Tellus (1986), *38A*, 25–41

Predictability studies with the ECMWF spectral model for the extended range: the impact of horizontal resolution and sea surface temperature

By U. CUBASCH and A. C. WIIN-NIELSEN,* *ECMWF, Shinfield Park, Reading, Berkshire RG2 3AX, England*

(Manuscript received December 27, 1984; in final form July 15, 1985)

ABSTRACT

Experimental 60-day forecasts have been run from observed data using the ECMWF operational spectral model with different resolutions. There still appears to be useful information beyond day 10 in the forecasts in spite of systematic errors. A more realistic sea surface temperature (SST) generally improves the forecasts in the long range. There are occasions when after a period of low predictive skill, the forecast recovers and reaches a level of skill similar to that found for short-range forecasts. The fact that this happens in the cases with the more realistic SST indicates that it is more than mere chance. The experiments, however, involve only a limited number of cases and are therefore not statistically significant.

1. Introduction

There is a growing body of evidence that in the long-term mean, the sea surface temperature (SST) influences the mid-latitude flow pattern (Cubasch, 1984; Shukla and Wallace, 1983). The predictability studies on a time range of up to 40 days described by Gilchrist (1977) indicated that dynamical models can provide a meaningful forecast for the largest waves, but it was unclear whether or not the forecasts improved using observed as opposed to climatological SST data. Medium-range forecasts show only a small response in mid-latitudes to changes of the SST (Arpe and Wallace, 1982). While the climate sensitivity studies had been carried out with low-resolution models (typically T21 or R15), the predictability studies have been performed with finer resolution models. In this study, we shall investigate the extent to which the horizontal resolution influences the impact of the SST on the forecast in the extended range, i.e., the transition period between a direct forecast and a climate simulation, i.e., from day 10 to day 60 in these experiments.

After a brief description of the model and the data used in these experiments (Subsections 2.1 and 2.2), the simulated long-term mean flow patterns of the atmosphere will be discussed (Subsection 3.1). This is followed by a suggestion about how extended range forecasts can be objectively verified (Subsection 3.2). Thereafter, two case studies will be discussed (Subsections 4.1 and 4.2).

2. The experiment

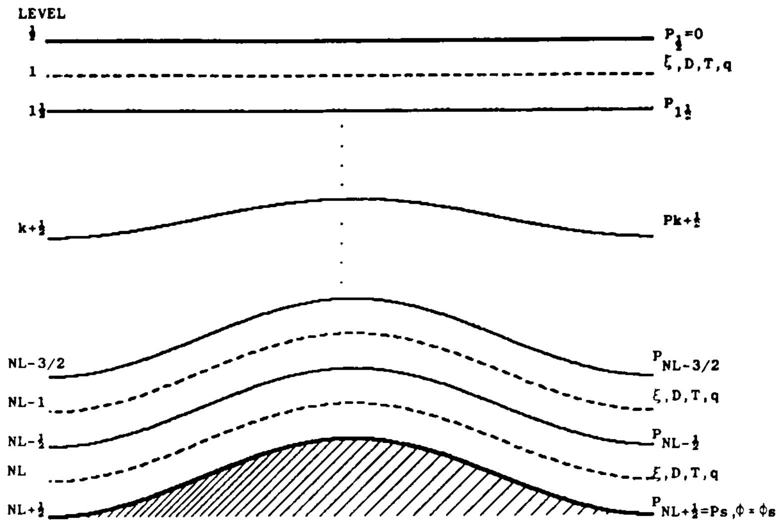
2.1. The model

The model used for all experiments is the operational ECMWF spectral model (spring 1984) (Louis, 1984); for further details see Fig. 1. It has been integrated to 60 days for the following horizontal resolutions: T21, T42 and T63 (i.e. triangular spectral truncation up to zonal wavenumber 21, 42 and 63). In all cases there were 16 levels in the vertical. All the models have been integrated with the same physical parametrization, which includes the diurnal cycle in the radiation calculations. All boundary fields were kept at their

* Visiting Scientist from Meteorological Institute, Copenhagen, Denmark.

$A_{k+\frac{1}{2}}$ (mb)	$B_{k+\frac{1}{2}}$
0	0
50	0
99	0.002
142	0.013
173	0.042
191	0.094
194	0.170
182	0.268
157	0.384
125	0.511
89	0.638
54	0.756
26	0.856
8	0.929
0	0.973
0	0.992
0	1

Vertical Coordinate Table (NL=16)



Disposition of variables in the vertical

	T21	T42	T63
Horizontal resolution Δx	5.625°	2.8125°	1.875°
Time step Δt	2400s	1600s	1080s

Dependent variables	ζ,D,T,q,ln(ps)	
Vertical coordinate	Hybrid, $P_{k+\frac{1}{2}}=A_{k+\frac{1}{2}}+B_{k+\frac{1}{2}}Ps$, details as above.	
Vertical representation	Finite-difference, energy and angular-momentum conserving.	
Horizontal representation	Spectral, with triangular truncation at wavenumber 63.	
Horizontal grid	96x192 points on a quasi-regular (≈1.875°) "Gaussian" grid.	
Time integration	Leapfrog, semi-implicit (Δt=20 min), time filter (V=0.06).	
Horizontal diffusion	Linear, fourth-order ($K=2\times10^{15} m^4 s^{-1}$).	
Orography	Grid-scale average from high resolution data set, enhanced by $\sqrt{2}$x(standard deviation of sub grid-scale orography), spectrally-fitted.	
Vertical boundary conditions	Kinematic.	
Physical parameterisation	(i)	Boundary eddy fluxes dependent on local roughness length and stability (Monin Obukov).
	(ii)	Free-atmosphere turbulent fluxes dependent on mixing length and Richardson number.
	(iii)	Kuo convection scheme.
	(iv)	Interaction between radiation and model-generated clouds. Albedo dependent on model snow cover.
	(v)	Large-scale condensation when grid-square saturated. Evaporation of precipitation.
	(vi)	Computed land temperature, no diurnal cycle.
	(vii)	Computed soil moisture and snow cover.
	(viii)	Fixed, analysed sea-surface temperature.

Fig. 1. The ECMWF operational forecast model.

initial value. Contrary to the operational forecasts, the orography used in these experiments was the "mean"orography (Tibaldi and Geleyn, 1981).

2.2. The data

The model integrations started from observed initial data in both sets of experiments, which for the first experiment were provided by the FGGE analysis. These data fell in the same time period as the ones used by Bengtsson (1981) for a case study of blocking. Two cases were considered:

(a) the SST was set to its climatological values (Alexander and Mobley, 1974): the results are marked "CLIM" in the figures;

(b) the SST was set to the 10-day mean value observed at the end of January and the beginning of February 1979 (FGGE data): the results are marked "OBS" in the figures.

The second experiment was started from operationally assimilated data for January 1983, which coincided with the mature phase of the 1982/83 El Niño event. Three cases were considered:

(a) with the climatological SST; marked "CLIM" in the figures;

(b) with the climatological SST plus the observed SST anomaly in the central Pacific for January 1983 (Reynolds, 1984); marked "POS" in the figures;

(c) with the climatological SST minus the observed SST anomaly in the central Pacific for January 1983 (Reynolds, 1984); marked "NEG" in the figures.

Case (c) was performed in order to determine the extent to which an erroneous SST influences the forecast quality and to obtain a better estimate of the natural variability of the model atmosphere.

3. Evaluation

3.1. The mean behaviour of the model atmosphere

A factor which limits the predictability in the extended range is the tendency of the model flow to become too zonal (Fig. 2). This behaviour seems to be independent of the model type, but dependent on the model resolution (Von Storch et al., 1985). Fig. 3 shows the deviation of the model climate from that observed during January for different model resolutions. The flow pattern for T21 and T40 have been obtained from integrations lasting several years performed with the predecessor of the current operational model (for further details see Von Storch et al. (1985)). The T63 field has been obtained by averaging all days beyond day 30 for all T63 runs analysed in this study. The difference between the climate of the old and new model at the same resolution appears to be small. However, the difference between the model flow and that observed is quite large for all resolutions, and it increases rapidly when going from T21 to T40; between T40 and T63 this increase is fairly small.

The zonalization tends to obscure any information still contained in an extended range forecast. Therefore it appears to be advantageous to compare anomaly pattern rather than the actual fields.

Contrary to what is found for the height field, the temperature field error hardly alters with increased resolution (Fig. 4). Note that the temperature is overestimated over the land and underestimated over the sea.

3.2. Objective verification of extended range forecasts

The anomaly correlation coefficient for daily forecasts (Hollingsworth et al., 1980) is defined using the deviations of the forecast and the observed state from the long term mean climate as depicted for example in Fig. 2. Additionally, a 10-day mean anomaly correlation has been used to objectively verify the results in the mid-latitudes (20.5° N to 82.5° N). This particular correlation is defined in terms of 10-day means instead of 1-day values. The calculation of the 10-day mean correlations is based on datasets truncated at T21. A truncation at a lower wavenumber would give higher scores, but does not alter the relative scores between different experiments.

Since it is not known at which threshold a 10-day mean forecast can still be considered useful, and since there exist no indications whether the 60% threshold used to define the limit of usefulness for a single day correlation has any meaning for 10 day mean forecasts, a limit of skill has been derived from multi-year climate forecasts.

It was achieved by asking how much skill the model climate possesses in predicting the 10-day means in the period considered for this predictability study. The correlation coefficients for these 10-day periods (12 cases) have been averaged and to this mean value has been added 1.96 times the standard deviation (Kreyzig, 1972). The resulting

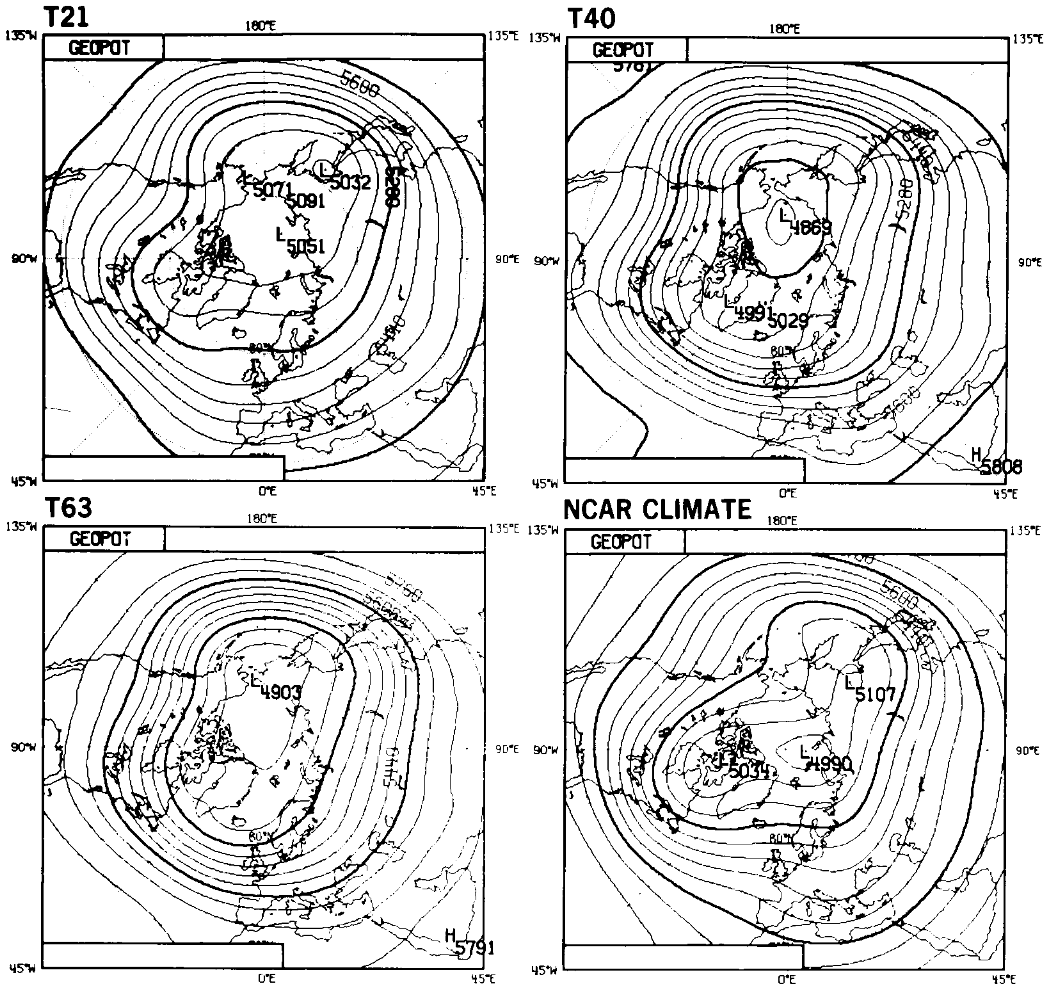


Fig. 2. The mean 500 mb height field (contour interval—80 m).

value gives an indication of what correlation coefficient could be achieved just by chance by the model climate. Any value higher than the threshold value would be generated by chance by the model climate in only 2.5% of the cases. This threshold value is designated as "NORM" in the diagram. Due to the similarity of the T40 and the T63 climate, the T40 climate has been used to define the threshold value for the T63 experiments as well. It was felt that using the T63 climate is inappropriate, since it is not independent because it is based on the experiments discussed here.

4. Results

The experimental results from the integrations for the experiments starting on 21 January 1979 and 1 January 1983 will be discussed separately.

4.1. The FGGE case

This period was originally selected because a warm SST anomaly stretched from the West African coast to the Caribbean Islands (Fig. 5), and this might have been responsible for the severe winter conditions in Europe during this time.

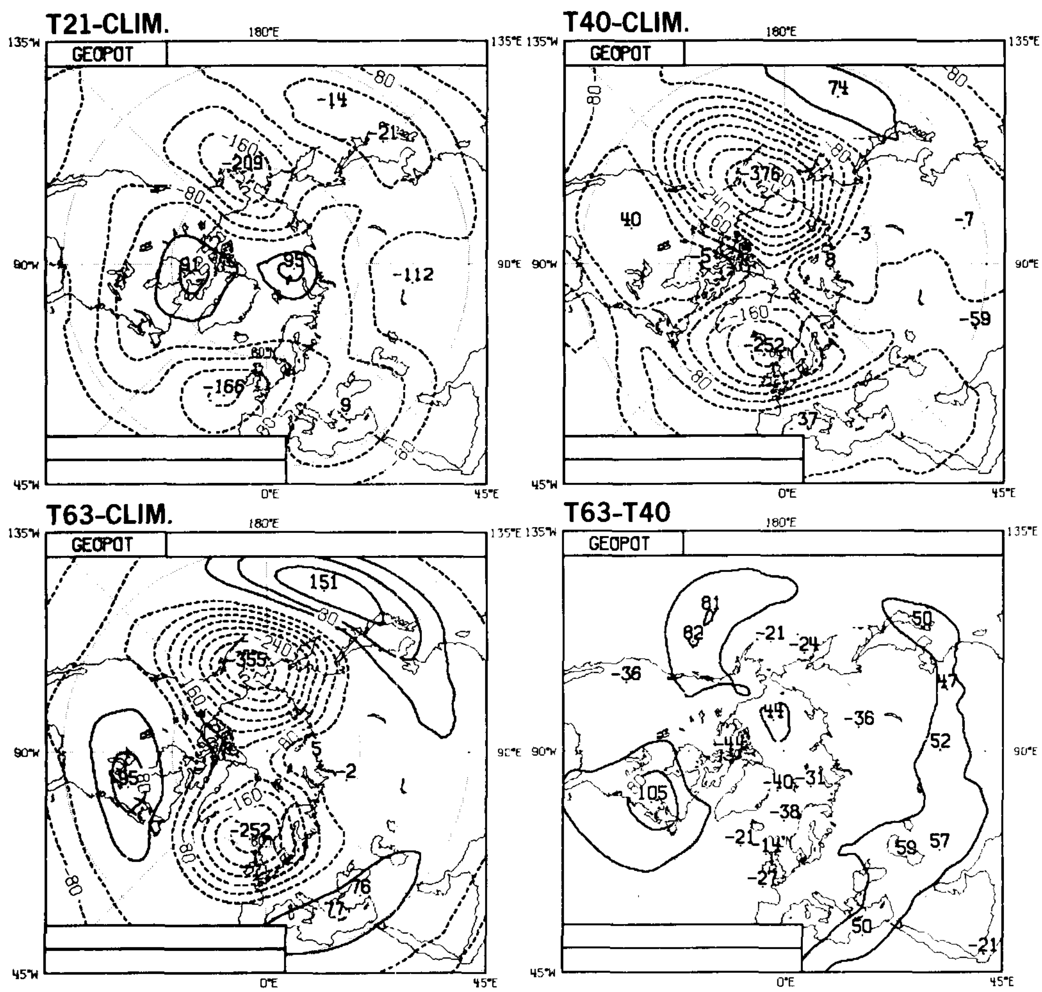


Fig. 3. The difference of the 500 mb height field from the observed January climate for different horizontal resolutions (contour interval—40 m).

However, temperature deviations from climate of similar magnitude occur during that period at various locations so that the impact of the Atlantic anomaly cannot be isolated. In particular the temperature increase east of New Guinea seems to have the largest impact on the rainfall and might be the true cause for any change in the circulation because it is situated in a highly sensitive area (Simmons et al., 1983). The mean response will therefore not be discussed further for this case.

The single day scores (Fig. 6) show for the longest waves a very small improvement of predictability (as measured by the 60% anomaly correlation) of 24 h with the T63 model when the observed SST is used; the T42 scores indicate a gain of about 36 h in predictability. The more realistic boundary condition allows the correlation coefficient to stay above the 60% threshold up to about day 14, which is quite remarkable. It is not clear why the T42 model performs slightly better than the higher resolution T63 model; it seems to be an exception rather than the rule (Jarraud et al., 1981).

For all resolutions, the 10-day mean scores (Fig.

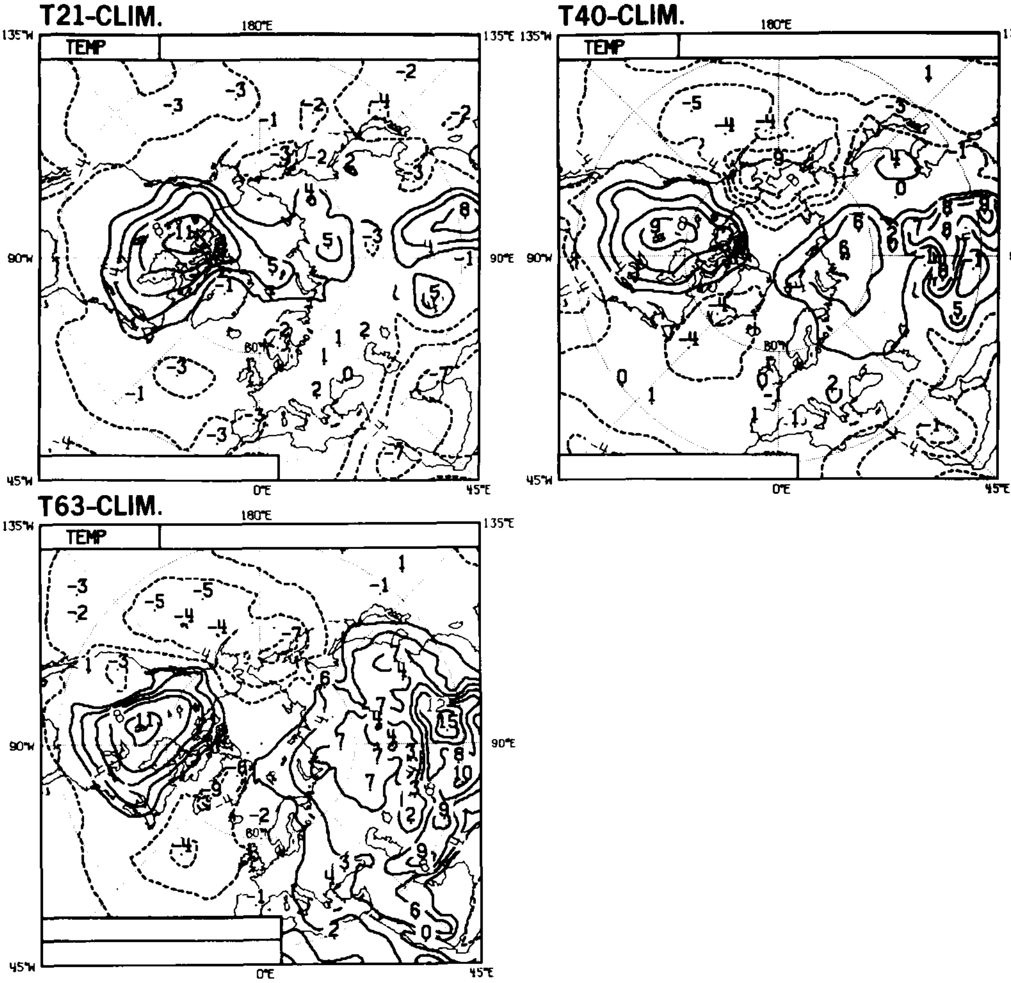


Fig. 4. The difference between the 850 mb temperature and the observed January climate for different horizontal resolutions (contour interval—2 °C).

7) show an improvement in the first 10 to 20 days and, for all resolutions except T63, even up to day 20 to 30 for the cases with the observed SST. The limit of predictability (according to the definition mentioned above) is reached at day 15 for T21 and at day 25 for T42 and T63. At the end of the forecasting period, the correlation increases again and reaches more than 60% for the T21 experiment at days 50–60 and 48% for T63 at days 40–50; again the observed SST gives the better scores (except at T42). This increase in the scores can also be found in extended range experiments performed by other groups (Miyakoda and Jin-Ping, 1982). One simple explanation might be that the ultra-long waves have been predicted with a wrong negative phase velocity. These waves would then have caught up with the observed long waves after a certain time. If a systematic error in the westward displacement takes place it would correspond to $-8°$ longitude/day for T63, i.e. 2 m/s at 45° N. However, such simple kinematic explanations are unlikely since the control experiment does not show the recovery to the same extent.

Figs. 8 displays the correlation coefficients for all 10-day means days 1 . . . 10, 6 . . . 15, . . .) of the runs with observed SST's against the integrations

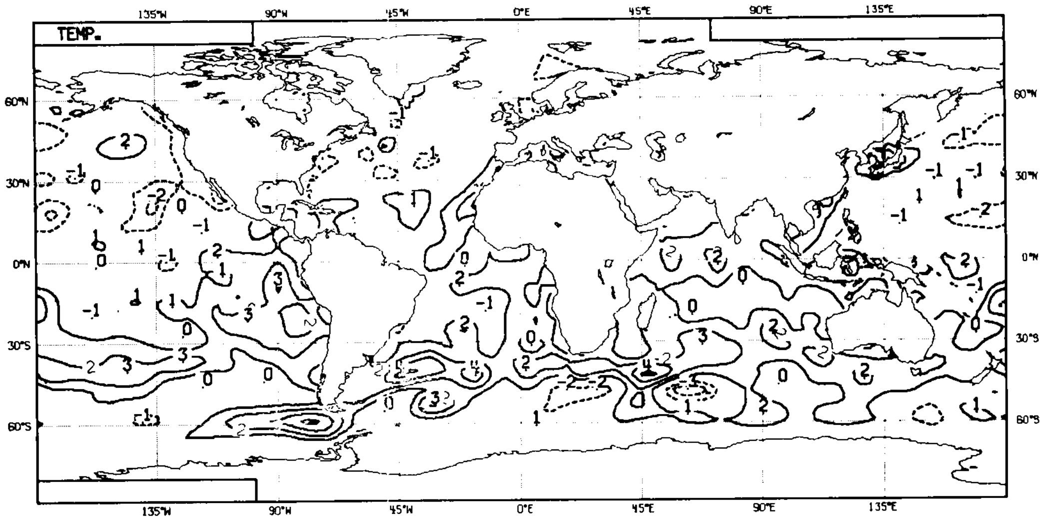


Fig. 5. The difference between the observed sea surface temperature during the first 10 days of February 1979 and the climatological February sea surface temperature (contour interval—1° C).

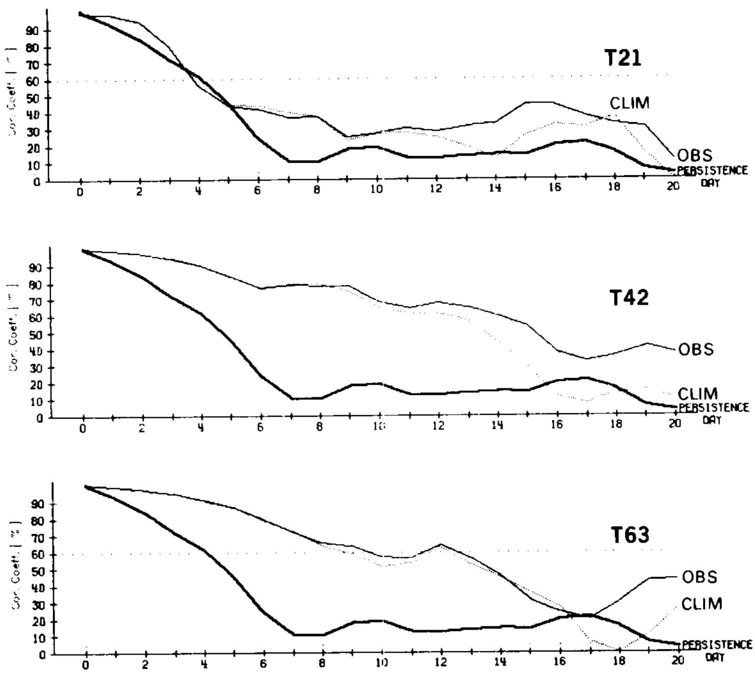


Fig. 6. The anomaly correlation for the 500 mb height field and wavenumbers 1–3 in the latitude belt from 20.5 to 82.5° N for the FGGE case.

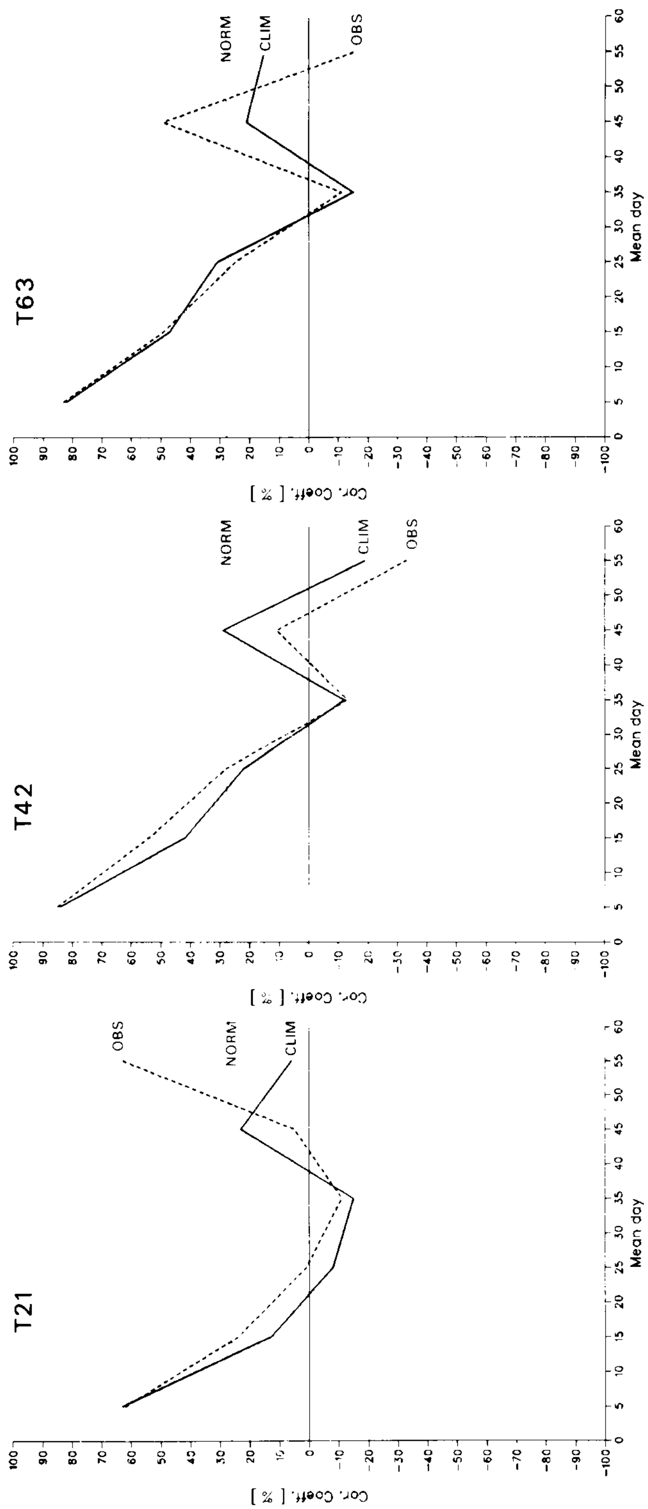


Fig. 7. The 10 day mean anomaly correlation for the 500 mb height field for the FGGE case.

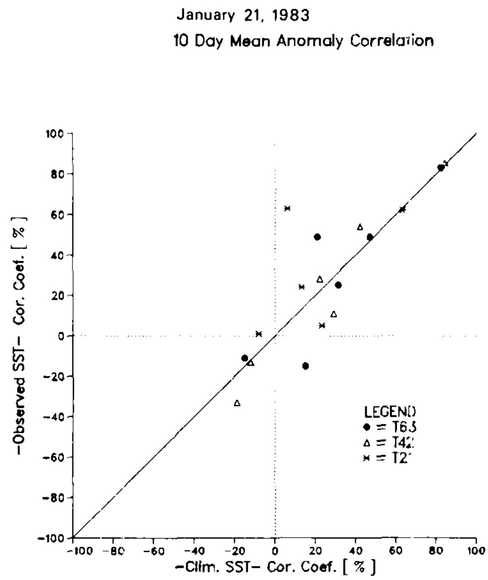


Fig. 8. Intercomparison of the skill of the 10-day means from the experiments with observed and climatological SST for the FGGE case.

with climatological SST's. Since correlations below 0% indicate forecasts without any practical use they should be ignored. From the remaining points in the scatter diagram it becomes clear that the integrations with the observed SST score higher than the ones with climatological SST. One can conclude that a more realistic SST distribution may have a beneficial impact on the quality of the forecast at the end of a medium-range forecast, as well as for extended range predictions.

4.2. The El Niño case

4.2.1. Long-term mean behaviour. In order to understand the results of this particular case it is necessary to investigate how well the anomaly is resolved at the various resolutions. As becomes clear from the temperature anomalies shown in Fig. 9, the interpolation to a T21 Gaussian grid reduces the observed maximum of the SST anomaly by about 1K, and even the T42 model experiences an underestimated temperature maximum. This fact has serious consequences for the strength and the location of the simulated rainfall, which is strongly dependent on the maximum of the absolute SST (Shukla, 1984, personal communication). The rainfall (Fig. 10) increases for the T21 simulation only by about 17 mm/day for the anomaly run while it reaches 36 mm/day in a 40 day mean for T63. This rainfall increase is larger than what one would obtain by simply increasing the horizontal resolution, as comparisons between the control experiments show. The location of the rainfall maximum for the higher resolution models is shifted towards the east.

Compared to the observed anomaly of the outgoing long-wave radiation (Arkin et al., 1983) which can be used to estimate the rainfall change in the tropical regions (Arkin, 1984), the rainfall band seems to be too confined to the equator in the simulations and a tail stretching from the equator eastwards to the southern hemisphere cannot be found in any simulations. The lack of this tail is a typical feature of the operational ECMWF model in the later stage of a medium range forecast (Molteni and Tibaldi, 1985).

As for the rainfall, the mean height field response is strongly dependent on the resolution (Fig. 11). In mid-latitudes the deviation in the 500 mb height field hardly exceeds 100 m for the T21 simulation, while it is more than twice as large for the higher resolutions and appears to be largest with the highest resolution. The T63 and T42 integrations simulate a deepening of the Aleutian low and confirm the observations of Bjerknes (1966, 1969), as well as the results of long integrations by Cubasch (1985). The T21 model produces a weak region of pressure increase in this region and it is suspected that this is noise rather than a true response to the anomaly since it also appears in the run with the negative anomaly. The response for the T63 simulation is of larger extent than for any other resolution. The response to the negative anomaly is not the inverse of the one of the positive anomaly; it appears to be rather similar but with decreased amplitude. A similar response to a negative anomaly has been reported by Cubasch (1985). The observed 500 mb height field anomaly for that time period bears an appreciable similarity to the T63 simulation for the northern hemisphere.

4.2.2. Predictability. The impact of the SST anomaly on the predictability appears to be small during the first 10 days for all resolutions for the single day anomaly correlations (Fig. 12). This confirms the results obtained by Arpe and Wallace (1982). The T21 run recovers after about day 14 to reach a correlation of about 70% at day 18. Between day 10 and day 40, the T42 experiment

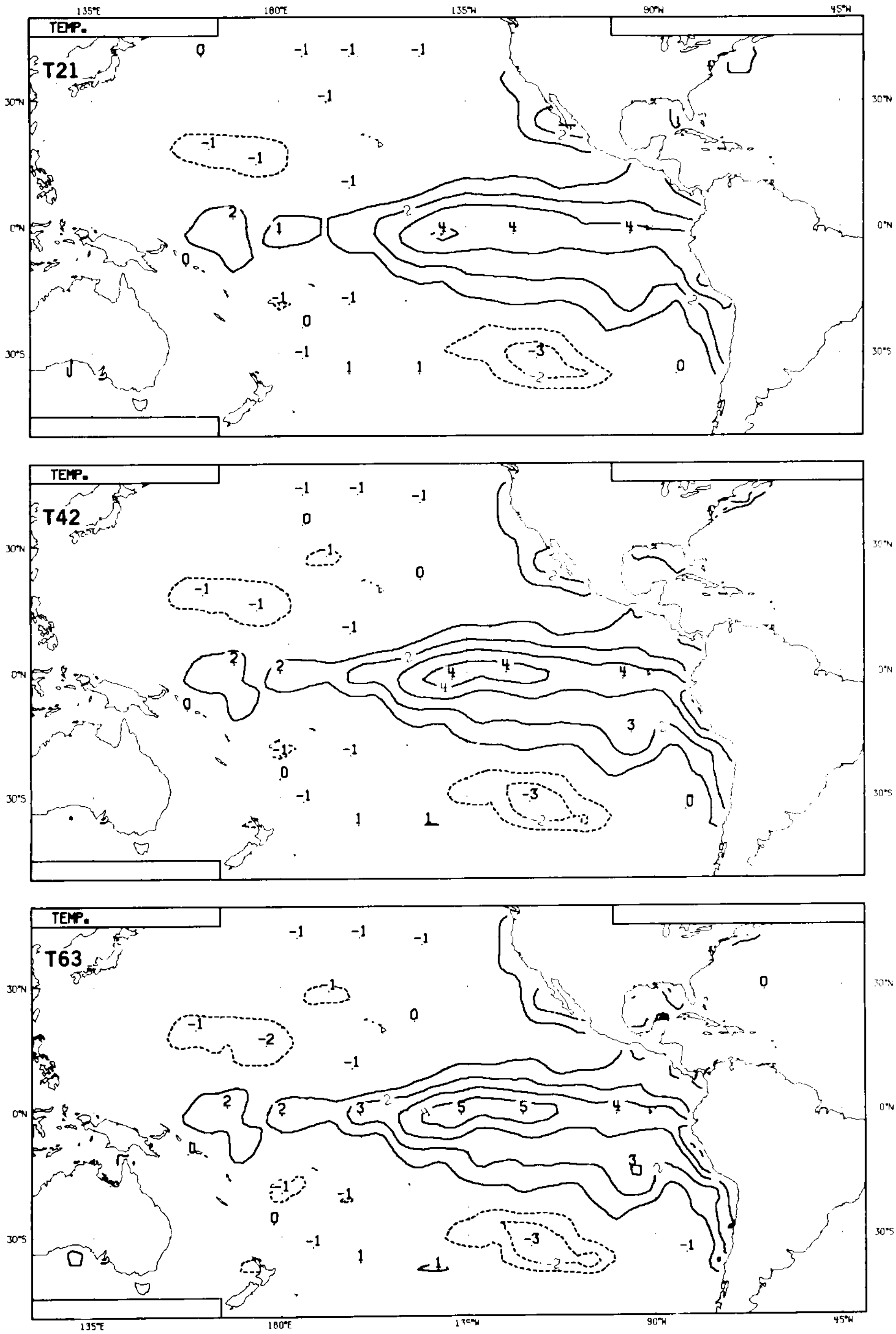


Fig. 9. The observed El Niño anomaly as seen by the different resolutions (contour interval—1° C).

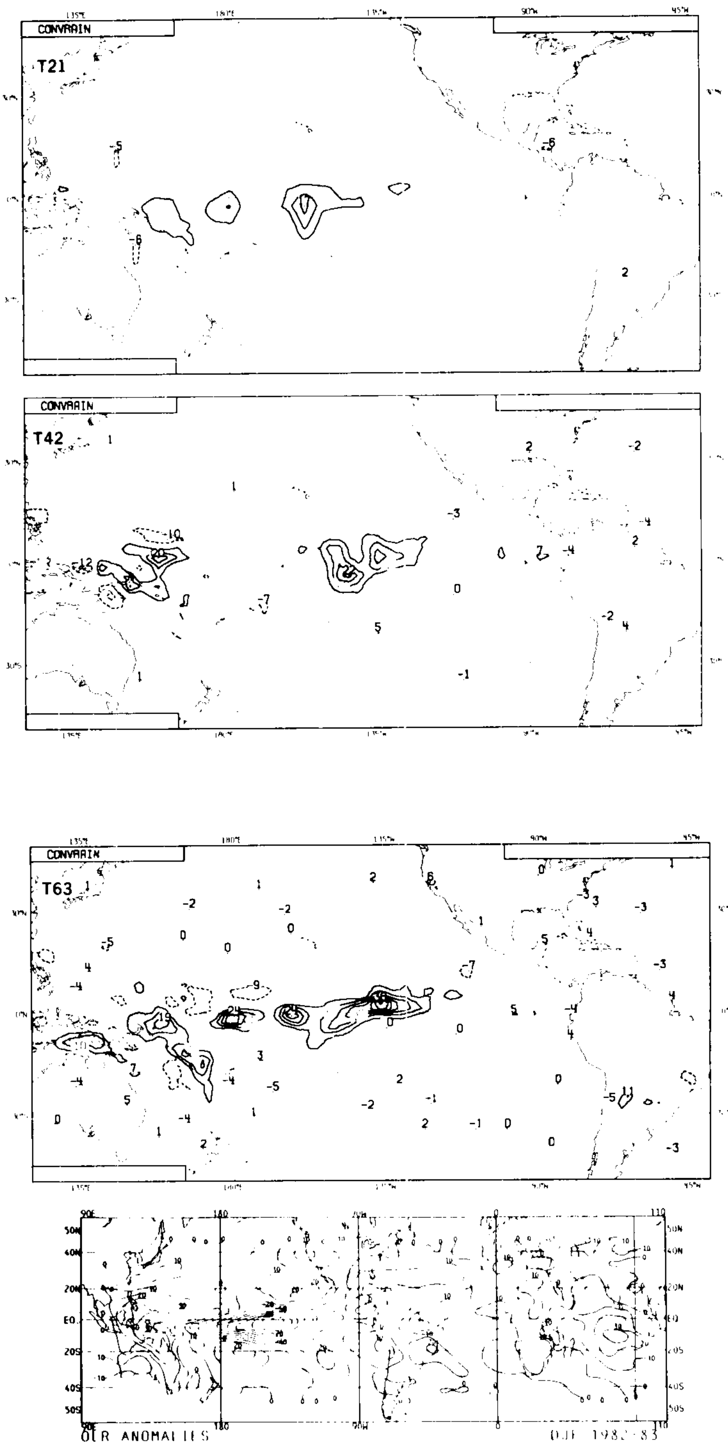


Fig. 10. The difference between the precipitation from the experiment with the positive El Niño anomaly and the control (contour interval—5 mm/day), and the observed outgoing longwave radiation anomaly (Arkin et al., 1983).

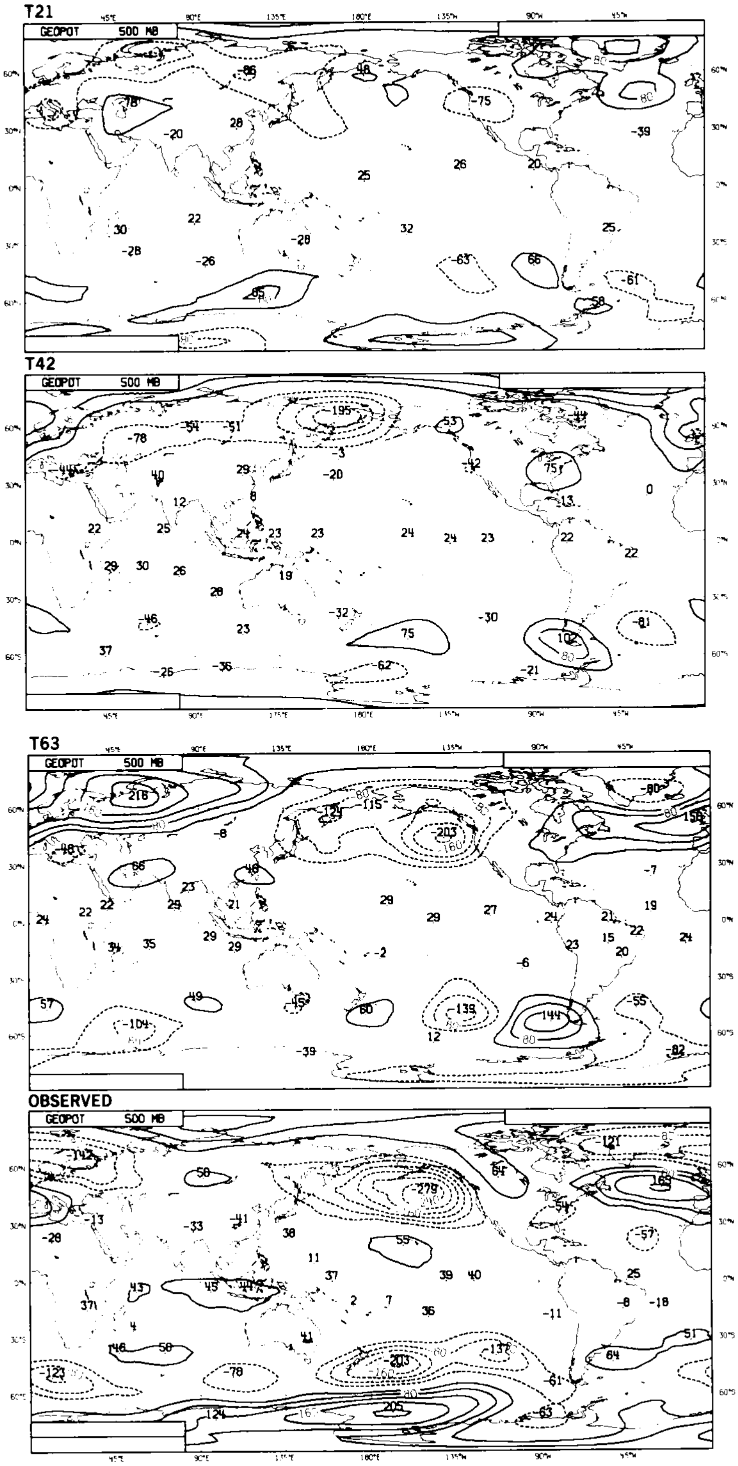


Fig. 11. The anomaly of the 500 mb mean height field averaged for days 11 to 40 for the El Niño case (contour interval—40 m).

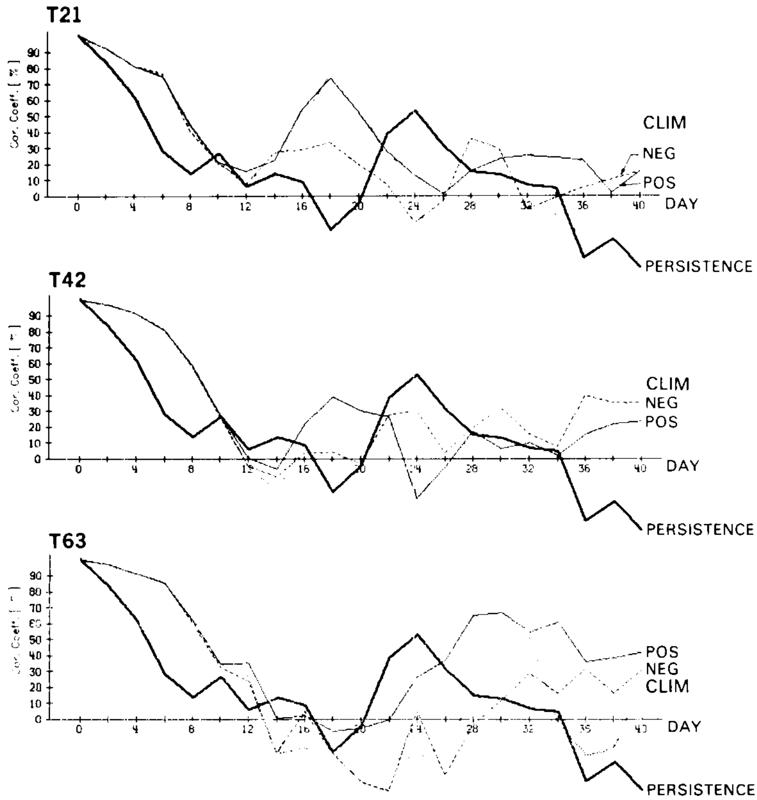


Fig. 12. The anomaly correlation for the 500 mb height field and wavenumbers 1–3 in the latitude belt from 20.5 to 82.5°N for the El Niño case.

shows no increase of the scores above 60%, while in the experiments with T63 the maximum correlation coefficient of about 65% is reached between days 28 and 30. This high score for T63 in the later part of the forecast must be more than mere coincidence since it happens only in the case with the observed positive anomaly, i.e., in the run with the most realistic boundary condition. The most striking aspect of this experiment as seen from Fig. 12 is that only the T63 integration is always better than the results for the negative anomaly and only for the T63 resolution are the positive anomaly results always better than the results obtained from the climatological SST.

Up to day 15, the 10-day mean correlation (Fig. 13) decreases rapidly for all resolutions, but during the same period the runs with the observed anomaly are better than the climatological experiments, which in turn score higher than the experiments with the negative anomaly as long as the score does not drop below the threshold values defined in Subsection 3.2. The T21 experiments stay above this threshold for almost all the forecast, while the T42 goes below it at about day 16 for the control and at about day 22 for the case with the observed anomaly. The T63 climatological experiment loses its predictability at around day 13 while the experiment with the El Niño anomaly just reaches the threshold value, but stays above it for the rest of the forecast. The difference in the correlation between the T63 control experiment and the anomaly experiment is about 30%.

To illustrate this difference, the 500 mb height anomaly between days 21 and 30 have been displayed together with the observed anomaly in Fig. 14. In the simulations, an area with large height field anomalies can be found over Europe. These are a reflection of the tendency of the model

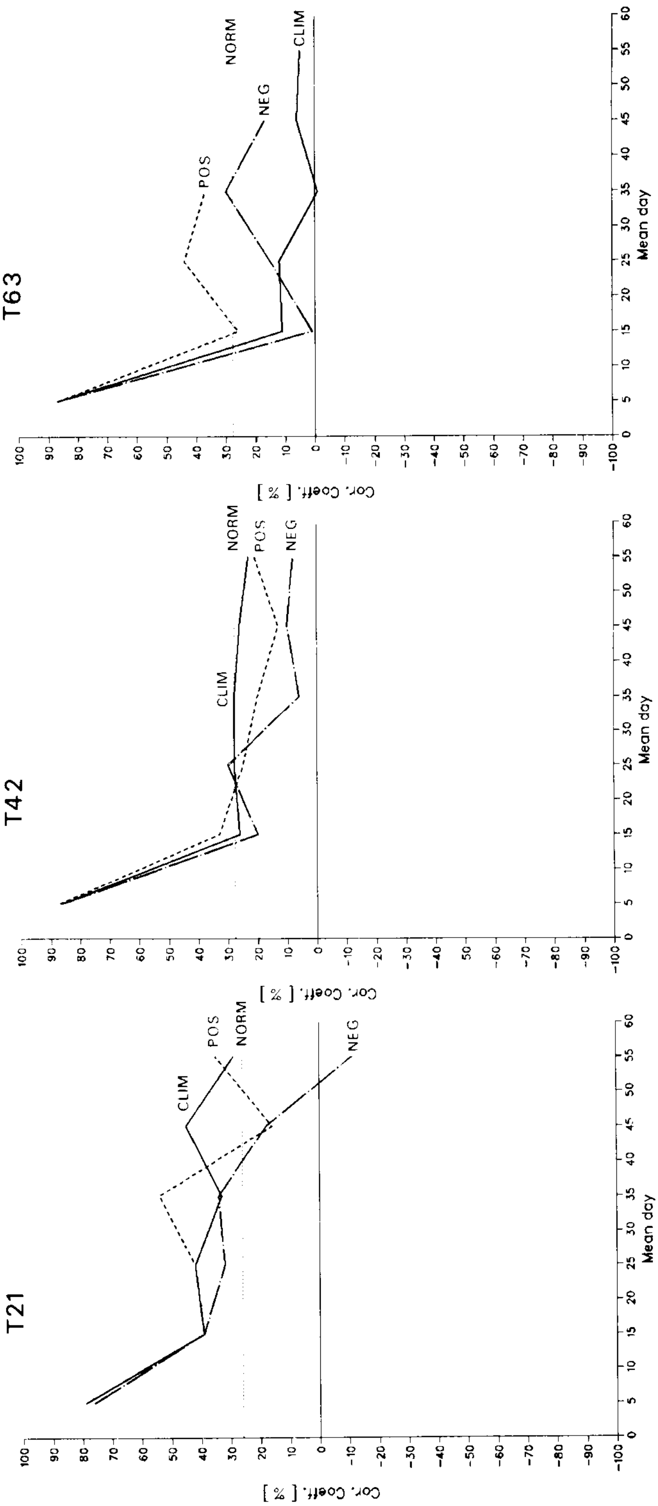


Fig. 13. The 10-day mean anomaly correlation for the 500 mb height field for the El Niño case.

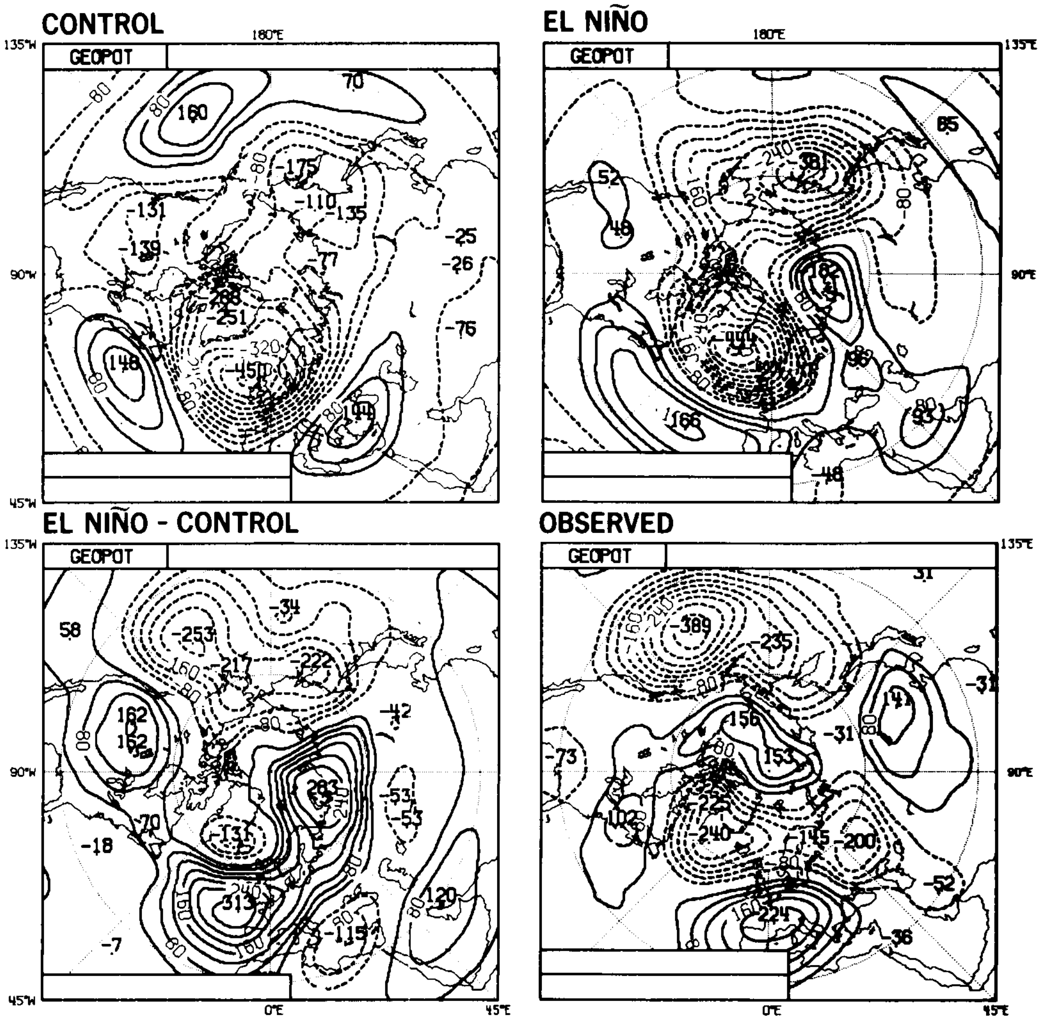


Fig. 14. The 500 mb mean height field anomaly for the El Niño case averaged for days 21 to 30 (contour interval—40 m); resolution: T63.

to zonalize the flow. However, the simulation with the SST anomaly generates a decrease of pressure in the north-east Pacific and a rise of the pressure over Europe, which coincides with observed pressure anomalies. A similar response of the model atmosphere had been found by Cubasch (1985) and resembles, at least in the Pacific region, the response of a linear model to tropical forcing. The scatter diagram (Fig. 15) shows a distinct improvement for the integrations with the more realistic SST anomaly, but mainly due to the results obtained from the T63 integrations. It might therefore be that the resolution used by Gilchrist (1977) was too coarse to exhibit any effect on the SST.

5. Concluding remarks

Before summarizing the SST experiments, it has to be stressed that all results apply for single case studies and their general validity has not been

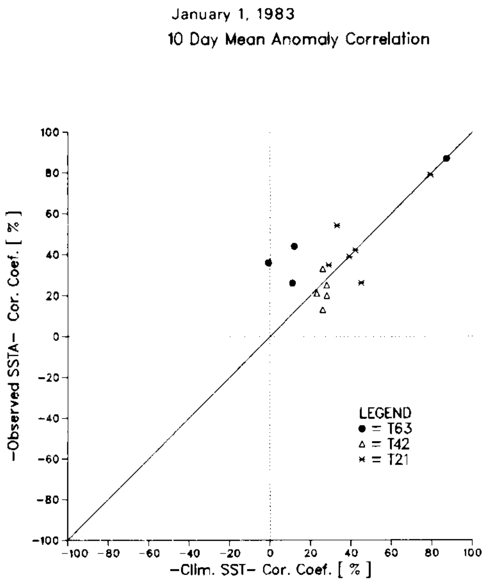


Fig. 15. Intercomparison of the skill of the 10-day means between the experiments with the positive anomaly and the climatology SST for the El Niño case.

established. However, they point in a direction along which such investigations could proceed.

- The SST has an impact on the forecast in mid-latitudes which becomes more important as the prediction progresses. The impact during the first 10 days seems to be only marginal.
- The low-resolution model has some difficulties in resolving SST anomalies. The impact of SST anomalies is therefore easier to detect in high-resolution models.
- The skill score of most experiments is higher when a more realistic SST distribution is used.
- The skill score of the experiments increases after about 15 days and can reach the level of direct predictability (i.e. a correlation coefficient of more than 60%). This means that some forecasts show predictive skill between day 20 and 30, even though they did not have any skill between day 10 to 20. The reasons for this behaviour are not known, but it appears to be systematic and can also be found in similar studies of other investigators. It is more pronounced using a more realistic SST distribution.
- The useful information contained in extended range predictions is obscured by systematic errors such as the flow becoming too zonal. This zonalization becomes worse with increased resolution. It seems, however, to approach an asymptotic value. The T63 model climate is about as zonal as that of a T42 model simulation.

REFERENCES

Alexander, R. C. and Mobley, R. L. 1974. Monthly average sea surface temperatures and ice pack limits on a 1 degree global grid. Rand Report R 1310–ARPA.

Arkin, P. A. 1984. An examination of the southern oscillation in the upper tropical and subtropical windfield. Ph.D. Thesis, University of Maryland, USA.

Arkin, P. A., Kopman, J. D. and Reynolds, R. W. 1983. 1982/83 El Niño/Southern Oscillation event quick look atlas. NOAA/National Weather Service, NMC/CAC, Washington, D.C. 2033.

Arpe, K. and Wallace, J. M. 1982. Tropical Pacific sea-surface temperature anomaly experiments with the ECMWF medium range forecast model. *Proceedings 7th Climate Diagnostics Workshop*, NOAA, 340–348.

Bengtsson, L. 1981. Numerical prediction of atmospheric blocking—a case study. *Tellus 33*, 19–42.

Bjerknes, J. 1966. A possible response of the atmospheric Hadley circulation to the equatorial anomalies of the ocean temperature. *Tellus 18*, 820–829.

Bjerknes, J. 1969. Atmospheric teleconnections from the equatorial Pacific. *Mon. Wea. Rev. 87*, 163–172.

Cubasch, U. 1983. Sensitivity of the ECMWF model to changes in resolution. ECMWF Workshop on "Intercomparison of large scale models used for extended range forecasts". 30 June–2 July 1982, ECMWF, Shinfield Park, Reading, UK, 435–453.

Cubasch, U. 1985. The mean response of the ECMWF global model to the El Niño anomaly in extended range prediction experiments. *Atmosphere-Ocean 23*, 43–66.

Gilchrist, A. 1977. An experiment on extended range prediction using a general circulation model and including the influence of sea surface temperature anomalies. *Beitr. Phys. Atmos. 50*, 25–40.

Hollingsworth, A., Arpe, K., Tiedtke, M., Capaldo, M. and Savijärvi, H. 1980. The performance of a medium-range forecast model in winter: impact of physical parameterization. *Mon. Wea. Rev. 108*, 1736–1773.

Jarraud, M., Girard, C. and Cubasch, U. 1981. Comparison of medium-range forecasts made with models using spectral or finite difference techniques in the horizontal. Techn. Rep. No. 23, ECMWF, Shinfield Park, Reading, UK.

Kreyzig, E. 1972. *Advanced engineering mathematics*. J. Wiley and Sons, N.Y., 866 pp.

Louis, J.-F. (Editor) 1984. The ECMWF forecasting model—physical parameterization. ECMWF Research Manual Vol. 3, ECMWF, Shinfield Park, Reading, U.K.

Miyakoda, K. and Jin-Ping, C. 1982. Essay on dynamical long-range forecasts of atmospheric circulation. *J. Meteorol. Soc. Japan 60*, 292–307.

Molteni, F. and Tibaldi, S. 1985. Climatology and systematic error of the ECMWF rainfall forecast. Tech. Rep. No. 51, ECMWF, Shinfield Park, Reading, U.K.

Reynolds, R. W. 1984. The 1982/83 El Niño sea surface temperatures. NOAA Climate Analysis Centre, NMC, Washington D.C., USA.

Shukla, J. and Wallace, J. M. 1983. Numerical simulation of the atmospheric response to equatorial Pacific sea surface temperature anomalies. *J. Atmos. Sci. 40*, 1613–1630.

Simmons, A. J., Wallace, J. M. and Branstator, G. W. 1983. Barotropic wave propagation and instability and atmospheric teleconnection patterns. *J. atmos. Sci. 40*, 1363–1392.

Tibaldi, S. and Geleyn, J.-F. 1982. The production of a new orography, land sea mask and associated climatological surface fields for operational purposes. ECMWF Tech. Memo. No. 4, ECMWF, Shinfield Park, Reading, U.K.

Von Storch, H., Roeckner, E. and Cubasch, U. 1985. Intercomparison of extended range January simulations: Statistical assessment of ensemble properties. *Beitr. Phys. Atmos. 4*, in press.