\(^{-1}\) wind over east Africa. In 1988, the forecast easterly winds increased in intensity, extending in both cases as far west and east as they did in 1985. It is worth noting that the stippled region in both forecasts for 1988 extended north of 15°N over east Africa and the Arabian sea. The forecast 20 m s\(^{-1}\) wind failed to reach as far north as this in any preceding year.

A number of points can be made comparing Figs. 15 and 16. Firstly, it is clear that there is a systematic bias in the geographical position of the strongest winds in the forecast fields. The forecast 20 m s\(^{-1}\) isotach is clearly shifted west compared with the analysed values. (As shown in T, there is also a bias in the 850 mb monsoon flow, with the low-level jet off the Somali peninsula and the curvature of the flow towards India being underestimated.) However, despite this systematic bias, the interannual variations in the forecast 200 mb winds have clearly been correctly predicted at least qualitatively, with a monotonic decrease in strength between 1985 and 1987, and an increase in 1988.

Although not shown here, we have studied 30-day mean forecast winds from other summer months, and, in addition, the 10-day mean winds comprising the monthly mean. In general, forecast winds from the later months in the season have not as high skill scores as those shown in Fig. 16. Moreover, for all months, the systematic errors have grown to a sufficiently large amplitude by days 21–30 that little predictive skill is left. In this sense, the skill of the monthly-mean fields is dominated by skill in the first 20 days.

(b) *Summer monsoon rainfall*

In this section, we show some regional maps over Africa and India of monthly mean predicted rainfall. In order to attempt to verify these rainfall amounts, we have taken, from the operational ECMWF forecast archives, accumulated 24-hour rainfall amounts
Figure 15. Analysis of 200 mb wind (m s$^{-1}$) averaged over 30 days from (a) 16 June 1985; (b) 15 June 1986; (c) 14 June 1987; (d) 19 June 1988. Contours every 10 m s$^{-1}$; stippled area shows wind speeds greater than 20 m s$^{-1}$.

30-day mean from mid June
Figure 16. 30-day mean prediction of 200 mb wind (m s⁻¹) from forecast initialized: on (a) 15 June 1985; (b) 14 June 1986; (c) 13 June 1987; (d) 18 June 1988. Contours every 10 m s⁻¹; stippled area shows wind speeds greater than 20 m s⁻¹.
averaged over thirty day-1 operational forecasts, within the period of the extended-range forecasts, and used that as 'truth'. According to Datta and Hatwar (1988) this 24-hour accumulated rainfall gives a reasonable approximation of monthly-mean monsoon rainfall over India, and values compare well with Jaeger's (1976) climatology. See also, below, comparison with station data. Molteni and Tibaldi (1985) found that, despite systematic error, ECMWF 24-hour forecast rainfall in tropical areas, averaged over monthly periods, retained a good deal of observed interannual variability in the absence of major model changes.

In Figs. 17 and 18 (top panels) we first show the monthly-mean predicted rain for mid June to mid July and for mid July to mid August, averaged over the four years. This is compared with the 'verifying' rain over exactly the same period (Figs. 17 and 18; bottom panels). Over Africa, it can be seen that the major maxima over the west coast and the Ethiopian highlands are well simulated. The northward displacement of the ITCZ between the two monthly periods is correctly captured.

Over India, rainfall is reasonably well simulated though, consistent with systematic errors in low-level wind, the rainfall maximum over the west coast is positioned somewhat too far north. The interior of northern India is also too wet, especially in the July forecasts, though the rain shadow effect to the east of the western Ghats is well simulated. The maximum over the Bay of Bengal is also predicted. The apparent weakening of 30-day mean forecast rainfall amount, over much of the Indian ocean, may be associated with problems related to the initial 'spin-up' of the model during the first few days of the forecast period, giving erroneously strong verifying rain.

Figures 19 and 20 show the monthly-mean 'verifying' rainfall anomalies (deviations from the corresponding 4-year mean) over Africa and India for mid June to mid July, and mid July to mid August (respectively). The corresponding 30-day mean predictions (deviations from the 4-year average 30-day mean field) are shown in Figs. 22 and 23.

As suggested above, it appears that African rainfall anomalies are generally correlated with the strength of the tropical easterly jet. During 1985, the 'verifying' rainfall showed positive values in the ITCZ over Africa during both 30-day periods. During 1986, as the jet weakened across western Africa (Fig. 15(b)), rainfall anomalies became negative in the west, but remained positive in the east. Again note the consistency between the two months. 1987 was a drought year across the Sahel, where rainfall anomalies were negative across most of the region. In 1988 anomalies were positive in the west and in the extreme east, where the excess rainfall in the Ethiopian highlands was associated with the well-publicized flooding in Sudan.

Before discussing the monthly forecast rainfall, it is worth comparing these 'verifying' day-1 values with observed data. In Fig. 21 we show the July mean anomalies for Africa (relative to the period 1931–60) analysed from station data by the Freie Universität, Berlin. Percentage anomalies are shown. The pattern for 1985 (Fig. 21(a)) shows mainly wet (>100%) values in the west and central Sahel and dry (<50%) values in the east (between 30 and 40°E). The pattern for 1986 (Fig. 21(b)) shows dry values in the west (between 15°W and 5°E) with an indication of wet values in the east (between 15 and 25°E). In 1987 (Fig. 21(c)) we find clearly the overall dry conditions, and for 1988 (Fig. 21(d)) values are dry in the extreme west and wet east of about 5°E.

It can be said that the pattern of our 'verifying' day-1 forecast anomalies are in fair agreement with the observed anomalies. On the other hand, the values for 1988 underestimate the return to more normal condition overall. As discussed above, this is consistent with the impact of the 1987 model change in surface parametrization.

The 30-day mean forecast rainfall anomalies from both integrations (Figs. 22 and 23) show anomaly fields which accord reasonably well with the 'verifying' rain over
Rainfall (mm day\(^{-1}\))
30-day mean from mid June

Figure 17. (a) 30-day mean simulated rainfall (mm d\(^{-1}\)), averaged over 8 T106 forecasts from mid June; (b) 24-hour accumulated rainfall (mm d\(^{-1}\)) averaged over the period for which the T106 forecasts in (a) verify. Contours: 2, 10, 20, 30 mm d\(^{-1}\); stippled area shows rainfall greater than 10 mm d\(^{-1}\).

Rainfall (mm day\(^{-1}\))
30-day mean from mid July

Figure 18. As Fig. 17 but for forecasts from mid July.
Africa. Note the consistency between forecasts initialized 24 hours apart, an indication of the relative unimportance of internal instabilities for these monthly tropical predictions. In 1985, both forecasts from each of the two monthly periods show a band of positive rainfall anomaly extending from the west coast across to about 40°E. These positive anomalies extend further north for the second 30-day period. During 1986, particularly for the July forecasts, anomalies generally showed negative values on the west coast and over the Gulf of Guinea, and generally positive values over east Africa. During 1987, all integrations correctly predicted negative rainfall anomalies over much of the ITCZ from west to east coasts. For 1988, the integrations show a partial, though insufficient, return to more normal conditions. Note, however, the indication of positive forecast
rainfall anomalies over the eastern Ethiopian highlands, particularly for the July forecasts.

Over India, the ‘verifying’ rainfall (Figs. 19 and 20) shows positive anomalies (for both months) over the west coast in 1985. 1987 was a drought year, as suggested by the extensive regions of negative anomalies. 1988 saw a return to more normal monsoon rainfall, though note the positive anomalies over the Himalayan foothills, presumably associated with the severe flooding of Bangladesh.

It can be seen that the predictions of Indian monsoon rainfall anomalies (Figs. 22 and 23) are not as accurate as for the African monsoon. Nevertheless, some of the more important features do appear to have been forecast for the first 30-day period (June–July). For example, during the drought year, 1987, there are indications of negative anomalies, at least to the north. During 1988, the strong positive anomalies over the Himalayan foothills also appear to have been correctly predicted. For the second 30-day
Figure 21. July total precipitation amounts in mm. Contours are anomalies and represent percentage from the mean related to the 1951–60 period (from Freie Universität Berlin analyses). Solid line: wet anomalies. Dashed line: dry anomalies.
Figure 22. 30-day mean forecast anomaly in mm day$^{-1}$ (departure from Fig. 17(a)) for each T106 forecast from mid June. Contours: ±1, ±10, ±20 mm day$^{-1}$. Negative anomaly, thick stipple; Positive anomaly, coarse stipple.
Figure 23: As Fig. 22 but for mid July.
period (July–August), on the other hand, the forecasts appear to be much poorer. For example, the 1987 drought is barely captured. (This is consistent with our comment earlier that the prediction of the tropical easterly jet was poorer for the July forecasts than for the June forecasts.) It is likely that the relative poorness of the Indian monsoon predictions is associated with the fact that the drift in climate of the model is more severe over the Indian ocean than over Africa. In addition, low-frequency variability associated with the 30–60 day oscillation could play an important role in monthly-timescale prediction of the large-scale flow over the Indian Ocean; less so over Africa where the amplitude of the oscillation is much weaker.

In conclusion, it would appear that some aspects of the large-scale tropical monsoon circulation, and even more regional rainfall anomalies have been predicted on a monthly timescale, particularly for the African region. This supports the results of Folland et al. (1986) and Owen and Folland (1988) that interannual and interdecadal variability in world-wide s.s.t. is responsible for drought over Africa, and that seasonal forecasts of African rainfall based on a knowledge of world-wide s.s.t. can be made. Indeed since the observed anomalies are fairly persistent from one month to the next, the 30-day anomalies can be treated approximately as proxies for seasonal timescale rainfall and wind anomalies. In this paper we shall not pursue further the question of the role that the boundary forcing (land processes and ocean temperatures) had in these forecasts: this will be the subject of a further investigation (see, for example, Palmer and Branković 1989).

6. CONCLUSIONS

We have studied interannual variability in extended-range skill of the ECMWF operational model over a set of monthly forecasts from April 1985 to September 1988. This period was characterized by a number of important changes to the model.

It is found that in the extratropics the changes in model formulation had a clear impact on the drift in climate of the model throughout the period of the extended-range programme. We note in particular an apparent improvement in the model’s ability to simulate blocking activity. On the other hand, in the tropics, the impact of model changes is less obvious, and forecast errors are dominated by their systematic component.

In the first winter forecasts there was evidence of real skill to day 15 and possibly to day 20 in middle latitudes. On the other hand, for the second and third winter, skill was confined to the medium-range forecast period. We have compared our results with those from other centres (National Meteorological Center, U.S.A. and the Meteorological Office in particular) and found that they are generally consistent. There were strong interannual variations in atmospheric predictive skill, and we found that this was correlated with a global version of the Pacific/North American index in the verifying analyses. We therefore conclude that interannual variability in extended-range skill was not primarily associated with any of the model reformulations. By contrast, it was shown that there appears to be some skill in the prediction of interannual variations in monthly monsoon winds and rainfall, particularly over Africa. On the other hand it was found that the large tropical systematic error in the model degraded tropical forecast skill significantly to the extent that most of the skill of monthly-mean fields was associated with the first 20 days of prediction.

For extratropical forecasting, more integrations are needed to establish the reliability of skill out to 15 days. As discussed in B, these integrations should be made in the context of multiple Monte Carlo (ensemble) integrations. On the other hand, for the tropics, multiple integrations are less important. Of more interest would be longer (say seasonal)
deterministic integrations with prescribed (observed and persisted) s.s.t.s. The ability to predict seasonal rainfall anomalies with a high-resolution NWP model using observed s.s.t.s would provide important motivation for the development of coupled ocean/atmosphere NWP models for operational seasonal tropical prediction. However, before such development takes place, significant reduction in model biases must be achieved.

REFERENCES


Blondin, C. and Böttger, H. 1987 The surface and sub-surface parametrization scheme in the ECMWF forecasting system: Revision and operational assessment of weather elements. ECMWF Tech. Memo. No. 135


ECMWF 1988 Reports of the working groups. Pp. 3–33 in 'Proceedings of the ECMWF workshop on predictability in the medium and extended range, 16–18 May 1988'


Newell, R. E. and Kidson, J. W. 1984 Assessment of the practical utility of extended range ensemble forecasts *ibid.*, 116, 89–126


Wallace, J. M. and Gutzler, D. S. 1981 Teleconnections in the geopotential height field during the northern winter. *ibid.*, 109, 784–812