TECHNICAL MEMORANDUM

299

Resolution studies

M. Miller

Research Department

September 1999

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

1. INTRODUCTION

1.1. Background

There have been several changes to the resolution of the ECMWF forecast models over the twenty years since operational forecasting began. In April 1983, the original 1.875O 15-level finite-difference model was replaced in operations by a T63 16-level spectral model, with the extra level in the planetary boundary layer. Since then, the resolution of the model used for data assimilation and the deterministic ten-day forecast has changed as follows:

- In May 1985, the horizontal resolution was increased to T106.
- In May 1986, the vertical resolution was increased to 19 levels, with three extra levels in the stratosphere and the top level raised from 25 to 10 hPa.
- In September 1991, the horizontal resolution was increased to T213 and the vertical resolution to 31 levels, with layer spacing approximately halved in the free troposphere.
- In April 1998, the spectral resolution was increased to T319, using the 'linear-grid' option in which the computational grid remained at the (~60 km) resolution used previously for T2131.
- In March 1999, the vertical resolution was increased to 50 levels, with resolution unchanged below 150 hPa, a layer spacing of about 1.5 km over much of the stratosphere and a top level at 0.1 hPa.

In addition, following commencement of ensemble forecasting with a T63 19-level model in December 1992, the resolution of the EPS model was increased to T159 (with linear-grid option) and 31 levels in December 1996.

The earlier changes in resolution and related numerical experimentation were documented quite extensively in a series of papers. Computational limitations in the past did not allow comprehensive investigation of the benefits of increased resolution for data assimilation, but the sensitivity of the ten-day forecasts to resolution was quite widely examined. Objective verifications such as presented in Fig 1.1 (from ECMWF/SAC(87)5) indicated a clear improvement in forecast accuracy as resolution was increased. A larger improvement was seen near the surface than at 500 hPa, and this improvement was rather systematic early in the medium range, particularly in summer. The impression from subjective synoptic assessment was more decisive in terms of the benefit from increasing horizontal resolution. The more intense and highly structured weather systems found at higher resolution tended to bias objective scores in favour of lower resolution once general error sources had caused a serious degradation in the synoptic-scale evolution of the forecast. More rapid eddy growth rates in the higher resolution models increased their realism, but also meant that the doubling times of small errors became shorter.

The move to T213 31-level resolution was accompanied by introduction of a semi-Lagrangian scheme for treating advection in the model. Teething problems in the formulation of the scheme and in cloud-radiation interaction in the upper troposphere led to initially disappointing performance beyond the early medium range. However,

1 In this paper, T159, T319, T511 and T639 are all 'linear grid' truncations i.e. the subscript 'L' is dropped for convenience.

a set of test cases reported by *Ritchie et al.* (1995) indicated clear medium-range benefit of this resolution increase once the problems had been addressed.

A number of recurrent topics emerge from the earlier horizontal resolution studies:

- particular sensitivity to resolution of tropical cyclone tracks and precipitation forecasts;
- benefits of improved orographic representation at higher resolution;
- increased synoptic-scale variance and gravity-wave activity as resolution is increased (and horizontal diffusion decreased);
- worsening of spin-up problems at higher resolution.

The studies of higher horizontal and vertical resolutions also exposed several instances of computational instability or noise, particularly when carried out using the longer time-steps enabled by the semi-Lagrangian method. They prompted a steady refinement of the numerical algorithms used in computing the model's tendencies from dynamical and parametrized physical processes.

Operational experience after the December 1996 increase in EPS resolution (and accompanying increase in ensemble size) showed that the change was substantially beneficial. Verification statistics such as Brier scores exhibited a marked and sustained improvement, and very positive responses from Member-State users were reported at an Expert Meeting in June 1997.

1.2 Comparison of T319 and T159 forecasts

Comparison of the extensive sets of archived T159 EPS control forecasts and T319 deterministic forecasts provide additional evidence of the sensitivity of forecasts to model resolution. Such comparisons do not measure any sensitivity of forecasts to the resolution used in the data assimilation. From April 1998 to early March 1999, the EPS control and main deterministic models had respectively T159 and T319 horizontal resolutions and both had 31-level vertical resolution. Fig 1.2 shows 500 and 1000 hPa height verification statistics computed over the extratropical Northern Hemisphere for the period from 1 December 1998 to 28 February 1999. The left-hand panels show root-mean-square errors. The errors of the T319 forecasts are smaller than those of the T159 forecasts out to day 6 or 7, with more advantage for T319 at 1000 hPa than at 500 hPa. The right-hand panels show the root-mean-square anomalies of the forecasts and verifying analyses, computed with respect to the three-month mean climate from ERA-15. They indicate that the model has a tendency to have too high a variance, more so at T319 resolution than at T159, particularly at 500 hPa. This contributes to the larger root-mean-square errors seen later in the medium range with T319. It is not necessarily a fundamental problem of higher resolution per se, but is a deficiency in overall performance of the T319 version of the model whose correction would result in a modest improvement in the scores of the Centre's deterministic forecasts. A further conclusion is that deficiency in the spread of EPS forecasts cannot be ascribed to lack of variance in the forecast model at T159 resolution.

1.3. Increased vertical resolution in the stratosphere

The addition of three extra levels in the model stratosphere in May 1986, and the data assimilation changes it enabled, resulted in a clear improvement in stratospheric analyses and forecasts. A distinct beneficial impact on tropospheric forecasts in the later medium range was seen in a small (by today's standards) set of 9 forecasts.

Aspects of the recent increase in vertical resolution and extent of the model stratosphere are discussed in the newsletter article reproduced in ECMWF/SAC/28(99)3c. The article discusses some of the substantial improvements in stratospheric analyses and forecasts, and also shows improvement in anomaly correlations of 500 hPa height, based on a large set of 154 forecasts.

The operational 50-level structure functions were derived in a 'boot-strap' procedure. The first batch of data assimilations were carried out using structure functions and balance operators in the J_b term specified from a first 50-level application of the "NMC method" of comparing two- and one-day forecasts verifying at the same time. Initial conditions for these "NMC' method" forecasts were constructed by blending 31-level ECMWF analyses and UKMO stratospheric analyses. In the first batch of data assimilations the specified forecast errors were reduced at upper levels to counter unrepresentative effects of incompatibilities between the ECMWF model and the UKMO analyses. In the "boot-strap" procedure a second set of statistics was derived using the forecasts from the first set of 50-level assimilations.

The positive tropospheric impact was clearly seen in the first batch of forecasts. Forecast performance over 81 cases using the second set of statistics was further improved in the stratosphere, but was very similar in the troposphere to that obtained using the first set of background statistics, at least for verifications over the complete extratropical hemispheric regions. Clear indication of sampling sensitivity is illustrated in Fig 1.3, which shows 500 hPa height anomaly correlations over the Northern Hemisphere for four consecutive 30-day periods run with the original background statistics. A substantial positive impact of the 50-level resolution is seen in two of the periods, with much smaller impacts (one positive and one mostly negative) in the other two.

Final testing of the 50-level resolution was carried out using the second (boot-strapped) set of background statistics and a new fast radiative transfer model for the assimilation of TOVS radiances. 60 cases run in research mode showed small positive impact on tropospheric forecast scores over the Northern Hemisphere, and a small negative impact over the Southern Hemisphere. A further 67 cases run in pre-operational e-suite mode showed small negative impact over both hemispheres (ECMWF/SAC/28(99)3b).

These results serve a general purpose of warning of sampling problems even when objective verification scores are based on several months of data assimilation. More particularly they show that considerably more experimentation would be required to give a reliable answer to the simple question as to how much the tropospheric forecasts are improved by increased stratospheric resolution. The rather disappointing tropospheric e-suite performance compared to the original set of 50-level experiments might be due to sampling, to the use of the boot-strapped set of background statistics, or to the new fast radiative transfer model. In the event, operational implementation was justified by the improved stratospheric performance and by the essential role for the 50-level system in the beneficial assimilation of raw radiances from the TOVS and ATOVS satellites, which became pressing following the failure of one of the two TOVS satellites in late Feb 1999.

1.4. Structure of the paper

As discussed above, there is no doubt that important and systematic improvements to many aspects of the forecast skill have been achieved through increases in accuracy of horizontal representations both through better numerical algorithms and by using higher resolution. The impact of higher resolution can be identified through the basic forecast model, whether in deterministic or probabilistic mode or in extended integrations, or through its use in the inner or outer loops of the assimilation system. The following sections consider each of these in turn using results from specific higher resolution studies. With the exception of results presented in Figures 2.1, 3.6 and in Section 3.3.3, the resolution of the forecasts and outer analysis loops are the same. Some of these studies are necessarily limited in their extent and duration at this stage but are generally considered a reliable indicator of more extensive experimentation that will be done prior to operational upgrades. For presentation purposes there are many references to results without supporting figures in order to keep this report to a manageable size.

In experimentation at higher resolutions there are several choices to be made, such as the model timestep, the horizontal diffusion coefficients and, for the current convection scheme, the relaxation timescale. For the forecast model and outer loops of 4D-Var, the semi-Lagrangian scheme makes the choice of model timestep less demanding, and timesteps of 15 or 20 minutes for the higher resolution, and 60 minutes otherwise are used in the following studies. The convective relaxation timescale, in these tests, was scaled with spectral truncation. It was decided to use inverse timescales of the same order of magnitude for the highest wavenumber in determining the horizontal diffusion, thus progressively reducing the amount of damping, for a given wavenumber, as the resolution is improved.

2. 4D-Var data assimilation

Three different model resolutions are used in the 4D-Var data assimilation system. The one at which the analysis is produced at the resolution of the outer loops of the incremental procedure, and is the one at which the background fields are compared with the observations; thus far it has been kept equal to the resolution of the deterministic medium-range forecast model. The inner loop resolution is the resolution at which the 4D-Var incremental minimization is carried out, and is the one at which dynamical and linearized physical processes are used by 4D-Var to enforce flow-dependency in the use of observations; it is a crucial factor in the computational cost of running the 4D-Var data assimilation. So far the inner loop vertical resolution has been kept equal to that of the outer loop, whereas the inner loop horizontal resolution is, at present, T63 (with a quadratic collocation grid). The third resolution is the one at which the analysis and background errors are estimated; presently it differs from the inner loop by being at T42. The same resolution has been used in most of the existing experimentation on the reduced rank Kalman filter (RRKF) for the estimation of the singular vectors related to short-range forecast error growth, and this is consistent with the current singular vector calculations of the ensemble forecast system.

The meteorological performance and cost of the assimilation system is a function of these three resolutions, the volume of data processed, and the number of iterations used in the minimizations. The current setup is a compromise between costs and benefits. There is little evidence that significant benefits would result from higher resolution in the estimation of analysis and forecast errors. Resolution may be more important in the context of the RRKF, but computer limitations have so far prevented investigations of RRKF singular vectors with

resolutions higher than T42L31. On the other hand, substantial evidence is presented below that the assimilation system benefits from higher-resolution inner and outer loops.

In a data assimilation system the quality of the analysis depends on the accuracy of the short-range forecasts that are used to produce background fields. This is particularly important in data-poor areas. In 4D-Var the quality of the background has an additional impact on the accuracy of the linearization of the evolution of short-range forecast errors: with the possible exception of discontinuities in the physical processes, the incremental 4D-Var and RRKF algorithms are statistically better if the linearization errors are smaller. This implies that higher resolution outer loops will provide benefits that go beyond the mere improvement of background fields.

2.1. Impact of the resolution of the outer loop

The physical realism of the model fields has a direct impact on the accuracy of the comparison between background and observations. A sufficient outer loop resolution is needed in instances where the observations are affected by small-scale processes such as cloudiness, coastlines, orography or land surface fluxes. Currently most conventional data and satellite radiances that are sensitive to land surface properties or clouds are not used in the assimilation because of representativeness problems that would need higher-resolution outer loops in order to be solved.

A demonstration of the importance of high-resolution outer loops has been provided as a by-product of many assimilation experiments that used a lower horizontal resolution (T159) than operations (T319) in order to save on computer resources. The negative impact of the lower resolution is most clear at low levels in the Northern Hemisphere and at all levels in the tropical regions, as shown (Fig 2.1) which compares T159 forecasts made from T319/63 and T159/63 4D-Var assimilations.

Conversely, one gets a beneficial impact of increasing the horizontal resolution of the outer loops from T319 to T511 or T639, e.g. Fig 2.2. How much of the impact is provided by the forecast model itself, or by a better use of observations can be estimated by running T319 forecasts from the high resolution analyses. This indicates that, as measured by standard scores, a good deal of the impact is from the high resolution outer loops, although the high resolution forecast improves all weather-relevant parameters also. There is a substantial increase in the amounts of low-level data used, particularly over land, presumably due to the reduction in representativeness problems.

2.2. Impact of the resolution of the inner loop

Changes to the inner loop resolution have a significant forecast impact also, with a direct impact on the use of observed data. The more realistic stratosphere brought by the L50 vertical geometry has allowed the use of more realistic radiative transfer models for satellite radiance data, a feature that was necessary for the introduction of ATOVS level 1C data. There are clear indications that the higher PBL resolution in the L60 vertical geometry improves the representativeness of low-level observations. However, the most striking impact is obtained by increasing the horizontal resolution of the inner loops. In the current operational system, the information brought by the observations is effectively filtered to T63 and no small-scale feature from the background can be corrected. Although a T63 truncation may seem to be acceptable as judged by spectra of mass and wind errors, it is much too low for humidity generally, and in the vicinity of fronts or in developing storms, for instance. Additionally,

it is crippling the use of observations with a spatial resolution better than about 200 km, such as most satellite data, in two ways. Firstly, if the observations depict a small-scale correction to the first-guess fields, either the correction will be smoothed out, or sampling effects will lead to unpredictable spurious analysis increments. Secondly, whenever observations are closer to each other than the points of the T63 quadratic collocation grid used by the 4D-Var observation operators, correlated interpolation errors will lead to the wrong weight being given to the observations in the analysis, because 4D-Var currently assumes no observation error correlation in space.

Another consequence of increasing the inner loop horizontal resolution is the reduction of model errors and incremental convergence problems in 4D-Var, resulting in theoretically better analysis increments. This issue is not simple, though, because error evolution is less linear on smaller scales: ultimately, there will be a physical limit to the resolution at which the 4D-Var algorithm is going to work. This limit will depend on the length of the 4D-Var time window and on the amplitude of background errors.

There is some concern about the speed of convergence of 4D-Var: if it is slowed down by the changes to the shape of the cost function brought by the additional resolution, it is possible that the cost of minimization becomes prohibitive. Fortunately, many weeks of experimentation are available with T106 inner loop resolution, with both 6-hourly and 12-hourly assimilation periods, and results point to a significant positive impact on the forecast quality. For consistency, Fig 2.3 shows the results for the periods used in the T639 experimentation, and it is noted that these are less positive than several other test periods. There are signs of a slightly degraded convergence as illustrated in Fig 2.4a, since the same number of iterations was used for each inner loop and a lower curve means a more efficient minimization. However, the practical consequences for operations are likely to be small.

Fig 2.4b shows statistics from an assimilation run for 2 weeks, without satellite data (except SATOB winds), with 4D-Var time windows of 24 hrs starting from 3 pm every day. The selected observations are the surface pressure reports (in hPa) from drifting buoys that were actually used; this dataset exhibits an approximately constant density and quality throughout the day. It is approximately U-shaped, which is known to be a sign of model error, however the curvature is small compared to the mean value, which suggests that model error effects on the optimality of 4D-Var may be negligible on average.

Some observations cannot be used properly by the analysis because of inconsistencies between the models used in the inner and outer loops; this causes convergence problems in the incremental 4D-Var. The main source of inconsistency is the difference in horizontal resolution. It is particularly detrimental for low-level data, including surface pressure observations, and for the humidity analysis. Increasing the inner loop resolution from T63 to T106 leads to a substantial reduction of these inconsistencies. The convergence of the incremental is worse with 12- than 6-hourly cycling 4D-Var, but both systems benefit from higher-resolution inner loops. Preliminary results suggest that convergence will not be a major issue for 24 hourly 4D-Var either.

3. Medium-range deterministic forecasts

3.1. Impact of resolution in both the assimilation and the medium-range forecast

The impact of better resolution in the deterministic forecast together with the combined effects of higher resolution in the analysis system has been tested by running the full analysis /forecast system with a resolution of T639 (i.e. double the current operational resolution) for the outer loops of 4D-Var and for the medium-range forecast, and with T106 resolution for the inner loops of 4D-Var. This set of experiments were all with 31 levels. No attempt at running with T63 inner loops (as currently used operationally) was considered as the large 'mismatch' in outer to inner loop resolution was not considered sensible. The use of T106 inner loops (with Eulerian advection) was a necessary choice prior to the availability of the T159 option with Semi-Lagrangian advection. Two periods have been completed, one for 22/08/98-15/09/98, and the other for 01/12/97-28/12/97. These two periods have also been repeated with only the T106 inner loop resolution enhancement (see Sect 2.2). The Autumn period was notable for its hurricane activity in the Atlantic basin, and this aspect is discussed in Sect 3.3.2. The anomaly correlations for the total set of T639 are given in Fig 3.1 which show an encouraging overall impact. Throughout this paper, only 500 hPa height scores are presented for midlatitudes, however as might be expected, the higher resolutions have greater positive impact nearer the surface. More detailed scores and discussion are reserved for the experimentation that resulted from a reappraisal of research plans when it was decided to concentrate on a somewhat lower horizontal truncation. Several periods of T511 (with T106 inner loops) have been run, together with some short period case studies. This most recent set of experiments all have 50 levels. In view of the demands on manpower and computer resources, no further substantial studies were made to separate the impacts of resolution between model and assimilation. The T511 with higher resolution inner loops is the prime candidate for operational consideration. Three periods amounting to more than thirteen weeks in total have been run, for 22/8/98-16/9/98, 1/2/99-04/3/99 and 20/4/99-24/5/99. The scores shown in Fig 3.2 are for the combined set in order to have as large a sample as possible, however on the hemispheric scale each of the three periods is similar. The scatter plots, Fig 3.3, indicate the improvements in the medium-range are statistically significant at the 0.1% level. Some of the more systematic benefits are seen in the Tropics (also at T639), and Fig 3.4 shows tropical results from the Apr/May period with the verification scores recomputed against the experiment's own analyses.

3.2. Model characteristics

It is important to monitor the impact of resolution changes on basic characteristics of the model. The levels of EKE and its evolution showed only small differences in the T319 and T639 except in the Tropics where some modest improvement was evident (not shown). As referred to earlier, the T639 and T511 models used horizontal diffusion coefficients on ζ, D and T with an inverse timescale for the highest wavenumber of the same order of magnitude as currently used for the T319 i.e. implying less damping for a given spacescale. Fig 3.5a shows that for the T511 model these diffusion values steepen the spectral slope of KE at scales less than about 200 km. The spectra when explicit horizontal diffusion is set to zero is shown in Fig 3.5b. Experiments with diffusion on vorticity only and with a longer inverse timescale (by a factor of 3), Fig 3.5c, show that the only damping apparent is for scales less than about 100 kms. These studies continue. A general examination of physical fields and fluxes shows improved detail without problems, including stronger mesoscale features such as sea breezes and orographically influenced flows. Verification of weather parameters such as two metre temperature, cloud, precipitation and wind are all improved (not shown).

3.3. Examples of synoptic impact

This section contains some specific illustrations of the beneficial impact that is realized by higher resolution forecasting. Rather than select individual medium-range forecasts from the longer test periods, the emphasis is on the challenge of improving severe weather prediction, a subject identified for its importance by forecasters and included as such in the Centre's long-term strategy document. A limited number of case studies had been carried out earlier, notably for the severe event that struck parts of the Iberian peninsula on 06 Nov 1997. This explosive development killed 31 people and caused extensive damage. Fig 3.6a shows the verifying surface pressure and wind field at 00 UTC on November 6 1997, while Fig 3.6d shows the corresponding PV field on the 305K isentrope. The Centre's T213 forecast (Fig 3.6b) was the only operational European forecast to give warning of the event at D+2.5. However, the forecast substantially under-estimated the intensity of the actual pressure field (Fig 3.6b) and of the PV field (Fig 3.6e). The forecast at T639 resolution from the same operational analysis shows a much more accurate prediction of the surface low (Fig 3.6c) and of the PV field (Fig 3.6f). In this case the impact of T639 was a marked improvement in both the weather elements and the track and lifecycle of the system.

3.3.1 A winter lee cyclone over the Mediterranean

An example of severe winter weather in the Mediterranean area occurred during one of the test periods of the T511L50 system. The MSLP for noon on Feb. 10 1999 is shown in Fig 3.7 together with the D+5 forecasts from the T319L50 and T511L50 systems, with a well-defined low centred over Italy. Fig 3.8a shows the corresponding winds at 10 metres. Both verifying analyses are shown so as to emphasise the enhanced flow details in the higher resolution analysis, as might be expected, there are better defined and more intense forecast features in the higher resolution system in the medium-range, and since both systems are fully coupled to the ocean wave model, the impact of the stronger winds in the higher resolution version can be seen in the wave heights, both analysed and forecast, and substantial increases of more than 1 metre in forecast wave heights are evident over a large part of the Mediterranean, which agree better with the verifying analysis, Fig 3.8b.

3.3.2 Atlantic hurricanes

The 1998 Atlantic hurricane season included several very intense hurricanes affecting USA and central America. One of them, hurricane Bonnie, was chosen to study the impact of horizontal resolution on analyses and forecasts. Special attention was paid to the assimilation of scatterometer wind data. Hurricane Bonnie was active from 19–30 August, reaching category 3 with maximum average 10 metre wind speeds of 100 knots and lowest MSL pressure of 955 hPa. It caused substantial damage and created great havoc when it made landfall in North Carolina, and half a million people had to be evacuated. Figure 3.9 shows the operational MSL analysis and all available scatterometer data near Bonnie 30 hours before it made landfall. It is clear that the scatterometer data give a good description of the hurricane. We examine how well the assimilation system was able to utilize the satellite information and whether higher horizontal resolution is beneficial. The operational configuration consisted of T319 for the model and outer assimilation loop and T63 for the inner loop, T42 was used for background error calculations. The notation T319/T63/T42 will be used. Two higher resolution assimilation experiments were studied, with the only modification made in the first of these increasing the inner assimilation loop to T106 and performing background error calculations at T63. The increase in resolution of the background

error calculation is believed to have only a minor influence on the results, and the assimilation experiment is really a test of the impact of improved inner-loop assimilation resolution. The second higher resolution assimilation experiment increased the model and outer assimilation loop resolution from T319 to T639. This experiment was performed before it was decided to target T511 for the high resolution model, but results and conclusions are similar for the case studies here.

During 25 August 1998 hurricane Bonnie was 'captured' twice by ERS scatterometer data. Figure 3.10 shows the MSL analysis and the data actually used by the assimilation system for the two assimilations, T319/T106/T63 (Fig 3.10a/b) and T639/T106/T63 (Fig 3.10c/d). The black dot marks the officially reported position, which in this case is wrong! The ERS scatterometer data show that for both analyses the ECMWF analysed position is excellent. ERS scatterometer data is thinned to a 100x100 km resolution to reduce problems related to correlated data and to have a data resolution more in agreement with the assimilation resolution. In the vicinity of Bonnie all data is used, but near its centre some data is, by default, rejected if data or first guess winds are above 25 m/s. Hurricane Bonnie is more intense in the T639 assimilation, and in addition, more of the scatterometer data has been assimilated in the T639 assimilation. At T639 there is a better circular representation of Bonnie. In the T319 assimilation, several winds are rejected including one not used because the first guess wind is above 25 m/s even though the observation clearly represents the structure near the centre correctly, whereas at the higher resolution all winds are used (with two treated as slightly suspect). While the T319 blurs the structure and extends the high wind region too far out from the core region, the T639 represents the hurricane better as a more intense and smaller scale feature, with better use of the data in the vicinity of the hurricane centre.

It is important that the assimilation system identifies smaller-scale features in order that proper intensification can take place in the forecast. Fig 3.11a-c shows that in the above case study the analyses were almost identical at 0 UTC 25 August 1998 (with the same MSL central pressure) for the operational T319/T63/T42, T319/T106/T63 and T639/T106/T63. By 18 UTC (Fig 3.11d-f), the T319/T106/T63 has 1.5 hPa lower central pressure than the operational resolution, whereas the T639/T106/T63 has a 10.5 hPa lower MSL pressure. This reflects the improved assimilation as detailed above. Forecasts of Bonnie with T319 and T639 from the associated assimilation experiments show that tropical cyclones are clearly represented better at higher resolution. This is seen in the representation of the cloud band structure, the presence of a distinct hurricane eye, and a more intense wind field with a more correct description of high and low wind extremes around the tropical cyclone. The precipitation increases little but has a more well defined structure (not shown). Fig 3.12 shows some of these features for a 60 hour forecast valid six hours after the analysis shown in Figure 4.. In the T639 forecast the descent in the hurricane eye is clearly visible with the associated ascent around the core also better described. The cloud bands are also in better agreement with the satellite images (not shown). Winds reach 95 knots in the T639 forecast compared to 70 knots in the T319 forecast.

The T639 system does not just intensify cyclones and small scale features. In many cases the overall state of the atmosphere is better described by the T639 assimilation leading to different and improved synoptic scale flows. Fig 3.13 shows a 156 hour T319 forecast from the T319/T106/T63 analysis and a T639 forecast from the T639/T106/T63 analysis together with the verifying analysis (T319/T106/T63). In this case, the T319 forecast produces the intense cyclone while T639 correctly keeps it weaker.

3.3.3 The Brig floods

The Brig floods on 24 September 1993 fit into a typical pattern for heavy precipitation events on the southern side of the Alpine mountain range. Favourable conditions for convectively enhanced precipitation include sufficient low level moisture supply in a southerly airflow impinging onto the southern side of the Alps with a vertical stratification favourable for deep convection. The evolution of the southerly flow is the result of a southward penetration of a trough and its associated positive PV anomaly over Spain into the western parts of the Mediterranean.

To study the impact of higher resolution on a case from 6 years ago, initial conditions from a more recent analysis system were generated, and a period of 5 days before and including the rain event was re-analysed using 4D-Var with a model resolution of T319 for the outer loop and T106 for the inner loop. The synoptic situation of a high upper level PV anomaly, extending from southern Britain to the east coast of Spain is shown in Figure 3.14. The two forecasts for a range of 3 days (T319, lower panel and T639, middle panel, and verifying on 23 September) generate a rather weaker Mediterranean PV anomaly. However, both forecasts generate a similar moist southerly flow extending from the southern part of the Mediterranean Sea to the Alps as shown by a T639 forecast averaged over 36 to 72 hours, Fig 3.15. The details of the flow pattern close the Alps are however very much dependent on the horizontal resolution. Figure 3.16 compares the day-3 forecast of the low level wind, humidity and sea level pressure from the two forecasts in the vicinity of the Alps. The higher resolution model with a more realistic orography is better able to maintain high pressure near the Alpine ridge, and a stronger low level flow carrying moisture from the south and southeast into the Piedmont area and towards the Alps.

The rainfall in the Alpine area for this event is shown in Fig 3.17 as an average over two forecasts over a range of 48 to 72 hours, and has more structure for the high resolution model run. Areas of large enhancement correspond reasonably to areas where very large amounts of precipitation were observed. The differences in the rainfall pattern seem to be linked to the different moisture advection and a steeper slope of the orography in the T639 model. The verification statistics for the selected area confirm that the higher resolution model produces better precipitation forecasts, and both the bias and the standard deviation of model values compared to observations have been reduced. A secondary maximum of precipitation occurred close to Genoa of more than 100 mm/day and was also captured by the higher resolution forecasts; these predicted around 50% of the maximum coming from convective precipitation, hence simulating quite well the convective enhancement of rainfall during this flooding event.

4. Ensemble prediction system

Following the very positive results in the EPS of the 1996 resolution increase discussed in Section 1, preliminary results of the impact of a further increase of the horizontal resolution from T159 to T255 on the performance of the ECMWF Ensemble Prediction System (EPS) are now presented. (The impact of the inclusion of more vertical levels in the lower troposphere, will be the subject of future investigation.) Both the T159L31 and the T255L31 ensemble configurations have been run with 50 perturbed members starting from the same set of initial perturbations generated using T42L31 singular vectors, and including a stochastic simulation of random model errors due to parametrized physical processes with the same settings.

As part of an ongoing collaborative study, experiments have been run for 5 cases of heavy wintertime precipitation over the central part of the United States, which had a poor EPS precipitation forecast between forecast days 4 and 5 (see Table 2 for the initial dates). Ensemble performance is measured by the ensemble spread, by the percentage of times the analysis lies outside the ensemble forecast range, and by the error of the ensemble-mean forecast. The skill of probabilistic predictions of precipitation amounts, and of 500 hPa geopotential height and 850 hPa temperature anomalies are measured by the Brier Skill Score (which is essentially the mean-square error of the probability forecasts) and by the area under the Relative Operating Characteristic curve (ROC curve, which depends on the ratio of the hit and false alarm rates of the system). Due to the limited number of cases, the ensemble performance is assessed only over the Northern Hemisphere extra-tropical region (North of 30□N, hereafter NH).

4.1. Impact on 500 hPa geopotential height and 850 hPa temperature prediction

On average for the 5 cases, the T255 ensemble shows a 4% increase in spread from forecast day 2 onward for the 500 hPa geopotential height field, and a 2% increase in spread for the 850 hPa temperature (not shown). This increase in ensemble spread is good since it reduces the level of spread underestimation in the current system, and Table 1 shows the reduction in the percentage of times the analysis lies outside the EPS forecast range.

Fig. 4.1 shows the area under the ROC curve for predictions of temperature anomalies larger than 8 deg, and Fig. 4.2 shows the area under the ROC curve for geopotential height anomalies larger than 100 m. Results indicate a positive impact of the resolution increase.

	fc d+3		fc d+5		fc d+7	
	T159	T255	T159	T255	T159	T255
500 hPa geop height	19.4	16.6	12.7	11.0	9.2	8.4
850 hPa temperature	22.5	20.0	13.7	12.5	10.9	9.9

Table 1 Percentage of times the analysis lies outside the EPS forecast range, for 500 hPa geopotential height and 850 hPa temperature predictions for the NH, at forecast days 3, 5 and 7.

4.2. Impact on precipitation prediction

Figure 4.3 shows the area under the ROC curve for the probabilistic prediction of 5 and 20 mm/12h of precipitation, verified against 0-12h and 12-24h forecasts given by the ECMWF operational T319 forecast. This verification field can be considered a good approximation of the observed precipitation field. Results show that the resolution increase has a positive impact, indicating a gain in predictability of about 6 hours during the whole forecast range for the 5 mm/12h threshold, and a gain of about 12 hours during the first 4 forecast days for the 20 mm/12h threshold.

Since the 5 cases were selected for precipitation situations over central United States, precipitation prediction has also been verified for this region (20°N-50°N, 110°W-80°W; work is also in progress to compare the two ensembles with an analyzed field of precipitation. Figure 4.4 shows the area under the ROC for central United

States. Results confirm a positive impact of the resolution increase, although they show more variability than the hemispheric results since the area of interest is smaller.

Denoting by BS255 the Brier Score of the T255 ensemble, and by BS159 the Brier Score of the T159 ensemble, then a measure of the impact of the resolution increase on the ensemble performance is given by

$$BSS_h = \frac{BS_{255} - BS_{159}}{0 - BS_{159}}$$

i.e. by the Brier Skill Score BSSh of the T255 ensemble computed using the T159 ensemble as reference instead of climatology. A positive (negative) BSSh indicates that the T255 (T159) ensemble performs better. Table 2 shows the values of BSSh for predictions of more than 10 and 30 mm/24h in the area marked on Fig 4.5. Results indicate that for the lower threshold (10 mm/24h) the T255 ensemble performs better in three of the five cases, while for the higher threshold it performs better in four of them. Since the positive values are in absolute terms larger than the negative values, results indicate that on average the T255 ensemble performs better.

Figure 4.5 shows maps of probability of more than 10 and 30 mm/24h predicted by the two ensemble configurations for 5-day forecasts from 12GMT of the 7th of December 1998. The T159 ensemble gives no indication of the possibility of more than 30 mm/24h of precipitation in the area where it occurred, while the T255

Initial date	Area ave prec	BSS_h			
		5 mm/24h	10 mm/24h	20 mm/24h	30 mm/24h
98.11.08	2.4 mm/24h	-0.014	-0.038	-0.014	-0.018
98.12.07	3.8 "	0.249	0.132	0.085	0.073
98.12.28	5.0 "	0.097	0.175	0.181	0.110
99.01.18	6.6 "	-0.018	-0.017	-0.025	0.114
99.01.25	5.7 "	0.141	0.145	0.075	0.015

Table 2 Area average precipitation over central United States (20°N-50°N, 110°W-80°W), and Brier Skill Score BSS_h of the T255 ensemble (computed with the T159 ensemble as reference, see text for details) for the probabilistic prediction of more than 10, 20 and 30 mm/24h between forecast days 4 and 5.

Figure 4.6 shows maps of probability for a second case, with predictions started at 12GMT of the 28th of December 1998 and valid for the 2nd of January 1999. For this case, the T159 ensemble gives a 2-30% probability of more than 30 mm/24h of precipitation in part of the area where it occurred, and the better performance of the T255 ensemble is due to a better prediction of the area of intense precipitation.

4.3. Summary

For this limited sample, the increase in a horizontal resolution reduces the amount of spread underestimation, reduces the percentage of time the analysis lies outside the ensemble forecast range, and increases the gain in skill

of the ensemble-mean with respect to the control forecast. The largest positive impact of resolution has been detected for precipitation prediction. Considering the area under the ROC curve as a measure of skill, results indicate a gain in predictability of 6-12 hours on average over the Northern Hemisphere for 5-20 mm/12h thresholds. These studies are currently being extended to the European region prior to any consideration of operational implementation.

5. Longer integrations

The role of resolution in extended integrations where it is normal to consider temporally averaged fields is somewhat different from that in the forecast problem. In longer integrations the impact of resolution is more apparent in the model's ability to describe precipitation and cloud distributions, other orographically-dominated processes, blocking frequency, transient eddy fluxes and associated storm tracks. These, in turn, are linked with the model's systematic errors, a key issue in the skill of seasonal prediction. However, it is well accepted that a significant component of a model's systematic error is directly attributable to deficiencies (often subtle) in the parametrization of physical processes, and hence the following results partly reflect the state of development (i.e. the model cycle used) of the physics, and as will be seen, are relatively insensitive to resolution.

5.1. Seasonal forecasts

Sets of 6-member ensembles with horizontal resolutions of T63, T159 and T319 were run for one winter (1987/88) and one summer season (1987). The model cycle was 18R6 with 31 levels in the vertical. Winter runs were initiated on and around 1 November and the summer ones on and around 1 May. For each resolution, seasonal ensemble averages were computed and compared with ERA-15 fields. Whereas one winter and one summer are by no means sufficient to evaluate the models fully, they nevertheless give an insight on the influence of horizontal resolution on simulations for seasonal timescales. During the summer of 1987, a moderately strong El Niño event in the equatorial Pacific was observed, which subsequently weakened and almost disappeared in the following winter. The following provides a brief summary of some basic characteristics of these simulations.

5.1.1 Upper air fields

Fig 5.1 shows the DJF 1987 500 hPa height ensemble mean errors at each resolution. The mainly negative mean height errors that dominate over the northern hemisphere are generally reduced in both T159 and T319 when compared with T63. However, the westerly bias over the Atlantic has increased in T319 relative to T63 (but is reduced relative to T159). Compared with T63, errors in the tropics and southern mid-latitudes have also been reduced in both T159 and T319 models, but increased over Antarctica. In summer, a negative error over Antarctica in T63, and to a lesser extent in T159, switches to a larger positive error in T319 (not shown).

For both seasons, the winter hemisphere errors in zonally averaged u-wind deteriorate somewhat as the model resolution increases (Fig 5.2). Whereas in the northern hemisphere this error increase is only marginal, it is more marked in the tropics and southern hemisphere. However, errors in the summer hemisphere reduce as resolution increases. In the tropics, the wind error represents an increase in easterlies near the tropopause in T319; similarly, a large positive error between 25°S and 45°S seen in all resolutions is due to an overestimation of the subtropical jet. However, it is noted that in the zonally-varying wind error maps (not shown), the maximum error at 200 hPa is largest at T63 (20 ms-1), and smallest at T319 (16 ms-1).

The marked loss of zonally averaged eddy kinetic energy found in many low resolution GCMs is also found here at T63 in both hemispheres and both seasons (not shown). In the northern hemisphere, this loss has been gradually reduced as the model resolution increases. In the southern hemisphere, there is a net gain in eddy kinetic energy in T319 model relative to ERA, and the maximum error is confined to midlatitudes. This energy increase at T319 is larger in the southern summer (DJF) than in winter.

The erroneous wintertime cooling of the northern high latitude troposphere is gradually reduced from T63 to T319 (Fig 5.3 left column). In summer, the largest erroneous warming is found in T319 lower troposphere (Fig 5.3 right column). Thus, in both seasons, there seems to be a monotonic increase in the northern hemisphere lower tropospheric temperature as the model resolution increases; this has a beneficial effect on mean error in northern winter and a deteriorating effect in northern summer. A similar feature is also seen in the southern hemisphere, however, with much smaller amplitude. It is not clear whether these errors at low levels are linked to land surface processes since at 850 hPa, temperature errors are almost equally spread over both land and sea points. In addition, in both winter hemispheres, the errors in poleward eddy heat flux are lowest with T319 resolution in the tropospheric mid latitudes, the region where this flux is strongest (not shown).

5.1.2 Tropical circulation

Some details of the 500 hPa spatial distribution of the tropical vertical velocity are shown in Fig 5.4. In winter the best-defined area of ascent over the Pacific warm pool region is found in T319 with descent in the central equatorial Pacific in between the two branches of rising motion of the ITCZ and SPCZ respectively. In summer (right panels), T319 displays the most structured and detailed ascent over the Indian monsoon region, including the Bay of Bengal. The Atlantic ITCZ is strongest with T319 resolution as is the area of rising motion off the eastern coast of the U.S. over the Gulf Stream. This may reflect the better representation of hurricanes in a high resolution model. Inspection of velocity potential maps at 200 hPa for the N Hemisphere summer shows that the upper tropospheric divergence gradually increases with resolution over S Asia and all resolutions are stronger than in ERA-15.

5.1.3 Surface fields

The seasonal mean precipitation for the European region gradually increases with resolution both over most of the continent and parts of eastern Atlantic. In some cases the increase in precipitation is related to better resolved orographic features (e.g. the Alps, the Caucasus, the Norwegian, Scottish and Albanian mountains). In both seasons, large scale precipitation is responsible for most of the increase. The maximum in monsoon precipitation over the western India and the Bay of Bengal is almost doubled in T319 when compared with T63. Such a substantial increase is due to both better resolved orographic effects and an enhancement of the low level jet off the Somali coast in T319.

There is no major change in the net surface thermal balance over sea points as the resolution increases. In winter, the mean deficit is 12.6, 10.4 and 11.6 Wm-2 for T63, T159 and T319 respectively, while in summer there is an excess of 19.4, 21.2 and 23.9 respectively. However, for T319 a strong compensating effect takes place since the surface latent heat flux, thermal radiation and surface solar radiation change relative to the lower resolutions. Since the surface sensible heat flux remains essentially unchanged when compared with the other resolutions, the

overall effect is that the net surface flux in T319 is rather similar to that in T63 and T159. The surface wind stress in the north Pacific winter storm track is monotonically enhanced from T63 to T319 (not shown).

5.1.4 Summary

Based on this limited sample, the general impression is that there are no particularly unexpected differences in simulated climate between the three resolution considered and no clear pattern emerges as the horizontal resolution changes. For most parameters, circulation and error patterns are similar, while differences among the resolutions are sometimes (but not always) due to improved orographic representation. In many cases there are no monotonic changes as the resolution increases from T63 to T319 and it is unclear whether this is the consequence of a rather limited set of experiments. More extensive seasonal forecast experimentation with the coupled model of T159 has yet to be done.

5.2. AMIP-2 experiments

In the context of the model's use in ERA-40 and seasonal forecasting and also to assess the longer timescale stability of physical processes such as atmosphere-land interactions or in the stratosphere, 20-year T63 and T159 runs with model cycle 18R6, 60 levels in the vertical and observed SST have been completed, as part of the AMIP-2 initiative. The runs extend from December 1978 till December 1998 (February 1999 for T159). For the period December 1978 - March 1996, SST and sea ice extent were provided by PCMDI; from April 1996 onwards, the SST was taken from ECMWF archive. Setting up and running 20-year experiments required significant research and technical effort, and some modifications to the model code were necessary to define additional fields normally not computed or post processed by the IFS. Also, many technical changes were required, e.g. computation of monthly means, new field codes, treatment of accumulated fields, etc. Monitoring diagnostics were performed during the running of the experiments. For most fields monthly averages were computed and saved; however, some basic fields are saved at more frequent intervals. The task of analysing the results of both runs and comparing with ERA-15 is in progress, and only some basic diagnostics are presented here.

In both experiments, no major drift in the evolution of model fields has been detected. Fig 5.5 shows time evolution of 200 hPa zonally averaged monthly mean u-wind for T159 resolution. The model integration was stable, well within climatological bounds. Similar evolution is seen for the T63 run as well (not shown). In addition, from Fig 5.5 it can be seen that the model response to the forcing imposed by a large scale SST anomalies during the period 1979-1998 was an appropriate one. Interannual variation of zonal wind is quite realistic, in particular, for years with strong ENSO forcing. The zonal wind in the climatological jet stream region (around 30oN) has been substantially increased in both amplitude and extent, in strong El Niño years 1982/83 and 1997/98, whereas a much weaker zonal wind is seen in the strong La Niña year 1988/89.

5.2.1 Upper air climatology

Fig 5.6 shows the DJF 500 hPa height errors in the T63 and T159 climate. The T159 error in the Alaskan region is larger than in T63, indicating an enhancement of the climatological ridge over the northeastern Pacific. The other three regions of positive differences are related to a somewhat weaker climatological troughs over the Atlantic, northern Europe and eastern Asia in T159 relative to T63 model. The variance of the interannual 500

hPa height field in T159 is generally higher than in T63 indicating stronger interannual variation in the higher resolution model, and the value of 60 dam2 over the northern Pacific in T159 can be compared with the ERA-15 value of 71 dam2 in the same region.

5.2.2 Precipitation climatology and anomalies

The difference in the model horizontal resolution influences precipitation in two ways. In the tropical belt, Fig 5.7, precipitation differences (larger than 1 mm/day) exhibit relatively large-scale features; outside the tropics they are on smaller scales and related mainly to the regions with marked orography such as S. America, the Rockies and the Tibetan plateau. In the tropical Pacific, there is less rain in T159 than in T63 along or near the equator and more rain at the edges of convergence zones. It is interesting to note that in T159 during DJF (top panel) there is generally less precipitation over the Andes and more over the Amazon basin. In the T159 in JJA there are more marked contrasts in the Indian monsoon region, much less rain in the E Pacific and more rain over the southern Andes.

6. Summary and concluding remarks

The results of these studies can be summarised as follows:

a) Increasing the resolution of the inner loops of the 4D-Var assimilation process is unambiguously beneficial in providing better analyses through improving the 4D-var incremental algorithm and its use of data, which contains information on smaller scales. It also impacts positively on convergence properties.

b) Increasing the resolution of the outer loops of the 4D-Var process is also clearly beneficial as it enables more information to be extracted from data especially at lower levels, and potentially from satellite radiances influenced by land surface properties.

c) Increases in forecast resolution alone provide very positive impacts on the prediction of smaller-scale and/or rapidly developing systems which, together with orographically-dominated processes, characterise the majority of severe weather events.

d) The combination of increases in all components of the medium-range forecast system has been shown to give a substantial (and statistically-significant) positive impact on all aspects of deterministic global weather prediction.

e) Although more testing is needed, results so far indicate that an increase in EPS resolution from T159 to T255 will bring further benefits to those seen when changing from T63 to T159.

f) Horizontal resolution increases improve some aspects of the model systematic errors in extended integrations, but also indicate that model 'climate' drift is dominated by other factors such as those related to the representation of physical processes and their interactions. Comprehensive studies of horizontal resolution increases on the seasonal forecasting coupled system have yet to be done.

g) The work reported here required relatively little manpower because of a) the flexibility and efficiency of the Centre's software and b) the absence of model and assimilation problems encountered when extending the systems to higher resolutions. This confirms the stability and robustness of the current numerical algorithms and parametrization formulations.

This collection of results from the Centre's forecasting systems specifically addresses the impact of refining the spatial resolution in various ways; for the most part the impact of spatial resolution is reasonably detached from issues of temporal resolution through the use of Semi-Lagrangian techniques. From first principles refinements of spatial resolution provide more accurate numerical calculations and hence are inherently a desirable step. In practice, in a complex modelling environment such as a global forecasting system this needs to be studied and demonstrated. The results discussed here show the benefits of improved resolution in a variety of ways.

There is no doubt that further improvements will be forthcoming as experience is built-up at these enhanced resolutions, and in particular, it will be necessary to pay particular attention to the coupling of the physics and dynamics, to the accuracy of the current semi-implicit technique, and also to the behaviour of the incremental 4-DVar method. The very preliminary results of the impact of model resolution on extended-range prediction clearly need further analysis, and this should provide invaluable input to the continuing challenge of minimising model systematic errors at all prediction ranges. This work is in progress, as are related studies with the fully-coupled system.

Taking into consideration the planned timetable of scientific and computer upgrades, future work will require the testing of the combination of the T511 forecast model with 60 levels, together with the 12 hr 4D-Var assimilation with outer loops at T511 and inner loops at T159 (semi-Lagrangian). However, current experience suggests that this should not present any major scientific or technical problems.

Central to the Centre's 4-yr plan and long-term strategy are progressive improvements in horizontal (and vertical) resolution of all components of the forecast system, and it is considered that the results presented are fully supportive of these plans, both from a scientific and technical viewpoint.

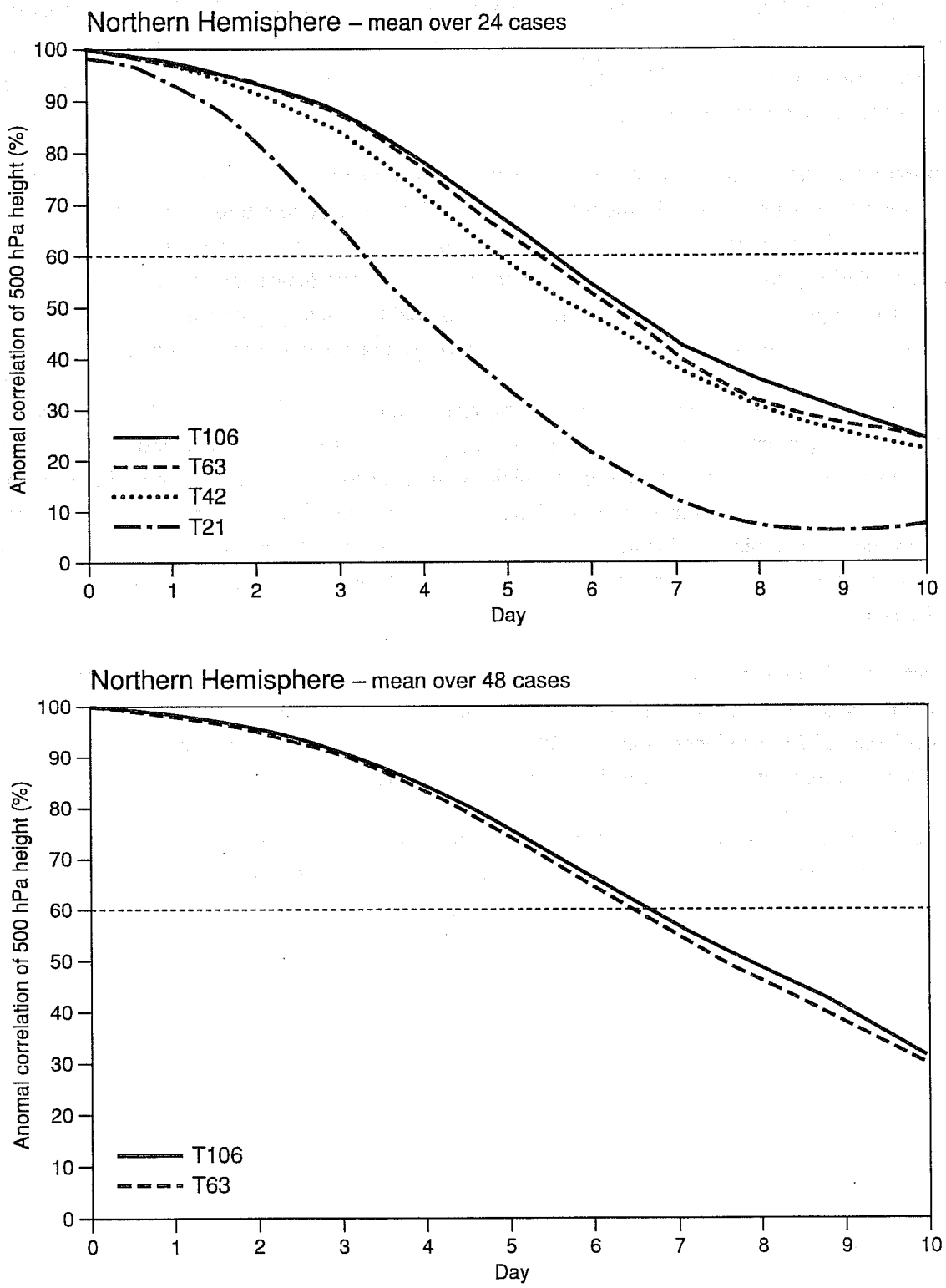


Fig 1.1 Anomaly correlations of 500 hPa height for the extratropical Northern Hemisphere averaged over sets of 24 (upper) and 48 (lower) cases, for spectral truncations T21, T42, T63 and T106 as indicated in each panel.

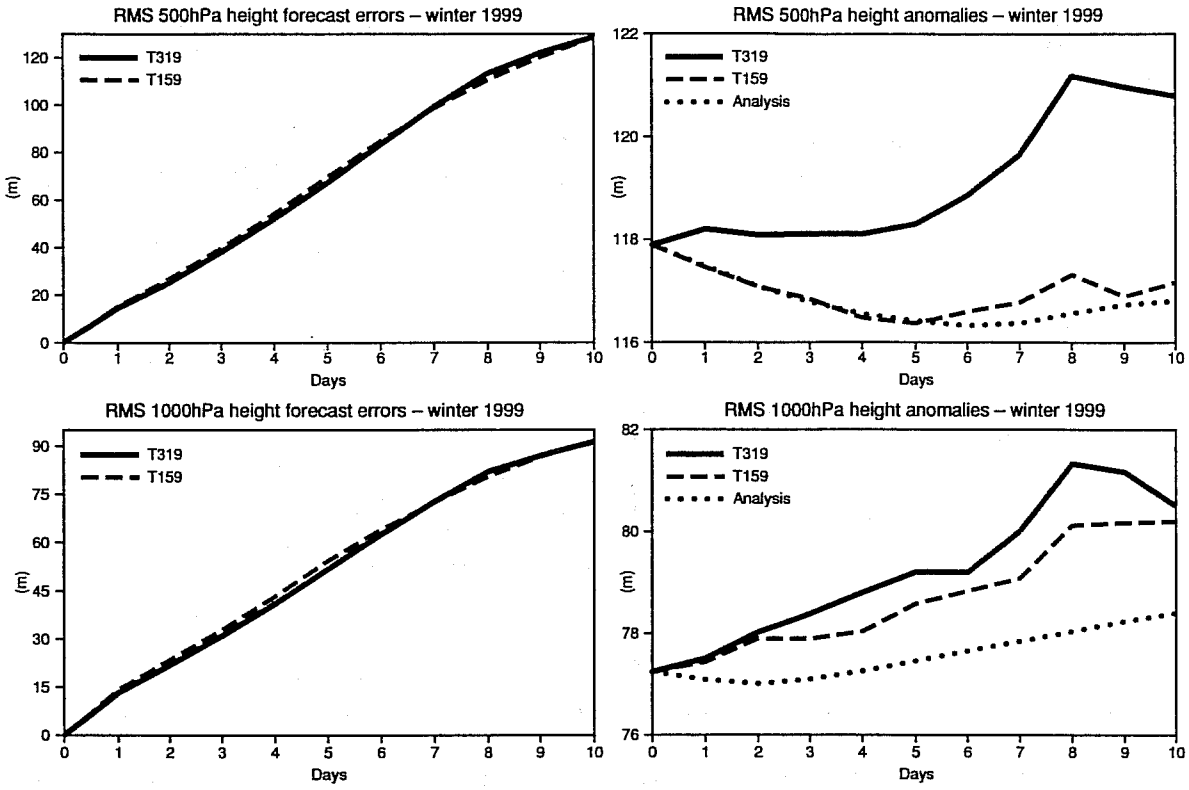


Fig 1.2 Root-mean-square errors (left) and anomalies (right) of 500 hPa (upper) and 1000 hPa (lower) height, computed over the extratropical Northern Hemisphere for the period 1 December 1998 to 28 February 1999. Errors and anomalies are shown for T319 (red, solid) and T159 (blue, dashed) resolutions, and the anomalies of the variational analyses are also shown.

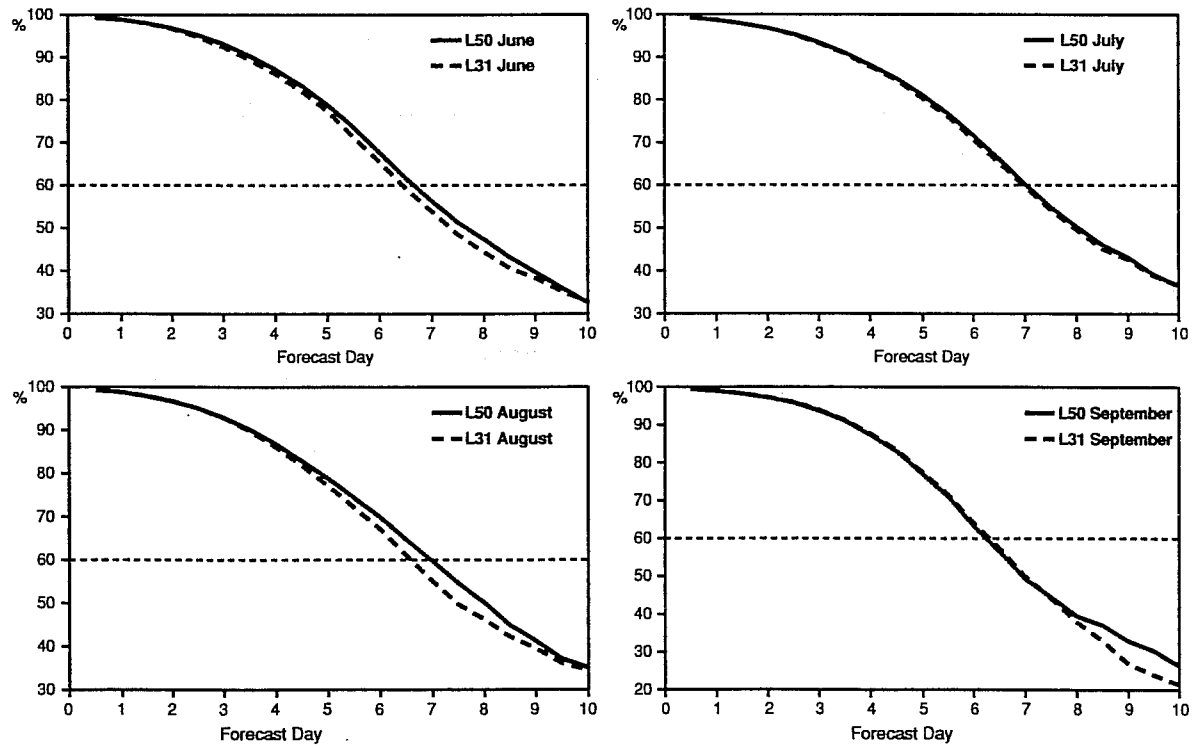


Fig 1.3 30-day means of anomaly correlations of 500 hPa heights from 50-level (red, solid) and 31-level (blue, dashed) forecasts, for the periods 27 May to 25 June (upper left), 26 June to 25 July (upper right), 26 July to 24 August (lower left) and 25 August to 23 September (lower right), 1998.

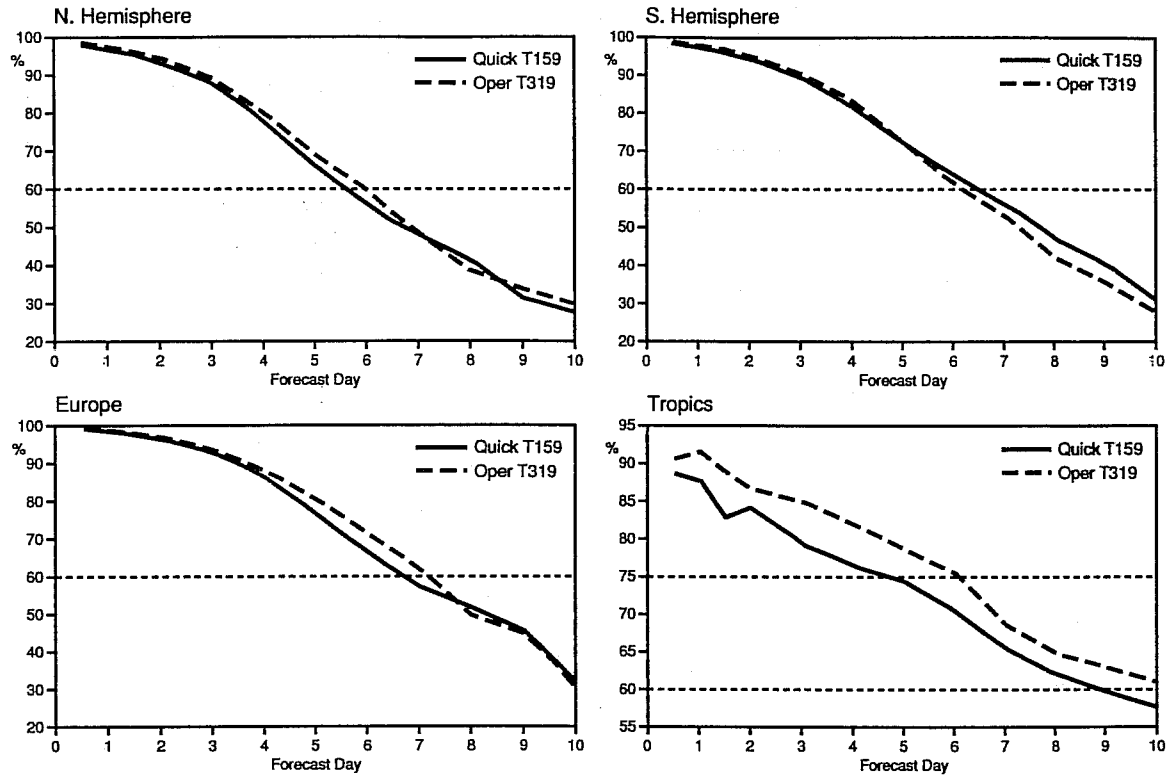


Fig 2.1 Mean anomaly correlations of the 1000 hPa geopotential for a 15-day period (15/8/98-29/8/98) for the operational analysis system (dashed) and the same system but with T159 used for the outer assimilation loops (solid).

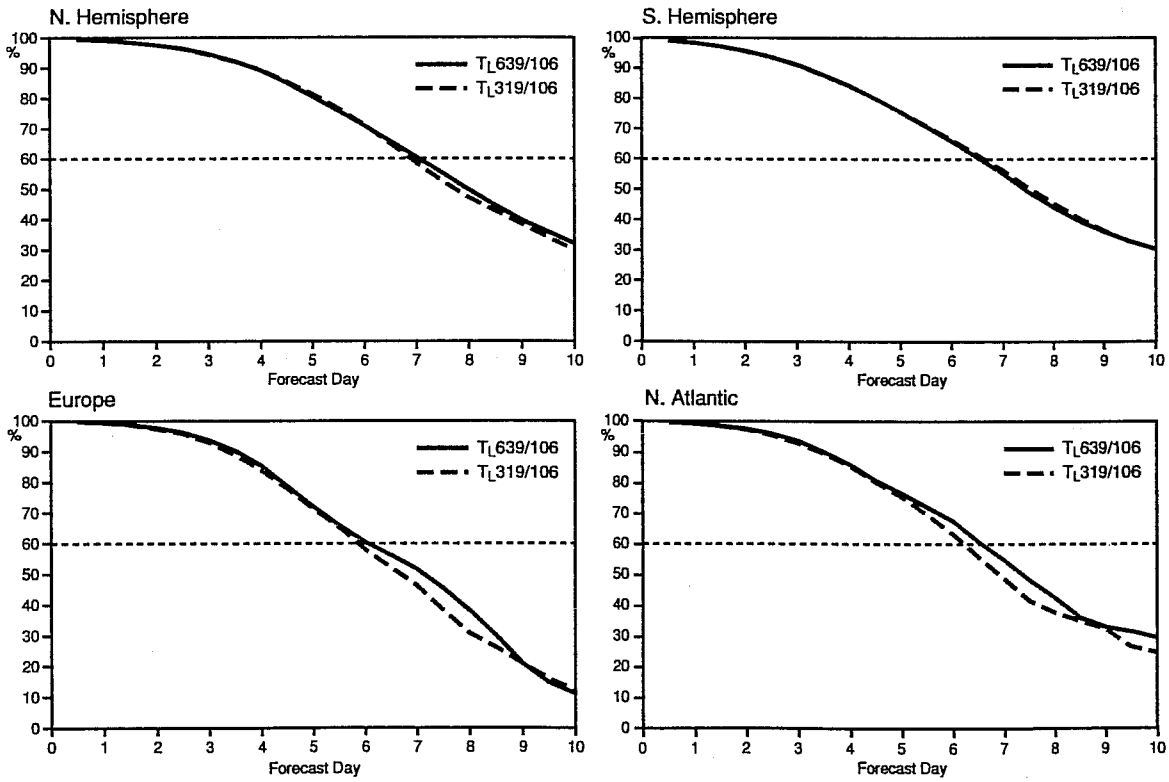


Fig 2.2 The mean anomaly correlations of 500 hPa height for a set of 53 forecasts with model at T639 and 4D-Var at T639/T106 (solid) and with model at T319 and 4D-Var at T319/T106 (dashed). See Section 3.1 for precise dates of forecasts. Areas are: N Hemisphere (upper left), S Hemisphere (upper right), Europe (lower left), N Atlantic (lower right).

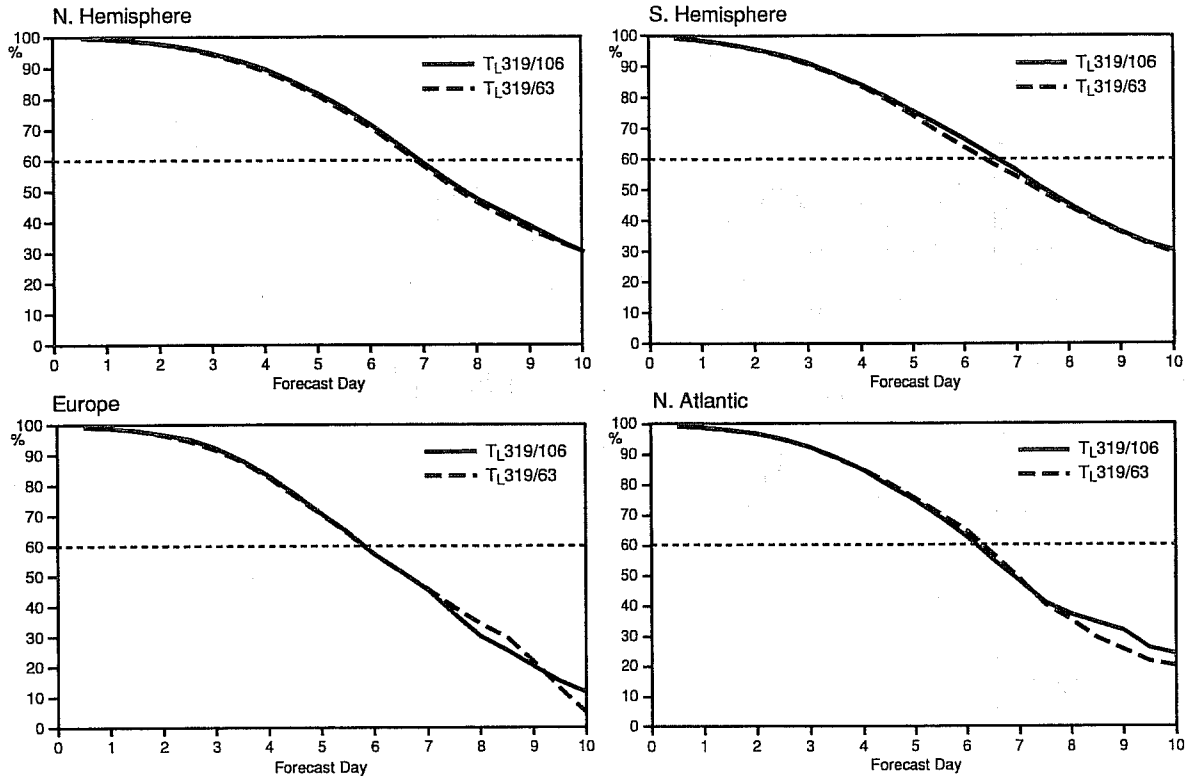


Fig 2.3 The mean anomaly correlations of 500 hPa height (for the same set of forecast dates as Fig 2.2) for T319 with inner loops at T106 (solid) and T319 with inner loops at T63 (dashed).

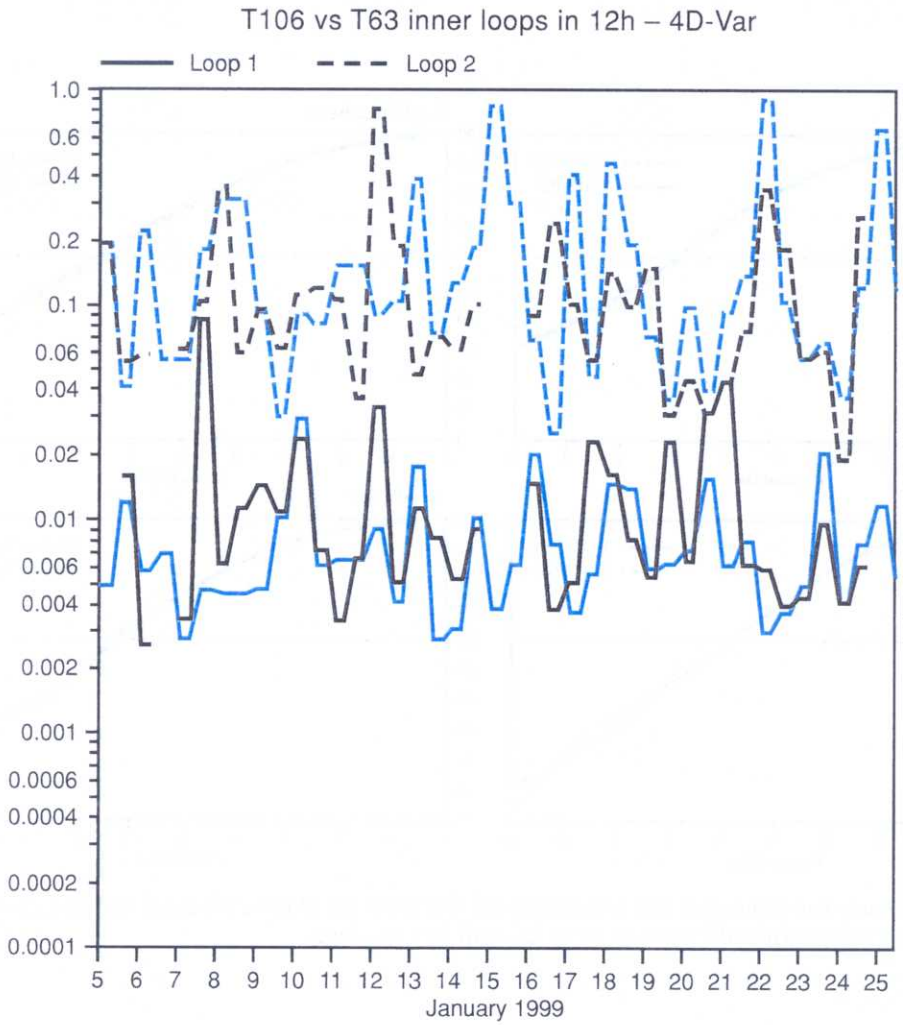


Fig 2.4a Squared relative decrease in the norm of the gradient of the 12-hourly 4D-Var cost function in the first (full curves) and second (dashed curves inner loops), for assimilations with T63 (blue) and T106 (black) inner loop resolution, with the same number of iterations of the minimization.

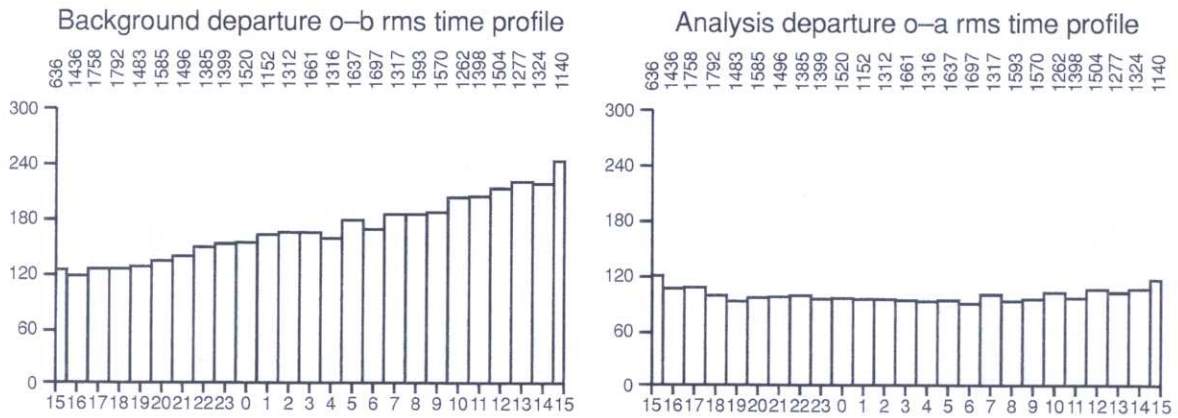


Fig 2.4b RMS fits in Pa of a 24-hourly 4D-Var global assimilation to DRIBU observations, stratified against the time of the day. The vertical numbers indicate the number of observations used in each bin. The left most histogram shows the rms departures from the background field at the appropriate time, so that its slope reflects the error growth in the forecast started from the previous analysis. The right-most histogram shows the rms departures from the model trajectory assimilated in 4D-Var.

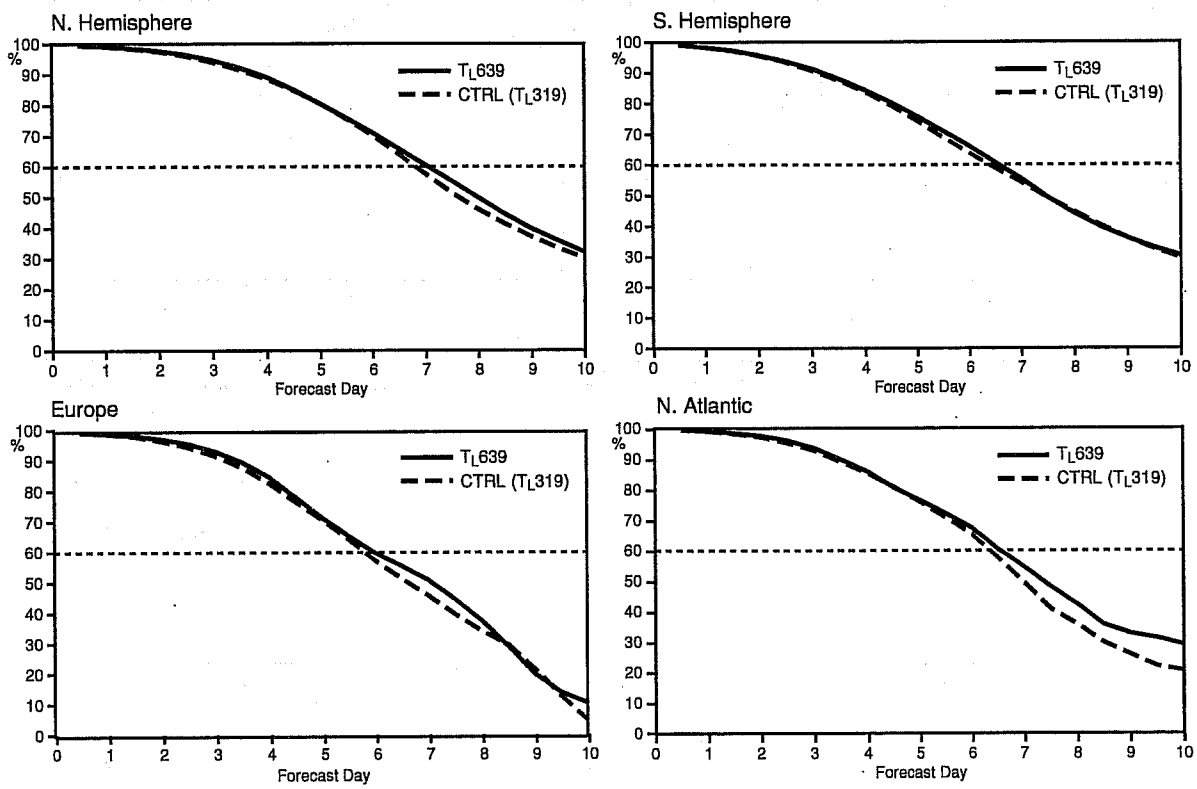


Fig 3.1 As Fig 2.2 but comparing T639 with inner loops at T106 (solid) and T319 with inner loops at T63.

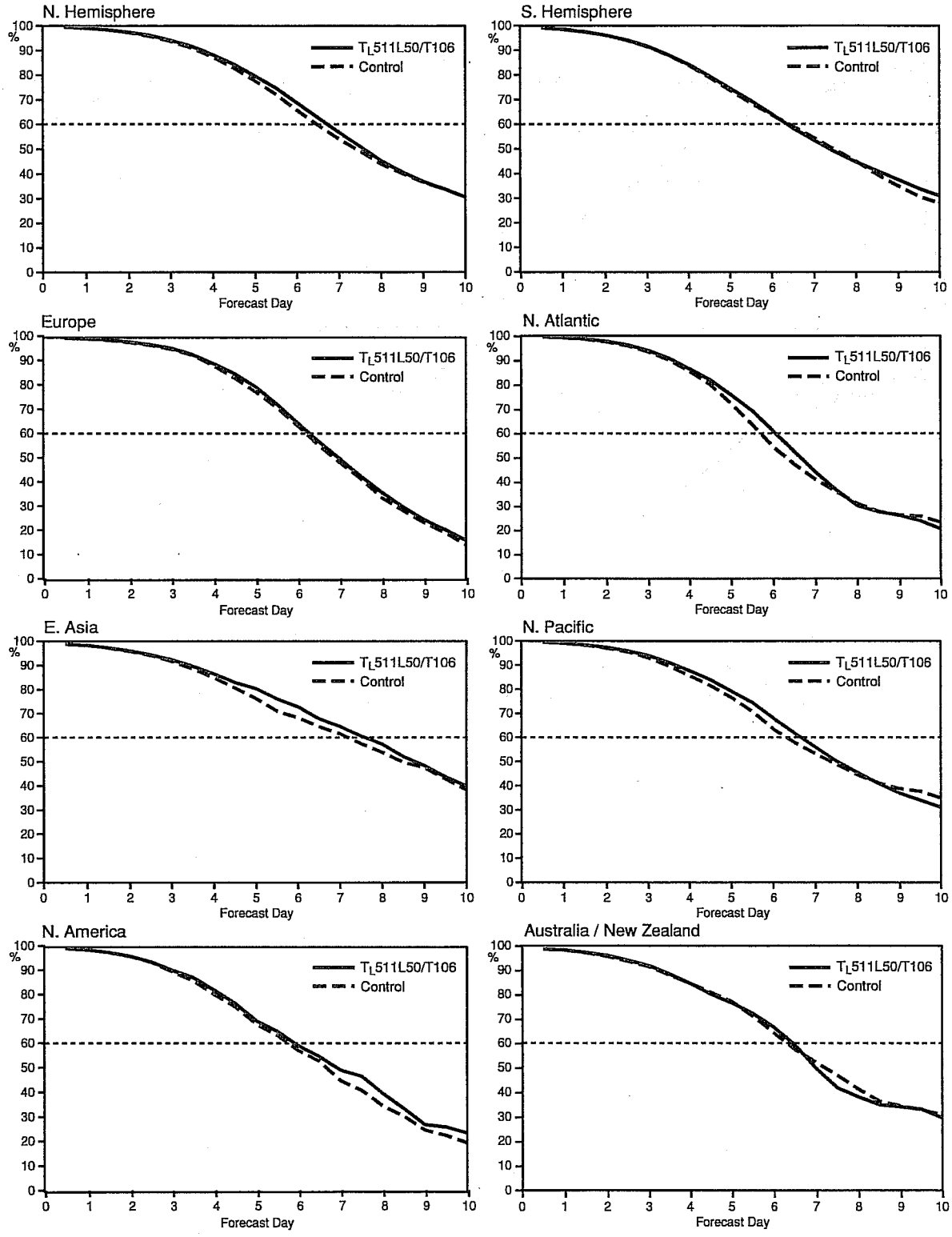


Fig 3.2 The mean anomaly correlation of 500 hPa height for a set of 93 forecasts. T511L50 with inner loops at T106 (solid), and T319L50 with inner loops at T63 (dashed). See Section 3.1 for precise dates of forecasts. Areas are: N Hemisphere (top left), S Hemisphere (top right), Europe (second row, left), N Atlantic (second row, right), E Asia (third row, left), N Pacific (third row, right), N America (bottom left), Australia/N Zealand (bottom right).

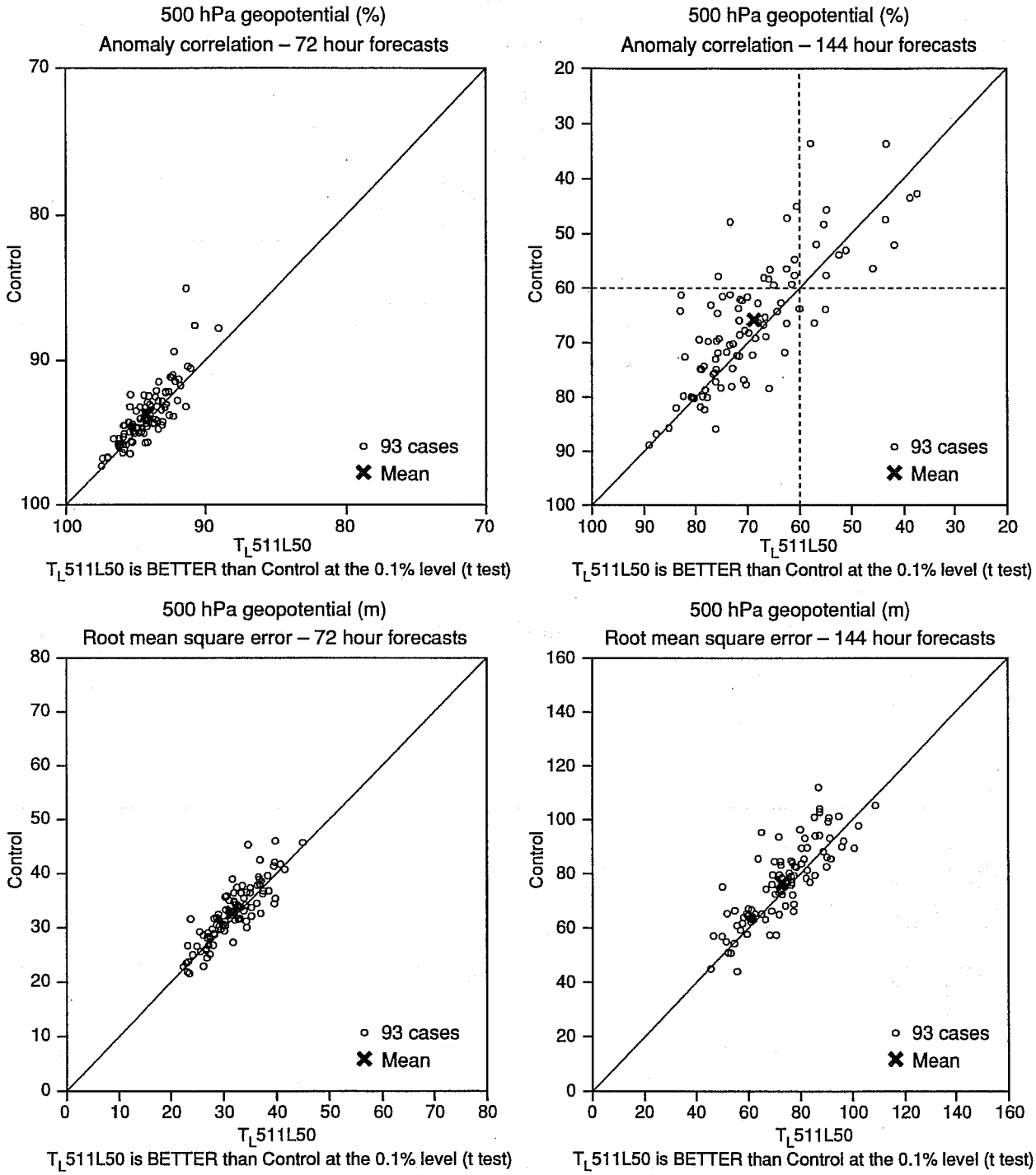


Fig 3.3 Scatter plots for the N Hemisphere of the set of 93 forecasts described in Fig 3.2 for 72 hrs and 144 hrs. Upper panels are anomaly correlation, lower panels are RMS error.

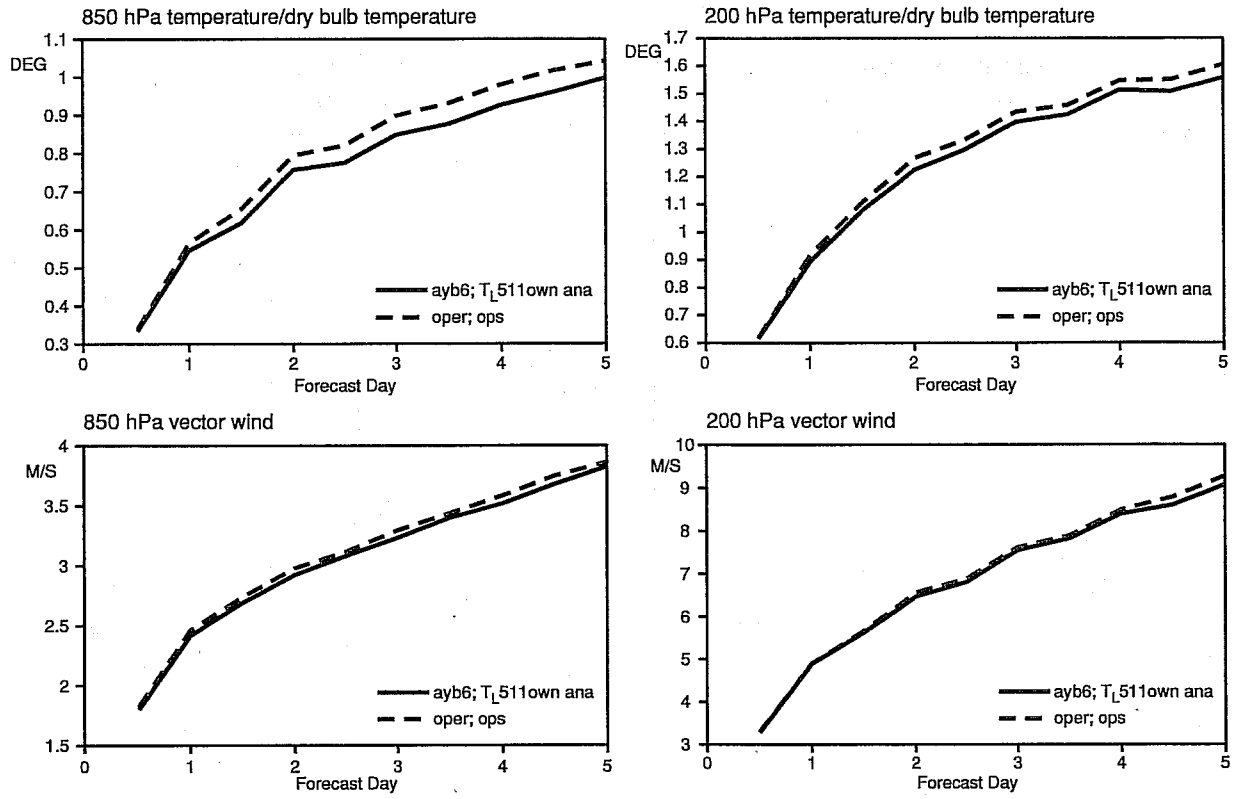


Fig 3.4 Evolution of RMS errors of temperature and wind for the Tropics for 15 cases in May 1999, with the errors computed against the respective analyses. 850 hPa temperature (upper left), 200 hPa temperature (upper right), 800 hPa wind (lower left), 200 hPa wind (lower right).

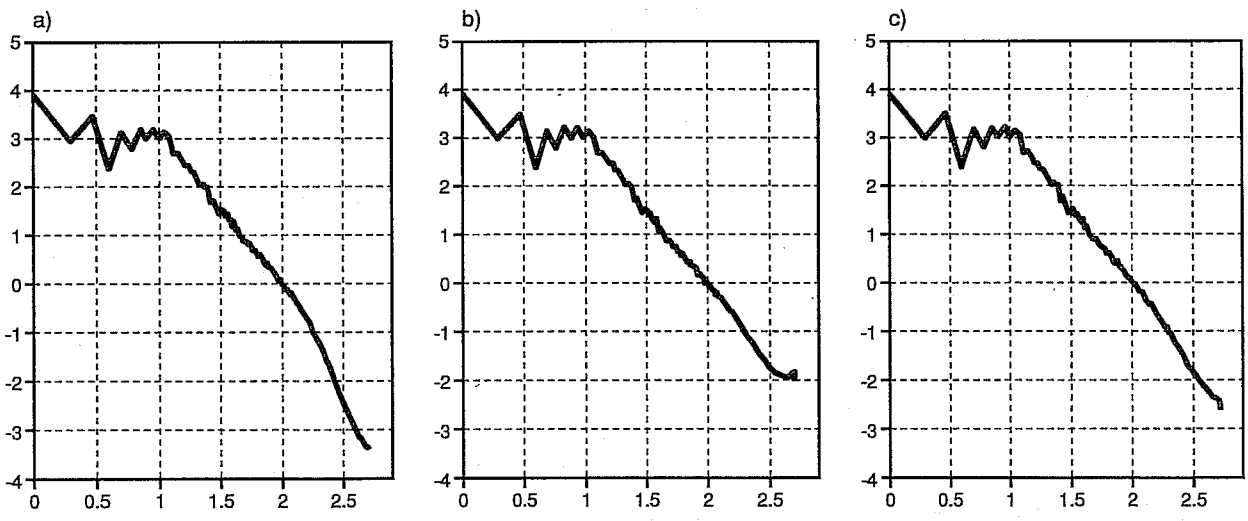
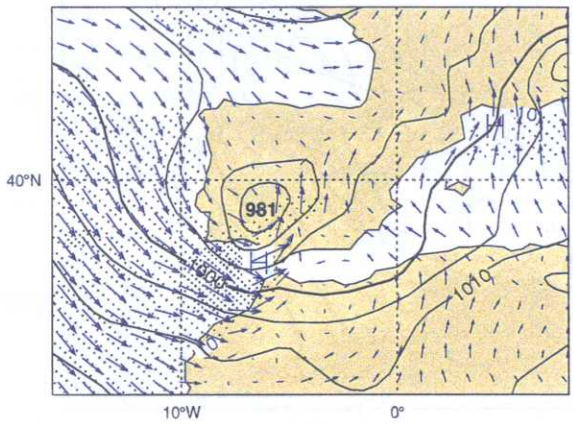
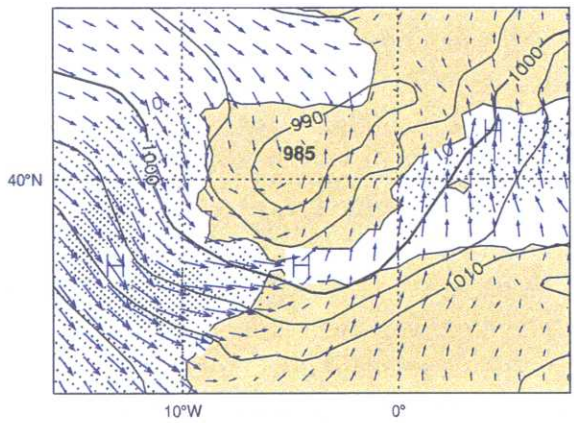


Fig 3.5 Eddy kinetic energy spectra at level 30 (~250 hPa) of three forecasts to 48 hours at T511L50 starting from the same initial conditions and differing only in the value of horizontal diffusion. a) uses the same horizontal diffusion on Vo, D and T with an e-folding time for the shortest scale of 900 s. b) has no horizontal diffusion, and c) uses horizontal diffusion only on vorticity with an e-folding time for the shortest scale of 2700 s. Axes are log (EKE) versus log (total wave number).

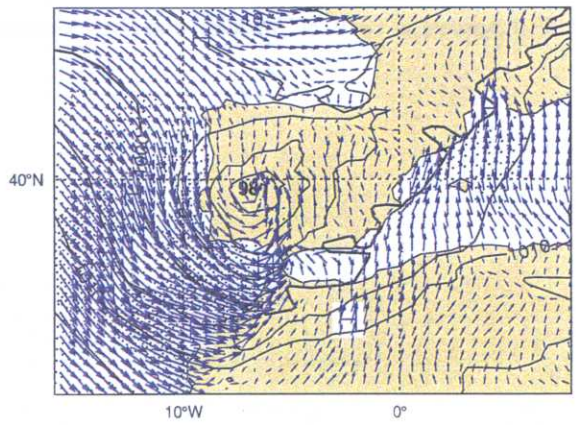
3.6a) **Analysis of pmsl & wind at 10 m, valid 971106 00 UTC**



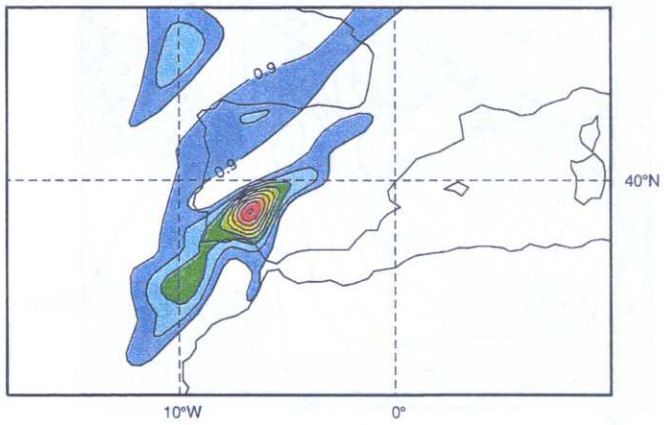
3.6b) **Ops (T213) 60 hour forecast, valid 971106 00 UTC**



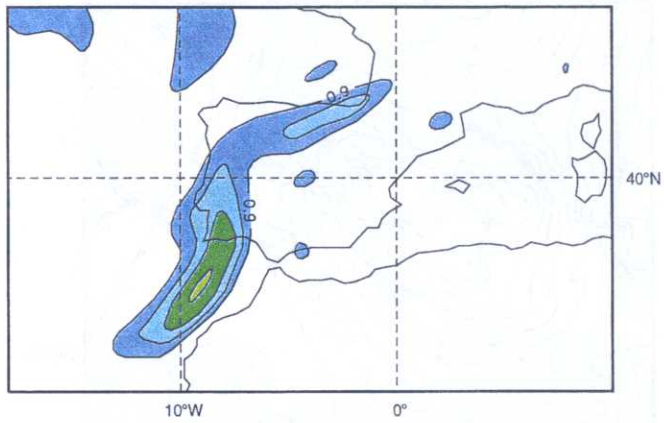
3.6c) **T_L639 60 hour forecast, valid 971106 00 UTC**



3.6d) **Analysis of PV on 305 K isentrope, valid 971106 00 UTC**



3.6e) **Ops (T213) 60 hour forecast, valid 971106 00 UTC**



3.6f) **T_L639 60 hour forecast, valid 971106 00 UTC**

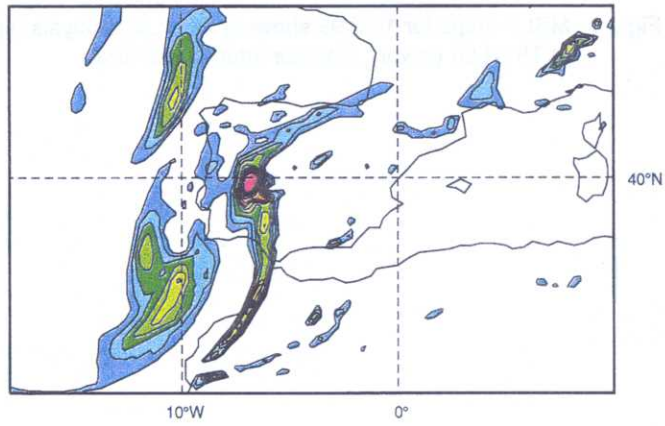


Fig 3.6

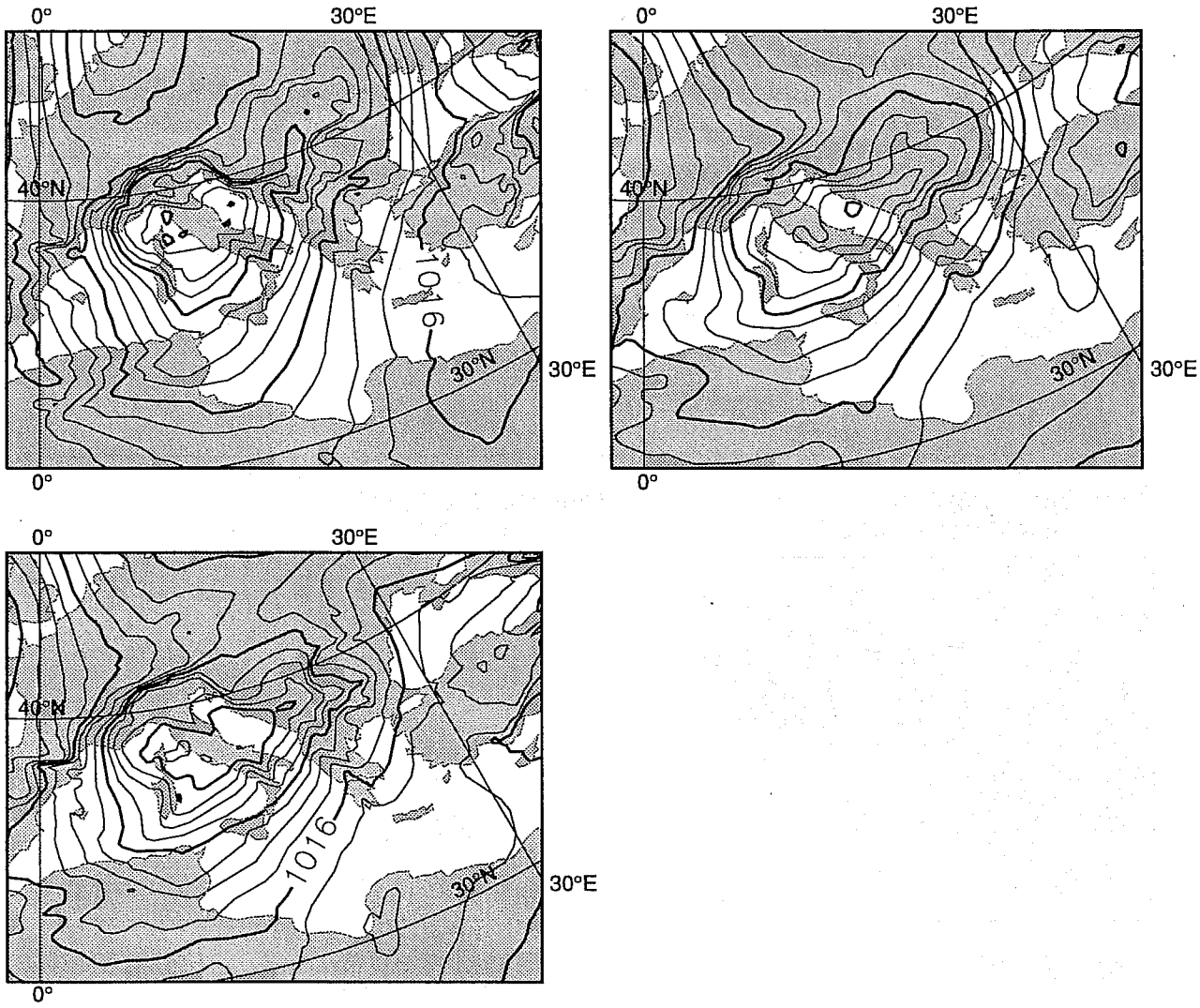


Fig 3.7 MSLP maps for 10/2/99 showing T511L50 analysis (upper), D+5 forecast at T319L50 (middle), and D+5 forecast at T511L50 (lower). Contour interval is 2 hPa.

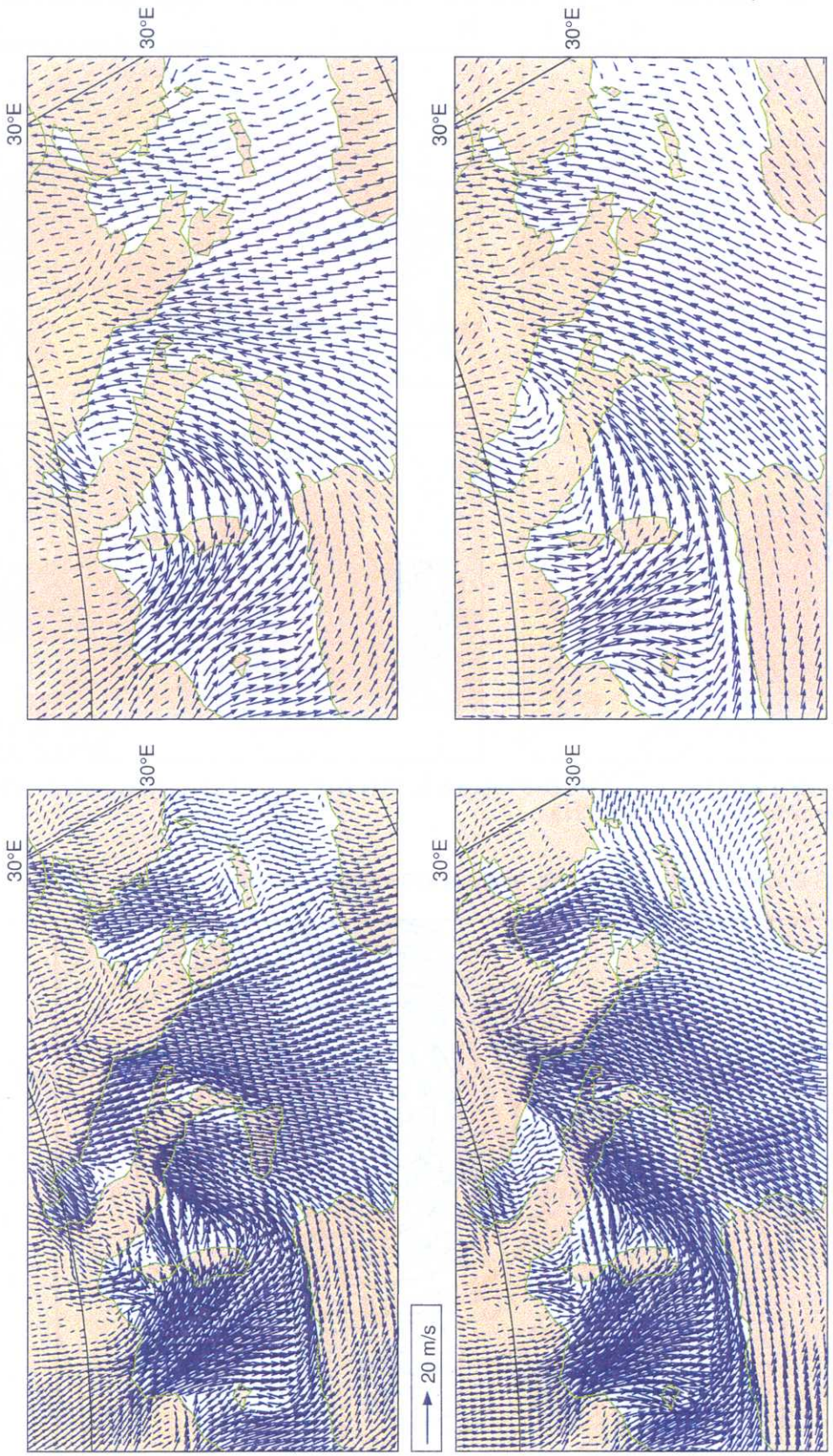


Fig 3.8a 10 metre winds for 10/2/99. T511L50 analysis (upper left), T319L50 analysis (upper right), D+5 forecast at T511L50 (lower left) and D+5 forecast at T319L50 (lower right).

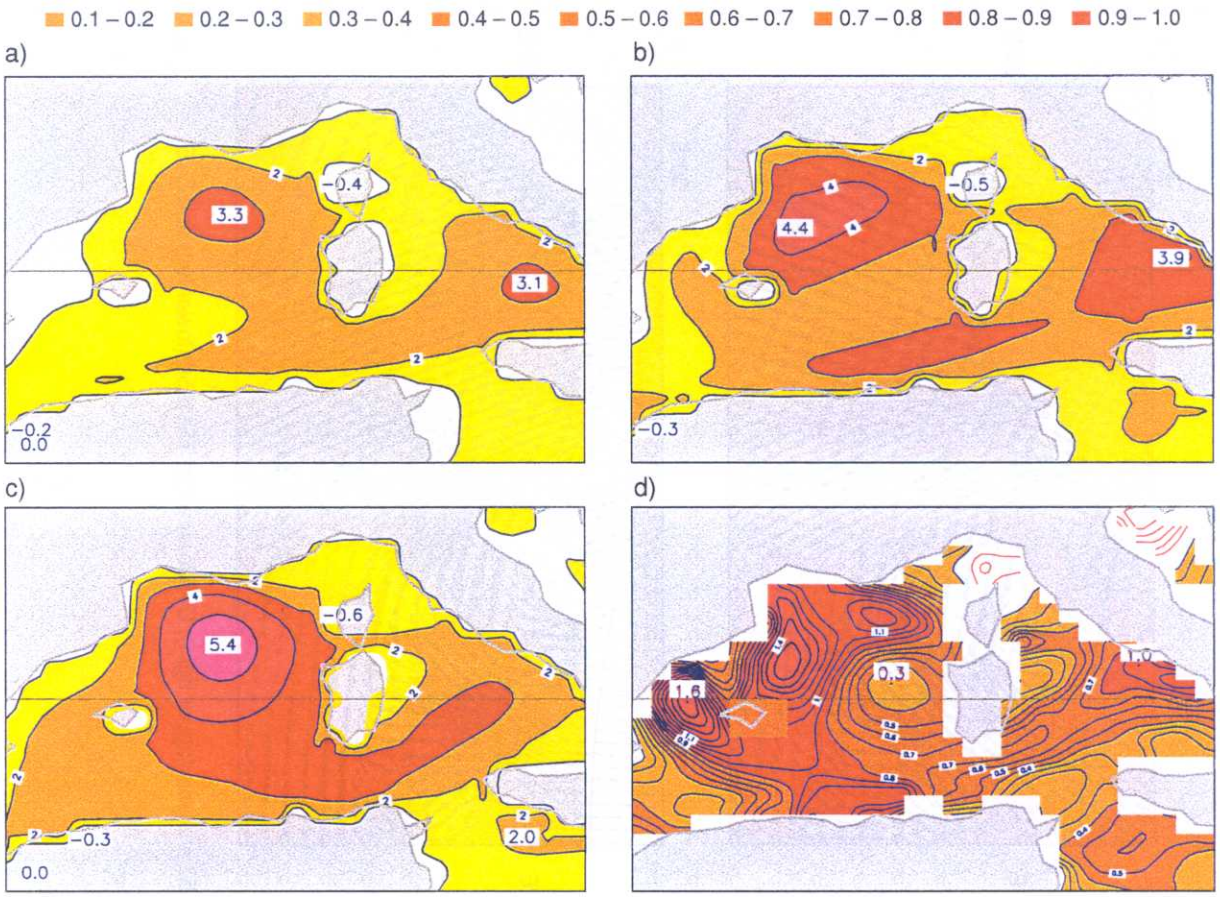


Fig 3.8b The corresponding wave heights a) T319 D+5 forecast, b) T511 D+5 forecast, c) verifying analysis (T319) and d) difference field for T511 minus T319.

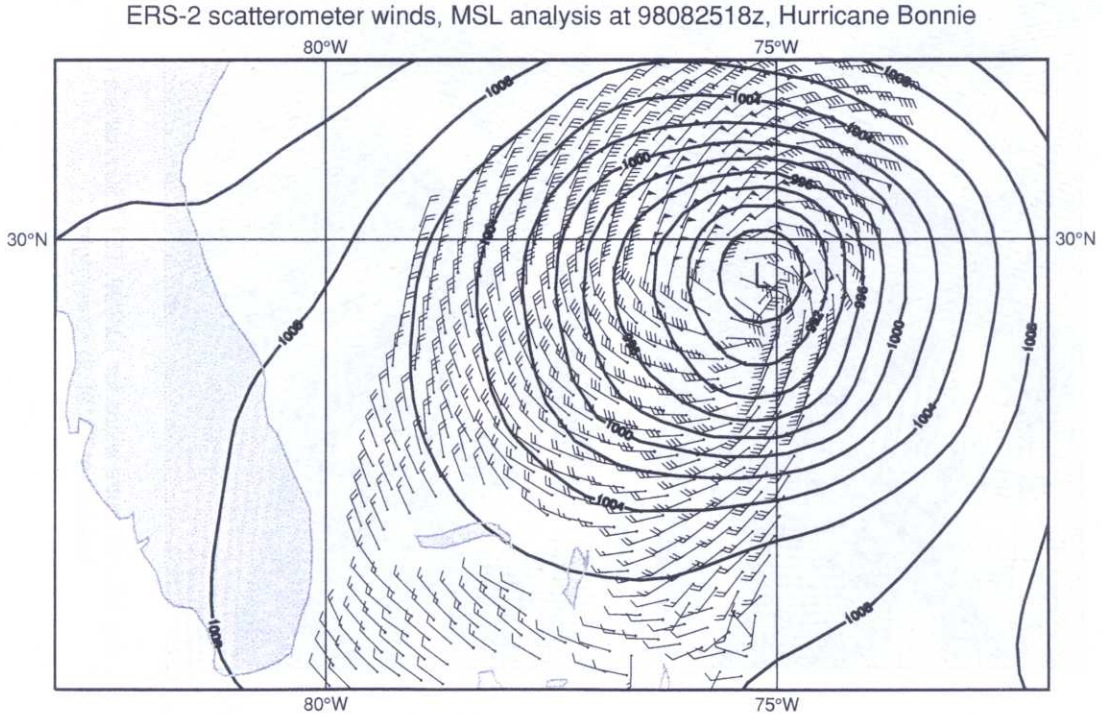


Fig 3.9 MSL analysis and all available ERS scatterometer data on 18 UTC 25 August 1998 in the vicinity of hurricane Bonnie east of Florida.

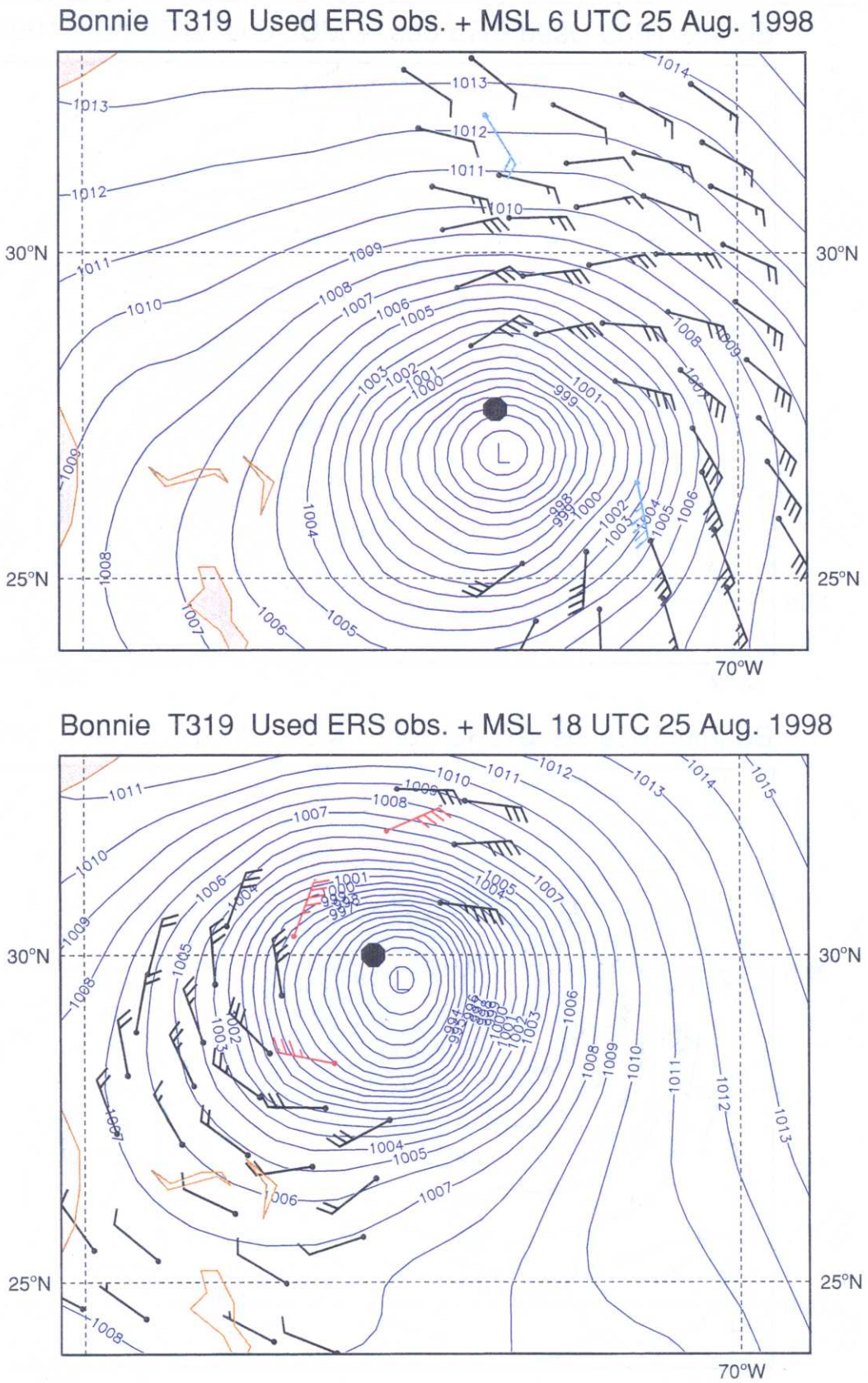
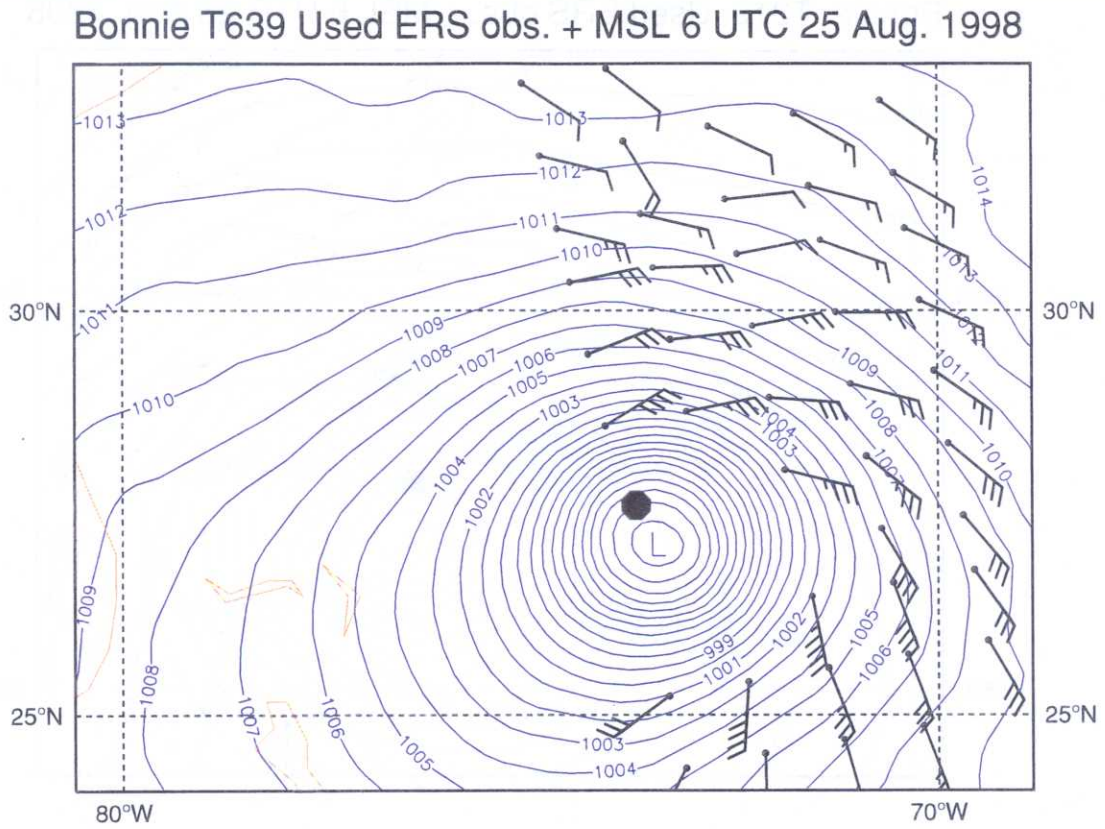


Fig 3.10a MSL analyses and ERS scatterometer data available to the T319/T106/T63 assimilation on respectively 6 UTC and 18 UTC 25 August 1998. The reported observed position is marked with a black dot. Black winds are fully accepted by the assimilation, turquoise are slightly suspect, while red winds are totally rejected.



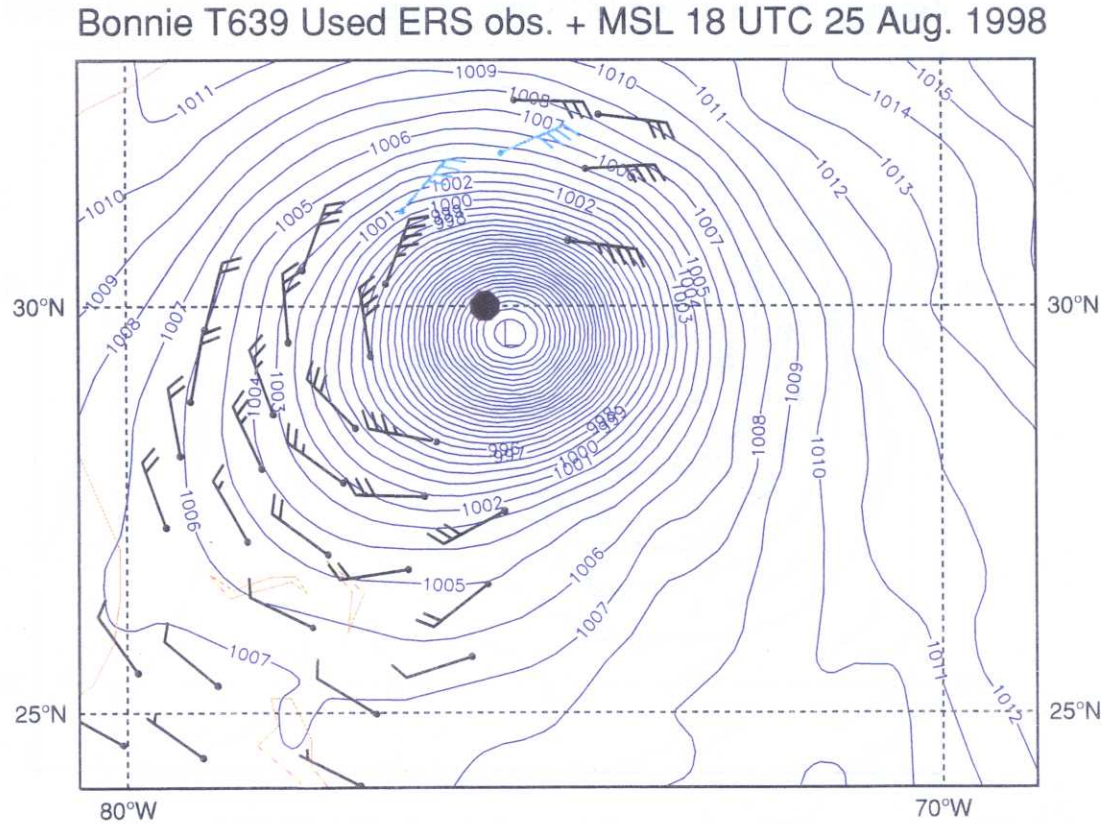


Fig 3.10b: As Fig a but for the T639/T106/T63 assimilation.

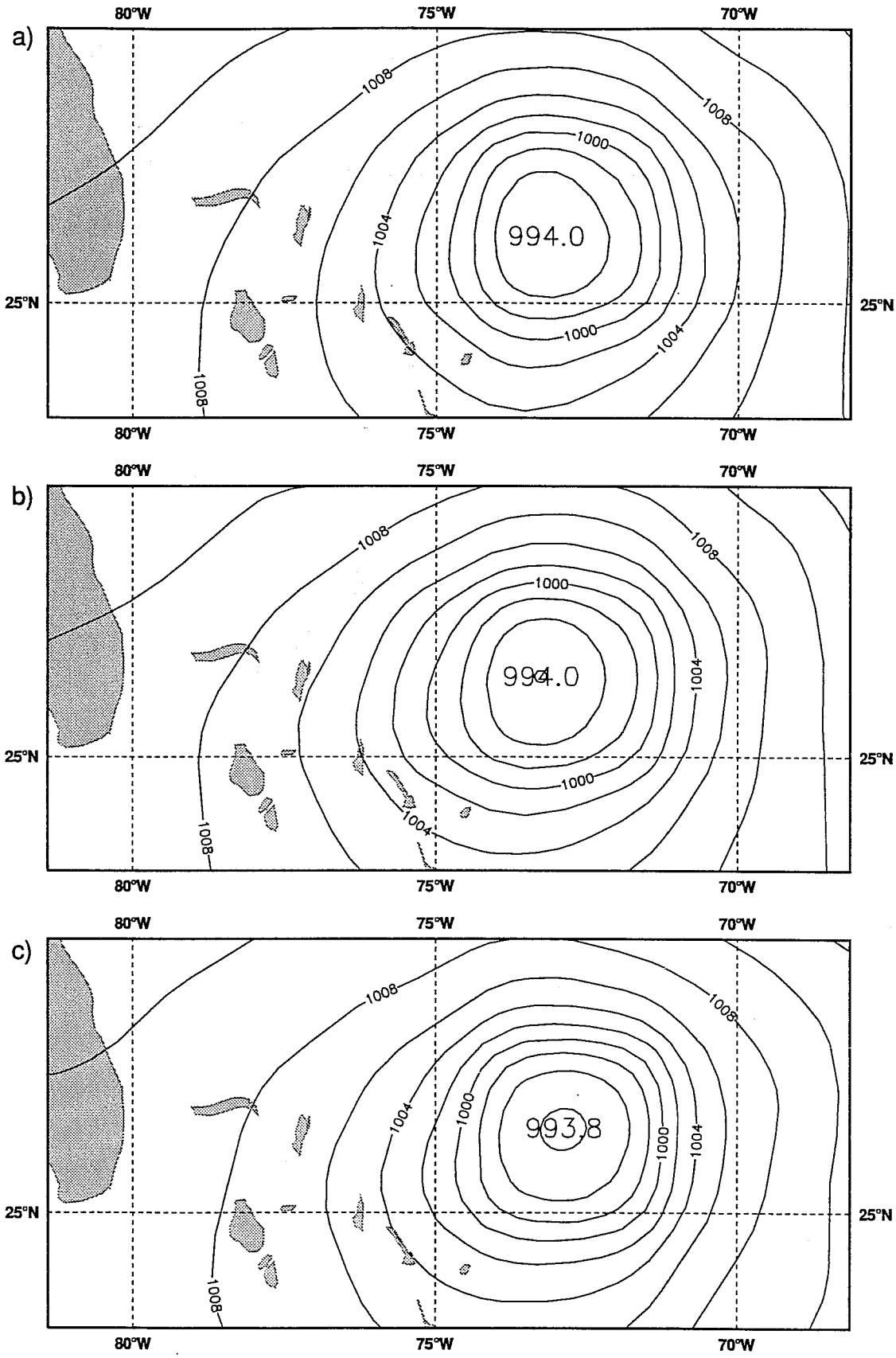


Fig 3.11a-c MSL analyses of hurricane Bonnie valid at 0 UTC 26 August 1998. (a) is the operational (T319/T63/T42). (b) is the high resolution inner assimilation loop experiment (T319/T106/T63). (c) also includes the high resolution for the outer assimilation loop and the first guess calculation (T639/T106/T63).

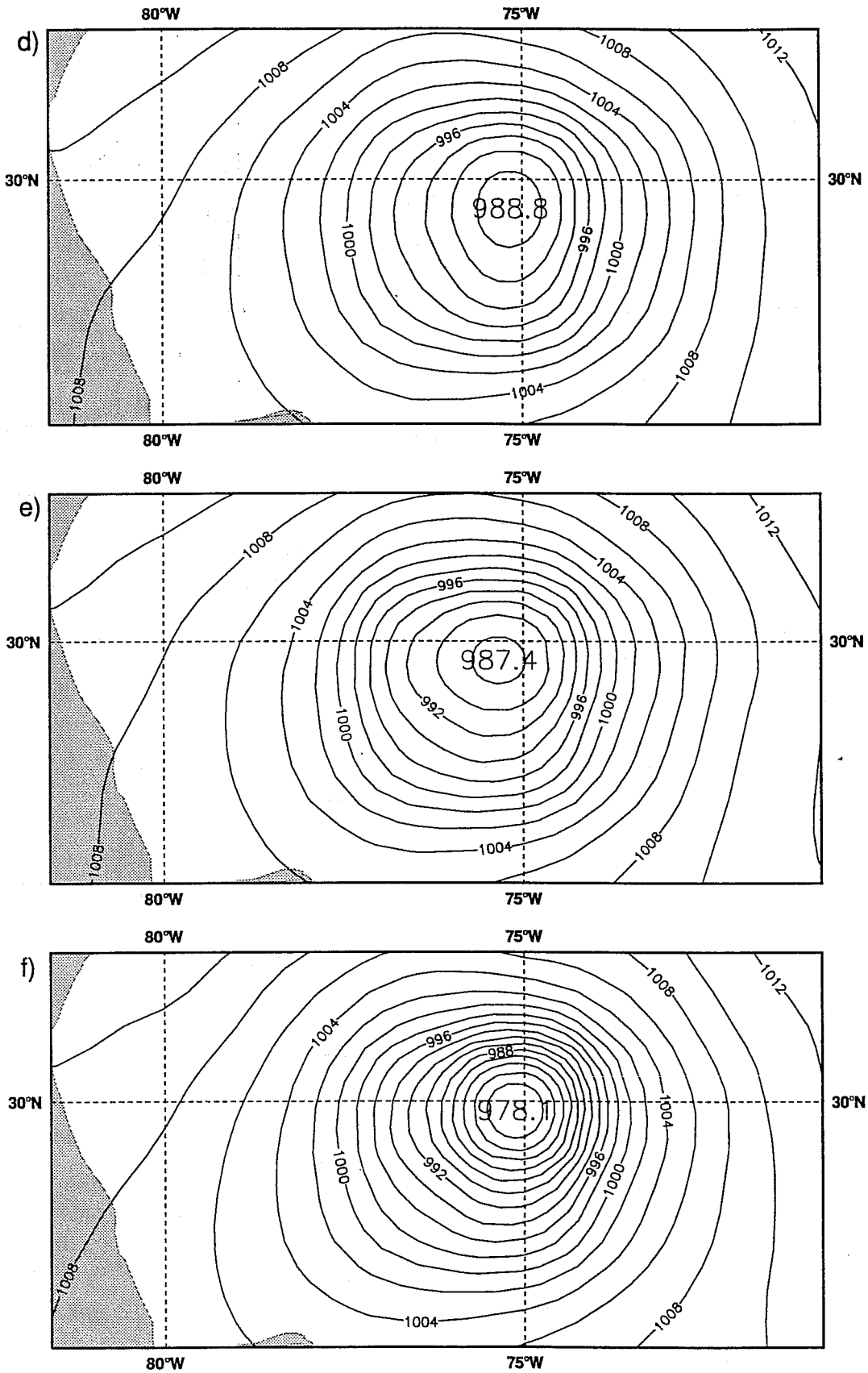


Fig 3.11d-f Like a-c but valid at 18 UTC 25 August 1998.

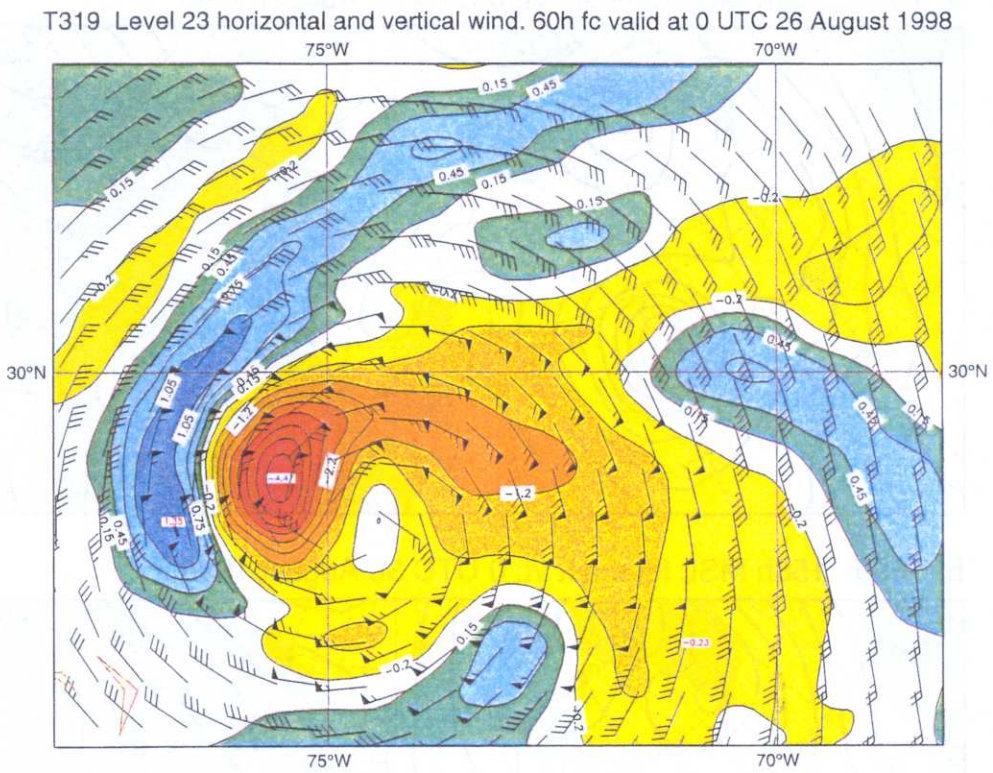


Fig 3.12a: Winds at level 25 (725 hPa) and vertical motion represented by shading (blue—green represent descent, red—yellow represent ascent) for a sixty hour forecast valid at 0 UTC 26 August 1998. Analysis in T319/T106/T63 and the forecast at T319.

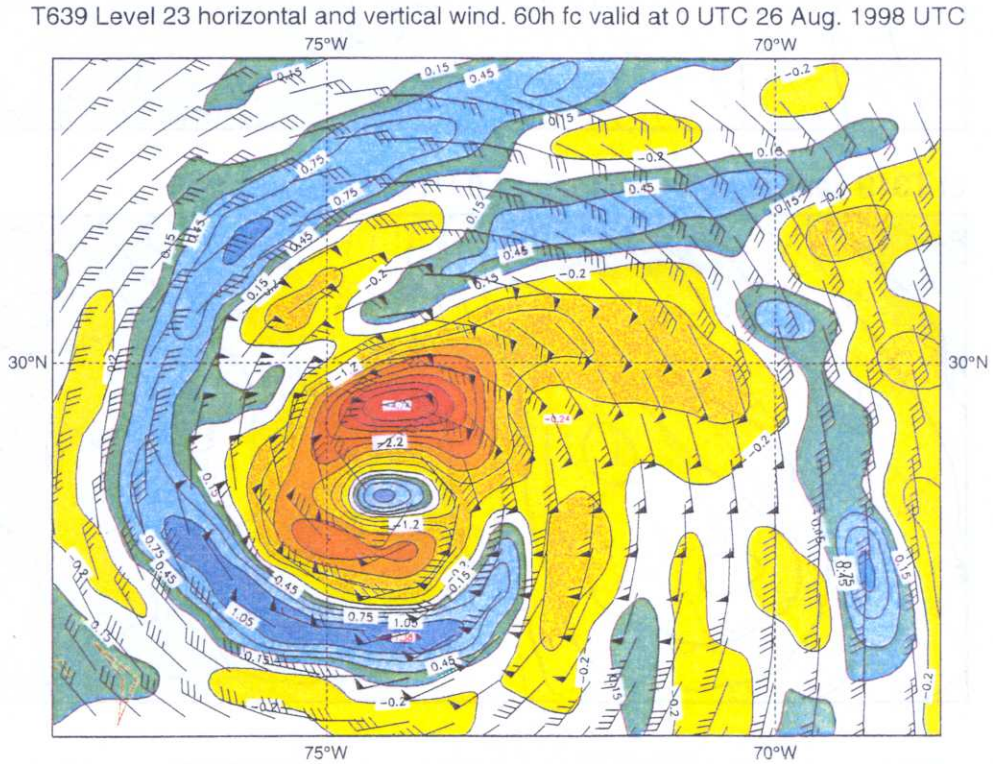
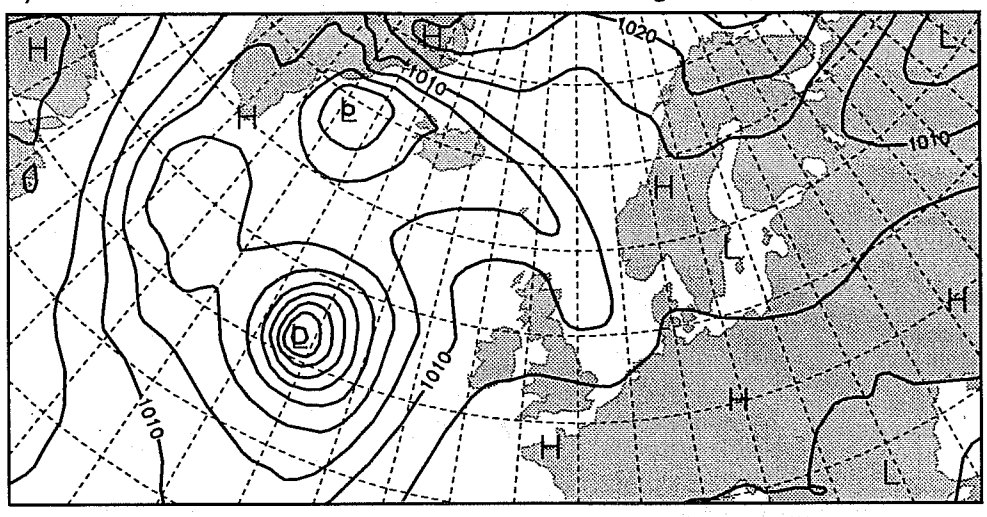
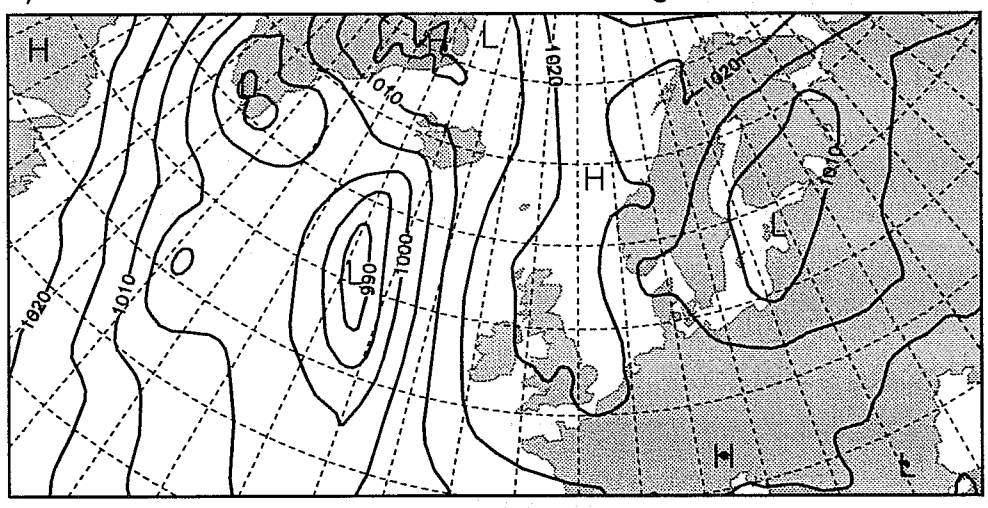


Fig 3.12b: Similar to figure 3.12a but based on the T639/T106/T63 assimilation and T639 forecast.

a) T319 156h MSL forecast vt: 0 UTC 30 August 1998



b) T639 156h MSL forecast vt: 0 UTC 30 August 1998



c) T319 verifying MSL analysis at 0 UTC 30 August 1998

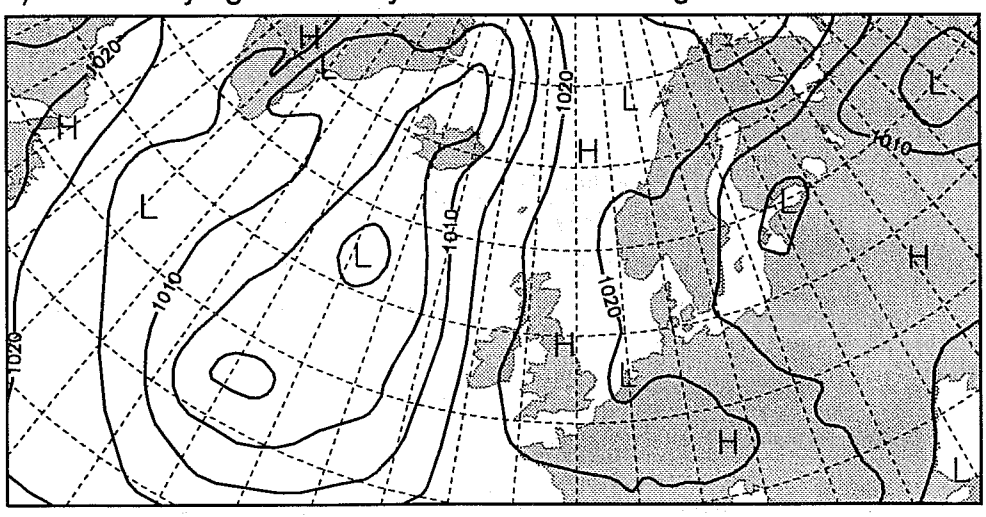


Fig 3.13 156 hour MSL forecasts valid at 00 UTC 30 August 1998 from the T319/T106/T63 assimilation (panel a), and the T639/T106/T63 assimilation (panel b). The verifying analysis is shown in panel c.

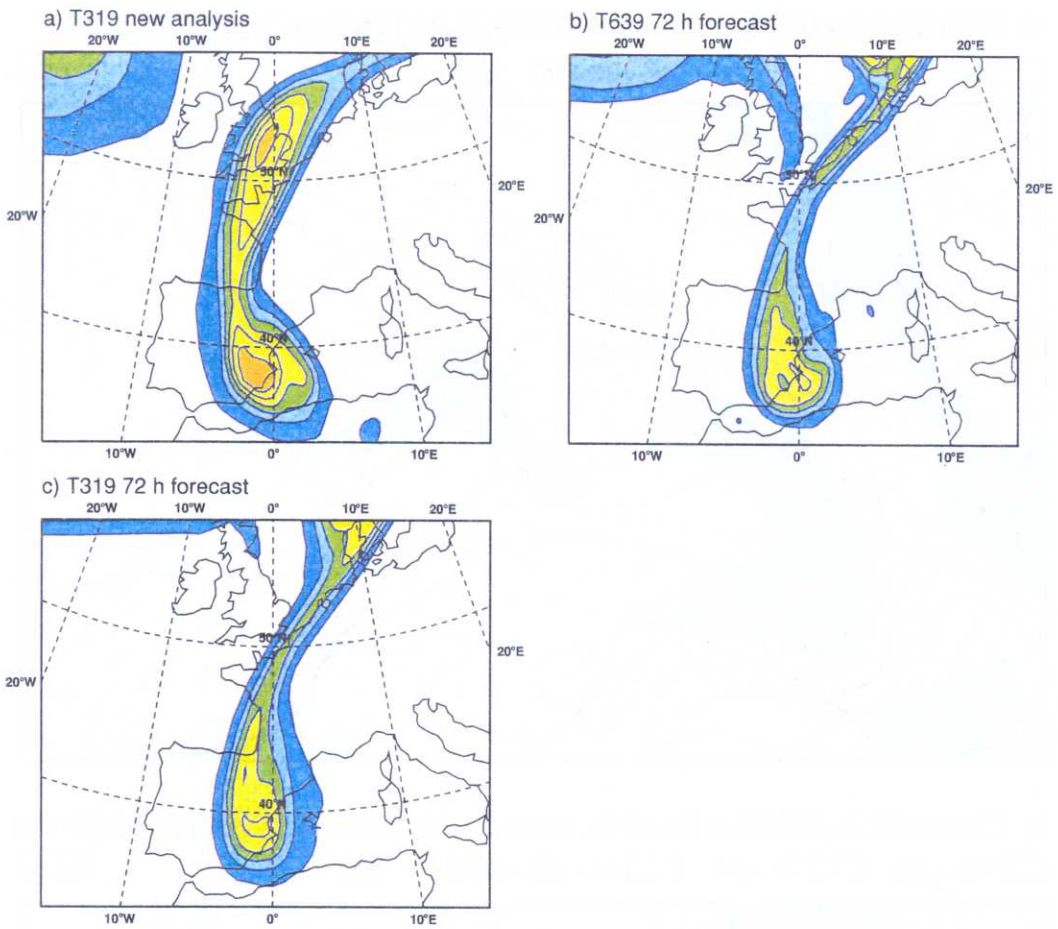


Fig 3.14 Potential vorticity on a potential temperature level of 325 K. T319 new analysis (upper panel), T639 72 hour forecast (middle panel) and T319 72 hour forecast (lower panel).

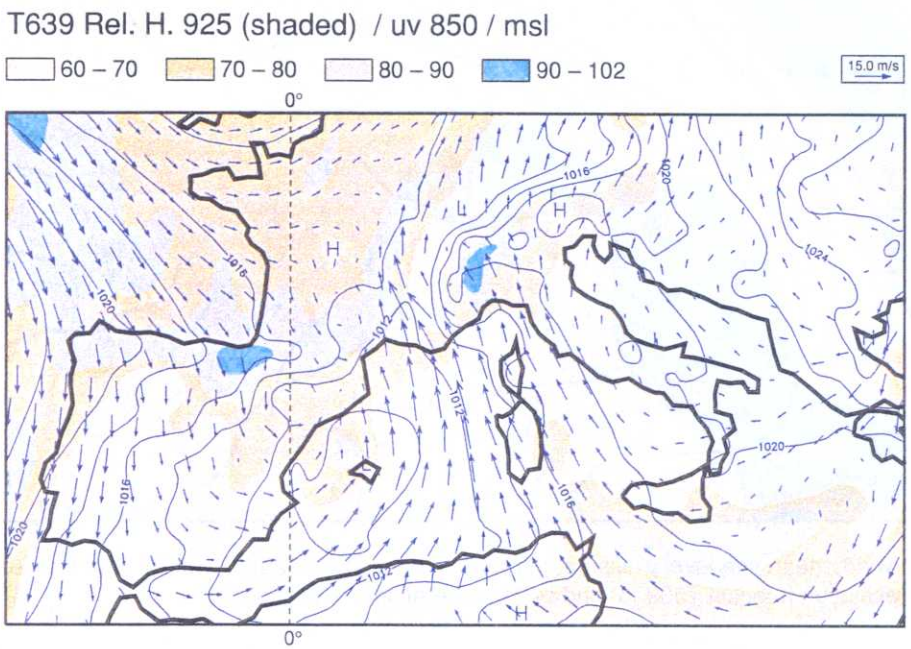


Fig 3.15 850 hPa wind, mean sea level pressure and relative humidity at 925 hPa, averaged over steps 36 to 72 of a forecast from 19930920.

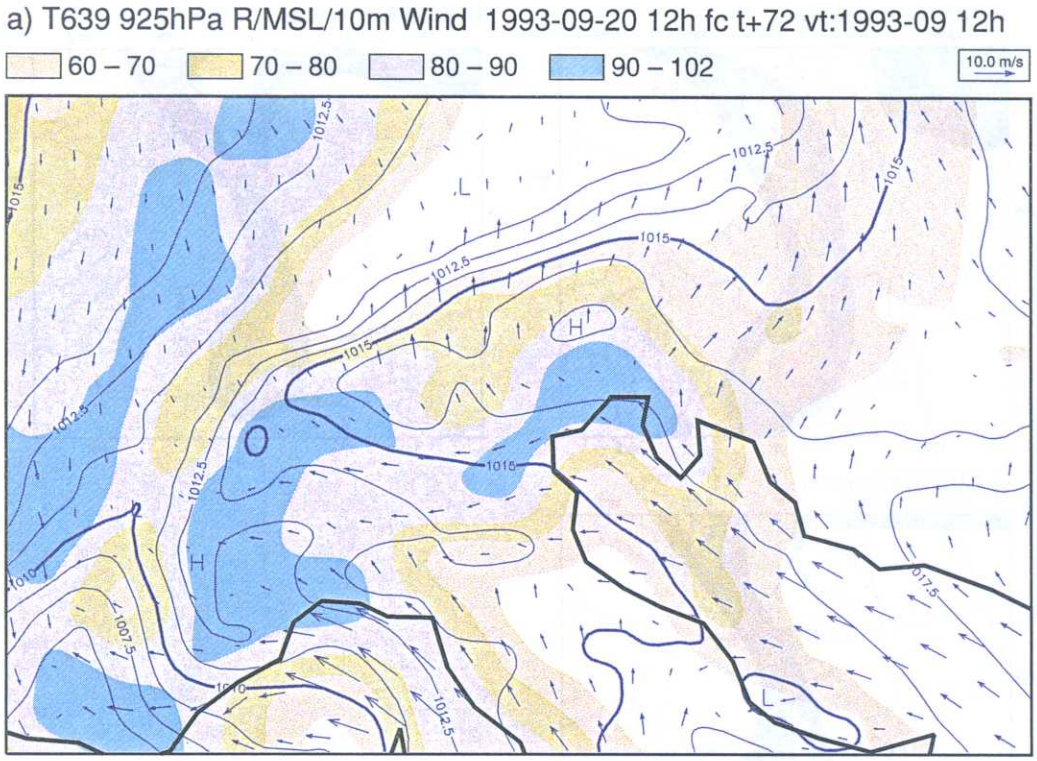
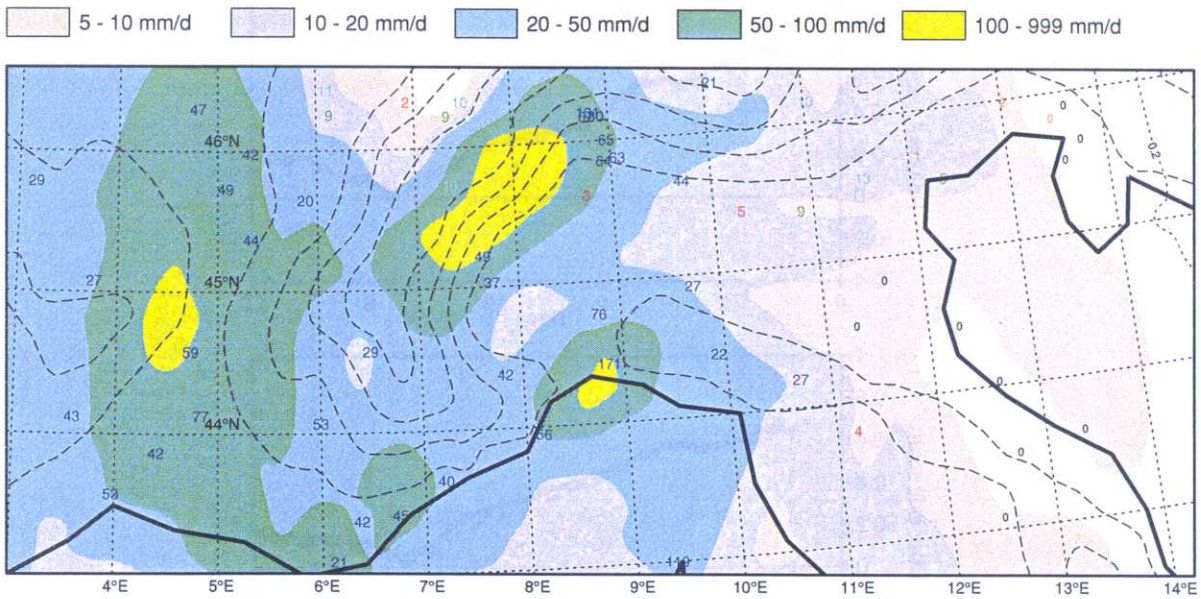


Fig 3.16 10 metre wind, mean sea level pressure and relative humidity at 925 hPa for 72 hour forecasts from 19930920. (a) High resolution forecast T639 (b) and current operational resolution T319.

a) High resolution forecast T639



b) Current operational resolution T319

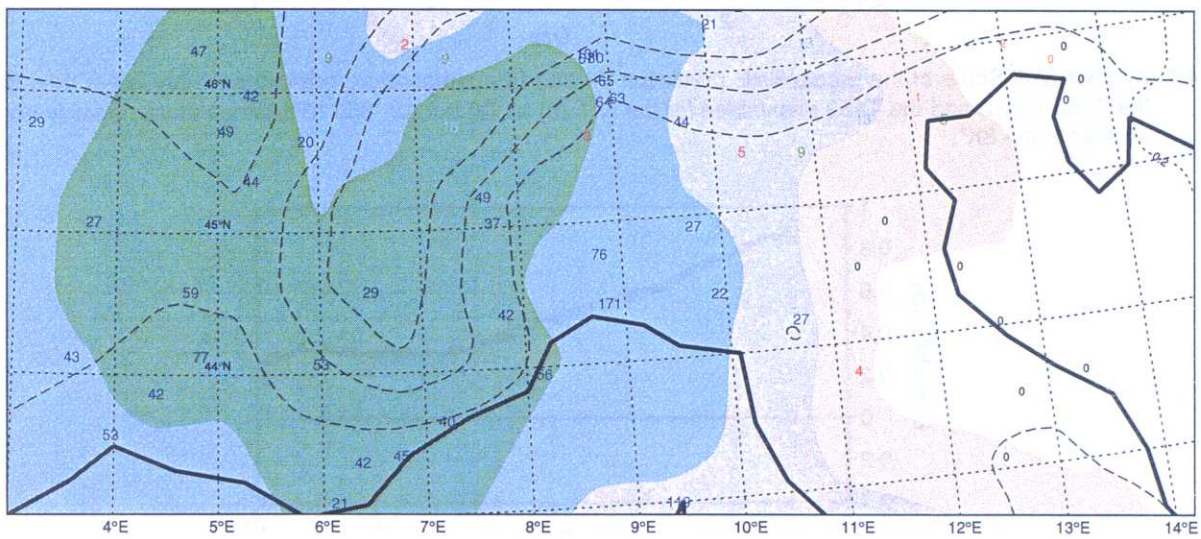


Fig 3.17 Verification of precipitation as an average over two forecasts from 19930920 and 19930921, range 48 to 72 hours. The numbers show the observed precipitation and the contours are the forecast precipitation (units: mm/day). Dashed lines are contours of the respective orography. (a) High resolution forecast T639 and (b) current operational resolution T319.

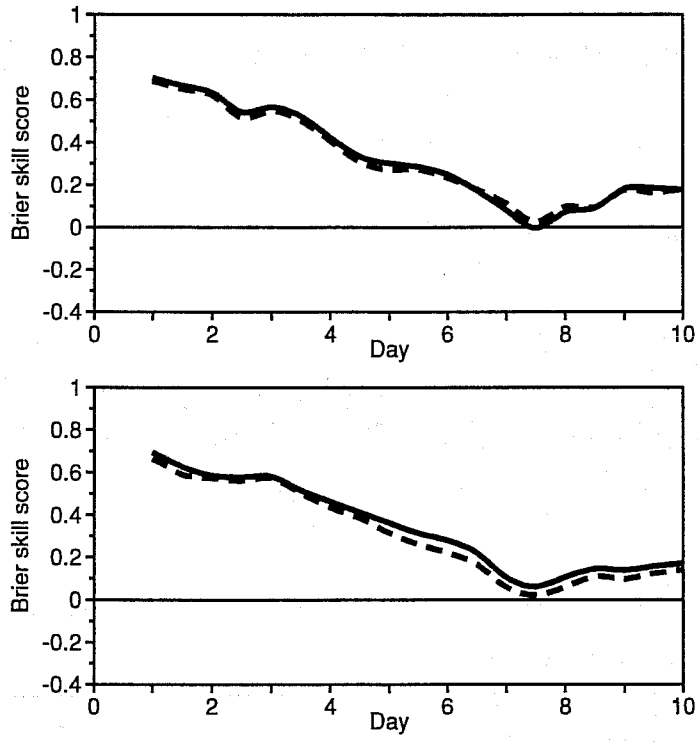


Fig 4.1 (a) Brier Skill Score of the probabilistic prediction of "850 hPa temperature positive anomaly larger than 8K" for the T255 (solid) and the T159 ensembles for the NH. (b) as (a) but for "850 hPa temperature negative anomaly smaller than - 8K".

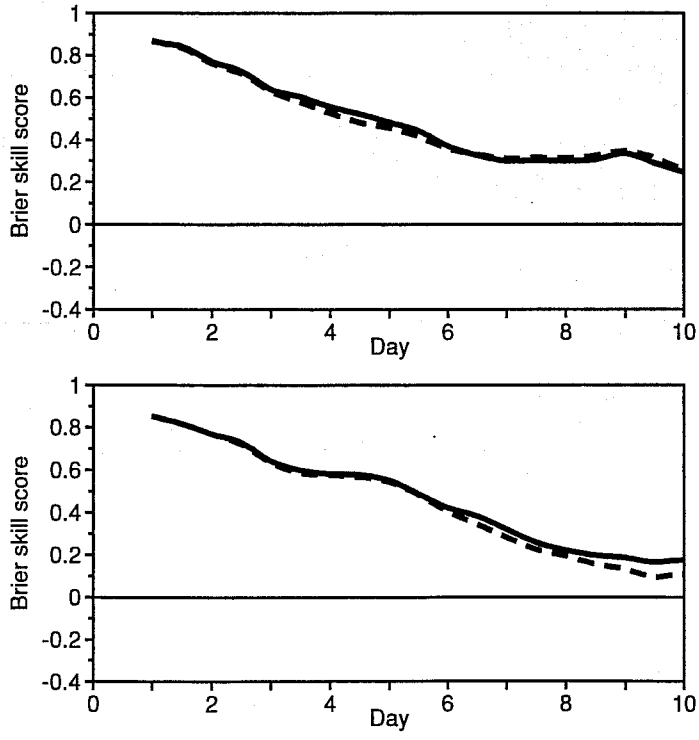


Fig 4.2 (a) Brier Skill Score of the probabilistic prediction of "500 hPa geopotential height positive anomaly larger than 100 m" for the T255 (solid) and the T159 (dash) ensembles for the NH. (b) as (a) but for "500 hPa geopotential height negative anomaly smaller than -100 m".Fig 4.1(a) Brier Skill Score of the probabilistic prediction of "850 hPa temperature positive anomaly larger than 8K" for the T255 (solid) and the T159 ensembles for the NH. (b) as (a) but for "850 hPa temperature negative anomaly smaller than - 8K".

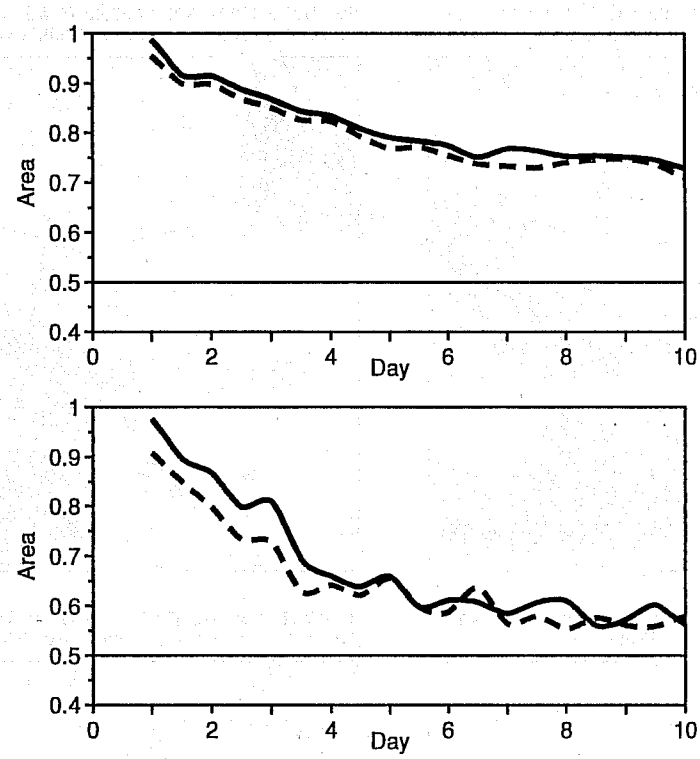


Fig 4.3 Top panel: area under the mean ROC curve (average of the 5 cases) for the probabilistic prediction of more than 5 mm/12h for the T255 (solid) and the T159 (dash) ensembles over the NH. Bottom panel: as top panel but for 20 mm/12h.

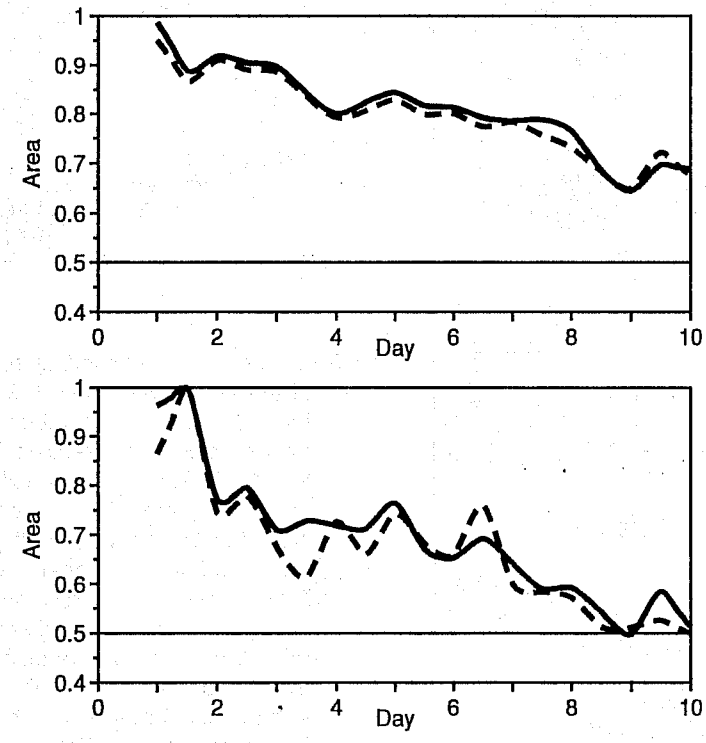


Fig 4.4 As Fig 4.3 but for central United States (20°N-50°N, 110°W 80°W).

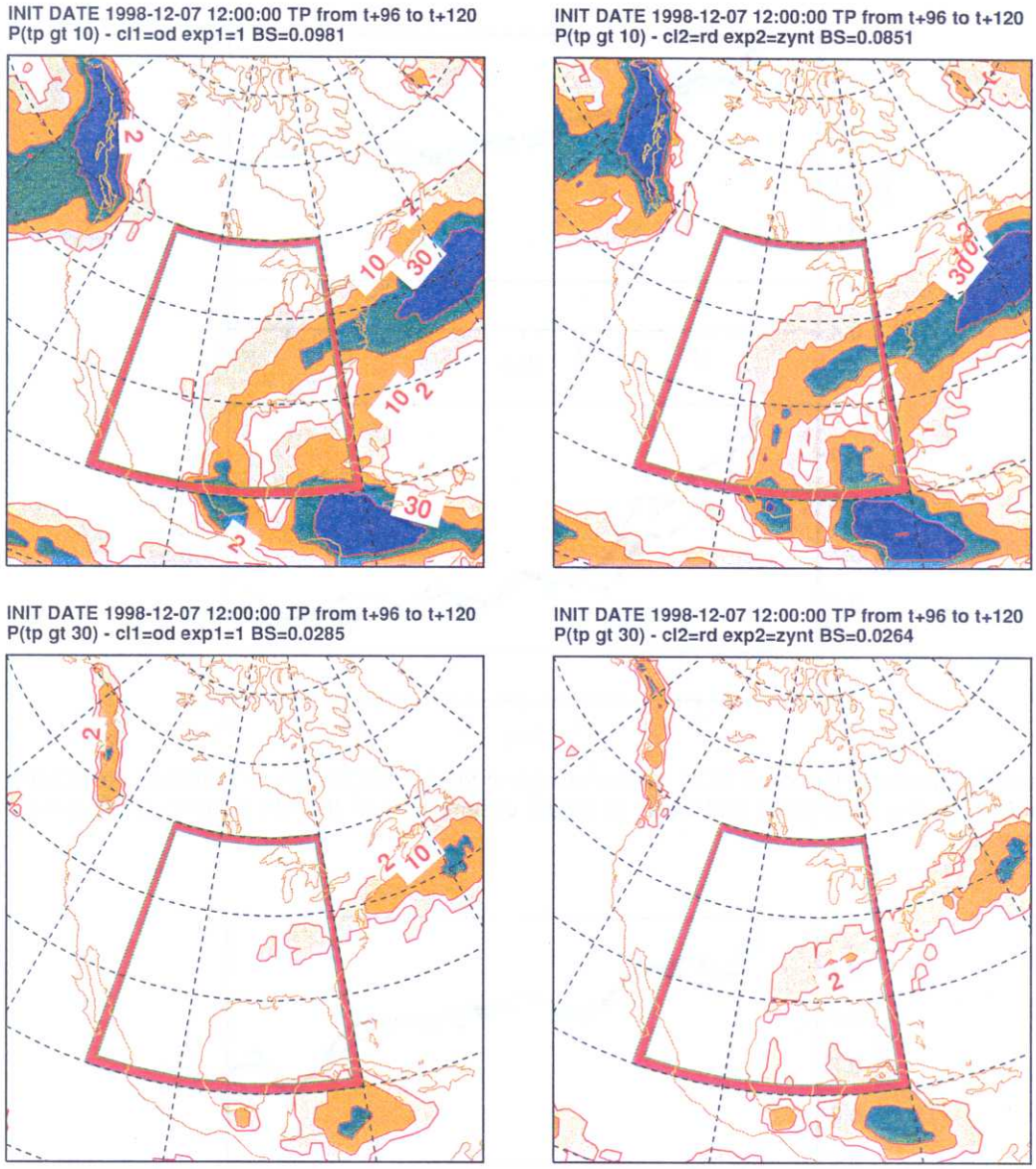
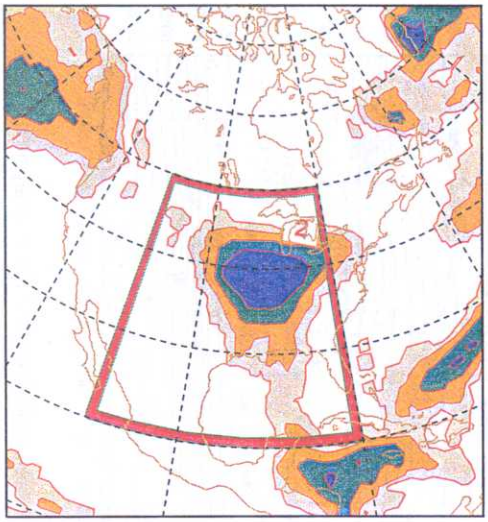
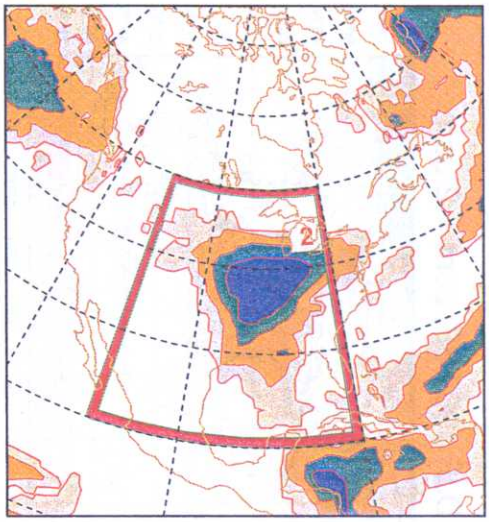


Fig 4.5 Ensemble predictions started at 12GMT of the 7th of December 1998 and valid for the 12th of December, based on total precipitation forecasts from t+96h to t+120h. Top row: probability of more than 10 mm/24h of precipitation predicted by the T159 (left panel) and the T255 (right panel) ensembles. Middle row: as top row but for more than 30 mm/d of precipitation. Bottom row: precipitation verification field, defined as the 0-24h T319 forecast started at 12GMT of the 11th of December. Contour isolines for 2, 10, 30 and 60% for probabilities, and for 2, 10, 30 and 60 mm/24h for precipitation verification.

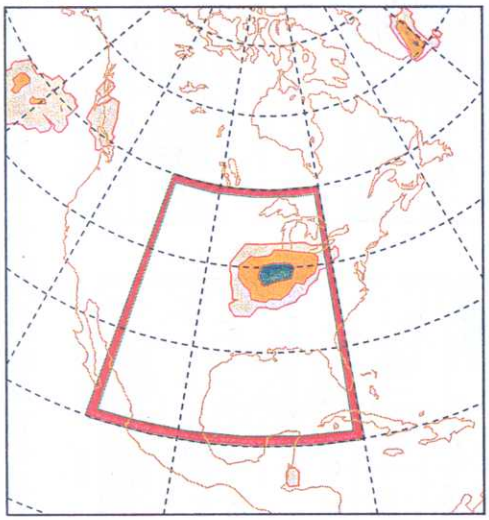
INIT DATE 1998-12-28 12:00:00 TP from t+96 to t+120
P(tp gt 10) - cl1=od exp1=1 BS=0.0751



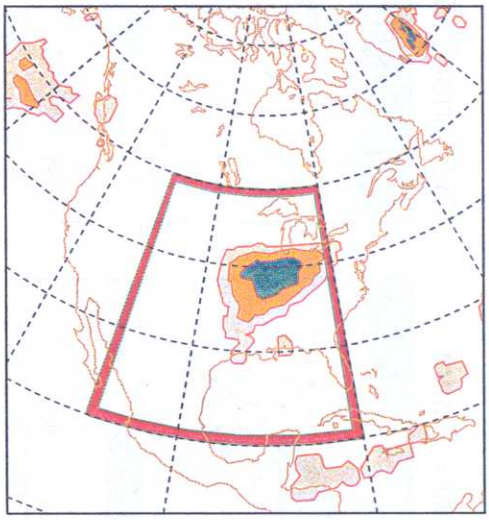
INIT DATE 1998-12-28 12:00:00 TP from t+96 to t+120
P(tp gt 10) - cl2=rd exp2=zynt BS=0.0619



INIT DATE 1998-12-28 12:00:00 TP from t+96 to t+120
P(tp gt 30) - cl1=od exp1=1 BS=0.0399



INIT DATE 1998-12-28 12:00:00 TP from t+96 to t+120
P(tp gt 30) - cl2=rd exp2=zynt BS=0.0355



VERIF DATE 1999-01-01 12:00:00 TP from t+0 to t+24
VERIFICATION - mean abs precip=5.006

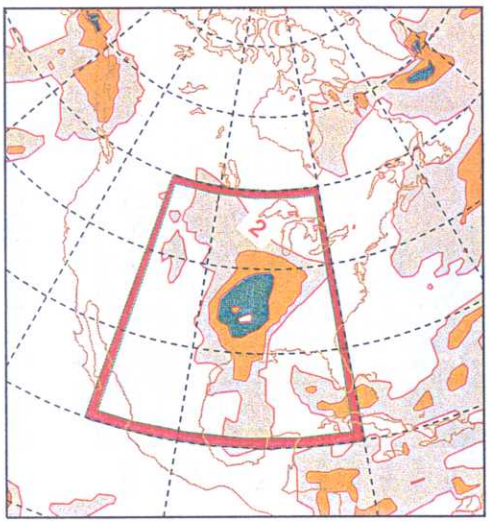


Fig 4.6 As Fig 4.5 but for ensemble predictions started at 12GMT of the 28th of December 1998 and valid for the 2nd of January 1999.

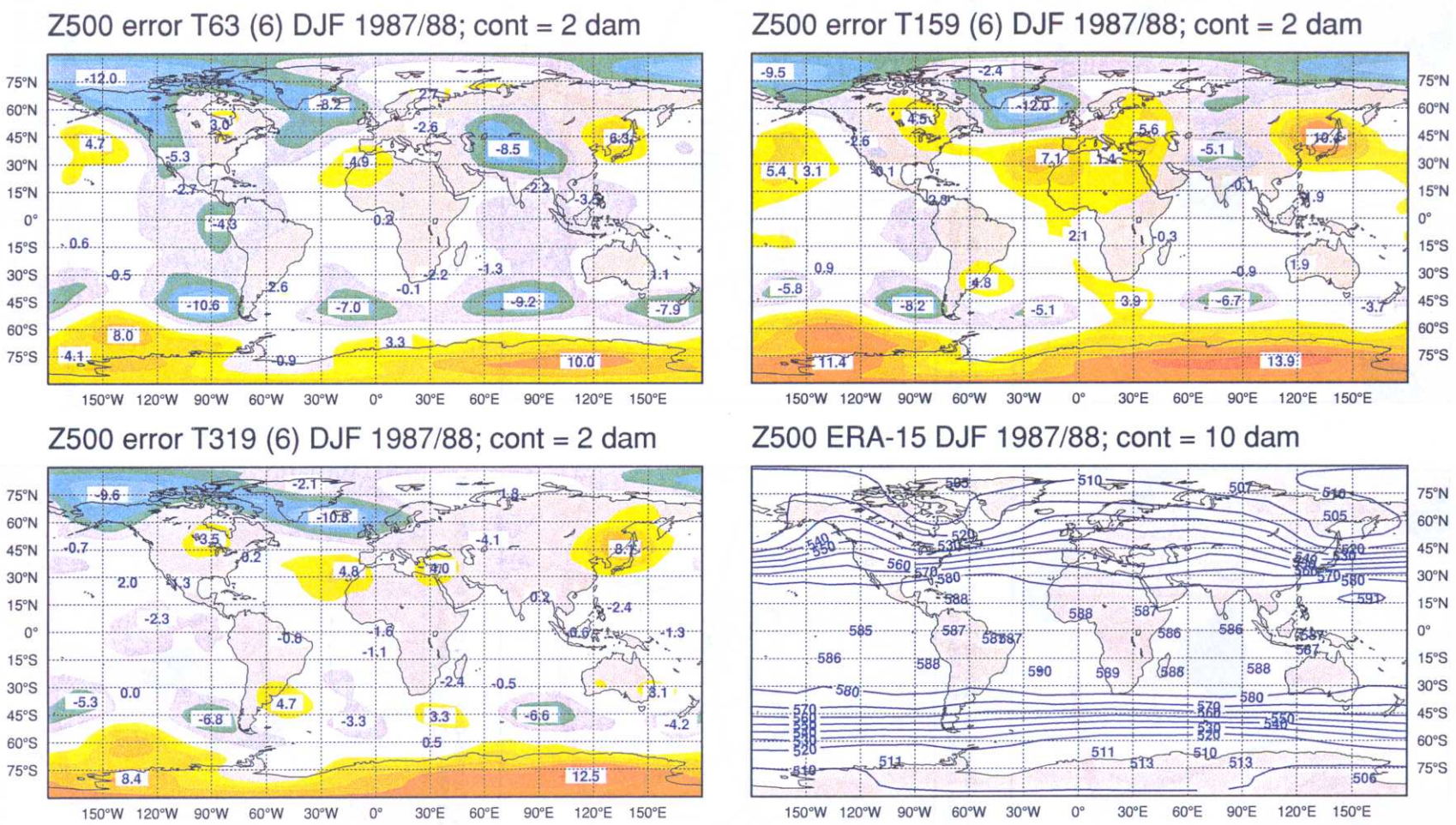


Fig 5.1 The ensemble mean 500 hPa height errors (in dam) for DJF 87/88 with respect to ERA-15 for the three resolutions and the reanalysis full field.

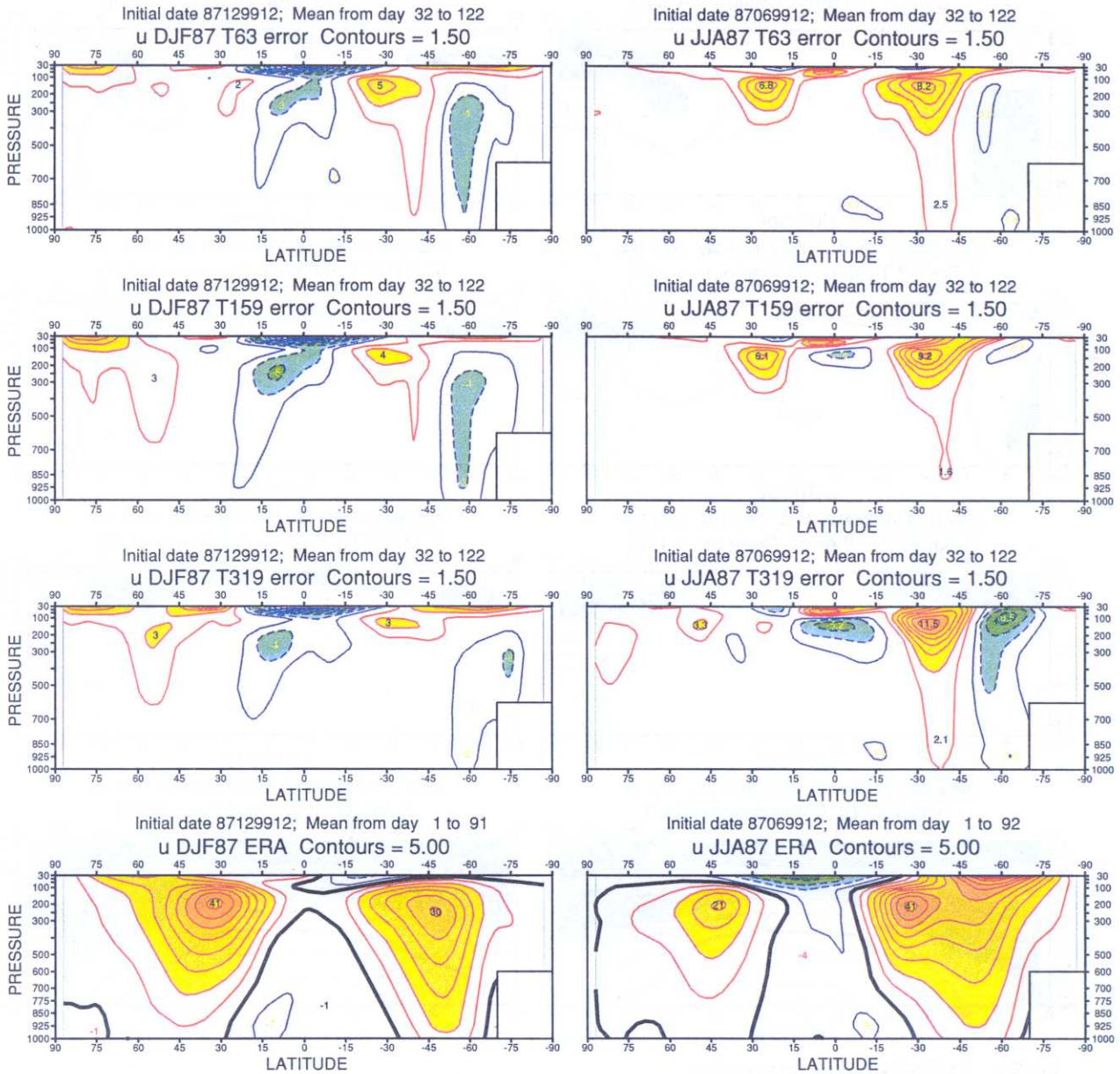


Fig 5.2 Latitude-pressure cross-sections of the zonally-averaged zonal wind error (in m/s) at each resolution for DJF 87/88 on left and JJA 87 on right together with full field from ERA.

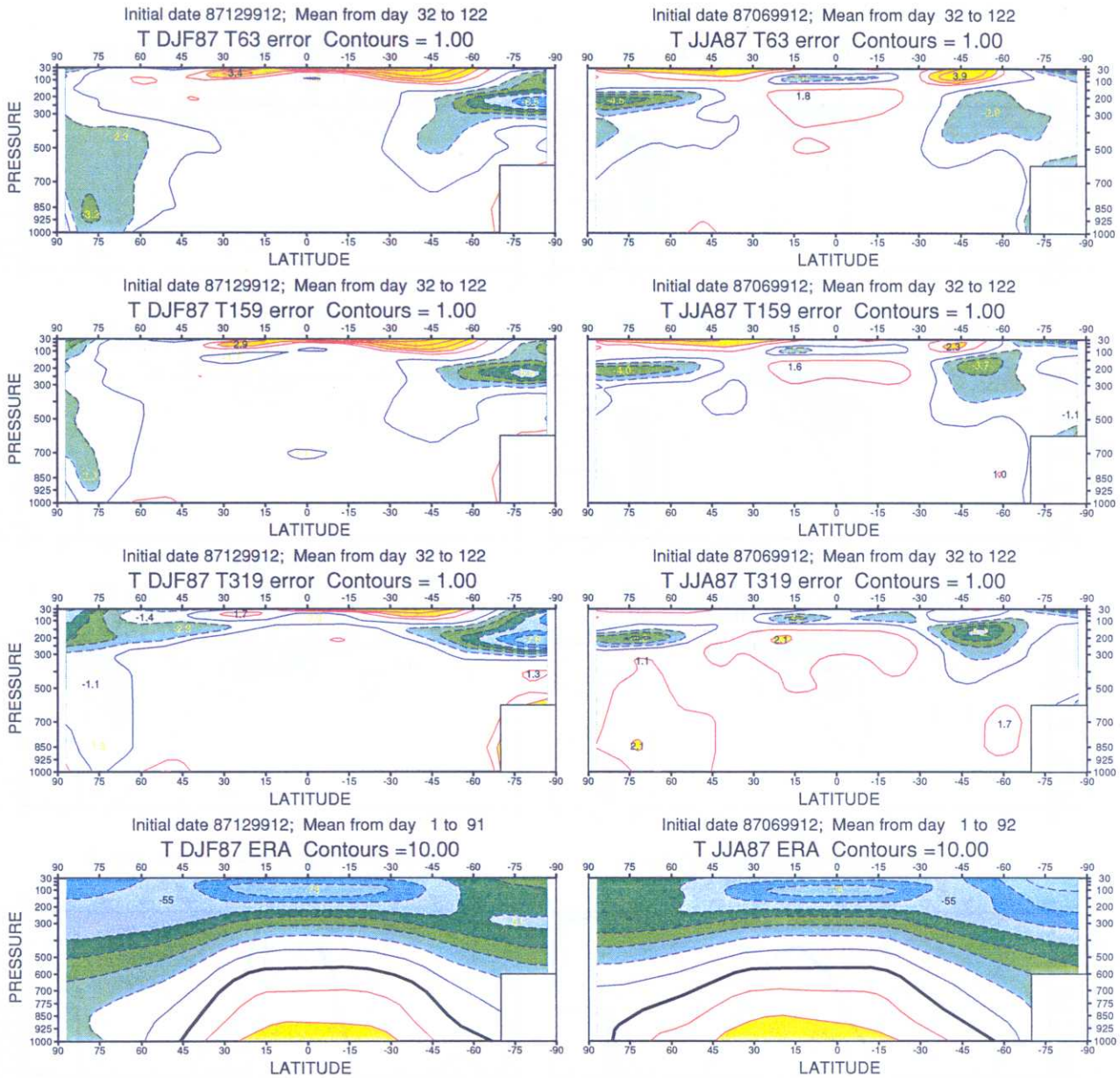


Fig 5.3 As Fig 5.2 but for temperature.

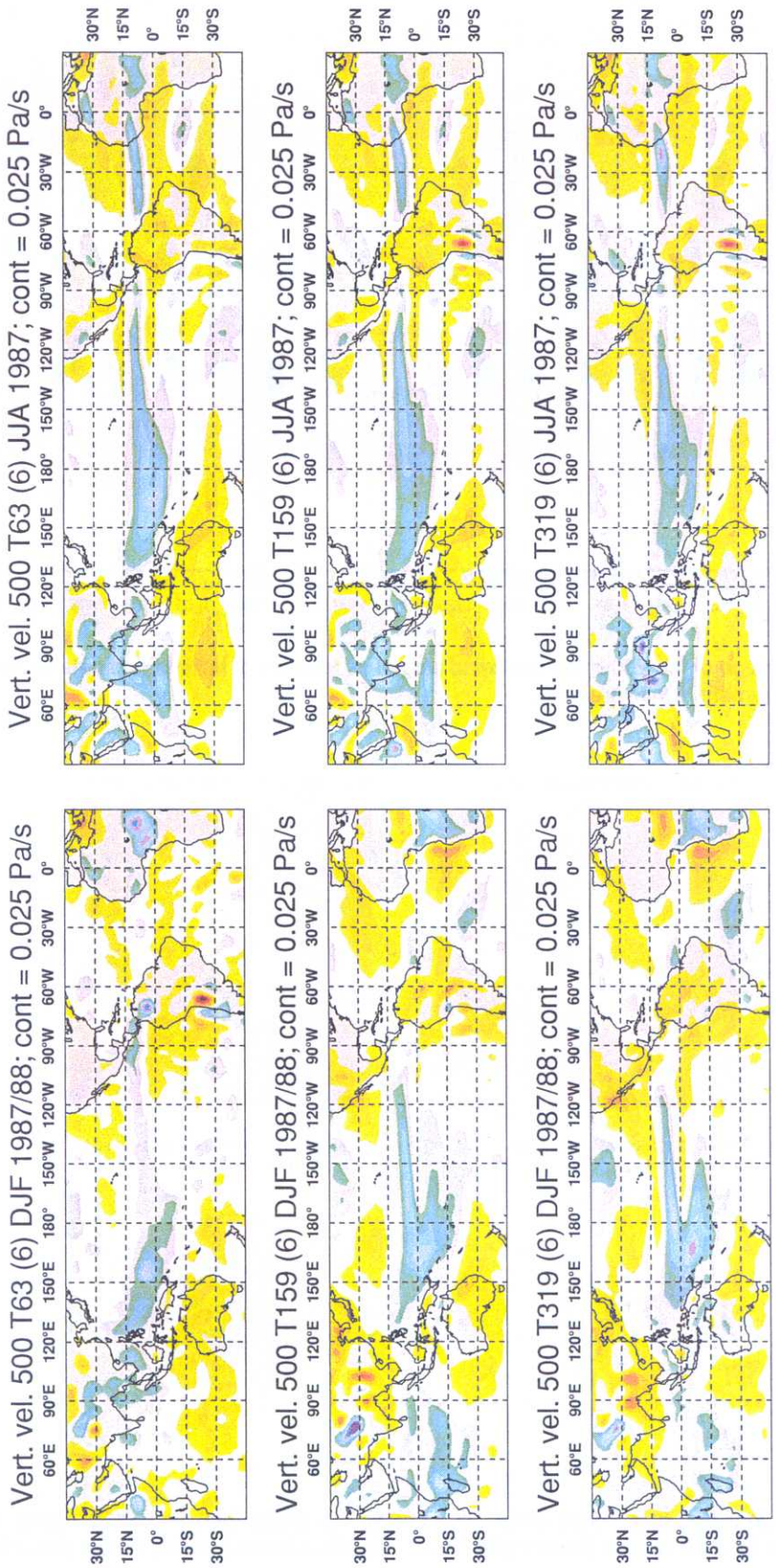


Fig 5.4 The ensemble mean vertical velocity (in Pa/s) at lower latitudes at each resolution for DJF 87/88 on left and JJA 87 on right.

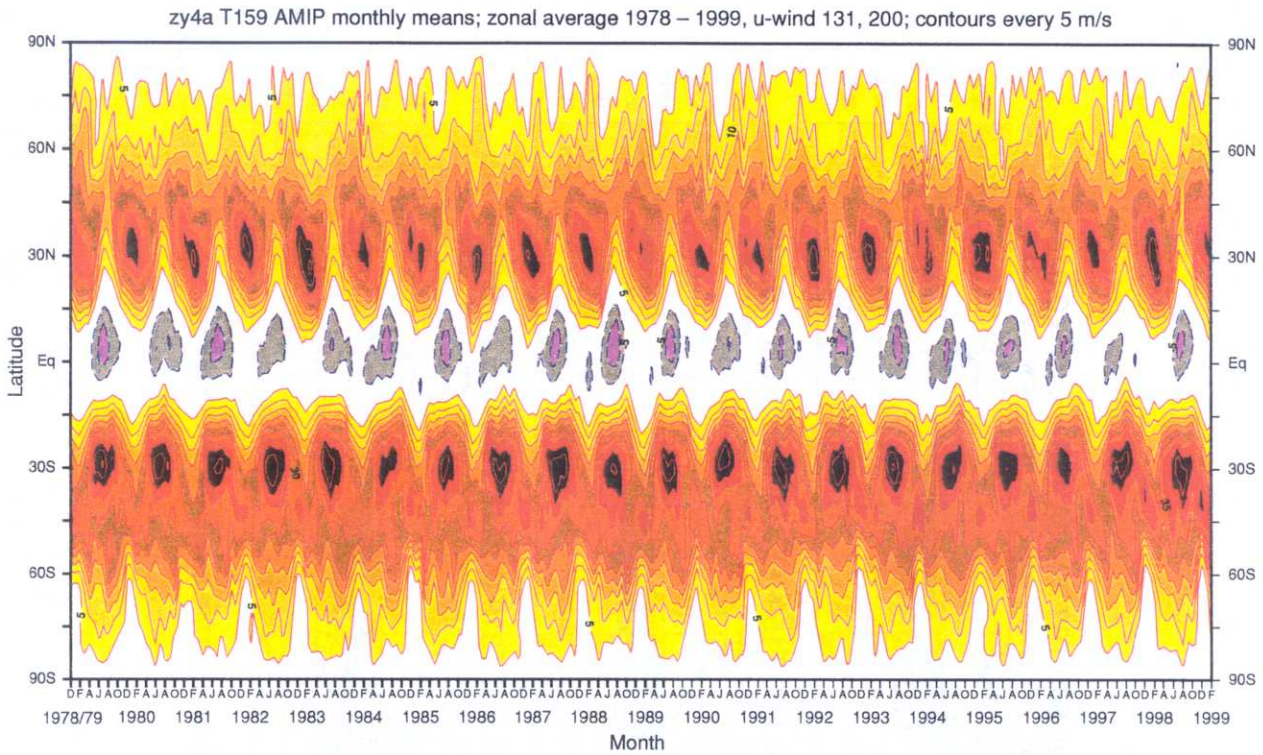


Fig 5.5 The zonal average of the 200 hPa zonal wind plotted as function of latitude and time for the AMIP-2 T159L60 experiment. (Units are m/s).

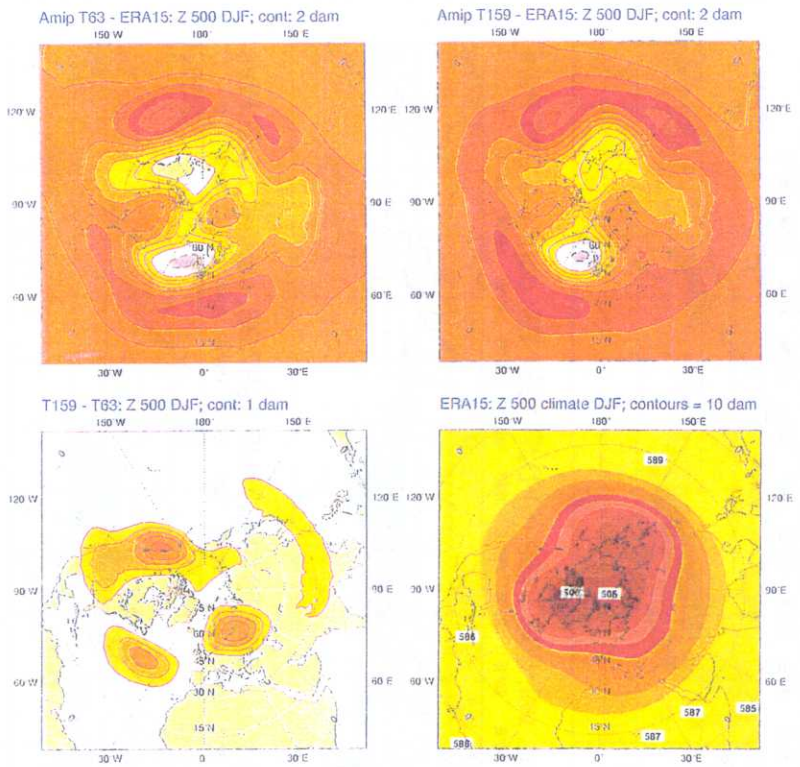


Fig 5.6 The DJF 20 year 500 hPa height errors (in dam) at T63 (upper left) and T159 (upper right) with respect to the ERA-15 climate (lower right). The T159/T63 difference is also shown (lower left).

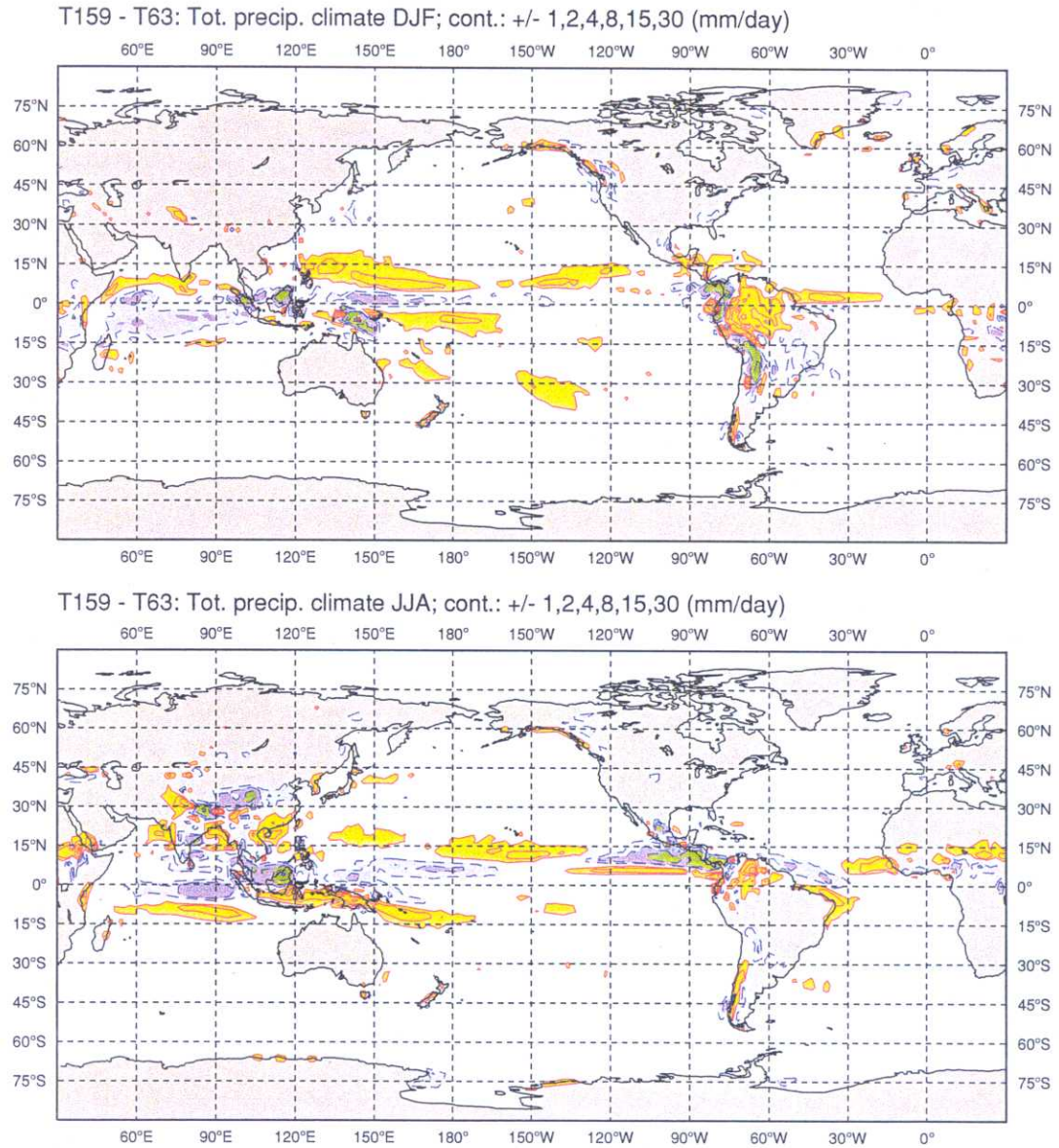


Fig 5.7 Difference fields (T159 minus T63) of total precipitation (in mm/day) for DJF (upper) and JJA (lower). Yellow/red are positive values, blue/green are negative.