THE VERIFICATION AND DIAGNOSTICS PACKAGE AT ECMWF

Klaus Arpe

ECMWF

1. INTRODUCTION

When planning future diagnostic investigations of the atmosphere or its simulations by models, we need to know which data and tools are available to carry them out. Therefore a short overview about the ECMWF archives and the use of them will be given.

2. OPERATIONAL ARCHIVES

Table 1 shows the data which are archived operationally and will be kept for a very long time. They are observed data, the analysed and forecasted fields. We distinguish between upper air fields, which are the mass-, wind- and humidity fields, and surface fields, which are mainly connected with subgrid scale parameterizations. The upper air fields are archived in spectral form. For analyses a T80 and for forecasts a T40 truncation is used.

The surface fields are archived in the grid in which they were generated because general ways of compressing the data to a smaller set of coefficients were not known. In Table 1 it is said that all surface fields are archived for the analyses, this is not really true because most of these fields are not analysed. Instead the values of the first guess made by the model are taken unless real analysed data are available.

The time interval for archiving analysis data is 6h and for forecast data 12h.

The analysis data are archived only after applying a normal mode initialisation (Temperton and Williamson, 1979). This causes some problems because the observed data and the analysed fields may differ considerably even over data dense areas due to the normal mode initialisation. For instance, differences of 3 mb in the surface pressure have been found over Europe and the divergent wind in the stratosphere is often changed to a different pattern.

These archived data can be used for producing a lot of maps and diagnostic parameters and packages to do so are available at ECMWF. Tiedtke (1980) will give some examples.

When we had to decide which presentation or resolution should be used for archiving the analysis and forecast data, quite severe restrictions were given by the vast amount of data which are produced each day. The optimum compromise seemed to be a spectral representation packed in 15 bit words with a T40 truncation for forecast data. We have not had time to look at all consequences of our choice of archive, but a few deficiencies were already noticed.

Table 1. **List of 'archived files' during an operational run**

1. Raw Data for the period 0 - 24h

   Surface Reports
   Upper Air Reports
   Satob Reports
   Bathy/Tesac Reports
   Temp Reports
   Pilot Reports
   Satem Reports

2. Analysis made at 00Z, 06Z, 12Z, 18Z

   Surface fields (grid point)

   Mean Sea Level Pressure
   Surface Temperature
   Soil wetness
   Snow depth
   Large scale rain
   Convective rain
   Snowfall
   Boundary layer dissipation
   Sensible heat flux from the surface
   Latent heat flux from the surface
   Surface stress
   Surface net radiative flux
   Top net radiative flux

   Upper air fields (spectral T80)

   Parameters: U,V,W,T,Q,Z
   Levels: 1000, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30

3. Forecasts - 12 h intervals

   Surface fields (grid point) of same parameters as in 2.

   Upper air fields (spectral T40)

   Parameters: U,V,W,T,Q,Z
   Levels: 1000, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50

In Fig. 1, lower lines, the spectrum in zonal wavenumber domain for the 500 mb height field variance is shown comparing a T40 and T80 representation. The impact of the meridional smoothing by a T40 representation can be seen from wave number 17 onwards. In Fig. 1 a global and monthly average is shown but spectra of small latitudinal belts and single cases show similar differences. If higher derivatives of the fields are used, as e.g. geostrophic kinetic energy or geostrophic enstrophy (middle and upper pairs of lines), the impact can be seen at much lower wave numbers.

On the other hand, it seems to be not worth using a higher than a T80 truncation, at least for the 500 mb height fields, because of the packing in 15 bit words. The range of coefficients lies between -300 m and +300 m (the area mean is treated separately) and from that a smallest increment in packing the coefficients is 0.02 m. This is a reasonable accuracy because the observed data are much lower in accuracy. This limit is, however, reached in the spherical harmonics coefficients of 500 mb height fields for most indices higher than 60. The coefficients of these short waves are then set to an increment nearest to zero. When going back to grid points from these truncated coefficients extra noise is introduced and may become important for derived quantities like geostrophic enstrophy, where the noise can reach a level which is of the same magnitude as the values of the longer waves.

The impact of truncation can also be seen in maps. In Fig. 2 we compare a 5-day mean 700 mb vertical wind field which is truncated in three different ways. The top panel shows the field using a T40 spectral representation with only 20 zonal wave numbers, the middle panel is the same field but with a Fourier series using 20 waves and a meridional 1.875° grid resolution; and the lower panel is again the same field but with a T80 spectral representation with only 50 zonal wave numbers. The low resolution map at the top is probably the best for looking at synoptic scales. The differences to the other maps are mainly in meridional resolution, especially near the equator over Africa. The middle panel, which has the highest resolution in meridional direction, shows a strong local circulation with downward motion of more than 50 mPa/s (5×10^{-4} mb/s) and upward motion of 100 mPa/s (10×10^{-4} mb/s) already one grid point to the poles. The strong local circulation, which might be an important deficiency in the model, would be concealed when using only a T40 representation. In the lower map showing the high resolution spherical harmonic representation this local circulation can be seen but not as clearly as in the middle panel.

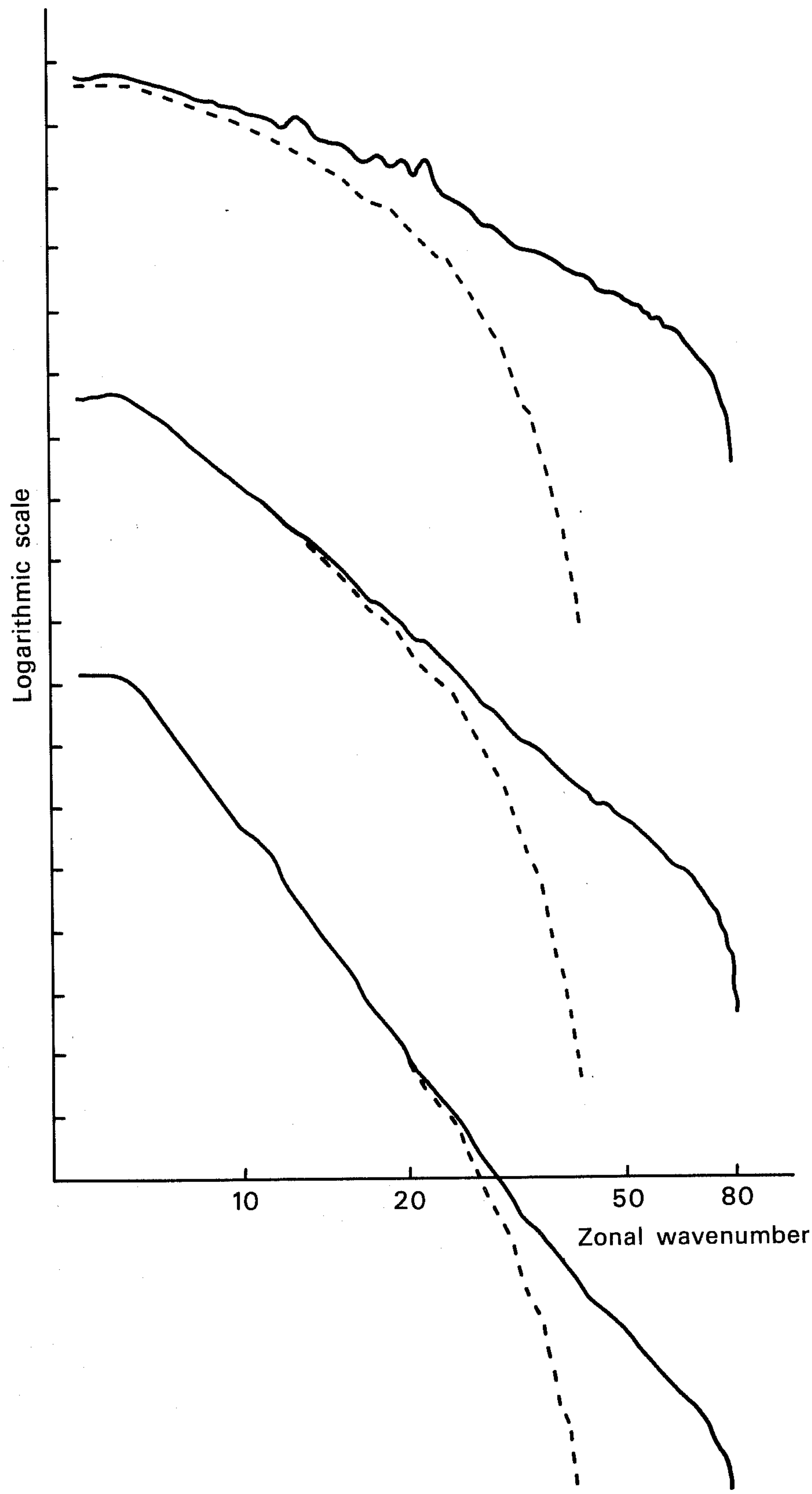

Fig. 1 Mean zonal spectra of the 500 mb height fields and their derived quantities of all November 1979 12Z-analyses. The analysed fields have been truncated by spherical harmonics using either triangular 40 (dashed lines) or triangular 80 (solid lines) representation.
Bottom pair of lines: variances of height
Middle pair of lines: geostrophic kinetic energy
Top pair of lines: geostrophic enstrophy

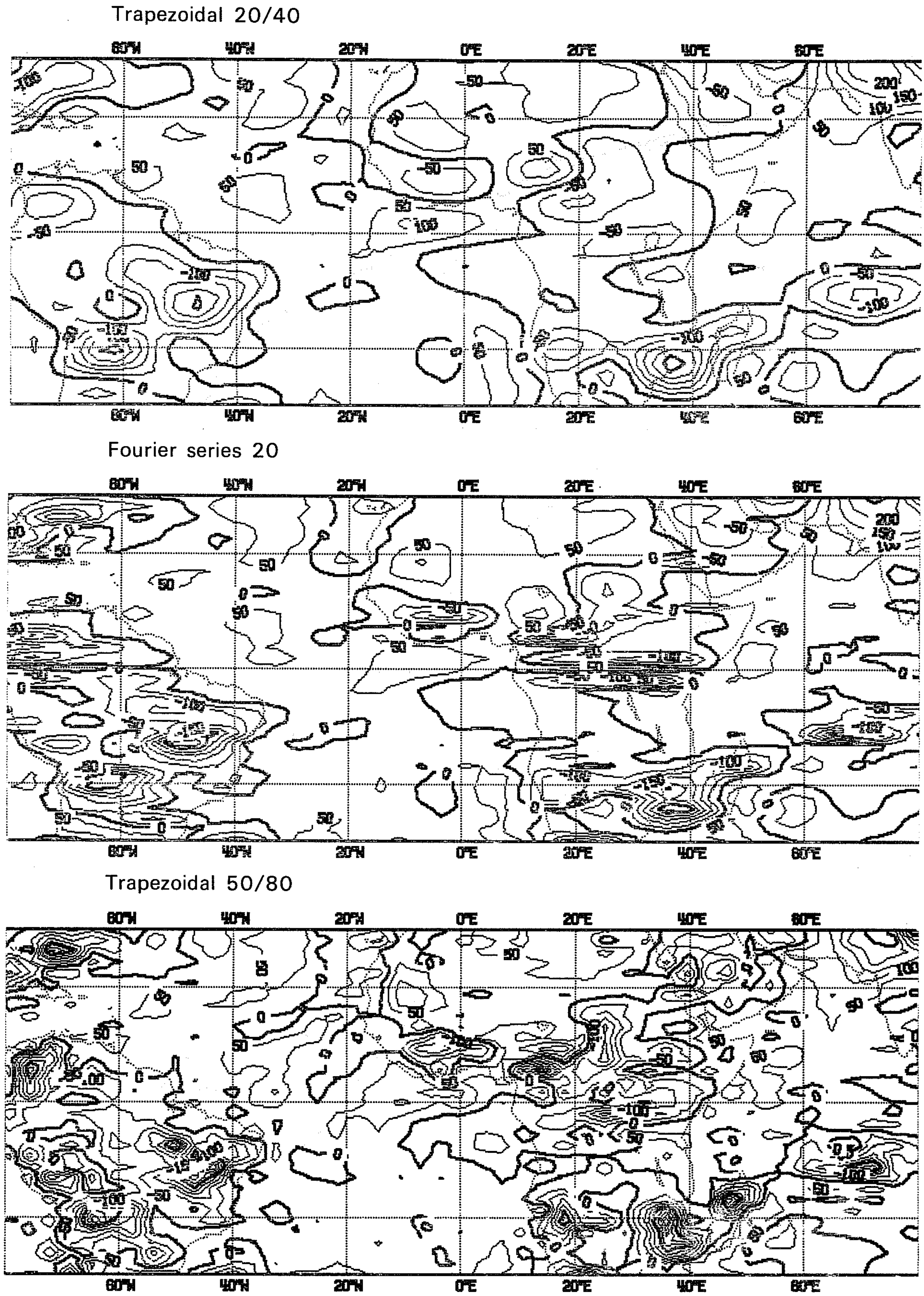

Fig. 2 Mean vertical wind (10^{-3} Pa/s) at 700 mb of a forecast from the 31 Jan 1980 12Z, averaged between day 3 and 7. Differences between the three maps are due to different truncations.
Bottom panel: trapezoidal 50/80 spherical harmonics
Middle panel: zonal wavenumber 20 fourier series and model's grid resolution in north-south direction.
Top panel: trapezoidal 20/40 spherical harmonics

3. SPECTRA, COSPECTRA AND ENERGETICS ON LATITUDE-PRESSURE CROSS-SECTIONS

Partly to overcome some of the problems mentioned above but more for convenience and economical reasons a further set of data is created each day. The parameters are given in Table 2. These spectra, cospectra and energetic properties are calculated on the horizontal grid of the model but using data which have been interpolated to pressure levels. As no horizontal smoothing or packing has been done, these data allow some insight into contributions also by the highest wave numbers. Fig. 3 shows as an example three terms of the eddy kinetic energy budget for the shortest waves (wave numbers 16 to 96). The transfer between the zonal flow and these small scale eddies (CK, lowest panel), is one order of magnitude smaller than both other terms and can therefore be neglected. The top panel shows that the eddy kinetic energy of shortest waves is created by conversion from eddy available potential energy ($\overline{\omega '\alpha '}$) and then transferred by non-linear interactions from these short waves to longer waves (LK, middle panel). These budget terms are an average of 21 forecasts of the 10th day and are also available for analyses and other forecast periods.

This set of data can in principle be reproduced from the operational archives, except the contribution by the very short waves, and is therefore kept for a short period only, but will be used to create monthly means or means of other length.

4. BUDGET CALCULATIONS IN THE MODEL'S GRID FOR ANALYSES AND FORECASTS

During the interpolations from sigma-surfaces to pressure levels, inconsistencies may be introduced into the mass- and wind fields (see e.g. Mahlman and Moxim, 1976). These inconsistencies are especially obvious in budget calculations over limited areas. As it is impossible to archive all data in the model's grid, it was decided to do some budget calculations for some selected areas while the original history files of the forecasts or initialised analyses are still available and to be kept for shorter periods to calculate e.g. monthly means.

Table 3 gives an overview over the available data. They are complete enough for calculations of the budgets of temperature, kinetic energy and humidity. The resulting residuals can be interpreted as being due to sub-grid scale processes either as observed in the analyses or parameterized in the model (for examples see Savijärvi, 1980).

Table 2. Available latitude-height cross-section

1. Spectra (zonal mean, wavenumbers 1 to 15, sum of higher wavenumbers)

   geopotential height
   temperature
   u - zonal wind component
   v - meridional component
   ω - vertical wind component in pressure coordinates
   q - specific humidity

2. Cospectra in four wavenumber groups (1-3, 4-9, 10-15, higher wavenumbers)

   $\omega\phi$ - vertical geopotential flux
   $v\phi$ - meridional geopotential flux
   $v\omega$

3. Cospectra in five wavenumber groups (same as 2 plus zonal mean)

   uv - meridional momentum flux
   $u\omega$ - vertical momentum flux
   Tv - meridional flux of sensible heat
   $T\omega$ - vertical flux of sensible heat
   qv - meridional flux of humidity
   $q\omega$ - vertical flux of humidity

4. Energetics in four or five wavenumber groups (same as 2 or 3)

   KE - kinetic energy
   AE - available potential energy
   C - conversion between KE and AE
   CK - transfer between eddy and zonal kinetic energy
   CA - transfer between eddy and zonal available potential energy
   LK - transfer between different wavenumber groups of kinetic energy

produced for initialized analyses each 6h and for forecasts at 12h, 24h, 48h, 72h...240h.

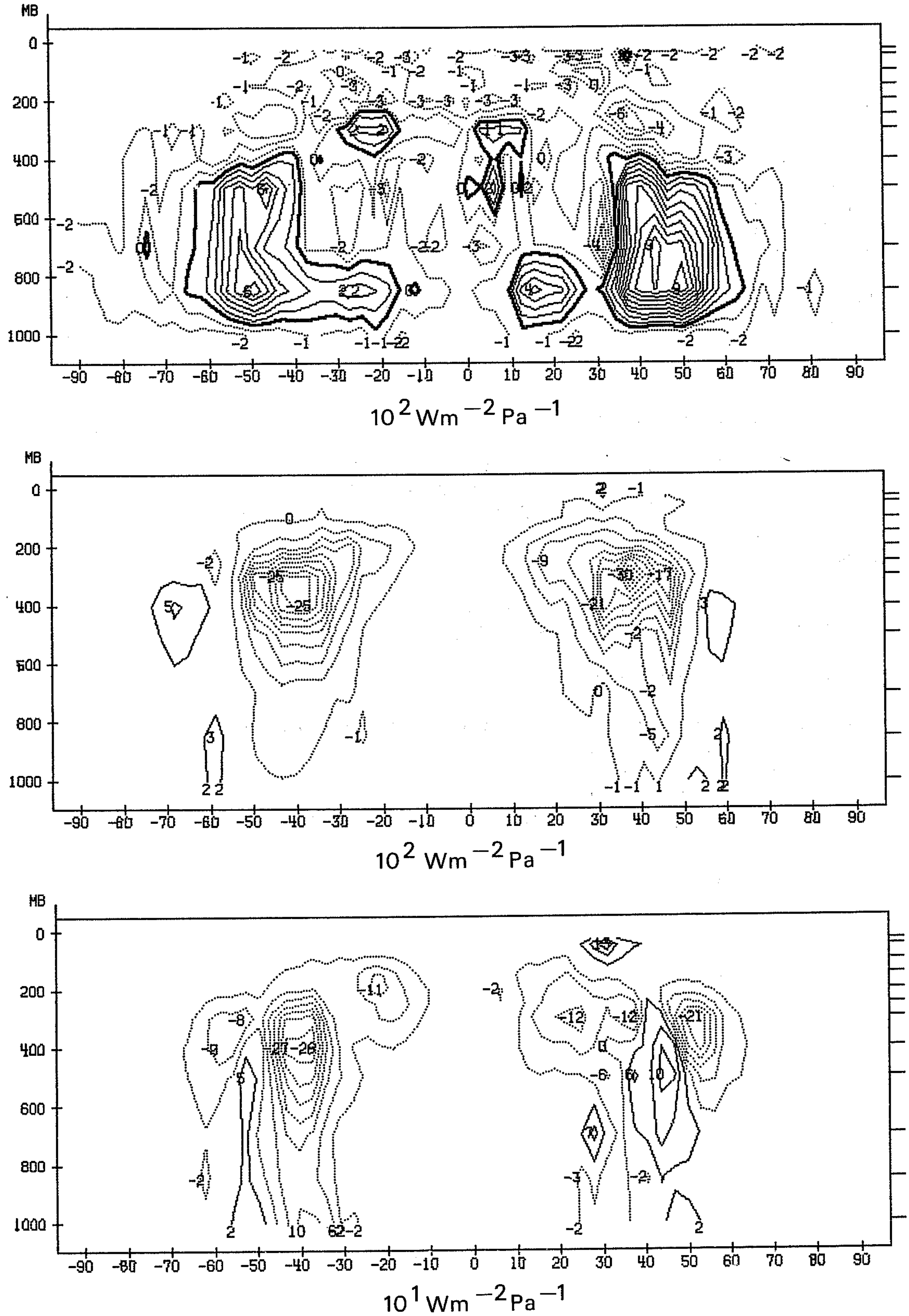

Fig. 3 Mean sources (positive values) and sinks (negative values) of eddy kinetic energy with zonal wavenumbers shorter than 15 due to following processes.

Bottom panel: transfer from zonal kinetic energy, CK

Middle panel: transfer from eddy kinetic energy of longer waves, LK

Top panel: conversion from available potential energy, C.

21 cases of forecasts of the 10th day have been averaged. Note the different units.

Table 3. Model grid diagnostics (sigma levels)

1. Zonal means

Temperature
specific humidity
3 wind components
kinetic energy
generation of KE : $\mathbb{V}$. grad ϕ
conversion: $\alpha.\omega$ on sigma levels
fluxes of sensible heat, humidity and momentum
  Tv, $T\omega$, Qv, $Q\omega$, uv, $u\omega$

2. Budget terms

2.1 for the following areas:

Arctic and Antarctic	75 - 90 N and 75 - 90 S
Mid latitude belts	25 - 75 N and 25 - 75 S
Tropics	25 S - 25 N
Global land areas	
Global ocean areas	
Europe (10W, 70N, 40E, 35N)	
Northern Atlantic (60W, 70N, 10W, 30N)	
North America (140W, 70N, 80W, 30N)	

2.2 of following budget terms:

horizontal and vertical flux divergence of temperature,
humidity, kinetic energy,
and geopotential
source terms $\alpha\omega$ and $\mathbb{V}$. grad ϕ
area means of temperature, humidity, kinetic energy

Table 4. **Tendencies of momentum, temperature and humidity**

Zonal means over land, sea, and over both areas are calculated due to the following processes:

1. Diabatic processes parameterized in the model

   Condensation due to cumulus convection
   Evaporation due to cumulus convection
   Condensation due to large-scale convection
   Evaporation due to large-scale convection
   Net-radiative heating
   Horizontal and vertical diffusion of temperature, humidity and momentum separately
   Dissipation of kinetic energy due to horizontal or vertical diffusion

2. Adiabatic processes

   Tendencies of momentum due to horizontal and vertical advection, coriolis-and pressure gradient forces separately
   Horizontal and vertical temperature advection separately
   Energy conversion
   Horizontal and vertical humidity advection separately

3. Means of the temperature, horizontal wind and humidity fields.

5. TENDENCIES OF MOMENTUM, TEMPERATURE AND HUMIDITY IN THE MODEL

The effects of sub-grid scale processes can be seen only as vertical integrals when using the archived data. For example we get from the condensation processes only the amount of rain reaching the ground. It is already quite difficult to verify such data but, nevertheless, it would be very interesting to know also their vertical distribution at least from the model's simulations of the atmosphere. Therefore another set of data is saved for a shorter period. Table 4 gives an overview over the available terms.

To keep the amount of data at a reasonable level, only zonal means over land, sea or both areas are saved for all latitudes and vertical levels of the model.

It is important to do the time averaging in relatively short intervals to get a reasonable representation of sub-grid scale processes. The archiving can be done in a much coarser interval. At the moment the averaging is done from data in an interval of 3 hours and the archiving is done each day of the forecast.

To give an example, we can see in Fig. 4 from two comparative forecasts the vertical and meridional mean distribution of the heating of the atmosphere due to two different parameterizations of cumulus convection.

For completeness also tendencies due to adiabatic processes are added to this data set. In these adiabatic terms there is a small overlapping with data mentioned above, but easier handling of the plotting programs seemed to be worth doing this double archiving.

6. SUMMARY

At ECMWF a good selection of data in the form of observations, analyses, forecasts and derived quantities is available, although not perfect. The diabatic processes are mainly represented as model output. From analyses we can get some ideas about diabatic processes from residuals of budget calculations or by assuming that short range forecasts with our model would give data similar to analyses.

Although more data in the archive may be wanted, it is already now an enormous task to digest what is available.

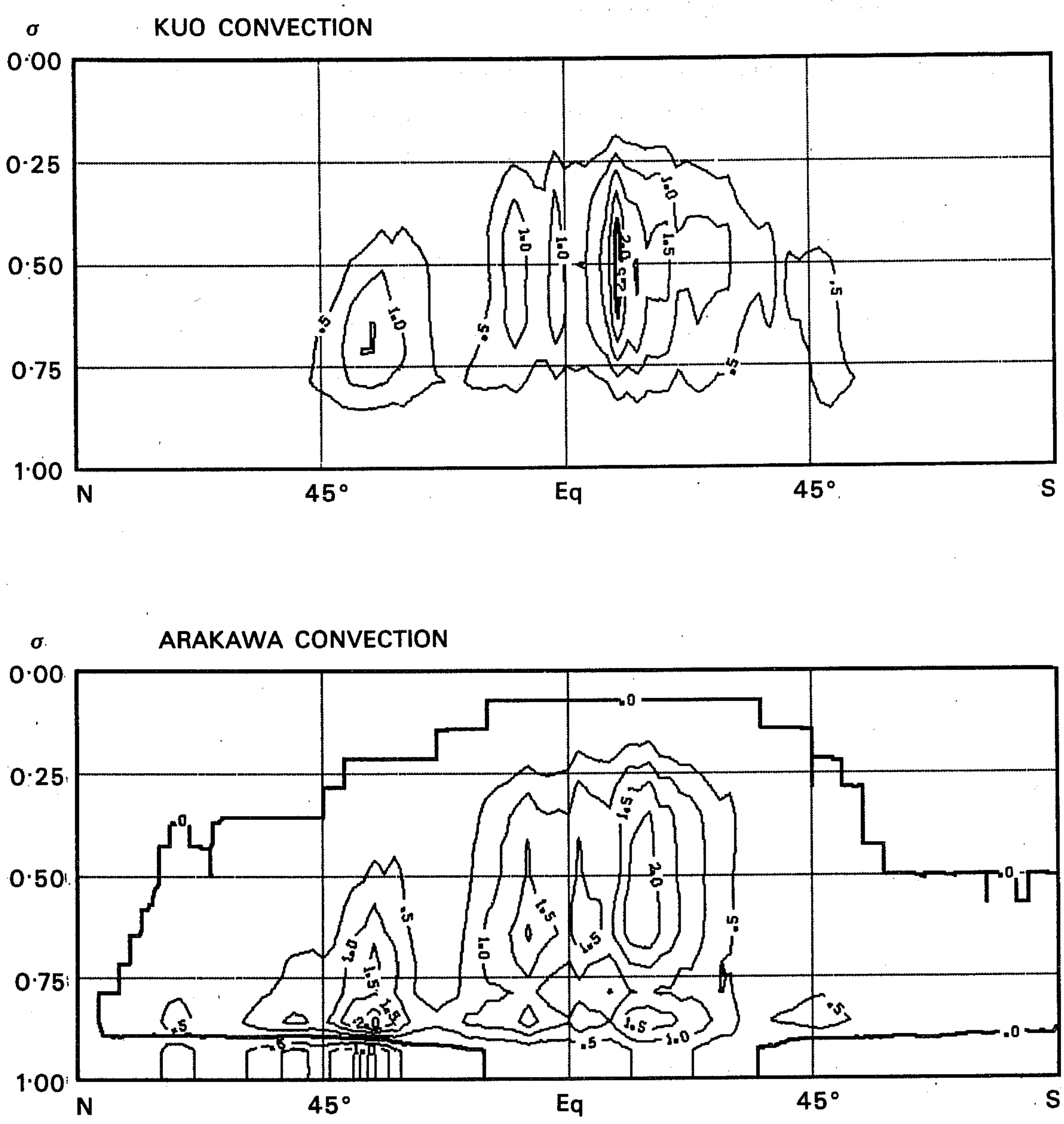

Fig. 4 Zonal mean heating (K/day) of the atmosphere due to cumulus convection parameterization, averaged over 10 days of a forecast. Two comparative forecasts using the parameterization by Kuo (top panel) or Arakawa (bottom panel) are shown.

References

Mahlman, J.D. and W.J. Moxim 1976 A method for calculating more accurate budget analyses of sigma coordinate model results. Mon.Wea.Rev. 104, 1102-1106.

Savijärvi, H. 1980 Energy budget calculations and diabatic effects for limited areas computed from ECMWF analyses and forecasts. Workshop on diagnostics of diabatic effects, ECMWF.

Temperton, C. and D.L. Williamson 1979 Normal model initialization for a multi-level grid point model. ECMWF Technical Report No.11, pp.91

Tiedtke, M. 1980 Diagnostics of diabatic processes in global numerical experiments at ECMWF. Workshop on diagnostics of diabatic effects, ECMWF.