U S E S O F S A T E L L I T E D A T A

I N S T U D I E S O F

S T R A T O S P H E R I C D Y N A M I C S

by

A. J. CRANE

A Thesis submitted for the degree of

Doctor of Philosophy

in the University of Oxford

Wolfson College

and

The Department of Atmospheric Physics

August 1977

ABSTRACT

Observations of the temperature structure of the stratosphere made by the selective chopper radiometer on the Nimbus 5 satellite are used to analyse the energetics of the sudden warming of January/February 1973.

A method of retrieving vertical profiles of zonal Fourier coefficients of temperature from the Fourier coefficients of the measured radiances is described. Some retrievals are compared with conventional observations and quite good agreement is obtained.

Comparison of independent estimates of the mean meridional circulation shows that the retrieved temperature and height fields are not sufficiently accurate to give precise values for the zonal mean vertical and meridional velocities. The implication of this result for the energy budget is demonstrated.

The sudden warming is analysed first in terms of wave structure and changes in temperature and zonal wind. No evidence for an upward propagating temperature or geopotential disturbance is found. Wave amplifications are observed to occur simultaneously at all levels or to propagate downwards. Largest temperature changes occur in the upper middle stratosphere and maximum zonal flow accelerations in the upper stratosphere.

The energy cycle of the lower middle stratosphere is found to be in agreement with observations of previous warmings except in that little increase in eddy energy occurs during the event. A marked baroclinic energy cycle below 10 mb in high latitudes is shown to enhance vertical energy propagation prior to the warming, leading to increased eddy available potential energy between 10 and 2 mb. The variation with latitude of the energetics

during the warming is shown to be significant. In the upper stratosphere barotropic conversion from zonal to eddy kinetic energy dominates in mid-latitudes causing deceleration of the zonal flow there first. In high latitudes the deceleration of the zonal flow occurs through the action of the induced mean meridional circulation, while the major source of eddy kinetic energy here is convergence of the vertical eddy energy flux.

Examination of the relationship between the latitudinal distributions of the vertical eddy energy flux and the zonal flow reveals that maximum upward propagation of energy is centred consistently to the north of the polar night jet until the onset of the warming when coincidence occurs.

Some aspects of the observed warming are compared with numerical simulations of sudden warmings. Although the 1973 event is characterised by a wavenumber one disturbance some striking resemblances are found in two numerically simulated wavenumber two type warmings.

ACKNOWLEDGMENTS

I should like to thank my supervisors, Dr. Bob Harwood and Professor John Houghton, for accepting me as their research student and for their support and encouragement during the course of this work.

Other members of the Atmospheric Physics Department gave valuable help, notably Drs. John Barnett, John Pyle and Clive Rodgers. Bob Wells and Trevor Jones were always ready to give assistance with computer programming difficulties.

I enjoyed useful discussions with Professor Isamu Hirota and Mr. Alan O'Neil in the later stages of the project.

I am indebted to all present and past members of this department and other institutions whose efforts led to the successful operation of the Nimbus 5 SCR, without which this study would not have been possible. Thanks are also due to Ernst Klinker and other staff of the Meteorology Department of the Free University of Berlin for providing the 50 mb height fields.

Financial support was given by the Natural Environment Research Council and Wolfson College, to both of whom I am grateful.

Finally I wish to thank Ann Read for her excellent preparation of the typescript, and Grant Guenther for the use of his typewriter which enabled me to present equations in a readable form.

C O N T E N T S

		Page
ABSTRACT		i
ACKNOWLEDGMENTS		iii
INTRODUCTION		1
CHAPTER 1 -	THE CLIMATOLOGY OF THE STRATOSPHERE	5
CHAPTER 2 -	RETRIEVAL OF TEMPERATURE PROFILES FROM REMOTE SOUNDINGS OF THE ATMOSPHERE	19
CHAPTER 3 -	RETRIEVAL OF FOURIER COMPONENTS OF TEMPERATURE	35
CHAPTER 4 -	THE EQUATIONS OF ATMOSPHERIC ENERGETICS	81
CHAPTER 5 -	VERTICAL MOTION AND THE MEAN MERIDIONAL CIRCULATION	98
CHAPTER 6 -	A SYNOPTIC DESCRIPTION OF THE JANUARY - FEBRUARY 1973 SUDDEN WARMING	119
CHAPTER 7 -	THE ENERGETICS OF THE 1973 SUDDEN WARMING	166
CHAPTER 8 -	VERTICAL ENERGY FLUX, THE KINETIC ENERGY BUDGET AND THE SUDDEN WARMING MECHANISM	203
CHAPTER 9 -	COMPARISONS BETWEEN OBSERVATIONS AND ANALYTICAL AND NUMERICAL THEORETICAL SIMULATIONS OF PLANETARY WAVE PROPAGATION AND SUDDEN WARMINGS	231
CHAPTER 10 -	SUMMARY AND CONCLUDING REMARKS	257
APPENDIX 1		261
APPENDIX 2 -	CALCULATIONS OF EDDY CORRELATIONS	263
APPENDIX 3 -	DETAILS OF THE RADIATIVE HEATING SCHEME	265
REFERENCES		274

INTRODUCTION

This thesis reports an investigation into the dynamical processes and energy exchanges which took place in the northern hemisphere stratosphere during the sudden warming of January/ February 1973. The warming phenomenon has been the subject of much research effort, both observational and theoretical, since its discovery in 1952. Before the advent of satellite observations research was necessarily confined to the lower levels, upper stratospheric data from rocket-sondes being sparse. The selective chopper radiometers on the Nimbus 4 and Nimbus 5 satellites provided the first global and daily coverage of the upper stratosphere. Using this data it has been possible to study for the first time the whole depth of the stratosphere on a daily basis during a major sudden warming. The aims of this study, therefore, were (a) to compare the behaviour of the lower and middle stratosphere during the 1973 warming with the results of previous studies, thereby giving support or otherwise to their fairly consistent findings and perhaps providing new information on the less certain characteristics of the events; (b) to examine for the first time the sudden warming processes of the upper stratosphere and its coupling with lower levels; and (c) to compare the observed behaviour, especially that of the upper stratosphere, with numerical and theoretical simulations of sudden warmings.

Chapter 1 sketches the way in which radiative processes determine the basic structure of the stratosphere and how this is disturbed by wave motions. Some basic theoretical ideas concerning the way in which the larger scale waves influence the dynamics of the stratosphere are presented with observational evidence.

Since satellite observations are effectively weighted mean temperatures of rather thick atmospheric layers, they cannot be used directly in studies of energetics. It is necessary to retrieve temperature profiles from the observations so that temperature and wind fields at specific levels may be obtained. Since the retrieval problem is generally under-constrained, additional information, such as atmospheric statistics, must be found to supplement the observations. During a sudden warming the atmosphere is very far from its mean state and the choice of suitable statistics is difficult. A retrieval method designed for situations in which the large scale planetary waves are the dominant features was therefore developed. This is described in Chapter 3 after the principles of remote sounding of the atmosphere from satellites and retrieval of temperature profiles have been outlined in Chapter 2.

Chapter 4 is devoted to a discussion of the equations describing energy conversion and generation in the atmosphere. The concept of available potential energy is introduced, in terms of which the energy equations are then formulated. The applicability of this concept to the region of the atmosphere considered in this particular study is discussed.

Solution of the thermodynamic equation was chosen as the most suitable method of obtaining vertical motion fields. Radiative processes become increasingly important higher in the stratosphere so must be included in this calculation. The radiative scheme used is described briefly in Chapter 5, with further details given in Appendix 3. Zonal averaging of the vertical motions and use of the continuity equation enabled a mean meridional circulation to be calculated. An independent estimate of the mean meridional circulation was obtained from the zonal mean eastwards

momentum equation. Much of Chapter 5 deals with a comparison of these two circulations. Possible reasons for the disagreement found are suggested and the implications of the evident incompatibility between the heat and momentum budgets for the energy budget are demonstrated.

The main results of the study are presented in Chapters 6 - 9. In Chapter 6 the evolution of the temperature and wind fields during the sudden warming are described. A number of papers on sudden warmings have dealt with the question of whether warmings propagate upwards or downwards. This topic is again taken up. The differences between the heat and momentum budgets of the middle and the upper stratosphere are examined.

The results of previous studies of the energetics of sudden warmings are reviewed briefly at the beginning of Chapter 7 after which the energetics during the 1973 warming of a near-hemispheric region of the stratosphere are discussed and compared with previous events. The overall effect of the warming event in terms of transfer of energy between zonal and eddy forms and between different layers of the atmosphere is illustrated. Comparison is made between the distribution of energy amongst its various forms during this warming period and the energy distributions during relatively quiet winters in the northern and southern hemispheres.

The process whereby kinetic energy is transferred from one atmospheric layer to another by means of the eddy 'pressure-work' effect was again shown to be of major importance at the onset of a sudden warming. This process is studied in greater detail in Chapter 8, where it is shown that its effect was very dependent on height and latitude. This result naturally leads on to an examination of the latitudinal

distribution of the kinetic energy budgets which indicates that the increases in eddy kinetic energy and the deceleration of the zonal flow were effected by a different mechanism in middle latitudes from that at high latitudes. From these investigations a model for the warming mechanism is postulated.

Analytical and numerical studies of vertical wave propagation and sudden warmings have increased our understanding of the dynamics of the winter stratosphere by directing attention to the apparently more important features of sudden warming events. Until the advent of satellite observations of the upper stratosphere, however, only some aspects of the hypothesised mechanisms could be tested against observations. An important part of this study was, therefore, to see to what extent the observations support or cast doubt upon the theoretical predictions regarding the behaviour of the middle and upper stratosphere during a sudden warming. To this end the relationship between the zonal flow and the vertical wave energy flux preceding and during the warming event is examined in the light of the wave propagation models of Dickinson (1969), Matsuno (1970) and Simmons (1974). Also discussed are some numerical simulations of sudden warmings which bore strong resemblances to the 1973 event.

A summary of the main results and conclusions which may be drawn from this study form the final chapter.

Chapter 1 THE CLIMATOLOGY OF THE STRATOSPHERE

The atmospheric state is determined by complex interactions of many physical and dynamical processes, but of these the three which are most important in explaining the large scale features are the absorption of solar radiation and the effects on the air flow of the Earth's rotation and of longitudinal variations in surface topography. A simple treatment of solar radiation absorption and infra-red emission by the atmosphere is sufficient to describe its basic temperature structure. The corresponding distribution of pressure gives rise to motions which are profoundly influenced by the Coriolis acceleration due to the Earth's rotation, and predominantly zonal (westerly or easterly) flow occurs. Upon this flow are superimposed large scale wave motions governed mainly by the variation of the Coriolis acceleration with latitude and variations in the surface features. The strong influences of mountain ranges and the different thermal properties of oceans and land masses are manifested in the semi-permanent, stationary wave patterns which dominate the northern hemisphere winter climate.

The following brief survey of the general circulation is based on these fundamental aspects. As this thesis is concerned with a sudden warming event particular emphasis is given to the winter stratosphere.

1.1. THE VERTICAL TEMPERATURE STRUCTURE OF THE ATMOSPHERE

The Earth's atmosphere may be divided into four layers on the basis of the vertical distribution of temperature. Figure 1.1 shows the mean temperature profile between the surface and 120km and the nomenclature used to describe these layers and

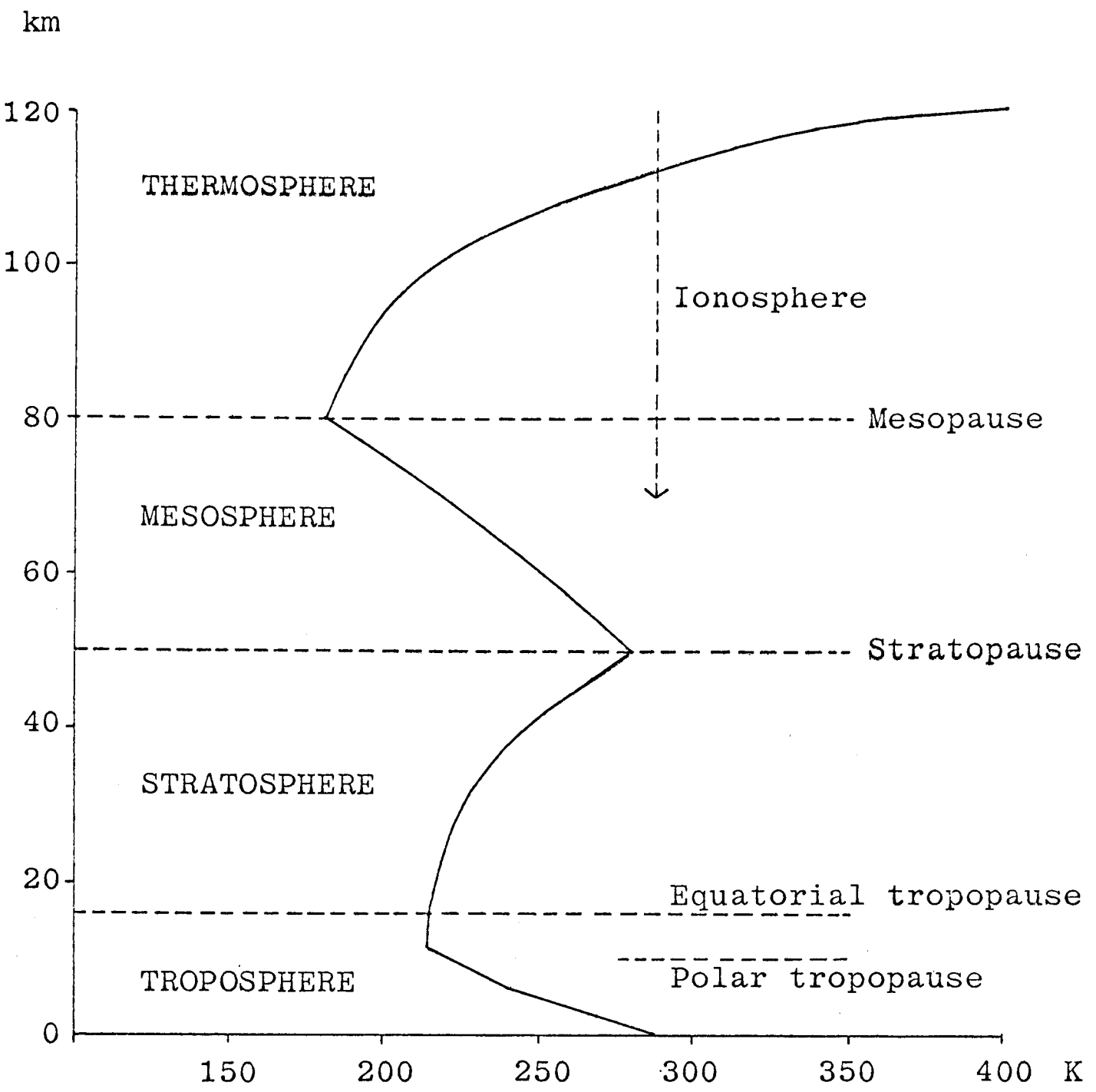


Figure 1.1. A schematic representation of the vertical temperature structure of the atmosphere with the nomenclature used to describe the different layers and their boundaries

their boundaries. This structure is due principally to the differential absorption of solar radiation. Let us consider, therefore, the processes which contribute to the depletion of a solar beam as it passes through the atmosphere. The very high temperatures of the thermosphere are caused by absorption of far ultra-violet wavelengths by atomic oxygen. As the radiation of these wavelengths is depleted, and density increases, heating is rapidly reduced giving the lowest atmospheric temperatures near 80km. The dissociation of molecular oxygen by radiation of wavelength less than 0.25μm leads to the formation of ozone from atomic and molecular oxygen. Ozone photochemistry is very complicated and the equilibrium concentration depends on many reactions involving the allotropes of oxygen and the trace atmospheric constituents, principally the oxides of nitrogen and hydrogen. A good approximation to observed concentrations can be obtained, however, from the five reactions between O, O_2 and O_3 of the 'oxygen-only' scheme of Chapman (1930). The ozone mixing ratio has a maximum around 30km but the very strong absorption by ozone of radiation in the Hartley and Huggins bands (0.18-0.29μm) gives heating which is most intense at 50km and which may be as large as 16K day^{-1} over the summer pole. Below 25km ozone is protected from these wavelengths and its distribution is controlled by atmospheric dynamics. In the upper stratosphere and mesosphere concentrations close to photochemical equilibrium occur. In the lower stratosphere weak absorption of solar radiation in infra-red wavelengths by carbon dioxide and water vapour, and in the Chappuis band of the visible spectrum by ozone, gives heating rates of only a few tenths of a degree per day. As a result a second temperature minimum occurs between 10 and 17km (the

tropopause) below which the temperature profile is determined largely by heat exchange with the Earth's surface. Direct heating in the troposphere is of relatively little importance. London (1957) calculated that of the energy in the solar beam 3% is absorbed in the stratosphere and above, 14.5% in the troposphere, 47.5% at the Earth's surface and 35% is reflected back to space.

Energy loss by thermal emission is mostly due to carbon dioxide at 15μm and ozone at 9.6μm in the stratosphere and by clouds and water vapour in the troposphere. There is significant thermal exchange between different atmospheric levels and especially between the troposphere and the Earth's surface.

The balance between solar heating and thermal emission varies considerably with latitude and season. Figure 1.2 is a latitude-height section of the net radiative heating at the solstices, while Figure 1.3 shows the temperature distribution that would exist in the stratosphere and mesophere at the solstices if radiative equilibrium prevailed. The lowest temperatures would occur during the polar night and the highest in the region of maximum ozone heating over the summer pole. Comparison with the observed latitude-height section of longitudinally, or zonally, averaged temperature (Figure 1.4) indicates that dynamical processes in the stratosphere transport heat from the sources over the summer pole and equatorial regions to the sink over high latitudes in the winter hemisphere, thereby reducing the meridional temperature gradient. The cause of the cold summer and the warm winter in the mesosphere is uncertain but may be related to a large scale inter-hemispheric circulation with ascent over the summer pole and descent over the winter pole.

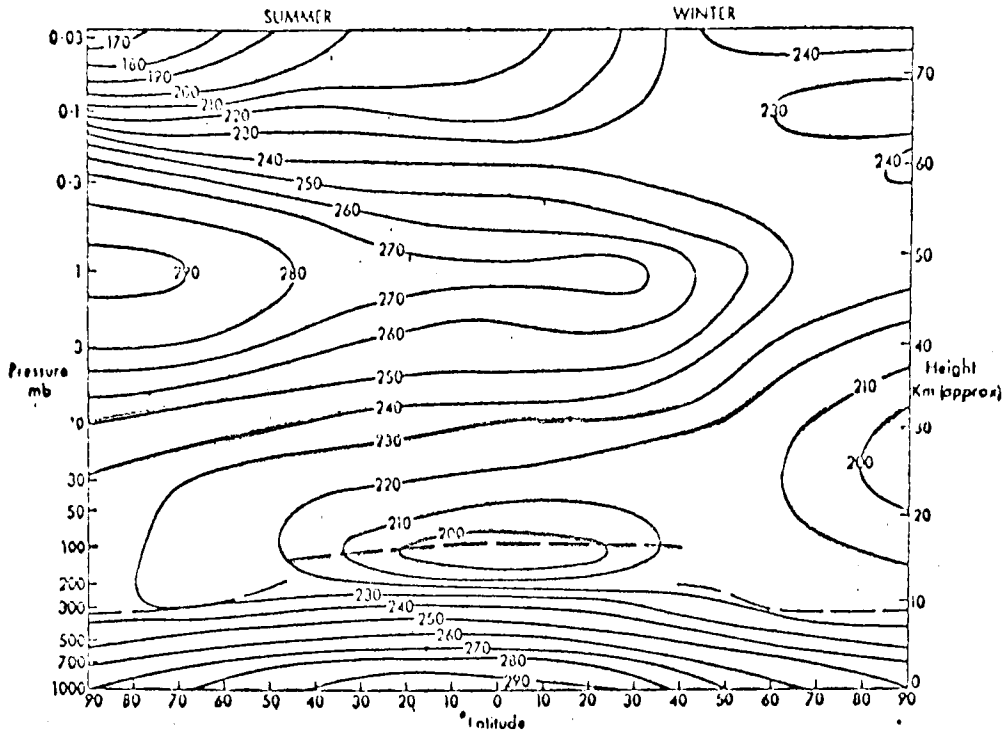


Figure 1.4. Representative zonal mean temperatures at the solstices (K) (taken from Murgatroyd, 1970)

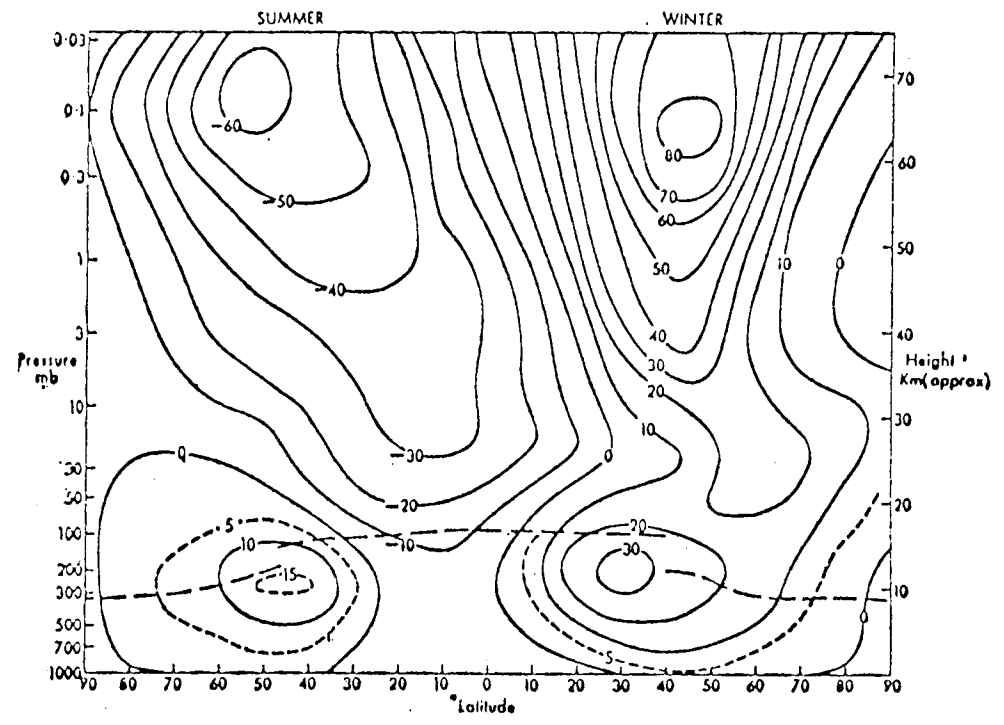


Figure 1.5. Representative zonal mean zonal winds at the solstices ($m\,s^{-1}$) (taken from Murgatroyd, 1970)

1.2. THE ZONAL MEAN ZONAL WIND

Pressure variations arising from the temperature distribution lead to atmospheric motions which are modified by the effects of the Earth's rotation (the Coriolis acceleration), friction and gravity. The equation of motion is:

$$\frac{d\underline{V}}{dt} = -2\underline{\Omega} \wedge \underline{V} - \frac{1}{\rho}\underline{\nabla}p + g + F \qquad 1.1$$

where $\underline{V}$ is the velocity, $\underline{\Omega}$ the Earth's angular velocity, ρ the density, p the pressure, g the acceleration due to gravity and F is the frictional force per unit mass. Away from the Earth's surface F is usually small and for the large scale, quasi-static, quasi-horizon tal systems which dominate the atmospheric circulation we may approximate equation 1.1 by

$$f\underline{k} \wedge \underline{V}_g = -\frac{1}{\rho}\underline{\nabla}_h p \qquad 1.2$$

where $f = 2|\Omega| sin\phi$ and is the vertical component of the Earth's vorticity (the Coriolis parameter), ϕ is latitude, $\underline{k}$ is a unit vector in the vertical and $\underline{\nabla}_h$ is the horizontal gradient operator. $\underline{V}_g$ is known as the geostrophic wind. The extent to which $\underline{V}_h$, the actual horizontal wind, is approximated by $\underline{V}_g$ is expressed by the Rossby number, R_o where

$$R_o = |d\underline{V}_h/dt| / |f\underline{k} \wedge \underline{V}_h|$$

For large scale motions R_o is typically 0.1. At the equator the Coriolis acceleration vanishes and thus the geostrophic approximation does not hold within about 10^o of the equator.

A scale analysis of the vertical component of equation 1.1 shows that for large scale motions

$$\frac{1}{\rho}\frac{\partial p}{\partial z} = -g \qquad 1.3$$

to within 0.1%. From the eastward component of equation 1.2,

the hydrostatic equation 1.3 and the equation of state

$$p = \rho RT \qquad 1.4$$

we obtain the equation:

$$\frac{\partial u}{\partial \ln p} = -\frac{R}{f}\frac{\partial T}{\partial y} \qquad 1.5$$

where u is the west-east, or zonal, wind and $\partial T/\partial y$ is the south-north temperature gradient on an isobaric surface. Equation 1.5 is the zonal component of the thermal wind equation and relates the vertical shear of the zonal wind to the latitudinal temperature gradient. Figure 1.5 (page 10) is a representative latitude-height section of the zonally averaged zonal wind at the solstices. The relation between the temperature and wind distributions of Figures 1.4 and 1.5 is seen to be well described by equation 1.5.

1.3. VARIATIONS IN THE ZONAL MEAN PATTERN

The zonal mean temperature and wind distributions at the solstices show the annual extreme conditions. The transition from summer to winter circulation in the stratosphere and mesosphere includes a period around the equinoxes of rather weak westerly flow in both hemispheres with the equator slightly warmer than both poles. In addition to this annual cycle of the circulation other regular oscillations are observable in the zonally averaged fields. A semi-annual oscillation in the tropics replaces the annual oscillation with the easterlies of the summer hemisphere encroaching into the winter hemisphere near the solstices and westerly flow prevailing at the equinoxes. Hopkins (1975) shows a maximum amplitude of $25ms^{-1}$ for the semi-annual wind oscillation which is accompanied by a temperature oscillation in phase with the sun of about 1K. The half year wave has also been detected in mid and high latitudes (van Loon

et al, 1972) although here it is weaker. A strong wind oscillation with a period of approximately 26 months - the Quasi-bienniel oscillation - exists in the lower stratosphere with an amplitude of about 20ms^{-1} between 22 and 30km altitude above which it becomes modulated by the semi-annual wave. Much research is currently being carried out in the field of equatorial waves of which Holton (1975) gives a concise account.

The advent of satellite measurements of the atmosphere has led to the discovery of other features of the zonal mean temperature field. Fritz and Soules (1970) were first to show that winter warming in polar regions of the stratosphere is accompanied by cooling in the tropics and low latitudes of the summer hemisphere. Similarly stratospheric warming often occurs simultaneously with mesopheric cooling (Labitzke, 1972). These large scale compensatory changes in the zonally averaged field are related to the behaviour of planetary waves which are the main topic of the remainder of this chapter.

1.4. LONGITUDINAL VARIATIONS, OR EDDIES, IN THE STRATOSPHERE

In middle and high latitudes waves around latitude circles play a vital role in the general circulation, transporting heat and momentum from low to high latitudes. The waves are essentially Rossby waves which owe their existence to the latitudinal variation of the Coriolis effect. In the troposphere zonal wavenumbers between 5 and 10 predominate - the familiar trough-ridge patterns of 500 mb circumpolar analyses - whereas in the stratosphere and mesosphere quasi-stationary waves of wavenumber 1, 2 or 3 are very evident in the winter circulation but are absent in summer. These observations can be explained by the theory of Charney and Drazin (1961) that a quasi-stationary wave generated in the troposphere can propagate into

the stratosphere only when it has a phase speed which is westward relative to the mean zonal flow and less than a critical speed which decreases with increasing wavenumber. During the summer the mean flow is easterly and the quasi-stationary waves are thus trapped in the troposphere. Normally in winter the westerly flow is sufficiently strong for waves of wavenumber higher than 3 to have phase velocities greater than the critical velocities; thus only the longest waves may propagate into the stratosphere. Recent satellite observations of the mesosphere (Austen et al, 1976) indicate that wave amplitudes decrease with height above the stratopause.

In addition to the quasi-stationary waves travelling waves of varying lifetime exist in the stratosphere in winter (Hirota, 1968) and also in summer (Muench, 1968; Hirota, 1975) Harwood, 1975). The origin of these travelling waves is not well understood, although alone they are probably not important mechanisms for poleward heat transport in the winter hemisphere.

Eliassen and Palm (1961) showed that for quasi-stationary, quasi-geostrophic, adiabatic waves the vertical wave energy transport varies with height and latitude in proportion to the product of the zonal mean west-east wind and the zonal mean northward heat transport by the waves. Since in the northern hemisphere winter the heat flux is generally poleward, a consequence of the westward tilt of the waves with height, the Eliassen-Palm relation implies that in westerly flow wave energy is transported upwards. Hirota and Sato (1969) gave observational evidence of a periodicity of about two weeks in the zonal mean westerly wind and the amplitude of wavenumber one geopotential height, which suggests that the upward energy propagation responds to variations in the mean flow. Miller

(1974) discussed this periodicity in some components of the stratospheric energy budget showing that the periodic forcing of the troposphere-stratosphere energy transfer resulted from a 14-day cycle in the baroclinic energy exchanges in the troposphere.

The importance of variation in the vertical energy flux lies in the fact that it is a likely cause of stratospheric sudden warmings. These irregular wintertime features are characterised by large vertical energy fluxes from the troposphere and enhanced eddy poleward heat transfer which leads to large and rapid warming of the polar regions (perhaps 50K in 4 - 5 days) and a consequent breakdown of the polar vortex. The initiation of this process has been observed to arise from the interaction of a travelling and a standing wave leading to an amplified quasi-stationary wave (Hirota, 1968; Quiroz, 1975). The effects of the increased eddy heat and momentum fluxes on the zonal mean temperature and wind fields are opposed by an induced mean meridional circulation. Normally there is a balance between the eddy forcing and the action of the mean meridional circulation, but during the sudden warming event the balance is perturbed resulting in net polar warming and deceleration of the westerly flow. Although there are features common to all sudden warmings they differ in terms of the magnitude and spatial extent of the temperature and circulation changes. During some warmings easterly flow appears only in the upper stratosphere while in a very recent event (Quiroz, 1977) the circulation reversal descended to ground level. More detailed description of sudden warmings and wave-mean flow interaction relevant to our work will appear in later chapters.

Another class of atmospheric wave motions are the solar

tides. Of the various modes observed only the diurnal and semi-diurnal components have sufficiently large amplitudes to affect the meteorology of the stratosphere and mesosphere. They are responsible for a significant part of the wind variance but because of their large phase speeds interaction between these waves and the mean flow is inhibited. Above the mesosphere tidal oscillations become increasingly important in controlling the circulation. The diurnal tide is excited primarily by water vapour absorption in the troposphere while ozone heating in the stratosphere and mesosphere forces the semi-diurnal mode. A comprehensive review of atmospheric tides is given by Chapman and Lindzen (1970).

Like solar tides gravity waves, which are generated by tropospheric events or possibly wave interactions at higher levels, have little influence on the stratosphere but contribute to the dynamics of the upper mesosphere and thermosphere to a much greater extent since they amplify with height.

1.5. DIFFERENCES IN PLANETARY WAVE BEHAVIOUR BETWEEN THE HEMISPHERES

Since quasi-stationary planetary waves seem to have their origin in surface topographical features and land-sea differences it is not surprising that differences between the wintertime stratospheric circulations of the two hemispheres have been observed. In the northern hemisphere disturbances of both wavenumber 1 and 2 are usually present as a result of the Aleutian high pressure ridge while only wavenumber 1 has significant amplitude in the southern hemisphere. Much smaller westward tilts with height are observed in the southern hemisphere waves which probably explains why during all observed southern hemisphere sudden warmings polar heating by the eddies

has been insufficient to cause a circulation breakdown (Barnett, 1975). In general the northern winter is much more active and less predictable than its southern counterpart. The final spring warming, during which the circulation reversal to summertime easterlies occurs is usually much earlier in the southern than in the northern hemisphere at upper stratospheric levels (Labitzke and Barnett, 1973). Interaction between the hemispheres is largely confined to short periods around the equinoxes when the circulation may be a weak westerly flow at all latitudes in the high stratosphere. Propagation of planetary waves from the more active to the quieter hemisphere may then occur, the influence extending beyond 30^{o} latitude in some cases (Barnett, 1975b).

1.6. ATMOSPHERIC ENERGETICS

Much insight into the way in which planetary waves interact with the zonal mean state can be obtained by studying heat, momentum and energy budgets of the atmosphere. It is appropriate to note here the most fundamental aspects of the energetics of the features we have described.

Contrary to the behaviour of small scale turbulence it is found that throughout the troposphere and stratosphere the kinetic energy of the zonal flow is usually maintained by transfer from the kinetic energy of the wave motions. The main wave energy sources however are quite different in the different atmospheric layers. The tropospheric cyclone waves are essentially baroclinic, that is they amplify at the expense of the potential energy of the mean state arising from the equator to pole temperature gradient. In the lower stratosphere the meridional distribution of solar heating is effectively a sink of potential energy on account of the reversed meridional

temperature gradient there. Baroclinic waves do not amplify to any extent. Instead the planetary waves grow through the vertical energy flux from the troposphere. In this way the lower stratosphere is essentially 'driven' by the troposphere. In the upper stratosphere both radiation and vertical energy flux convergence may be energy sources, although the latter tends to be more important during disturbed winter periods.

Energy exchanges are smallest during the summer, when wave activity is low and the zonal circulation relatively weak, and largest during winter sudden warmings when their pattern is radically altered in connection with rapid wave amplification and the breakdown of the normal circulation. We shall see that in the early stages of a sudden warming barotropic conversion, in which kinetic energy of the mean flow is transferred to the wave kinetic energy, may play an important role.

Chapter 2 RETRIEVAL OF TEMPERATURE PROFILES FROM REMOTE SOUNDINGS OF THE ATMOSPHERE

Introduction

Global scale studies of the atmosphere require simultaneous measurements of the atmospheric state from locations evenly distributed throughout the world. Such an observational network is neither geographically nor economically feasible. Remote sounding of the atmosphere from satellites, first suggested by Kaplan (1959), provides a relatively much cheaper and practical alternative, and although a period of about 12 hours is required for global coverage from a single satellite, time scales in the upper atmosphere are usually sufficiently long for such observations to be regarded as simultaneous. The major problem with satellite measurements is one of resolution: a single measurement represents the average state of a region of atmosphere of rather large vertical extent, typically 10km.

In this chapter we shall briefly outline the general principles of remote sounding with reference to the Selective Chopper Radiometer and the processing of raw data into a usable form. The different ways in which satellite data may be used for atmospheric analysis will be discussed with particular attention given to retrieval of temperature profiles. This will prepare the ground for Chapter 3 devoted to the development of the retrieval method used in our research.

2.1. GENERAL PRINCIPLES OF REMOTE SOUNDING

Remote sounding of atmospheric temperature is achieved by measuring the intensity of radiation emitted in particular spectral intervals by atmospheric constituents of known distribution. If the chosen gas is of constant mixing ratio

and local thermodynamic equilibrium applies to the emissions at the levels being observed, the radiation emitted at a given frequency depends only on the temperature distribution. Emissions in the ν_2 vibration-rotation band of CO_2 centred at 15μm satisfy these requirements up to the mesopause and the band is not overlapped by strong emission bands of other atmospheric gases. It is thus very suitable for temperature sounding and has been employed by Nimbus satellite Selective Chopper Radiometers as well as other instruments. The 4.3μm CO_2 band below 35km and the 5mm oxygen band have also been used for temperature sounding but for the sake of clarity we shall refer only to the 15μm CO_2 band in our discussion.

Suppose we choose to detect radiation emitted at a particular frequency in the band at which CO_2 emits strongly. Very little of the radiation of this frequency emitted by CO_2 in the lower atmosphere reaches the instrument as it is mostly absorbed by CO_2 overlying the level of emission. Radiation of the same frequency emitted from the highest levels of the atmosphere will be detected but the intensity will be small owing to the low density at these levels. Thus the bulk of the detected radiation comes from intermediate levels. For frequencies at which the atmosphere is relatively transparent the peak of the weighting is at levels lower than for more strongly absorbing frequencies. Thus radiation detected at frequencies corresponding to spectral line centres originates high in the atmosphere whilst most of that detected at frequencies in the wings of the lines, where the absorption coefficient is smaller, originates lower in the atmosphere. The distribution of the weighting clearly depends on the variation with height of the atmospheric transmission at the particular frequency which depends primarily on the pressure

and, below 50km, only slightly on temperature.

Consider a horizontally stratified atmosphere of infinite depth. The intensity emitted in the vertical by a slice of atmosphere of temperature T, density ρ, vertical thickness δz at a depth $-z$, is given by Kirchhoffs law:-

$$\delta I_\nu = B_\nu(T)\kappa_\nu\rho\delta z$$

where κ_ν is the absorption coefficient and $B_\nu(T)$ the Planck function at temperature T and frequency ν. The proportion of this intensity observed at the top of the atmosphere ($z = 0$) is

$$\tau_\nu = exp\left(- \int_{-z}^{0} \kappa_\nu\rho \, dz \right)$$

Integrating over all slices the intensity emitted at the top of the atmosphere is

$$I_\nu = \int_{-\infty}^{0} B_\nu(T) exp\left(-\int_{-z}^{0} \kappa_\nu\rho \, dz \right)\kappa_\nu\rho \, dz$$

Since $d\tau_\nu = \kappa_\nu\rho \, exp\left(-\int_{-z}^{0} \kappa_\nu\rho \, dz \right) dz$, $\tau_\nu = 0$ at $z = -\infty$ and $\tau_\nu = 1$ at $z = 0$ we may write

$$I_\nu = \int_0^1 B_\nu(T) \, d\tau_\nu$$

It is convenient to use the variable

$$y = - ln\,(p/p_s)$$

where p_s is the surface pressure, as the vertical co-ordinate. We may then write

$$I_\nu = \int_0^\infty B_\nu(T) \frac{d\tau_\nu}{dy} \, dy$$

I_ν is thus the weighted mean of the Planck function intensity with weighting function $K_\nu = d\tau_\nu/dy$. The use of pressure rather than height as a vertical co-ordinate renders the weighting function more nearly independent of temperature (Houghton and Smith, 1970).

On account of technological limitations the narrowest spectral

band-pass obtainable with a conventional filter radiometer or spectrometer is very much larger than the widths of and spacings between the individual lines. The radiation received covers a frequency range in which absorption varies considerably so the depth of the region of origin is broad. However a large proportion of this frequency range is occupied by the line wings in which κ_ν is low and only slightly frequency dependent; the weighting function therefore has maximum weighting low in the atmosphere and is not much broader than the ideal monochromatic weighting function for the lower atmosphere. At frequencies where absorption is large κ_ν is highly frequency dependent and with conventional instruments it is not possible to separate these frequencies from neighbouring frequencies of low κ_ν. Weighting functions peaking high in the atmosphere cannot therefore be obtained (Houghton and Smith, 1970).

2.2. THE SELECTIVE CHOPPER RADIOMETER

The Selective Chopper Radiometer (SCR) launched on the Nimbus 4 satellite in April 1970 employed two techniques which improved the vertical resolution of low level weighting functions and made possible remote sensing up to 50km altitude. Both are based on the attainment of an effective spectral resolution comparable with line widths while collecting sufficient energy to make accurate measurements. This is achieved by detecting radiation over many frequencies all of which correspond to absorption co-efficient lying in a given narrow range. Thus while radiation from only a limited part of any individual line is observed, collection over the total bandwidth provides adequate energy.

If a cell of CO_2 is introduced into the path of radiation entering the radiometer, radiation of frequencies corresponding

to line centres is absorbed by the cell and the transmitted radiation is that emitted only in the wings of the lines. The height range from which it originates is therefore less than if radiation from the whole band were detected since the contribution from high levels is removed. Figure 2.1 compares weighting functions obtained for (1) the whole of a $5cm^{-1}$ frequency band, (2) the same band including selective absorption and (3) a monochromatic frequency in the wing of a spectral line. Channels 3 - 6 of the Nimbus 4 SCR employed the Selective Absorption technique. Figure 2.2 shows the weighting functions of all six channels.

The second innovation enables high level weighting functions to be obtained by chopping the incoming radiation between a cell containing CO_2 and a cell which is empty or contains CO_2 at a lower pressure. The signal contains alternately emission from the line wings only and emission from all parts of the spectral lines. The difference, obtained by phase sensitive detection, is the radiation absorbed by the higher pressure cell which is just that emitted from line centres. A weighting function peaking high in the atmosphere is thus obtained. Channels 1 and 2 of the Nimbus 4 SCR used this technique.

On the Nimbus 5 SCR, from which our data are obtained, selective chopping is carried out in a different manner. Four cells, containing different pressures of CO_2 and mounted on a filter wheel, are switched in turn in front of the detector. The signals obtained are chopped against a space signal so that stray radiation may be eliminated and amplification be made easier by virtue of the signals' A.C. characteristics. The differencing between pairs of channels corresponding to selective chopping is performed during data processing at Oxford. This

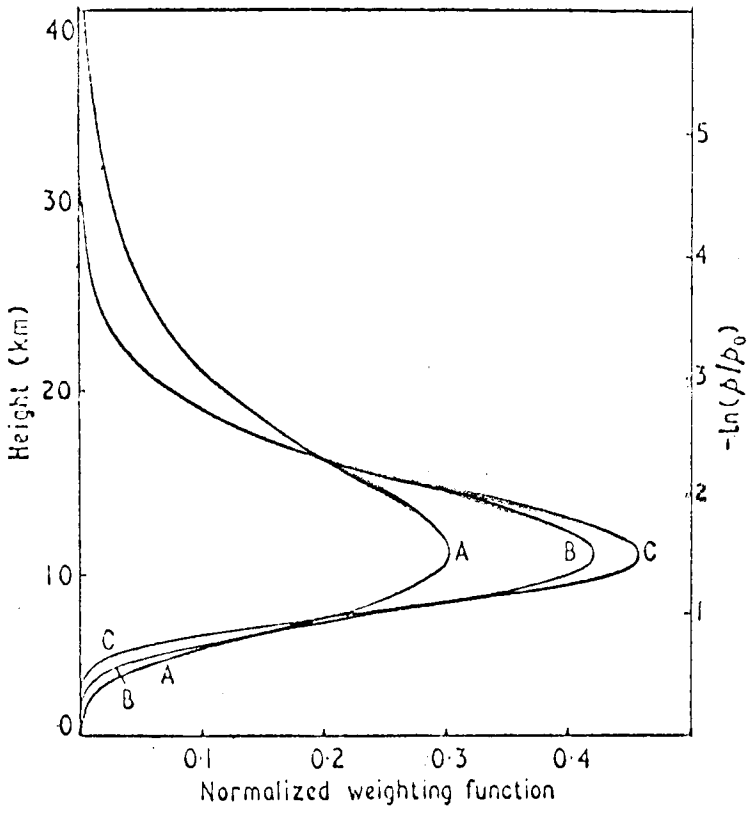


Figure 2.1. Demonstrating the effect of selective absorption: weighting functions for (curve A) a 5 cm^{-1} wide interval near 690 cm^{-1}, (curve B) the same interval but including a path of CO_2 and (curve C) an ideal weighting function for a monochromatic frequency in the wing of a spectral line. (Taken from Abel et al, 1970)

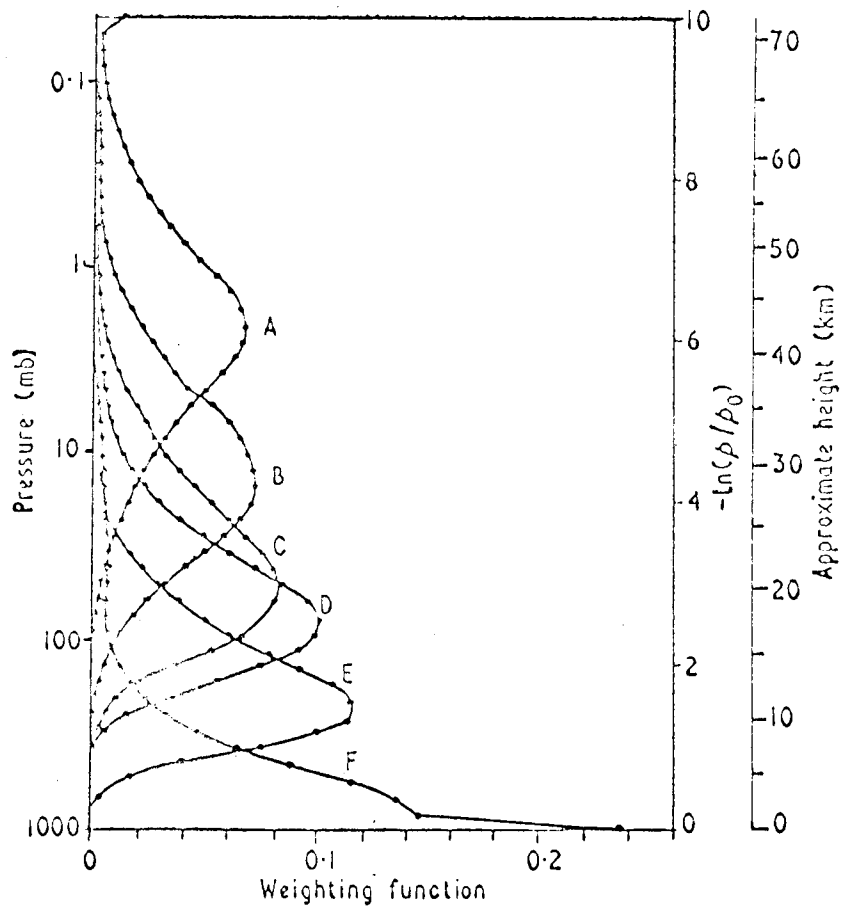


Figure 2.2. The weighting functions for the six Nimbus 4 SCR channels. (Taken from Abel et al, 1970)

differencing method has advantages over mechanical chopping since imbalance between the cells due to chopping asymmetries is removed. On the other hand differencing large signals after telemetry requires that the signals be measured and telemetered to a higher accuracy. Typically the differences are an order of magnitude smaller than the signals. Imbalances caused by deterioration of optical components and other effects are common to both chopping techniques. Since the higher the level of observation the smaller is the difference in transmission between the cells, there is an observational ceiling for the SCR at about 55km. This problem has been overcome by using a single cell Pressure Modulator Radiometer on Nimbus 6 which observes the upper stratosphere and mesosphere between about 40 and 90km (Curtis et al, 1974). The weighting functions of the Nimbus 5 SCR temperature sounding channels are shown in Figure 2.3.

2.3. THE SCR DATA

The retrieval method used (Chapter 3) employs zonal Fourier coefficients of the radiances observed in channels B12, B34 and A1 of the Nimbus 5 SCR. We shall discuss here the processing of the raw data and make an appraisal of likely sources of errors.

Nimbus 5 is in a sun-synchronous polar orbit between 80^{o}S and 80^{o}N crossing tropical and mid-latitudes close to local noon and midnight. The orbit ground tracks are spaced at about 27^{o} longitude and 13.4 orbits are made daily from which about 12 orbits' data are recovered and processed. The day and night observations are processed separately. In all channels each observation is a 16 second average of measurements made every 4 seconds. For the B difference channels (B12, B23, B34) smoothing over several observations is necessary; this is accomplished by applying a filter function to the observations equivalent to 4

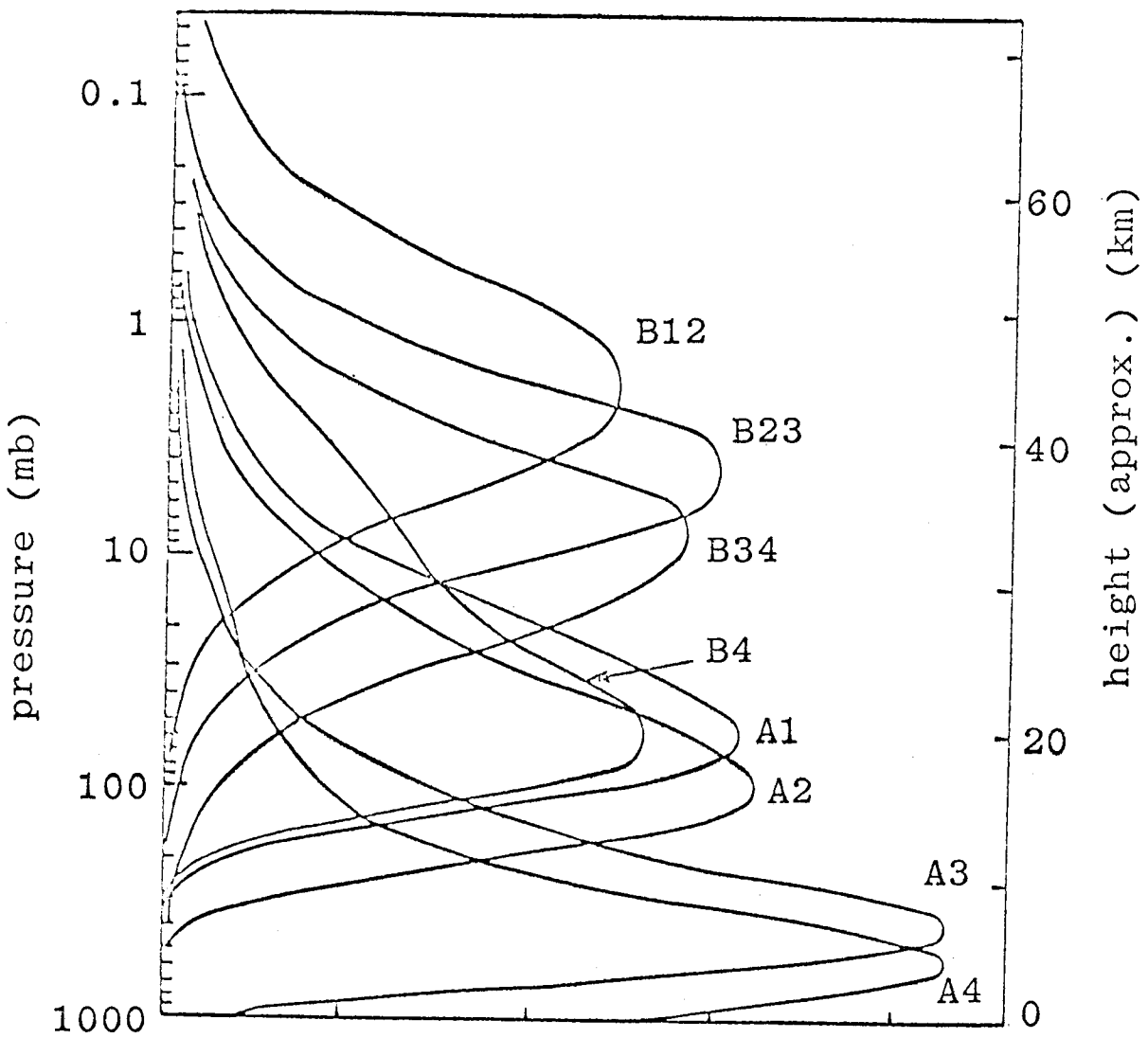


Figure 2.3. Weighting functions for the Nimbus 5 SCR. The B difference channels (B12, B23 and B34) are obtained by selective chopping. (After Ellis et al, 1973)

passes of a 5-point running mean. These values are further smoothed and then interpolated on to a 4^{o} latitude x 10^{o} longitude rectangular grid between 80^{o}S and 80^{o}N. The day and night grids are averaged to give a single mean grid. At points where only day or night data are obtained half of any zonal average day-night difference (calculated from points with both day and night data) is added or subtracted, whichever is appropriate. Barnett (1975b) shows that such differences are usually small although during active winter periods this is not so and the corrections are important at these times. The final grid is Fourier analysed with respect to longitude for the purposes of wave analysis.

A considerable amount of noise in the raw data is removed by the smoothing and gridding processes. Both by examining the smoothing functions of the gridding program and by running an experiment replacing real observations with random numbers of known mean and standard deviation, Barnett (1976) showed that the standard deviation of grid-point data is about one third that of orbit observations for unsmoothed channels; the reduction is likely to be a little more for B difference channels. By comparing the standard deviations of the Fourier components it was shown that there are 11.4 independent points around the globe compared with the 36 grid points: consequently only the first 5 or 6 zonal waves can be resolved. We describe a method of estimating the random noise in grid point and Fourier component radiances in Chapter 3 in connection with the retrieval method.

Systematic errors cannot usually be estimated from the observations themselves. There is a variety of causes of such errors including stray radiation, instrument temperature

variations, interference by other instruments on the satellite, errors in calibration and errors in the calculated weighting functions. Analysis of the housekeeping data enables many errors of this type to be discovered over a period of time so that the data may then be reprocessed. One example of such errors which was discovered from the observations was caused by very slow leaks of carbon dioxide and was manifested in slow drifts of global mean radiance during the first two years of B channel observations. This resulted in slow vertical drifts of the weighting functions. Once all possible corrections have been made total systematic error is thought not to exceed 1 radiance unit (Barnett et al, 1975). Comparisons with rocket-sonde observations over short periods indicate that the errors can be considerably larger if one assumes the rocket-sonde measurements to be correct, which is not of course wholly justified. Account is taken of such a comparison in our retrievals and is described in Chapter 3.

The time scales of stratospheric temperature changes are such that day-night mean radiances usually give a good representation of existing patterns. A comparison between B12 wavenumber 1 radiance amplitude at 60^{o}N calculated using firstly day-night mean radiances and secondly synoptic radiances calculated by Rodgers (private communication) indicates that even during sudden warmings the mean radiances describe well the synoptic situation.

2.4. USE OF SATELLITE DATA IN ATMOSPHERIC STUDIES

The way in which satellite observations are used in atmospheric analysis depends on the nature of the problem being investigated and may also be influenced by one's knowledge of the sensing instrument and its weighting functions and additional

information about the atmosphere. Fritz **et al** (1972) gave a concise review of this topic. Three methods have been widely used in the literature: (a) the derivation of radiance equivalent temperatures, (b) the derivation of temperatures at certain levels and layer thicknesses by regression techniques, and (c) the retrieval of temperature profiles.

By using the Planck function for the appropriate wavelength of observation the measured radiances may be converted to give equivalent black-body temperatures. These may be regarded as good indicators of the mean temperature of the atmospheric layer between the half-width levels of the weighting function. Much fundamental research on the broadscale features of the stratosphere has been achieved using radiance equivalent temperatures including studies of the evolution of sudden warmings (Barnett et al, 1973; Barnett, 1975), stratosphere-mesosphere interaction (Labitzke, 1972b; Austen et al, 1976), inter-hemispheric wave propagation (Barnett, 1975b) and high-low latitude temperature relationships (Fritz and Soules, 1970, 1972).

An extension of this method is to evaluate the precise layer within the domain of the weighting function whose mean temperature, and consequently thickness, is best represented by the radiance measurement. Quiroz and Gelman (1972) regressed radiances implied by a large sample of rocket- and radio-sonde profiles against various thicknesses within the weighting function spread. The coefficients from the best correlation may be used operationally to determine the particular thickness from radiances, and, provided base geopotential height fields are available, height fields and hence geostrophic winds may be determined for higher levels. Using linear combinations of channels a range of thicknesses may be estimated. Somewhat less

satisfactory results may be obtained by similar methods to give single level temperatures (Quiroz, 1974). Height and temperature data derived in this way have been used to give necessarily limited descriptions of sudden warmings in the upper stratosphere (e.g. Klinker, 1976). In order to study atmospheric energetics it is necessary to estimate vertical motions and energy fluxes. This requires a knowledge of temperature and height fields at reasonably close levels on a regularly spaced grid. For this reason regression methods are not adequate and retrievals of temperature profiles must be attempted.

2.5. RETRIEVAL OF TEMPERATURE PROFILES

The inverse problem of radiative transfer, namely the determination of the state of the atmosphere from measurements of the radiation it emits, is underconstrained since an attempt is made to describe a continuous variable from a finite number of measurements. As a first step the continuous variables - the Planck function profile $B(z)$ and the weighting function K_ν - are made discrete so that the problem is now to deduce the temperature at, say, 50 levels from a set of about 4-8 weighting functions. Since there is an infinite number of profiles which can give rise to the observed radiances, the profile fitting the observations, within experimental noise limits, which is in some way the 'best' of all these must be found. The form of the constraints used to obtain this solution depends on how the 'best' estimate is defined. In preparation for Chapter 3 we shall discuss here one type of linear constraint, based on the presentation by Rodgers (1976), to which the reader is referred for a comprehensive discussion of retrieval theory.

Rodgers (1976) defines a linear constraint as one which takes the same mathematical form as a linear, direct measurement -

it gives a value for a known linear function of the unknown profile together with an error covariance matrix for this value.[1] The linear constraint often takes the form of a guessed, forecast or climatological profile, with its covariance, and may be regarded as a "virtual" measurement. The virtual and the direct measurements are combined in the usual way of combining independent estimates by weighting each with its inverse covariance. Thus if we have two measurements of vector $\underline{B}$, $\underline{B}_1$ and $\underline{B}_2$, with error covariances H_1 and H_2 the best estimate $\hat{\underline{B}}$ is given by

$$\hat{\underline{B}} = (H_1^{-1}\underline{B}_1 + H_2^{-1}\underline{B}_2)(H_1^{-1} + H_2^{-1})^{-1}$$

with covariance

$$\hat{H} = (H_1^{-1} + H_2^{-1})^{-1}$$

The retrieval problem is a little different in that the direct measurements are not of the profile itself but are a set of radiances $\underline{I}$, with instrumental error covariance E, related to the profile $\underline{B}$ by the relation $\underline{I} = K\underline{B}$, where K is the discretised weighting function matrix. If our virtual (e.g. climatological) profile is $\underline{B}_0$, with covariance H, the best estimate of $\underline{B}$ is

$$\hat{\underline{B}} = (H^{-1} + K^T E K)^{-1}(H^{-1}\underline{B}_0 + K^T E^{-1}\underline{I})$$

with covariance

$$\hat{H} = (H^{-1} + K^T E^{-1} K)^{-1} \qquad \text{(Rodgers, 1976)}$$

where T represents the transpose of a matrix. This solution may be written in the computationally simpler form

$$\hat{\underline{B}} = \underline{B}_0 + HK^T(KHK^T + E)^{-1}(\underline{I} - \underline{I}_0) \qquad 2.1a$$

where $I_0 = KB_0$, with covariance

---

[1] The covariance matrix of a vector is analagous to the variance of a scalar quantity, the diagonal terms giving the variance of each element of the vector and the off-diagonals a measure of the correlation between the vector elements.

$$\hat{H} = H - HK^T(KHK^T + E)^{-1} KH \qquad 2.1b$$

This optimum solution $\hat{\underline{B}}$ is obtained by a variety of apparently different retrieval methods of which we give two examples for the purpose of illustration.

(a) Minimum variance solution

The method is to find the prediction matrix D in the equation

$$\hat{\underline{B}} - \underline{B}_0 = D(\underline{I} - \underline{I}_0) \qquad 2.2$$

where $\underline{B}_0$ is a statistical mean profile and

$$\underline{I}_0 = K\underline{B}_0$$

such that the variance of the error in $\hat{\underline{B}}$ is a minimum. Given a statistical ensemble of n profiles Bj, j = 1,n, the n radiance vectors, I_j, j = 1,n, implied by the B_j are determined, assuming a suitable noise covariance E, from

$$\underline{I}_j = K\underline{B}_j + E$$

The required D is that which makes

$$\sum_{j=1}^{n} (\hat{\underline{B}}_j - \underline{B}_0)^2$$

a minimum, where

$$\hat{\underline{B}}_j - \underline{B}_0 = D(\underline{I}_j - \underline{I}_0)$$

This is a multiple regression problem yielding the solution

$$D = HK^T(KHK^T + E)^{-1}$$

where H is the covariance of the statistical mean profile, which upon substitution in equation 2.2 gives equation 2.1a.

(b) Maximum likelihood solution

In this case we seek the most likely profile consistent with a direct observation and, for example, a statistical ensemble. This is obtained by maximising the conditional

probability density function $P(\underline{B}|\underline{I})$ of the solution $\underline{B}$ given the observation $\underline{I}$. Bayes theorem states

$$P(\underline{B}|\underline{I}) = P(\underline{I}|\underline{B})\, P(\underline{B}) \,/\, P(\underline{I})$$

Assuming the instrumental noise has a Gaussian distribution we may write

$$P(\underline{I}|\underline{B}) \propto exp(-\tfrac{1}{2}(\underline{I}-K\underline{B})^{T}E^{-1}(\underline{I}-K\underline{B}))$$

where E is the noise covariance. $P(\underline{B})$ will depend on the nature of the atmospheric statistical ensemble: if these are also Gaussian

$$P(\ \underline{B}) \propto exp(-\tfrac{1}{2}(\underline{B}-\underline{B}_0)^{T}H^{-1}(\underline{B}-\underline{B}_0))$$

where H is the covariance of the mean profile $\underline{B}_0$. $P(\underline{I})$ is not dependent on $\underline{B}$ so to maximise we set

$$\frac{\partial}{\partial \underline{B}} P(\underline{B}|\underline{I}) = \frac{\partial}{\partial \underline{B}}(P(\underline{I}|\underline{B})\, P(B)) = 0$$

For Gaussian statistics the most likely solution given by maximisation is again $\hat{\underline{B}}$ of equation 2.1a. The solution 2.1 is always a minimum variance solution but is only the most likely solution for normally distributed statistics. Depending on the nature of the statistics it may not be possible to determine the maximum likelihood solution by matrix methods for other distributions.

The equation 2.1a is most commonly solved by matrix inversion but an alternative method, giving the same solution, is that of 'sequential estimation'. Each channel is treated separately and the contribution to $\hat{\underline{B}}$ from each channel is added sequentially. The running total contribution to $\hat{\underline{B}}$ is used as the a priori estimate $\underline{B}_0$ for the next stage. On account of its computational simplicity the technique is well suited to retrieving profiles at adjacent points along satellite orbit

tracks, the previous retrieval being used as the a priori estimate. Determination of H is, however, difficult in this case.

Accuracy of the solution

The covariance $\hat{H}$ of the solution profile $\underline{\hat{B}}$ is given by equation 2.1b. The diagonal values of $\hat{H}$ give the simplest measure of the accuracy of $\underline{\hat{B}}$. Strand and Westwater (1968) and Rodgers (1970) assessed their retrievals by this 'residual variance'. Since the off-diagonal terms of H are generally non-zero, errors at different levels are correlated and $\underline{\hat{B}}_j$ is known to a greater accuracy than $\underline{\hat{B}}_j \pm \hat{H}_{jj}^{\frac{1}{2}}$

Such assessments of error are of little practical value however since solution errors arise as a result of the real atmospheric profile not being, for example, the most likely profile. Comparisons with coincident radio- and rocket-sonde profiles, when available, are a more useful indicator of the quality of retrieved profiles. We shall see in Chapter 3 that the vertical resolution of the observations is often not sufficient for features of small vertical scale to be represented in the retrieved profile. Poor retrievals may also result from the use of inadequate or inappropriate constraints. The choice of suitable constraints, or a priori information, is most difficult, and retrievals tend to be least accurate, during anomalous atmospheric situations. The sudden warming is such a situation and much of Chapter 3 is devoted to the way in which we attempted to produce a priori information which would be appropriate for temperature profile retrievals in the very disturbed period under investigation.

Chapter 3 RETRIEVAL OF FOURIER COMPONENTS OF TEMPERATURE

INTRODUCTION

Since our analysis of stratospheric energetics requires resolution of the temperature and height fields into zonal mean and eddy components it is desirable to retrieve profiles of the harmonic components of temperature directly from the harmonic components of radiance obtained by Fourier analysis of the observations with respect to longitude. Only the first three or four wavenumbers are important in the dynamics of the winter stratosphere (Barnett, 1973). We therefore need to retrieve the zonal mean and three or four pairs of sine and cosine Fourier coefficients for each latitude. Since Fourier components of radiance are linear combinations of the observations, the linear retrieval method outlined in Chapter 2 may be used to obtain Fourier component profiles which are linear combinations of the profiles which would be obtained from the radiance measurements at individual grid-points. Thus in theory the grid-point and Fourier component retrievals give the same results but considerable savings in computation time are achieved with the latter method as Fourier analysis is performed on the radiance observations rather than on the more numerous retrieved temperatures and only seven or nine profiles need to be retrieved compared with about 36 grid-point profiles.

The choice of appropriate constraints to be applied to the Fourier component retrieval is rather difficult since it is not possible to obtain adequate statistics of Fourier coefficient profiles from radio- or rocket-sonde data. In the absence of suitable statistical constraints we could, for example, resort to the 'minimum information' retrieval in which the matrix H of

equation 2.1a is the unit matrix, implying that all atmospheric levels behave independently. However the information that we do have on stratospheric planetary waves suggests that there is considerable correlation between different levels. It would therefore seem sensible to develop constraints on the basis of our current knowledge of stratospheric waves even though much of this derives from satellite observations and numerical and theoretical models of wave behaviour. The suggestion by Rodgers (private communication) that the observations themselves may be used as a constraint on the solution is particularly appropriate in the context of Fourier component retrieval, for which conventional observational statistics are lacking, and to the situation when the atmospheric state is very abnormal, such as during sudden warmings.

This chapter reports the development of a Fourier component retrieval method which was used to produce a data set for the analysis of the sudden warming of January - February, 1973. Since this is the first attempt at a retrieval of this sort our philosophy was that it should be relatively simple, and consequently a semi-empirical approach was adopted. We begin by illustrating how inappropriate a priori information may give rise to poor retrievals. The bulk of the chapter deals with the construction of a covariance matrix from synthetic Fourier coefficient statistics. The modification of this matrix for different latitudes and wavenumbers is described. We show how instrumental random noise in the radiance Fourier coefficients was estimated from two years of Nimbus 5 data. The retrieval of zonal mean temperature is treated in the following section. Finally we compare retrieved temperatures and geopotential heights with conventional observations made during the period.

3.1. AN EXPERIMENT TO DEMONSTRATE THE ROLE OF THE A PRIORI INFORMATION

In this section we illustrate the importance of using appropriate constraints by describing an experiment in which grid-point retrievals were performed on synthetic atmospheres using two different sets of a priori information.

The retrieval experiment employed two covariance matrices, C_1 and C_2. C_1 had constant leading diagonal terms, indicating constant atmospheric variance with height, and a Gaussian bell shape to describe the covariance between levels. C_2 was derived from a large sample of rocket-sonde observations. Two theoretical distributions of temperature on a height-longitude plane, A_1 and A_2, were constructed. A_1 consisted of a standard atmospheric temperature profile at all longitudes on which was superimposed a zonal wavenumber one wave with constant amplitude with height of 20K and a phase slope of 25^{o}W per scale height.[1] A_2 differed from A_1 in that the wave amplitude was the same function of height as the atmospheric variance in the matrix C_2, the maximum amplitude being 20K. Grid-point radiances implied by atmospheres A_1 and A_2 in channels B12, B34 and A1 of the Nimbus 5 SCR were evaluated and grid-point temperature profiles for each atmosphere were retrieved using equation 2.1a with covariance matrices C1 and C2 in turn. The observational error covariance matrix, E, was estimated from the instrumental random noise values given by Barnett et al (1975). The wave amplitude was determined by Fourier analysis of the grid-point retrievals for each covariance matrix - atmosphere combination. Note that the wave amplitude is independent of the a priori profile, $\underline{B}_o$, and an arbitrary

1. We shall often refer to 1 unit in $\ln(p_s/p)$ as a scale height, although it is not an exact distance (see section 2.1).

value for $\underline{B}_o$ may be used. Figures 3.1(a) and (b) compare the actual wave amplitude profile of A1 with those derived from the grid-point retrievals of A1 using (a) C1 and (b) C2, while Figures 3.1(c) and (d) are the corresponding comparisons for A2 using (c) C1 and (d) C2. The best agreement between retrieved and actual wave amplitudes was obtained using the covariance matrix whose variation of variance with height matched the height dependence of the atmospheric wave ((a) and (d)). Since it was possible that the different pattern of the off-diagonal terms rather than the different height dependence of the variance might have been the major cause of the different results a third covariance matrix C3 was tested. C3 was obtained by multiplying the n leading diagonal terms C_{ii}, i = 1,n of C1 by factors r_i, i = 1,n, in order that the variance would have the same height dependence as C2. Each of the off-diagonal terms C_{ij}, i,j = 1,n, of C_1 was, necessarily, multiplied by $(r_i^2 + r_j^2)^{\frac{1}{2}} 2^{-\frac{1}{2}}$. C3 thus had a similar covariance pattern to C1 with large correlations only between closely spaced levels. The wave amplitude profiles obtained using C3 in place of C2 were not significantly different from those shown in Figures 3.1(b), (d). It is therefore likely that the height dependence of the variance and of the covariance between closely spaced levels largely determined the shapes of the profiles in Figure 3.1. All the retrieved profiles of Figure 3.1 are, of course, possible solutions which give, allowing for instrumental noise, the measured radiances. Figures 3.1(b) and (c) show that if the constraints are too tight at one level, giving a low value, a correspondingly high value must be produced at another level in the weighting function spread for the observed radiance to be satisfied.

The choice of constraints is clearly critical. Ideally each

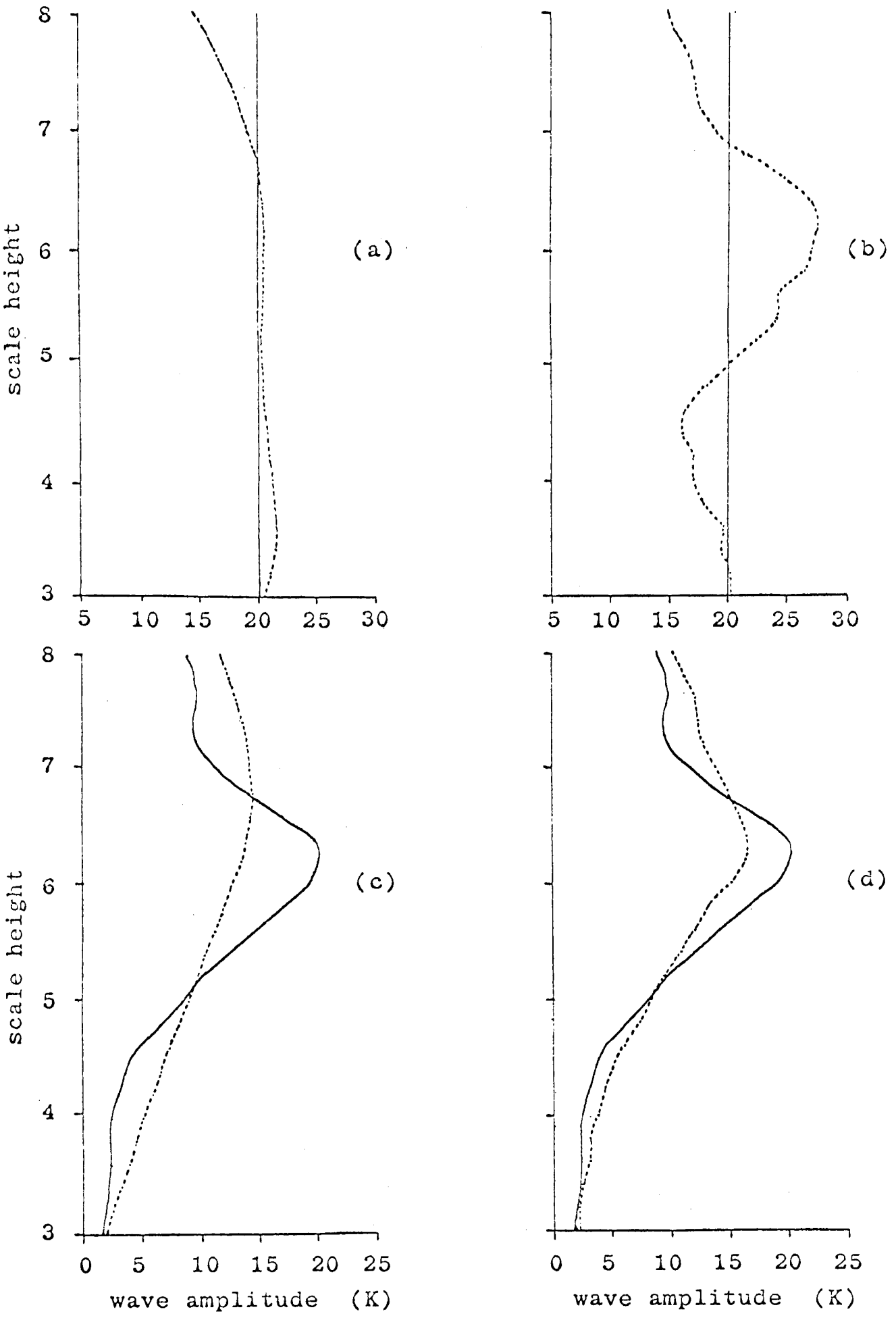


Figure 3.1. Comparison of actual and retrieved wave amplitudes using atmosphere - covariance matrix combinations A1-C1 (a), A1-C2 (b), A2-C1 (c), and A2-C2 (d). See text for details

profile should be retrieved with its own a priori information; such a technique may be employed in a forecasting situation using the forecast profile as the virtual measurement, or, as noted in Chapter 2, when retrievals are performed along orbit tracks, one retrieval being used as the virtual measurements of the next. Such a scheme is not applicable to Fourier component retrievals so the aim in developing the method described in this chapter was to include in the virtual measurements those characteristics of Fourier components which are thought not to change considerably from one profile to the next, but to accommodate, as far as possible, expected spatial and wavenumber variations.

3.2. SYNTHESIS OF A COVARIANCE MATRIX OF FOURIER COMPONENTS OF TEMPERATURE

(i) Properties of winter planetary waves

With the aid of Nimbus 5 SCR observations the properties of winter planetary waves which may be used to synthesize artificial statistics of Fourier components of temperature are now considered. Figure 3.2(a) shows the variation with time of the amplitude of wavenumber 1 radiance at 60^{o}N in channels B12, B34 and A1 (which peak at about 45, 33 and 20km) for large periods of the winters 1972-3, 1973-4. Although the wave amplitude does vary with height there is no long-term preferred level of maximum amplitude. The monthly mean latitudinal distributions of wavenumber 1 amplitude in November 1973 and January 1973 (Figures 3.2(b) and (c)) demonstrate that the level of maximum amplitude varies with latitude as well as with time. These figures also indicate that there may be a tendency towards similarity of amplitude in all channels in high and low latitudes. In constructing a set of wave statistics our first assumption,

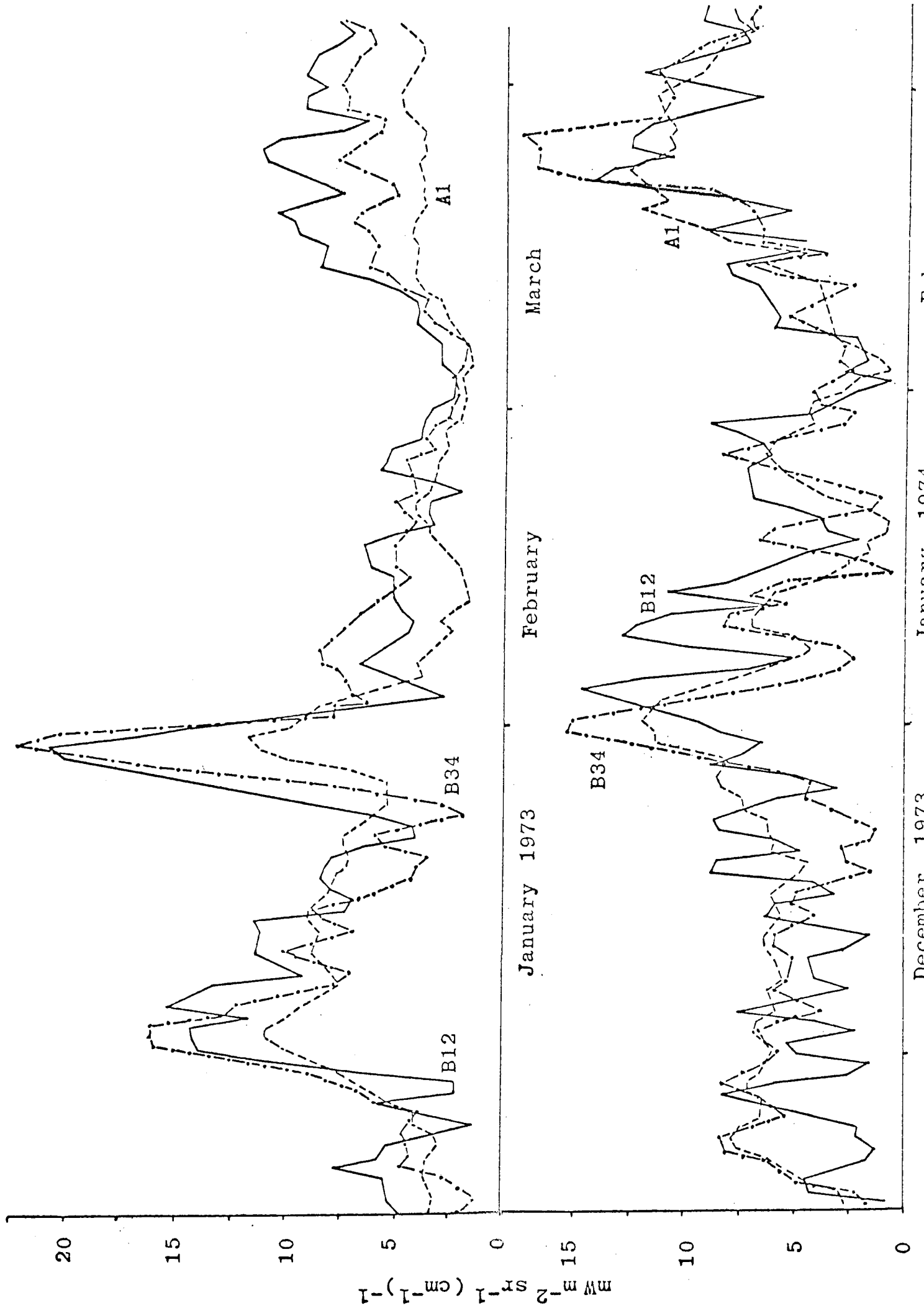


Figure 3.2a. The variation with time of the wavenumber one radiance amplitude at 60°N in channels B12, B34 and A1 during two winter periods

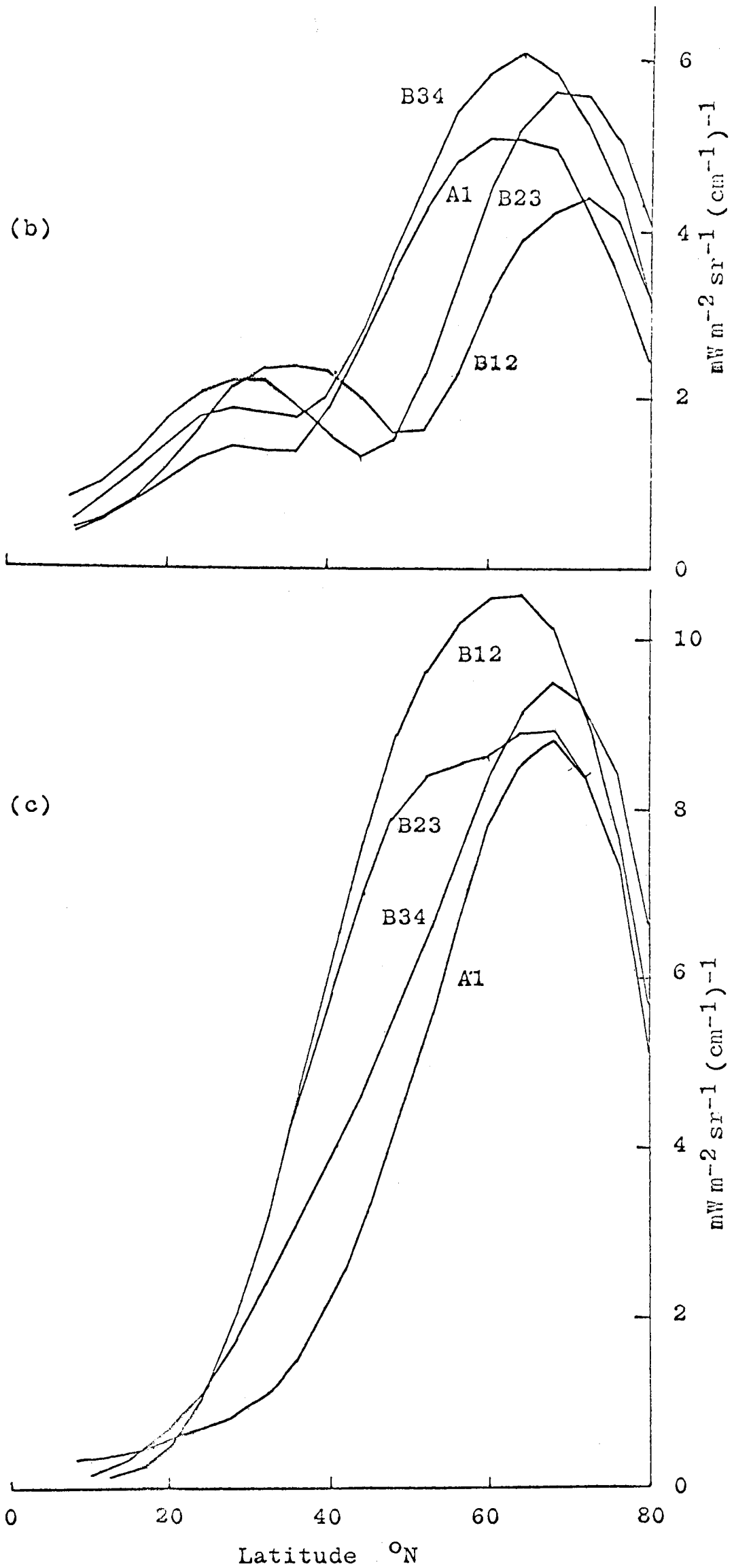


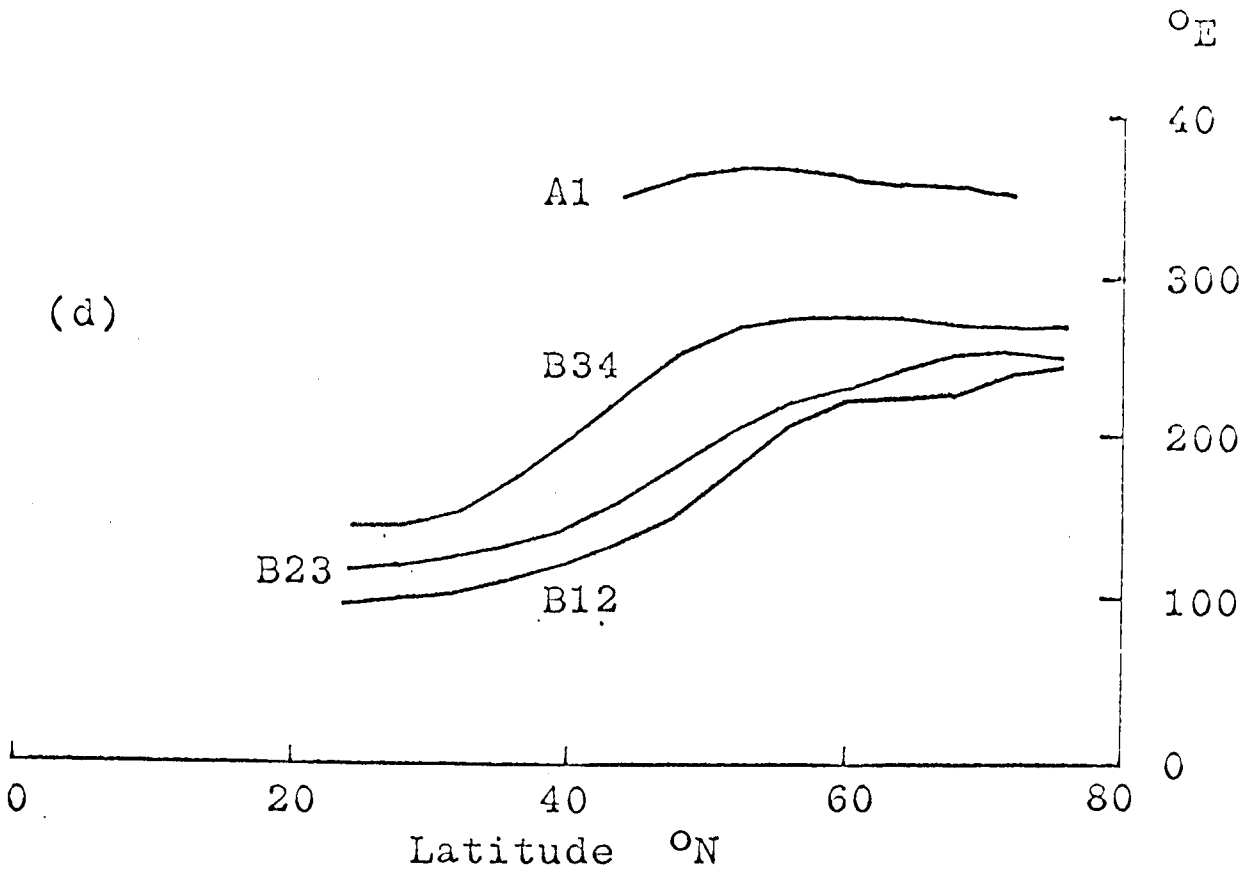
Figure 3.2b,c. Monthly mean distribution with latitude of wavenumber one radiance amplitude in channels B12, B23, B34 and A1 in November 1973 (b) and January 1973 (c)

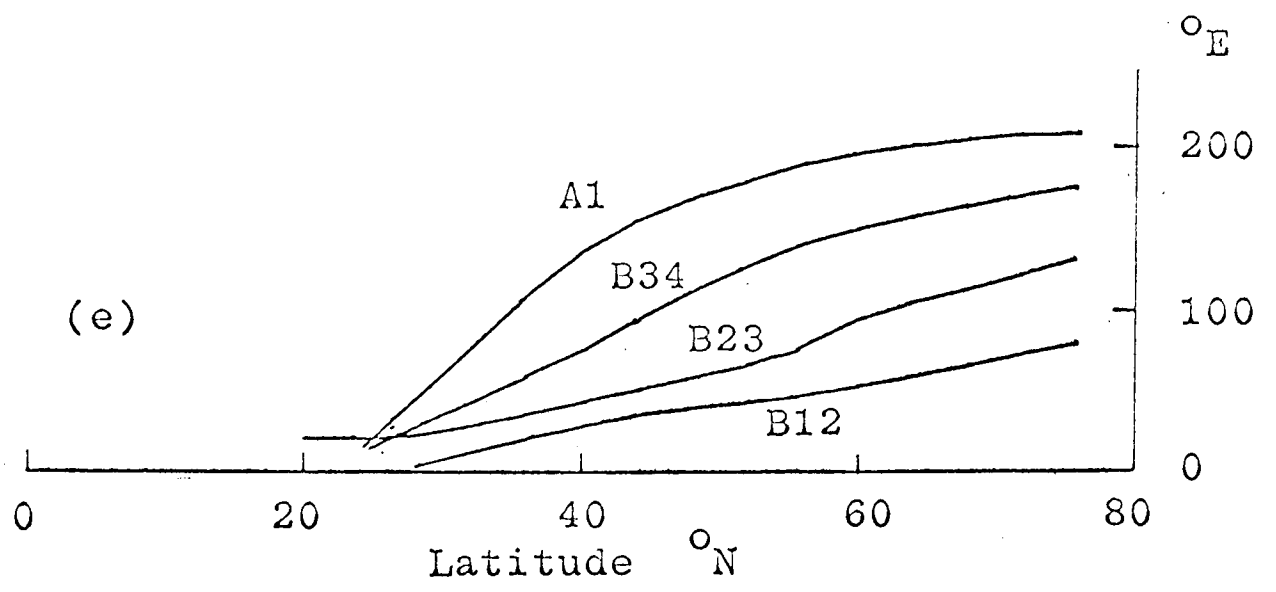
therefore, was that the wave amplitude is constant with height. During winter the planetary waves generally slope westward with height (see Chapter 1), and for simplicity it was assumedthat the phase change with height is linear. This is not an unreasonable assumption as is indicated by Figure 3.2(d) and (e), in which are shown monthly mean latitudinal distributions of wave phase for February 1973 (wavenumber 2) and January 1973 (wavenumber 1), and by Figure 3.2(f), which shows the positions of maximum and minimum radiance as observed in the stratosphere and mesosphere by the Nimbus 5 SCR and the Nimbus 6 Pressure Modulator Radiometer on 7th February 1976. Further evidence in support of these assumptions, at least for the low and middle stratosphere, is given by van Loon et al (1973), who presented height-longitude cross-sections of the January mean wavenumber 1 and 2 temperature waves at 65^{o} and 60^{o}N respectively for five winters, showing that amplitudes and phase tilts did not vary substantially between 50 and 10 mb.

(ii) Covariance of Fourier component profiles in a model winter stratosphere

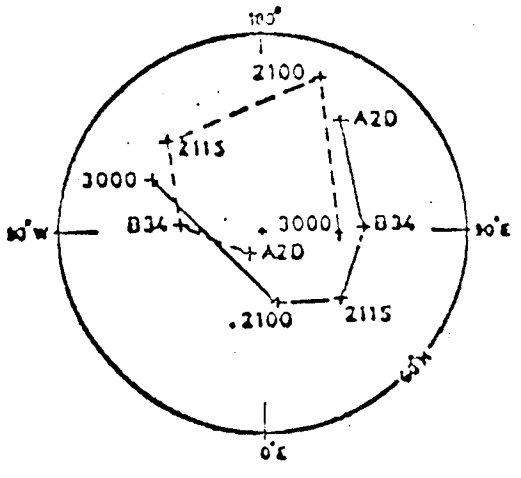
An atmospheric temperature profile covariance matrix is often determined from an ensemble of rocket-sonde profiles. An equivalent procedure for the Fourier component retrieval based on the assumptions made above will now be described.

Consider an ensemble of temperature waves which have uniform amplitude and slope with height. In addition suppose that the waves have the same zonal wavenumber and amplitude but that the horizontal phase and the rate of phase change with height varies from wave to wave. For any wave the rate of phase change with height may be represented by the vertical wavenumber k, or vertical wavelength $\lambda = 2\pi/k$. The horizontal phase, ϕ_o, is given





(f)



Height order of channels: A2D (lowest), B34, 2115, 2110, 3000 (After Austen et al, 1976)

Figure 3.2d-f. Monthly mean latitudinal distributions of wave phase (°E) in channels B12, B23, B34 and A1 for wavenumber 2, February 1973 (d) and wavenumber 1, January 1973 (e), and the positions of maximum (solid line) and minimum (dashed line) radiance observed by Nimbus 5 SCR and Nimbus 6 PMR on 7.2.1976 (f)

in terms of the longitude of a wave ridge. For an atmosphere consisting of a single wave of zonal wavenumber n, vertical wavenumber k, horizontal phase ϕ_o at z = o and amplitude A, the departure T' of the temperature from the zonal mean at a given level z and longitude ϕ is

$$T'(z,\phi) = A \cos n(\phi - kz - \phi_o)$$

Fourier analysis shows that the cosine and sine coefficients are given by

$$a_n = A \cos n(kz+\phi_o)$$

$$b_n = A \sin n(kz+\phi_o)$$

By integrating over all phases ϕ_o the mean values of a_n and b_n for the ensemble are zero. If we consider that the ensemble contains waves in the vertical wavenumber range $0 \le k \le k_c$, the covariance between the cosine coefficients at levels z_i and z_j may be written

$$H_{ij} = \frac{\int_0^{k_c}\int_0^{2\pi} A \cos n(kz_i+\phi_o) A \cos n(kz_j+\phi_o)\ d\phi_o dk}{\int_0^{k_c}\int_0^{2\pi} d\phi_o dk}$$

$$= \frac{A^2}{2\pi k_c}\int_0^{k_c}\int_0^{2\pi} \tfrac{1}{2}\{\cos n(kz_i+kz_j+2\phi_o) + \cos(nk(z_i-z_j))\}d\phi_o dk$$

The first term in the integral vanishes leaving

$$H_{ij} = \frac{A^2}{2} \sin(nk_c(z_i-z_j)) \,/\, nk_c(z_i-z_j) \qquad 3.1$$

The equivalent term in the sine coefficient covariance matrix is identical. It should be noted that the vertical wavenumber k corresponding to a given rate of phase change with height varies with zonal wavenumber, n; i.e. $k_n = nk_1$. Thus the sine and cosine coefficient covariance matrices given by equation 3.1 are valid for all wavenumbers for a specified wave tilt but not for a

specified limiting vertical wavenumber, k_c.

In order to test this form of covariance matrix the radiance Fourier coefficients implied in channels B12, B23, B34 and A1 by sloping waves of uniform amplitude and various rates of phase change with height were evaluated. Covariance matrices (equation 3.1) were evaluated for a range of cut-off slopes, i.e. a range of values of nk_c, and used to retrieve Fourier coefficient profiles from these radiances. The amplitude and phase profiles derived from them, which for convenience we shall refer to as 'retrieved amplitude and phase profiles', were compared with the original wave amplitudes and phases. It was found in general that waves whose rate of phase change with height was smaller than the specified cut-off slope were well retrieved while those with larger tilts were retrieved very poorly. An example is shown in Figure 3.3(a) of the retrieved amplitude and phase profiles obtained for waves of constant amplitude 8.5 units and vertical slopes corresponding to vertical wavelengths for a zonal wavenumber one wave of 40, 10 and 6 scale heights. The covariance matrix used had a cut-off slope corresponding to a vertical wavelength of 8 scale heights. The problem of retrieving waves with large tilt is not necessarily solved by using a covariance matrix with a greater range of allowed slopes, since the negative correlations between relatively close levels which are thereby generated produce spurious oscillations in the retrievals of the waves of longer vertical wavelength. This effect is illustrated in Figure 3.3(b) which shows the retrieved amplitude profiles of the three waves considered in Figure 3.3(a) obtained using a covariance matrix with a cut-off slope corresponding to a vertical wavelength of 1 scale height for a wavenumber one wave. It is possible, however, to construct, as described below, a

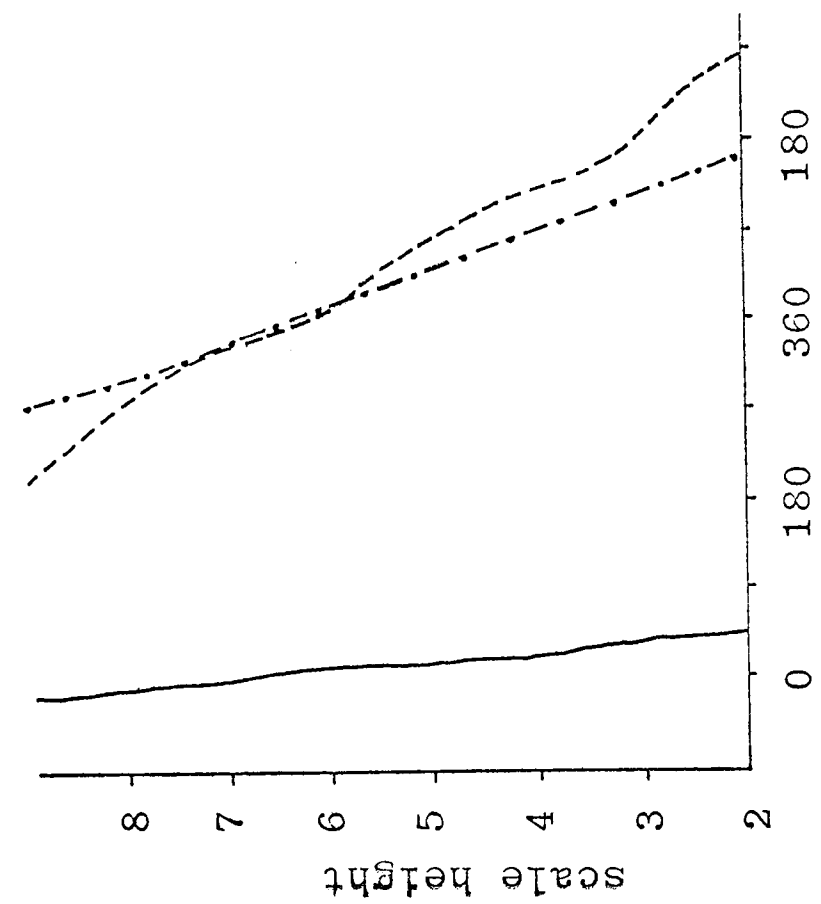
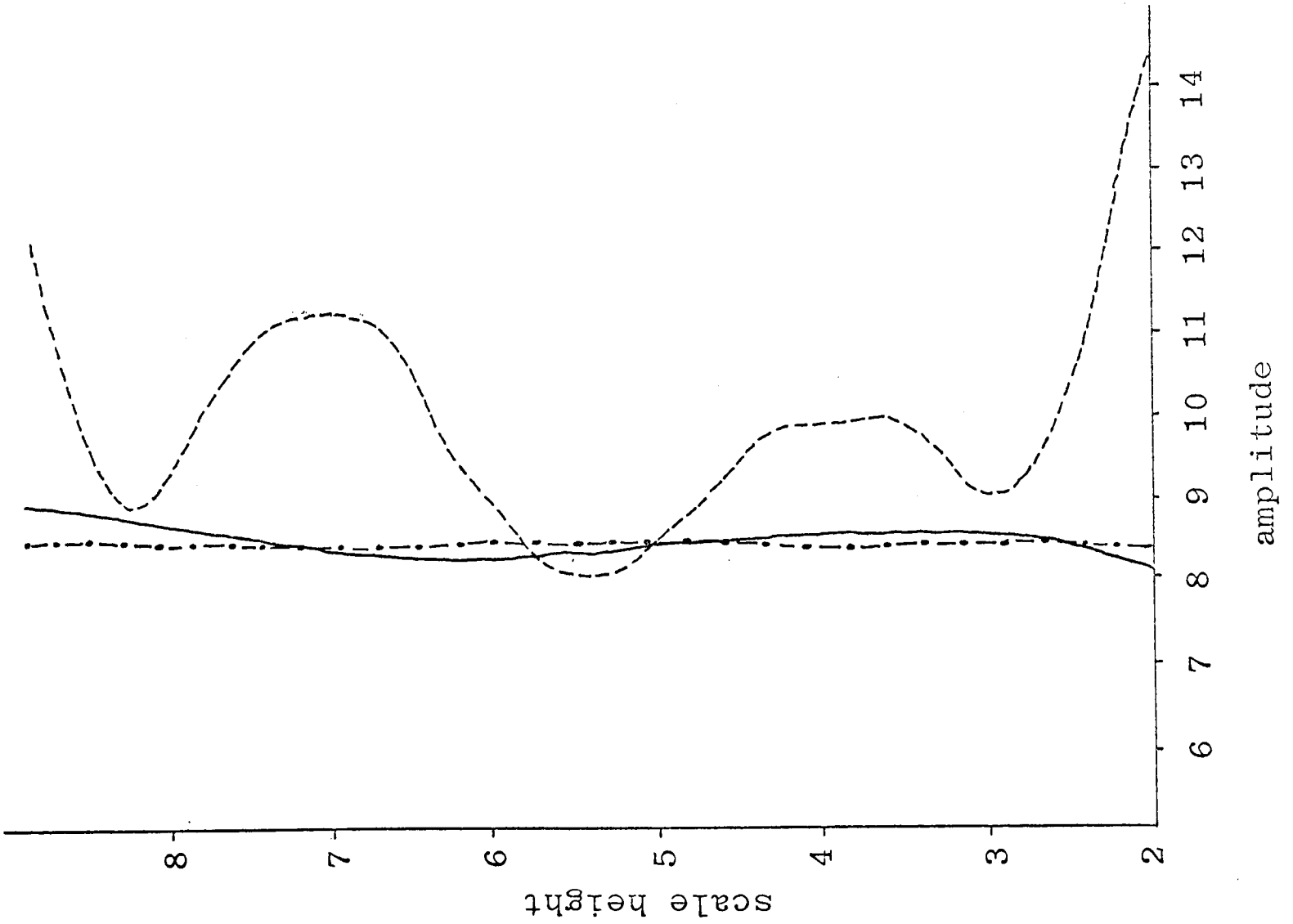


Figure 3.3a. Retrieved amplitude (left) and phase (above) profiles obtained for waves of constant amplitude (8.5 units) with height and uniform vertical slopes corresponding to vertical wavelengths for a zonal wavenumber one wave of 40 (———), 10 (-.-.-.) and 6 (- - - -) scale heights. The covariance matrix used had a cut-off slope corresponding to a wavenumber one vertical wavelength of 8 scale heights



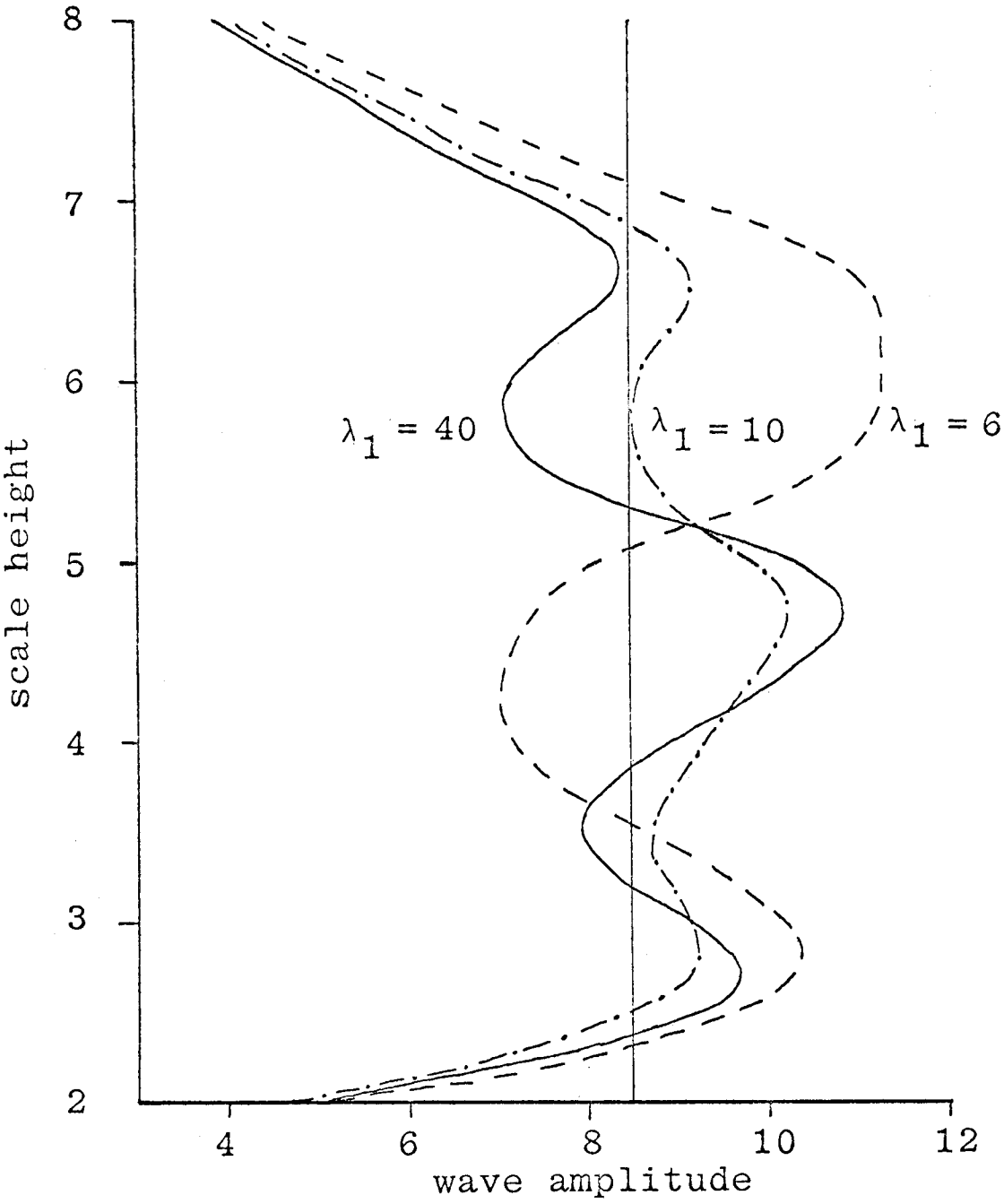


Figure 3.3b. Retrieved wave amplitudes as in Figure 3.3a but using a covariance matrix with a cut-off slope corresponding to a wavenumber one vertical wavelength of 1 scale height

covariance matrix in which this problem is less serious.

(iii) Modification of covariance matrices to permit retrievals of waves of all vertical slopes

Figures 3.4(a) and (b) are scatter diagrams which give an indication, albeit rather crude, of the range of wave slopes which occur in the winter stratosphere. The average wave amplitude calculated from radiances in channels B12, B34 and A1 is plotted against the phase difference between channels B12 and A1 for (a) wavenumber one and (b) wavenumber two at 60^{o}N. Each point is a daily value from the periods mid-December 1972 - March 1973, December 1973 - March 1974 and December 1974. A few occasions on which low amplitude, eastward sloping waves occurred have not been represented on the diagrams. It is seen that few waves of wavenumber one or two had phase differences of greater than 180^{o} between the peak levels of channels B12 and A1, which corresponds, approximately, to vertical wavelengths of 6 scale heights for wavenumber one and 3 scale heights for wavenumber two. In practice the wave slope may vary with height in which case sharper tilts would be obtained over shorter vertical distances. We therefore require a covariance matrix which retrieves well the components of waves with slopes corresponding to $n\lambda_n \geq 6$ scale heights but has the capacity to retrieve waves of shorter vertical wavelength without being detrimental to the longer vertical wavelengths. A covariance matrix in which covariance decreases in a Gaussian manner with increasing separation of the levels does not contain negative correlations. Thus by choosing a Gaussian function which gives correlation between closely spaced levels similar to those obtained by equation 3.1 using a cut-off slope corresponding to $n\lambda_c = 6$ scale heights our requirements should be met. Assuming a constant amplitude, A, with height, the

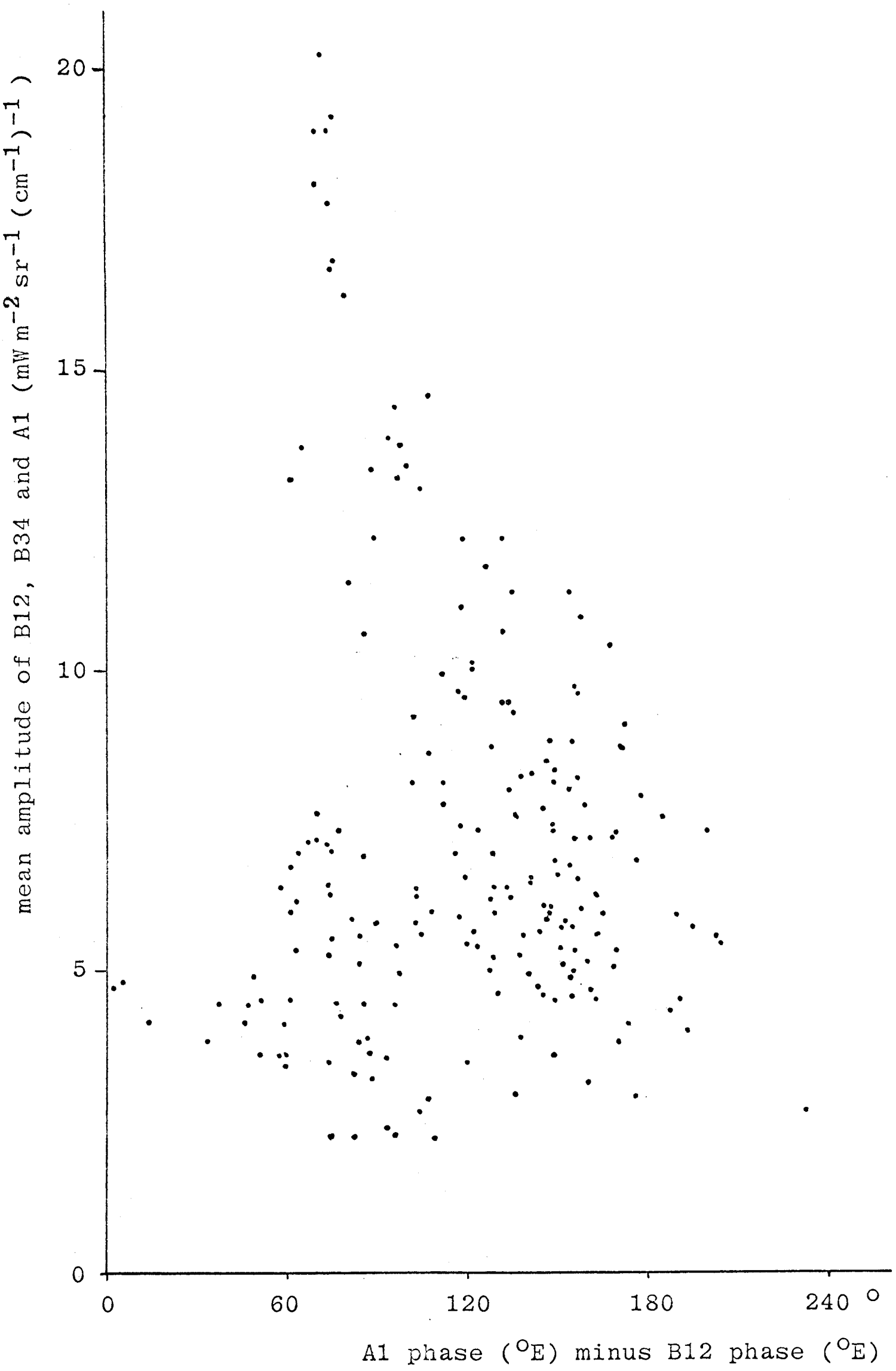
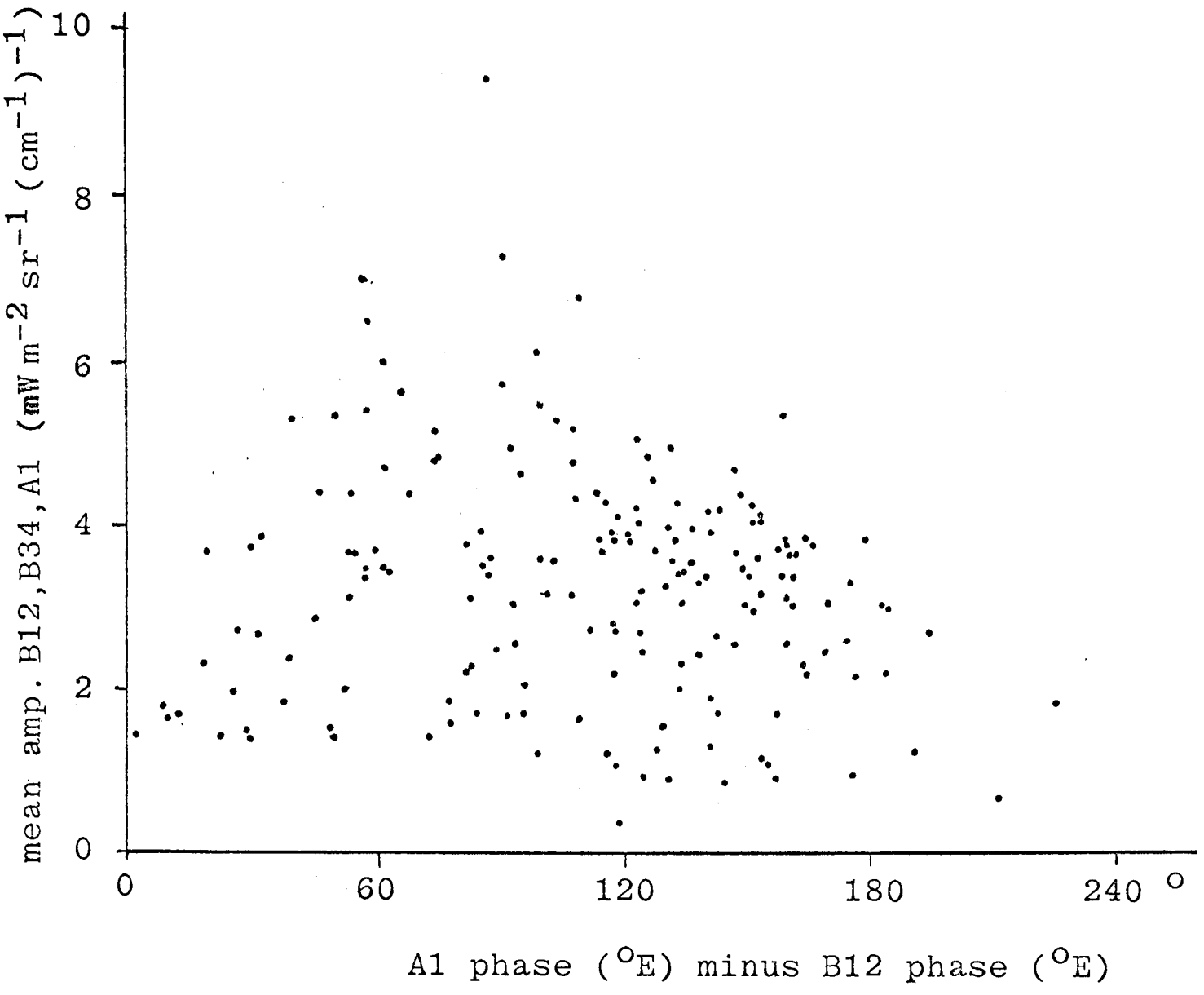


Figure 3.4a. See next page for caption



Figures 3.4a (previous page) and 3.4b (above). Scatter diagrams showing the mean wavenumber one (a) and wavenumber two (b) radiance amplitude of channels B12, B34 and A1 plotted against the difference in phase between channels A1 and B12 at 60°N. Daily values from the periods mid-December 1972 to March 1973, December 1973 to March 1974 and December 1974 are plotted

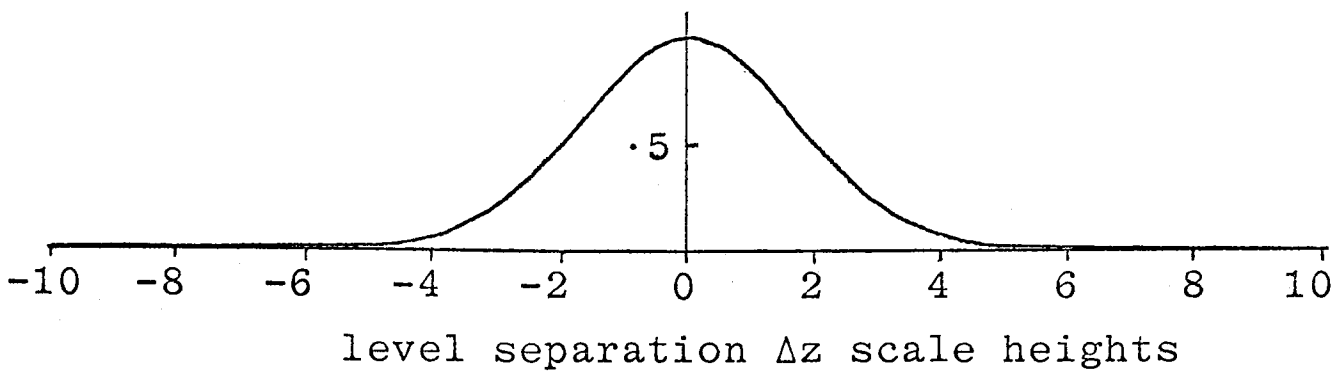
covariance between levels i and j is given by

$$H_{ij} = A \, exp\,(-\alpha(z_i - z_j)^2) \qquad 3.2$$

where α determines the width of the Gaussian bell. Figure 3.5 compares, as a function of level separation, the exponential of equation 3.2 for $\alpha = 0.17$ with $(\sin nk\Delta z)/nk\Delta z$ for $nk = 2\pi/6$ (see equation 3.1). Correlations between closely spaced levels are very similar in the two cases while the correlation falls to near zero at a separation of 4 scale heights in the Gaussian matrix. The results of using this matrix to retrieve the three waves considered above (Figure 3.3) are shown in Figure 3.6. It will be noticed that although the retrievals of the waves of small tilt are a little less smooth than those of Figure 3.3(a), there is much improvement over the profiles of Figure 3.3(b). The Gaussian covariance matrix was considered to be sufficiently general to give reasonably smooth retrieved profiles for the range of planetary waves likely to be found in the winter stratosphere.

(iv) Modification of covariance matrices for different wavenumbers and latitude bands

It is well known that winter stratospheric disturbances are much more pronounced in middle and high latitudes than in low latitudes and that wavenumbers three and four usually have smaller amplitudes than wavenumbers one and two. Accordingly the basic covariance matrix should be adjusted to suit different latitudes and wavenumbers. Rodgers (private communication) suggested how the satellite observations themselves may be used to modify an initial estimate of the atmospheric covariance. Suppose that a covariance matrix, X, of the observed radiances is derived from a set of observations - for example, one month's data over a given range of latitude. An atmospheric covariance, H,



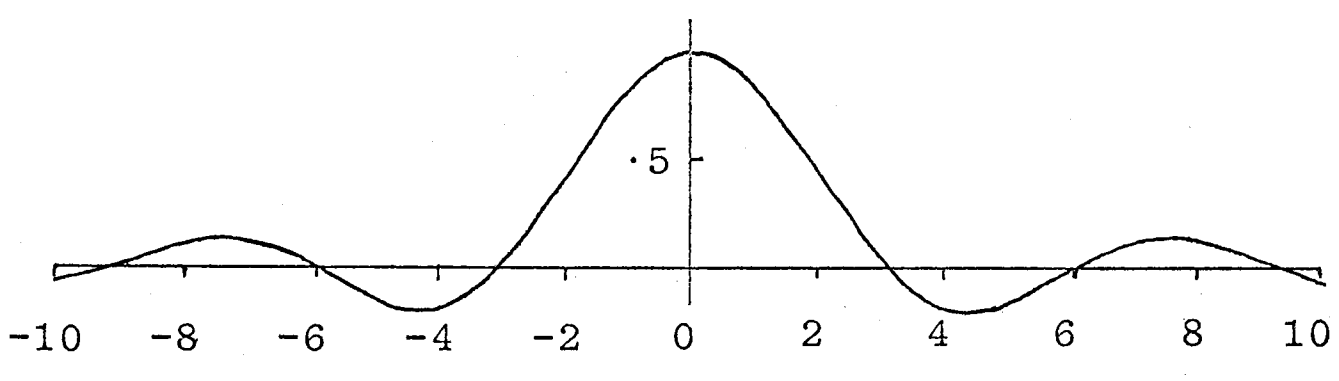
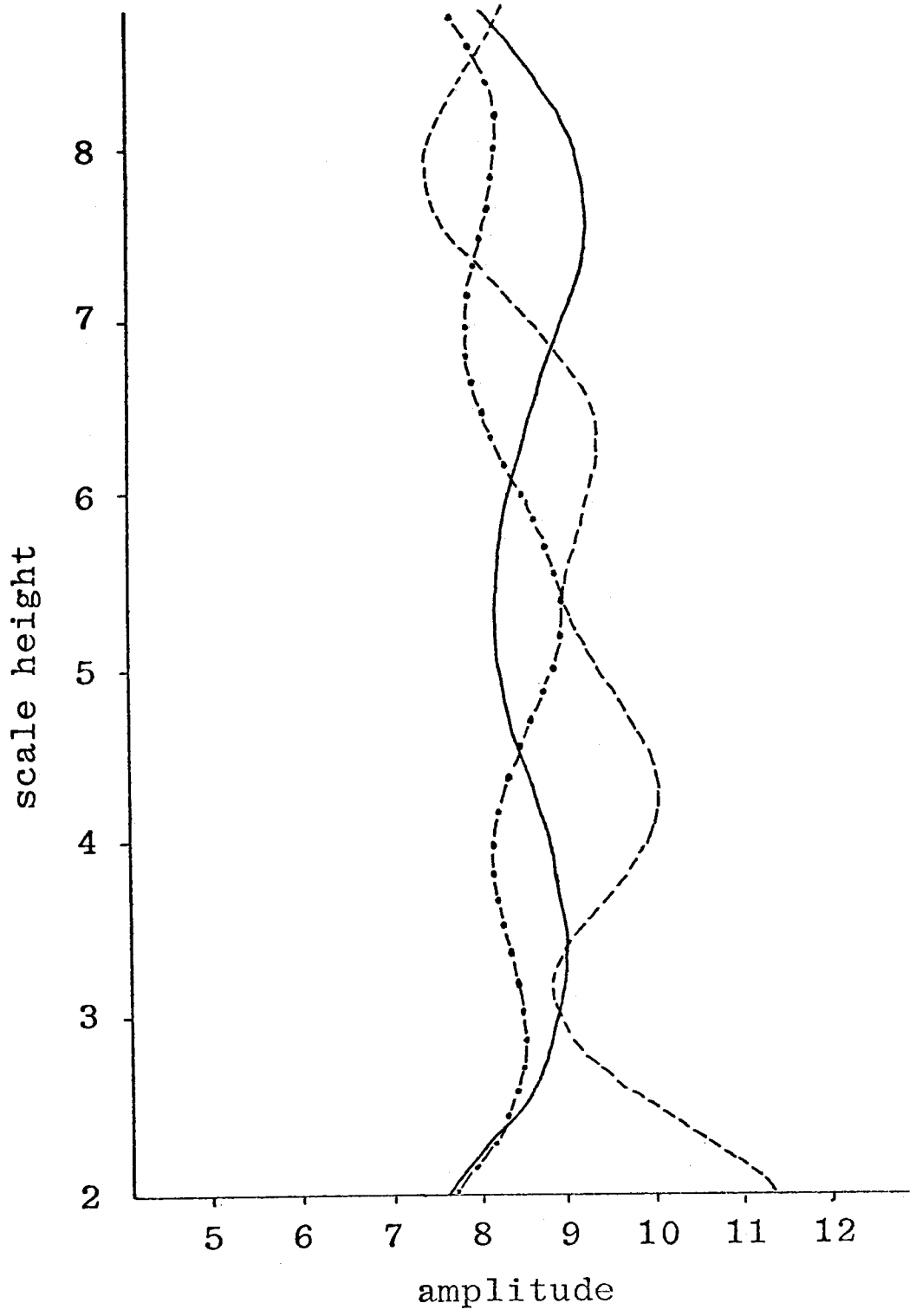


Figure 3.5. The functions *exp* $(-0.17(\Delta z)^2)$ (upper) and $(\sin \Delta z\ 2\pi/6)/(\Delta z\ 2\pi/6)$ (lower) plotted against Δz. The ordinate shows the correlation between levels of separation Δz implied by a covariance matrix of the given form



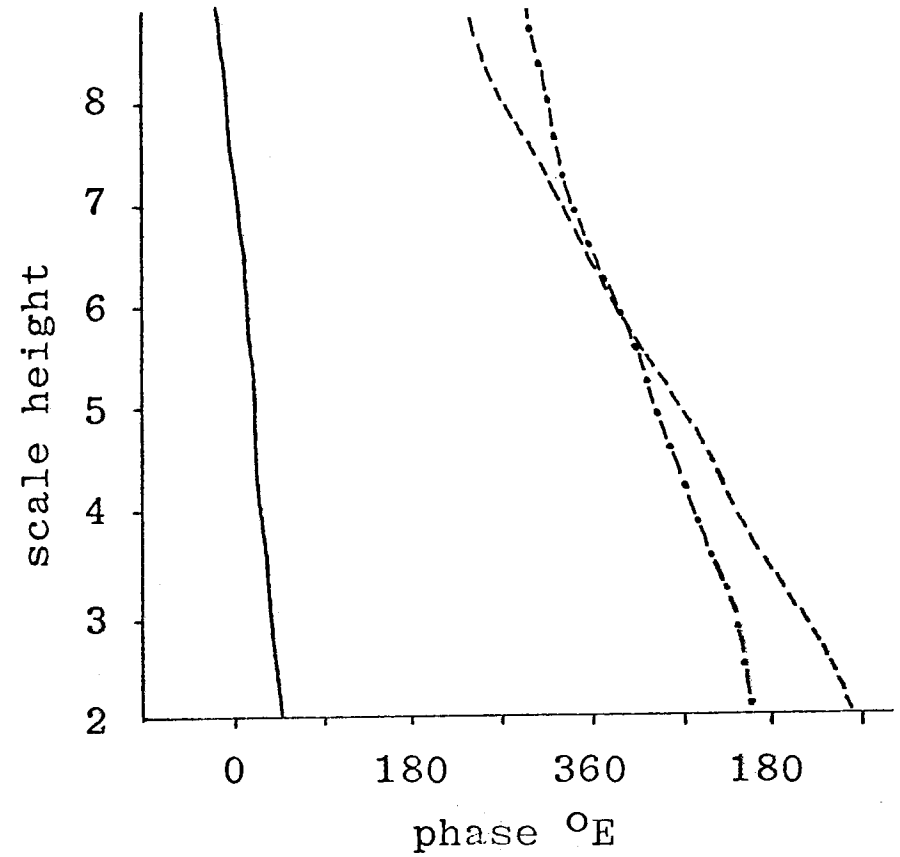


Figure 3.6. As Figure 3.3a but using a covariance matrix of the form given by $H_{ij} = A\ exp\ (-0.17(z_i - z_j)^2)$

corresponding to X must satisfy

$$KHK^T = X$$

where K is the weighting function matrix. There are of course an infinite number of covariance matrices H satisfying this equation. A first estimate of the atmospheric covariance, H_0, may be replaced, in the light of the measured X, by the H nearest to H_0, in the least squares sense, which is consistent with the observed X. Thus

$$\sum_{i,j}(H_{ij}-H_{0ij})^2$$

is minimised subject to

$$\sum_{i,j}(K_{pi}H_{ij}K_{jq}) = X_{pq}$$

where i,j are atmospheric levels and p,q are channels of the observing instrument. In our case the matrix H_0 would be the Gaussian covariance matrix described above and X the covariance matrix of radiance Fourier coefficients. There is of course a problem in deciding how large should be the data set used to determine X; the smaller the sample of observations the less general is the derived matrix H but the better are the retrievals from that sample.

Covariance matrices of sine and cosine coefficients of radiance are presented in Table 3.1. Each corresponds to a given wavenumber and latitude band. They were evaluated from SCR data for the period 11.1.73 - 14.2.73. It is seen that the wavenumber one coefficient observation covariance matrices for the 36-60°N and 56-80°N bands have similar variances in two channels but a very different variance in the third in each case. As a consequence direct modification of the Gaussian covariances by the method outlined above gave modified covariance matrices very different from the original Gaussian form. Moreover a consistent

TABLE 3.1

Covariance matrices of sine (upper) and cosine (lower) coefficients of radiance for wavenumbers 1 to 4 and latitude bands 12-40, 36-60 and 56-80°N derived from channel B12, B34 and A1 radiances during the period 11.1.73 - 14.2.73. Units: $(mW\,m^{-2}\,sr^{-1}\,(cm^{-1})^{-1})^2$

		12 - 40°N			36 - 60°N			56 - 80°N		
		B12	B34	A1	B12	B34	A1	B12	B34	A1
wave 1	B12	2.59	1.20	0.48	16.05	13.23	4.36	22.77	17.98	3.91
	B34	1.20	2.06	0.84	13.23	16.87	5.94	17.98	25.79	10.38
	A1	0.48	0.84	0.46	4.36	5.94	2.56	3.91	10.38	6.35
wave 2	B12	3.39	1.39	0.26	10.52	4.71	-0.22	7.13	3.40	-0.72
	B34	1.39	1.11	0.38	4.71	5.38	1.77	3.40	4.16	1.18
	A1	0.26	0.38	0.21	-0.22	1.77	1.50	-0.72	1.18	1.45
wave 3	B12	1.38	0.80	0.15	4.46	2.04	0.11	2.10	0.69	0.10
	B34	0.80	0.85	0.26	2.04	2.29	0.94	0.69	1.15	0.50
	A1	0.15	0.26	0.15	0.11	0.94	0.71	-0.10	0.50	0.49
wave 4	B12	0.16	0.09	0.02	0.58	0.21	-0.01	0.35	0.13	0.00
	B34	0.09	0.09	0.03	0.21	0.34	0.13	0.13	0.19	0.06
	A1	0.02	0.03	0.04	-0.01	0.13	0.14	0.00	0.06	0.08

		12 - 40°N			36 - 60°N			56 - 80°N		
		B12	B34	A1	B12	B34	A1	B12	B34	A1
wave 1	B12	5.43	2.40	-0.07	11.04	-0.29	-3.85	7.50	-1.28	-4.65
	B34	2.40	1.70	0.18	-0.29	6.05	5.38	-1.28	20.29	17.07
	A1	-0.07	0.18	0.34	-3.85	5.38	6.63	-4.65	17.07	17.55
wave 2	B12	3.57	1.31	0.32	14.43	4.53	-0.94	6.38	1.40	-1.11
	B34	1.31	1.25	0.46	4.53	3.97	1.21	1.40	2.66	1.15
	A1	0.32	0.46	0.26	-0.94	1.21	1.47	-1.11	1.15	1.48
wave 3	B12	0.84	0.36	0.09	3.16	1.29	-0.03	1.39	0.46	-0.31
	B34	0.36	0.62	0.26	1.29	2.35	0.98	0.46	1.22	0.47
	A1	0.09	0.26	0.14	-0.03	0.98	0.75	-0.31	0.47	0.50
wave 4	B12	0.27	0.09	0.02	0.65	0.19	0.05	0.23	0.14	0.03
	B34	0.09	0.14	0.06	0.19	0.49	0.18	0.14	0.36	0.09
	A1	0.02	0.06	0.04	0.05	0.18	0.16	0.03	0.09	0.09

irregularity in the retrieved wave phase was obtained when these covariance matrices were tested with real data, a feature not observed with the unmodified forms. It was therefore desirable to retain the Gaussian form of covariance matrix but at the same time make use of the information provided by the observation covariances. An alternative approach was thus adopted: for each latitude band - wavenumber combination the value, A, of the leading diagonal terms in the Gaussian matrix H_0 (i.e. the value of the atmospheric variance at all levels) was chosen such that the difference between KH_0K^T and the observation covariance matrix X was minimised in the least squares sense subject to H_0 retaining its specified Gaussian form; i.e.

$$\sum_{i,j} (K_{pi}H_{0ij}K_{jq} - X_{pq})^2$$

was minimised subject to

$$H_{0ij} = A \, exp\,(-\alpha(z_i - z_j)^2)$$

Table 3.2 gives the values of A so calculated for each wavenumber and latitude band. These values were employed in equation 3.2 to construct the 24 covariance matrices required for the retrieval of the coefficient profiles of the first four wavenumbers in the three latitude bands. α was taken as 0.17 (see section (iii) above).

3.3. RANDOM NOISE IN RADIANCE FOURIER COEFFICIENTS

It will be recalled that since waves of all horizontal phase may occur we may assume that the a priori Fourier coefficient profiles are zero. Thus $\underline{B}_o$ and $\underline{I}_o$ in equation 2.1a disappear and the retrieval equation takes the form

$$\hat{\underline{B}} = HK^T(KHK^T + E)^{-1}\underline{I}$$

The method of estimating the instrumental error covariance, E, is described in this section.

Correlation between the random error, or noise, in the

TABLE 3.2

VARIANCES OF THE PLANCK FUNCTION USED TO CONSTRUCT THE 24 COVARIANCE MATRICES USED FOR THE RETRIEVAL OF PROFILES OF SINE AND COSINE FOURIER COEFFICIENTS OF THE PLANCK FUNCTION

Units: $(mWm^{-2}sr^{-1}(cm^{-1})^{-1})^2$

Sine Coefficient

wave number	12 - 40°N	36 - 60°N	56 - 80°N
1	2.10	16.74	24.69
2	1.82	6.43	4.49
3	0.99	2.89	1.29
4	0.12	0.37	0.22

Cosine Coefficient

wave number	12 - 40°N	36 - 60°N	56 - 80°N
1	2.67	6.47	15.61
2	1.93	6.49	3.11
3	0.63	2.32	1.00
4	0.17	0.46	0.26

radiance measurements of different channels depends on instrument design. The retrievals performed for this study used channels B12, B34 and A1 of the SCR. A large degree of noise correlation would not be expected between these channels and the instrumental error covariance, E, was therefore taken to be diagonal. The random noise variance in the radiance Fourier coefficients of a given channel should be independent of zonal wavenumber provided that (a) the wavenumber is low enough to be resolved adequately, i.e. less than 6, and (b) the amplitude of the higher, unresolved wavenumbers are small enough not to produce significant "aliassing", i.e. the process whereby high wavenumber components appear as contributions to the resolved lower wavenumbers.

At the equator wave activity is often very small and a large proportion of the variance in the Fourier coefficients is then due to instrumental noise. Monthly values of the equatorial variance in the Fourier coefficients of radiance were evaluated from 14 months' data. Figure 3.7 shows the results for both coefficients of wavenumbers 1 - 4 in each channel for selected months: (a) channel B12 for January and February, 1973 and March and July, 1974; (b) channel B12 for March 1973 and June 1974; (c) channel B34 for January - March, 1973; (d) channel A1 for January and March, 1973 and July 1974. Figure 3.7 (a) shows months during which equatorial wave activity was high in the upper stratosphere, in sympathy with polar warmings occurring during these months, while (b), (c) and (d) show months which were quiet in the regions covered by the particular channels. It is apparent that most of the variance due to atmospheric variations is contained in wavenumber 1, and there is a tendency towards asymptotic values for higher wavenumbers which may be

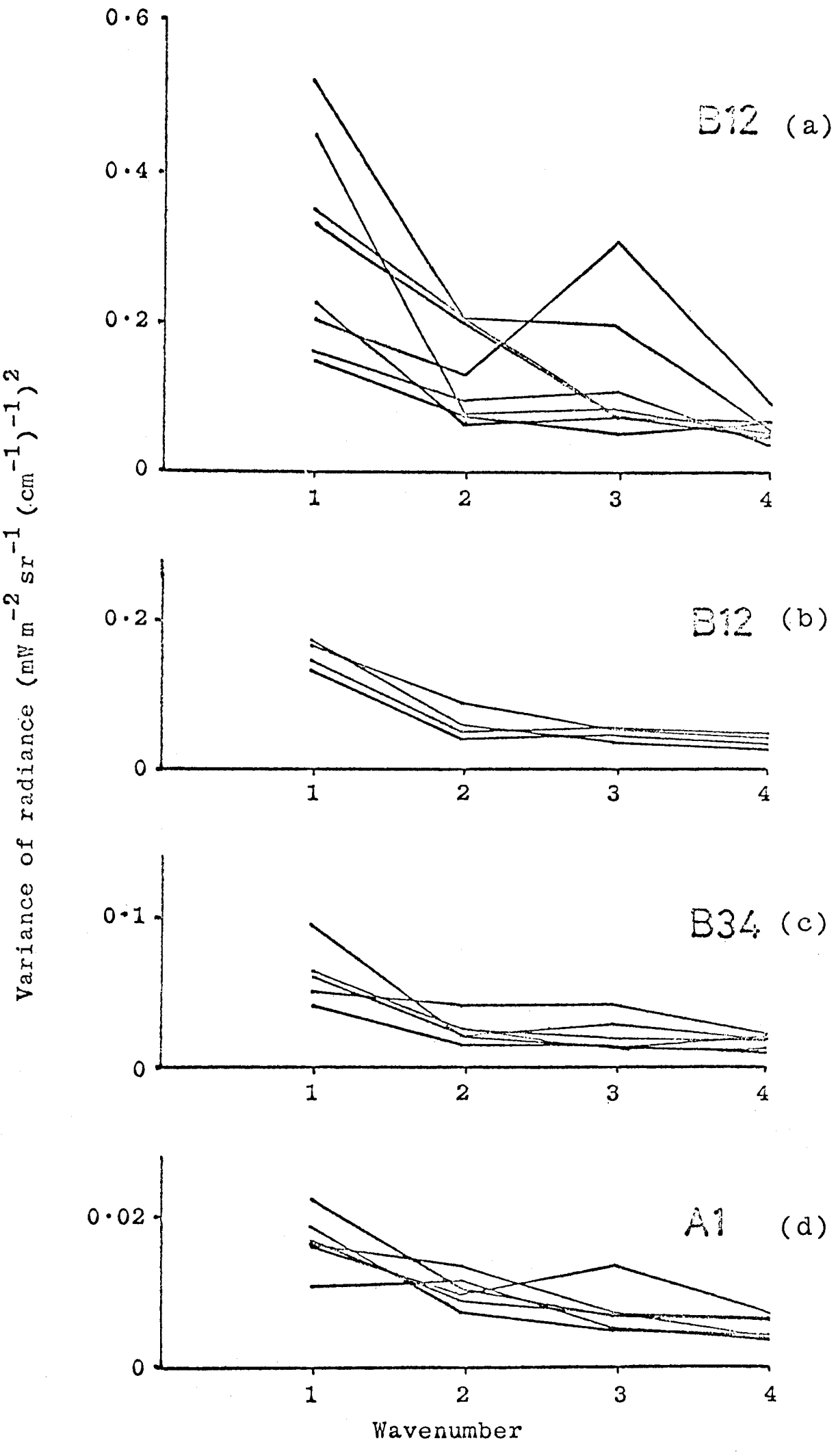


Figure 3.7. Variance of Fourier coefficients of equatorial radiance. Each curve shows the variances of the sine or cosine coefficients of wavenumbers 1 to 4 for a given month in a given channel (see text for details)

regarded as estimates of the variance due to random instrumental noise. The noise is very much lower in channel A1 than in the B channels. It is interesting to note that there was no increase in equatorial wave activity in the lower channels during months with polar warmings. Table 3.3 gives the instrumental random noise standard deviations estimated by this method.

Details of the Fourier component retrievals performed using the atmospheric and instrumental noise covariance matrices described in this and the previous sections are given in the final section of this chapter.

3.4. RETRIEVAL OF ZONAL MEAN TEMPERATURE

To retrieve the zonal mean Fourier component it is necessary to employ a non-zero a priori estimate of the profile. Climatological mean profiles were constructed for every 4^{o} of latitude between 12^{o} and 80^{o}N, using Groves' (1972) standard atmosphere for February 1st augmented by statistics from Labitzke and collaborators (1972) and Oort and Rasmussen (1971).

The form of the atmospheric covariance matrix was determined in the light of the covariance of the departure of the observed zonal mean radiance from the radiance implied by the a priori profile, i.e. the covariance of $\underline{I}-\underline{I}_{o}$ of equation 2.1a. These are given in Table 3.4 and show that the covariance falls to about one half of the variance for a separation equivalent to the spacing of the weighting function peaks, if one assumes that the radiance covariances are representative of the levels of peak weighting. For a Gaussian covariance matrix this gives a value of α in equation 3.2 of 0.25. The values of the variance, A, for each of the three matrices (one for each latitude band) were determined in the same way as for the sine and cosine coefficient covariance matrices. These are given in Table 3.5.

TABLE 3.3

STANDARD DEVIATION OF THE INSTRUMENTAL RANDOM NOISE IN RADIANCE FOURIER COEFFICIENTS OF CHANNELS B12, B34 AND A1.

Units: $mWm^{-2}sr^{-1}(cm^{-1})^{-1}$

B12	B34	A1
0.17	0.09	0.05

TABLE 3.4

COVARIANCE MATRICES OF THE DEPARTURE OF THE OBSERVED ZONAL MEAN RADIANCE FROM THAT IMPLIED BY A STANDARD ATMOSPHERE TEMPERATURE PROFILE FOR LATITUDE BANDS 12-40, 36-60 AND 56-80°N DERIVED FROM DATA FOR THE PERIOD 11.1.73-14.2.73.

Units: $(mWm^{-2}sr^{-1}(cm^{-1})^{-1})^2$

Latitude range (°N)		Covariance matrix		
		B12	B34	A1
12 - 40	B12	6.03	3.75	0.79
	B34	3.75	4.91	1.87
	A1	0.79	1.87	0.92
36 - 60	B12	14.25	10.91	4.03
	B34	10.91	11.79	5.65
	A1	4.03	5.65	3.54
56 - 80	B12	123.97	66.50	26.09
	B34	66.50	68.91	47.68
	A1	26.09	47.68	43.47

TABLE 3.5

VARIANCE OF PLANCK FUNCTION USED TO CONSTRUCT THE COVARIANCE MATRICES USED FOR THE RETRIEVAL OF THE ZONAL MEAN PLANCK FUNCTION PROFILE IN LATITUDE BANDS 12-40, 36-60 AND 56-80°N.

Units: $(mWm^{-2}sr^{-1}(cm^{-1})^{-1})^2$

12 - 40°N	36 - 60°N	56 - 80°N
5.8	16.2	119.9

The method of estimating instrumental error covariance used for wave components must be modified for zonal mean radiance on account of variance arising from seasonal trends in temperature. By reducing this to a minimum, the smallest monthly variance was taken to be the instrumental random noise variance. The effects of seasonal trends were minimised by the following technique:

Let the equatorial zonal mean radiance for day n of a month be

$$r_n = G_n + (an + b)$$

where G_n is the total random noise due to instrumental error and atmospheric fluctuations, a is the slope, or daily increment, of the seasonal trend, and b is an initial radiance value for the month. By taking G to be random we have made the assumption that atmospheric fluctuations with a time scale of more than one day, which would give rise to correlations between the daily values of G_n, are negligible. The monthly mean square difference between radiances on consecutive days is

$$\overline{(r_n - r_{n+1})^2} = \overline{((G_n + an + b) - (G_{n+1} + a(n+1) + b))^2}$$

$$= \overline{(G_n - G_{n+1} - a)^2}$$

Since G_n is random we may write

$$\overline{(r_n - r_{n+1})^2} = 2\overline{G^2} + \overline{a^2}$$

where $\overline{G^2}$ is the variance of the total random noise. Generally $\overline{a^2} \ll 2\overline{G^2}$ and may be ignored, so the random noise variance may be approximated by

$$\frac{1}{2N} \sum_{n=1}^{N} (r_n - r_{n+1})^2$$

where N is the number of days in the averaging period. Note that, as in the case of the equatorial variance of Fourier components, variance due to random atmospheric fluctuations is still present. Monthly values of the random noise, in terms of the standard

deviation, are given in Table 3.6; the values underlined, being the smallest, were taken as the standard deviations due to instrumental error.

It was noted in Chapter 2 that comparisons between radiance measurements and rocket-sonde observations have shown good agreement (Barnett et al, 1975). However in the first two months after the launch of Nimbus 5, during which the 1973 warming occurred, consistent differences were observed between the B channel (grid-point) radiances and radiances implied by coincident rocket-sonde profiles. Figure 3.8 shows these differences for channels B12 and B34. The mean differences over the part of this period (11.1.73 - 14.2.73) during which retrievals were performed were 4.34 and 1.81 radiance units for channels B12 and B34 respectively, and were found to be statistically significant. Channel A1 radiances were not significantly different from rocket-sonde measurements. As a result all zonal mean radiances in channels B12 and B34 were increased by 4.34 and 1.81 radiance units respectively. A systematic error of this sort does not affect the Fourier sine and cosine coefficients of radiance.

3.5. COMPARISONS BETWEEN RETRIEVALS AND CONVENTIONAL MEASUREMENTS

Profiles of the Fourier sine and cosine coefficients and the zonal mean of the Planck function were retrieved at every 4^{o} of latitude between 12^{o} and 80^{o}N for the period 11.1.73 - 14.2.73. At those latitudes spanned by two coefficient covariance matrices, i.e. 36^{o}, 40^{o}, 56^{o} and 60^{o}N, the profiles were retrieved using each matrix in turn. An average profile was then calculated for each of these latitudes so that the transition from one latitude band to another would be smooth. Levels in the vertical were spaced at 0.2 scale heights. Profiles of temperature were derived as follows.

TABLE 3.6

STANDARD DEVIATION OF THE RANDOM NOISE IN EQUATORIAL ZONAL MEAN RADIANCE OF CHANNELS B12, B34 AND A1 FOR EACH MONTH OF 1973 AND 1974. THE SMALLEST VALUES FOR EACH CHANNEL (UNDERLINED) WERE TAKEN AS THE STANDARD DEVIATIONS OF INSTRUMENTAL RANDOM NOISE.

Units: $mWm^{-2}sr^{-1}(cm^{-1})^{-1}$

	1973			1974		
Month	B12	B34	A1	B12	B34	A1
Jan.	0.38	0.22	0.12	0.24	0.20	0.13
Feb.	0.33	0.20	0.13	0.22	0.22	0.23
Mar.	0.23	0.13	0.10	0.32	0.18	0.17
Apr.	0.20	0.09	0.15	0.20	0.17	0.15
May	0.21	0.13	0.11	0.20	0.13	0.13
June	0.19	0.15	0.19	0.18	0.13	0.14
July	0.26	0.17	0.16	0.32	0.22	0.11
Aug.	0.27	0.18	0.17	0.20	0.19	0.10
Sept.	0.20	0.17	0.16	0.21	0.16	0.09
Oct.	0.21	0.16	0.13	0.21	0.12	0.14
Nov.	0.21	0.12	0.10	0.20	0.16	0.22
Dec.	0.25	0.24	0.20	0.33	0.24	0.10

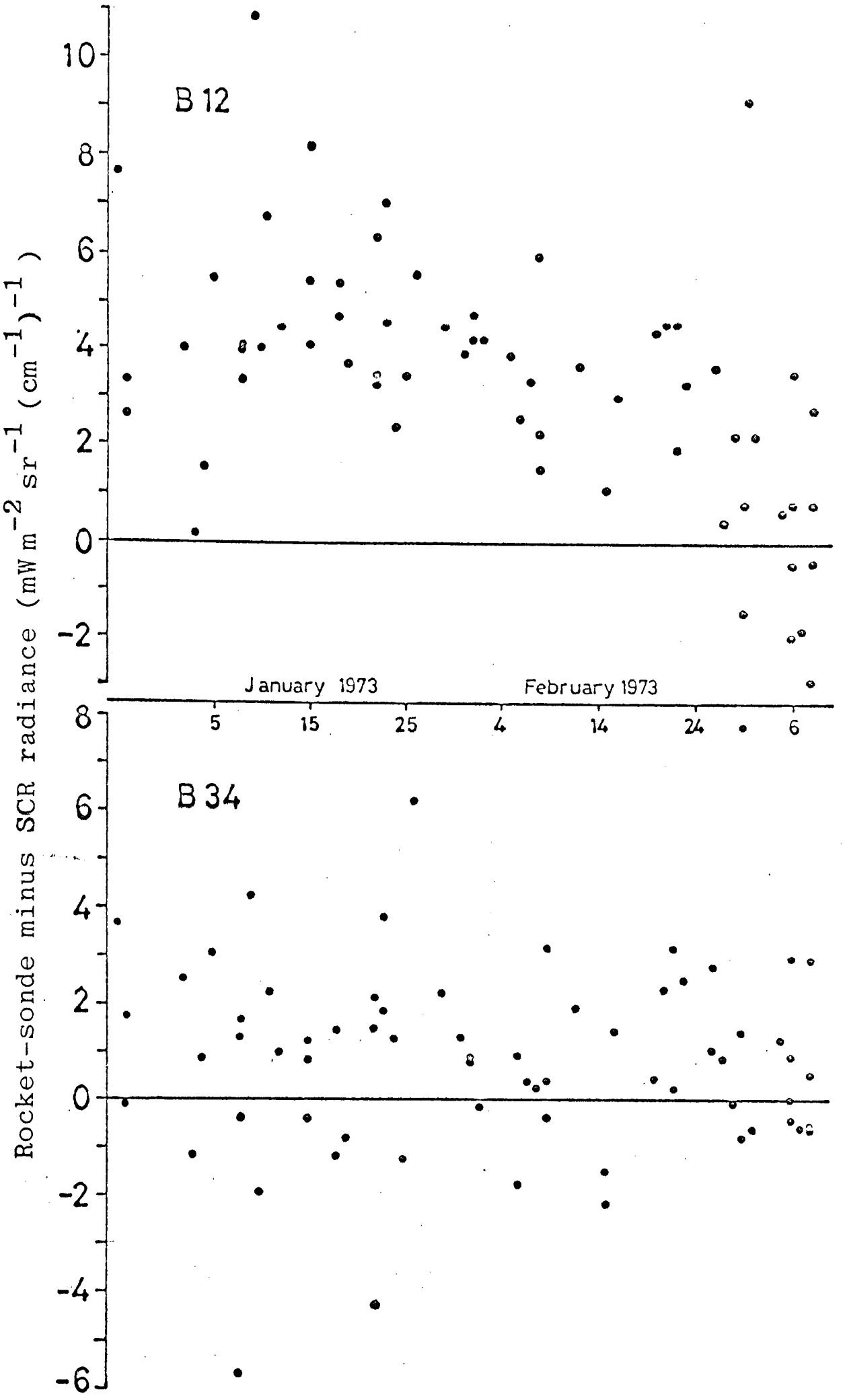


Figure 3.8. Differences between radiances implied by rocket-sonde measurements and radiances observed in channels B12 and B34 at nearest grid point. The rocket-sonde data were part of the data set used by Barnett et al (1975)

The Planck function intensity, B, per unit solid angle is related to the temperature, T, by

$$B(\bar{\nu},T) = (c_1\bar{\nu}^3)/(\exp(c_2\bar{\nu}/T) - 1))$$

where $c_1 = 1.1906 \times 10^{-8}$ $Wm^{-2}sr^{-1}(cm^{-1})^{-4}$, $c_2 = 1.4388$ cmK and $\bar{\nu}$ is the mean wavenumber for the spectral range of observation, 668 cm^{-1} for the ν_2 CO_2 band. The zonal mean Planck function profile was retrieved first and converted into temperature, $\bar{T}$, using this relationship. Temperature sine and cosine coefficients were derived from Planck function coefficients using the relation

$$\delta T = \delta B\,((dB/dT)_{\bar{T}})^{-1}$$

Profiles of harmonic components of geopotential height were obtained by adding 50 mb geopotential heights supplied by the Free University of Berlin to thickness, ϕ, evaluated by vertical integration of the retrieved temperature components:

$$\phi_b^{z'} = \frac{-R}{g}\int_b^{z'} T(z)\,dz$$

where R is the gas constant, g the acceleration due to gravity and $\phi_b^{z'}$ is the thickness between the 50 mb base level and level z'.

Comparisons between rocket-sonde profiles and temperature profiles reconstituted from the harmonic components at the nearest grid-points are shown in Figure 3.9 for (a) an almost isothermal profile, (b) a typical low-latitude winter profile and (c) a very unusual profile exhibiting very sharp vertical temperature gradients. Of ten such comparisons, (a) and (b) are a little better than most and (c) shows the worst agreement obtained. The difference between the two profiles in (c) is very evidently a result of the inherent vertical averaging in the radiance observations, the weighting functions peaking at about 6.5, 4.9 and 3.2 scale heights. Information about components of

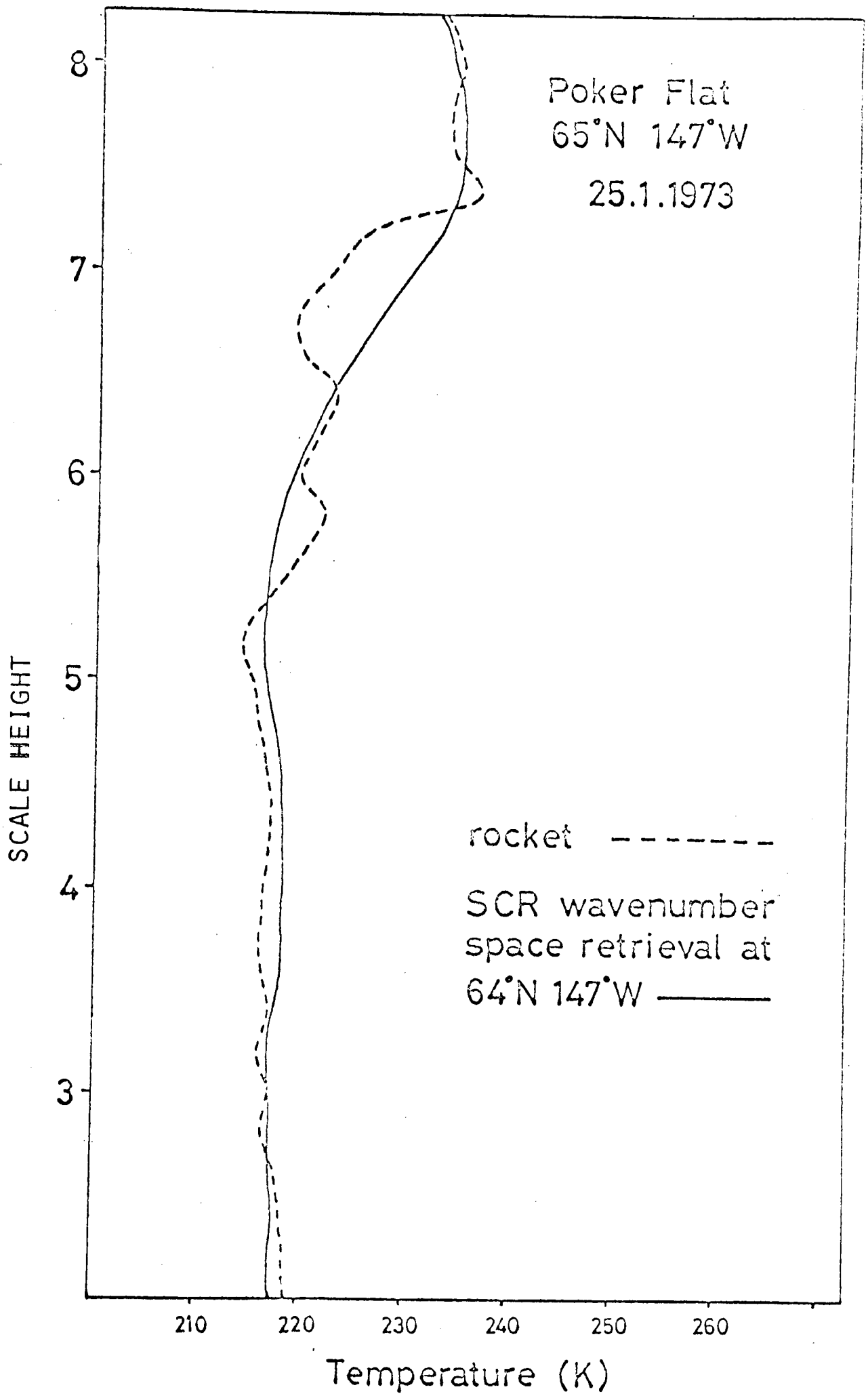


Figure 3.9a. Comparison between rocket-sonde profile and the profile derived from Fourier component retrieval for the nearest grid point for the date and location shown above

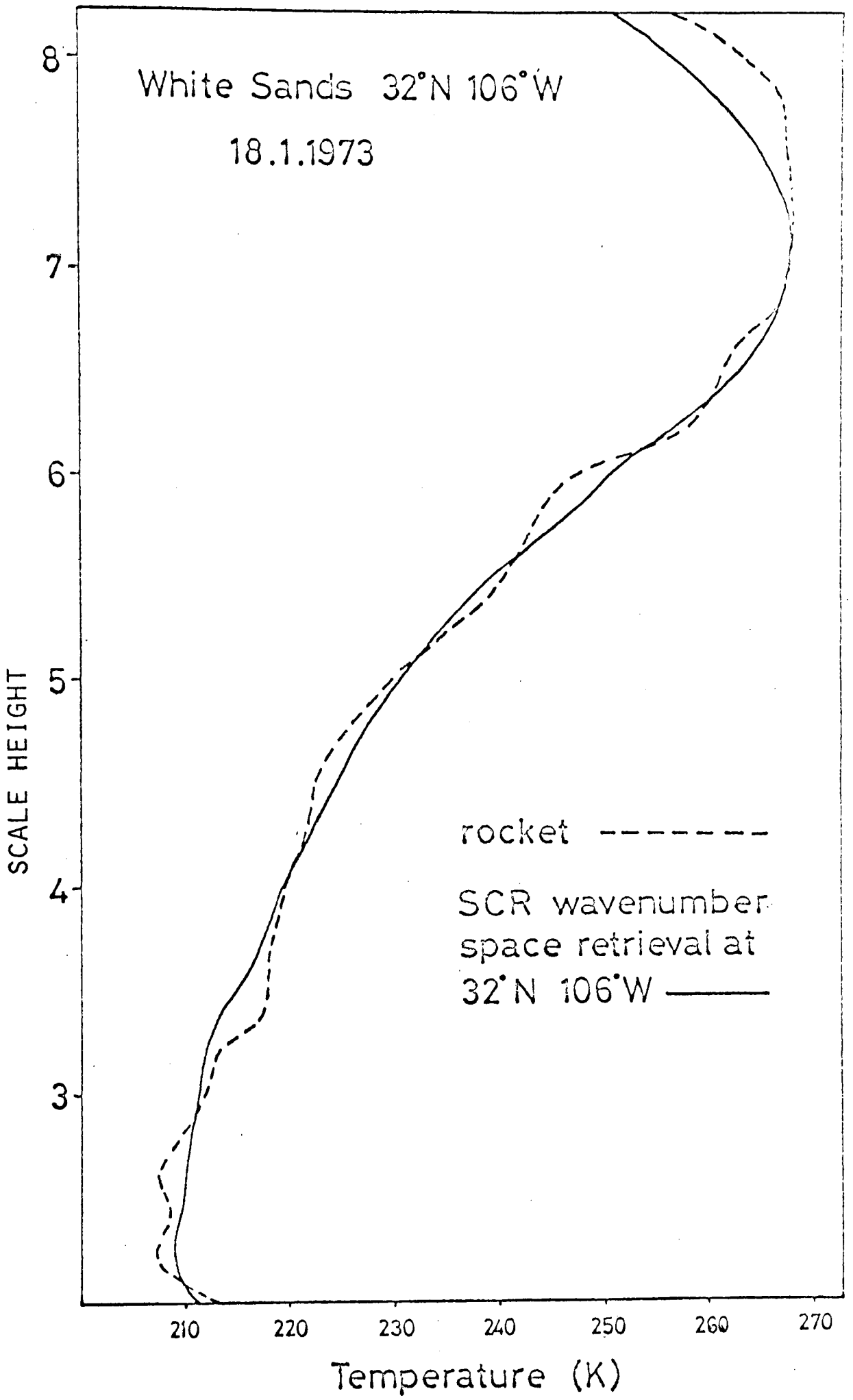


Figure 3.9b. Caption as for Figure 3.9a

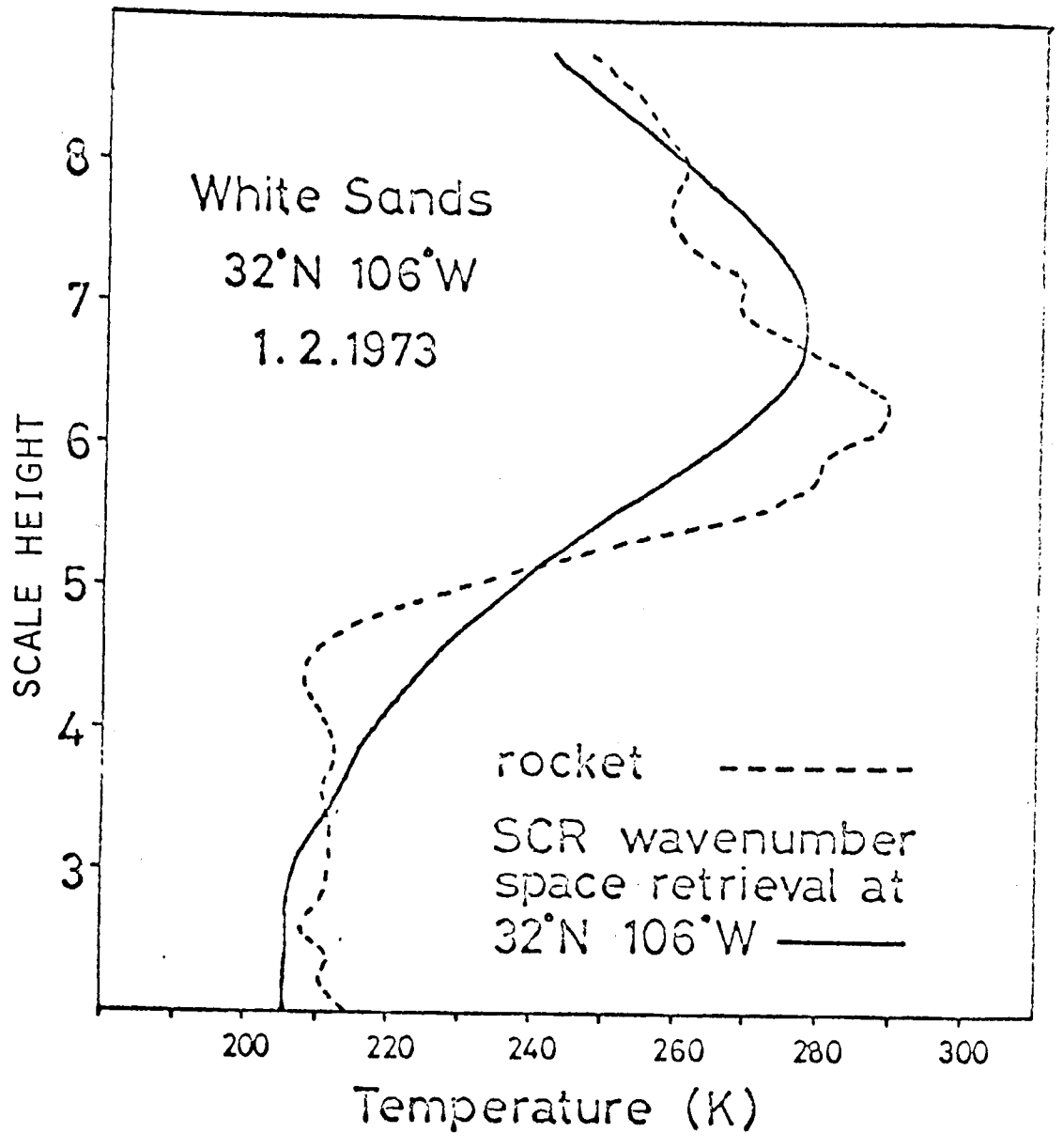


Figure 3.9c. Caption as for Figure 3.9a

the profile which are orthogonal to the weighting functions can only be provided by the a priori information which will often be inadequate for very unusual profiles such as shown in (c). Figure 3.10 compares the 50 mb retrieved temperature field with the analysis from the Free University of Berlin for three days during the sudden warming. The wave components have been well retrieved although the zonal mean temperature is rather low on 25.1.73 and 28.1.73. Retrieved temperatures and the Berlin analyses for 50 mb are further compared in Figure 3.11. Time-longitude sections of the departure of temperature from the zonal mean at 60^{o}N for the period 11.1.73 - 10.2.73 are presented. Agreement is generally good although a warm region of small longitudinal extent near 140^{o}E late in the period is not represented in the retrieval. Since only the first four zonal wavenumbers were retrieved features with a longitudinal scale of less than 45^{o} will not be resolved.

Finally we compare our results with circumpolar temperature and geopotential height analyses produced by NOAA (1975) in Figures 3.12 and 3.13. The NOAA analyses were constructed from both rocket-sonde measurements and satellite radiance observations made by the Nimbus 5 SCR and the NOAA 2 VTPR. The use in part of the same data makes this more an assessment of the retrieval alone than of both the retrieval method and the satellite observations. Apart from over the North Atlantic, Western Canada and the Baltic, the temperature fields again show very good agreement. Although the retrieved heights are within 1% of the NOAA analysis for much of the hemisphere, the small errors are sufficient to give a noticeable difference in the strength of the polar vortex since there is a change in height of less than 9% between the extreme height contours on this pressure surface. In a similar comparison

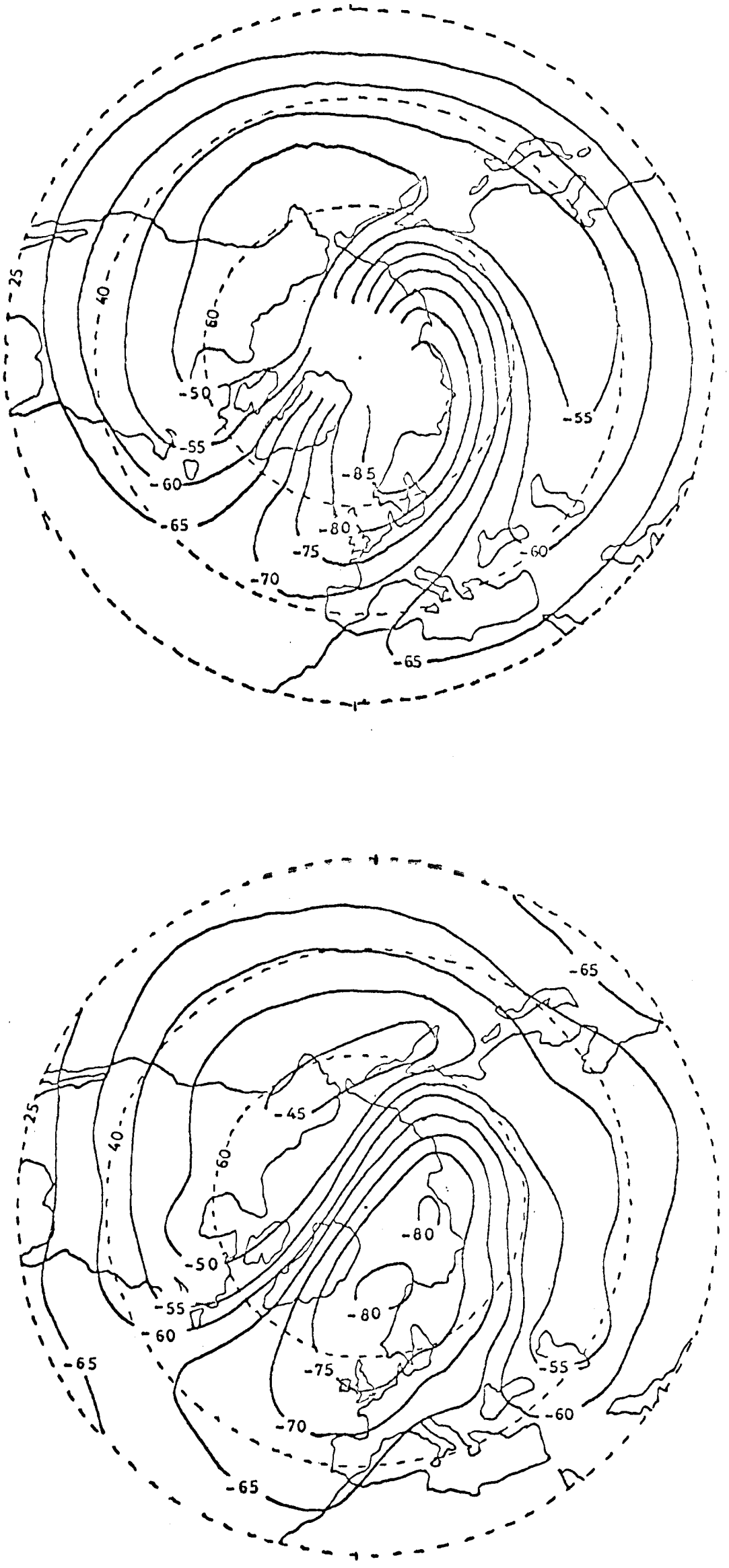
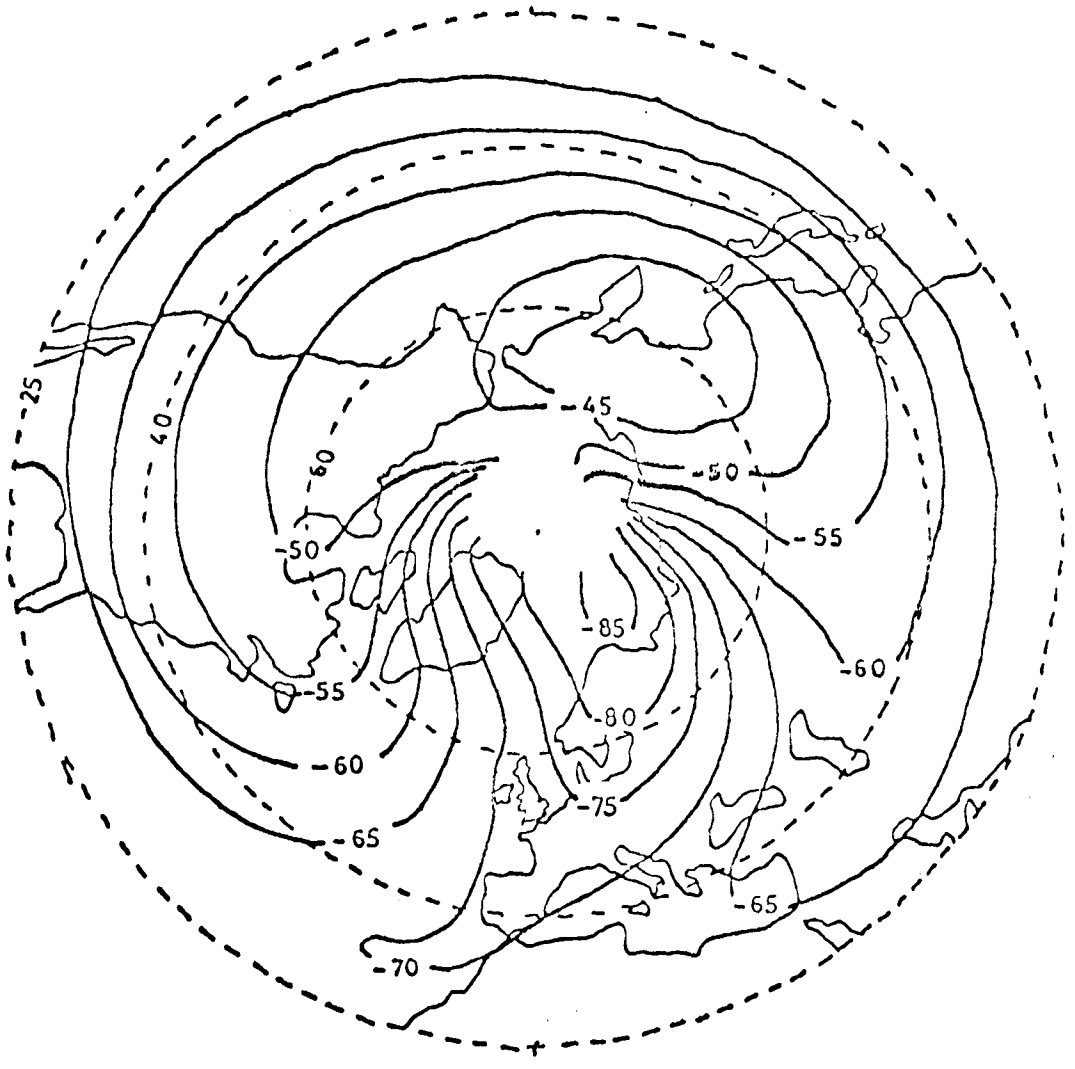


Figure 3.10a. Comparison between the 50 mb temperature fields given by Fourier component retrieval (upper) and the analysis of the Free University of Berlin (lower) for 25.1.73 (^{o}C)



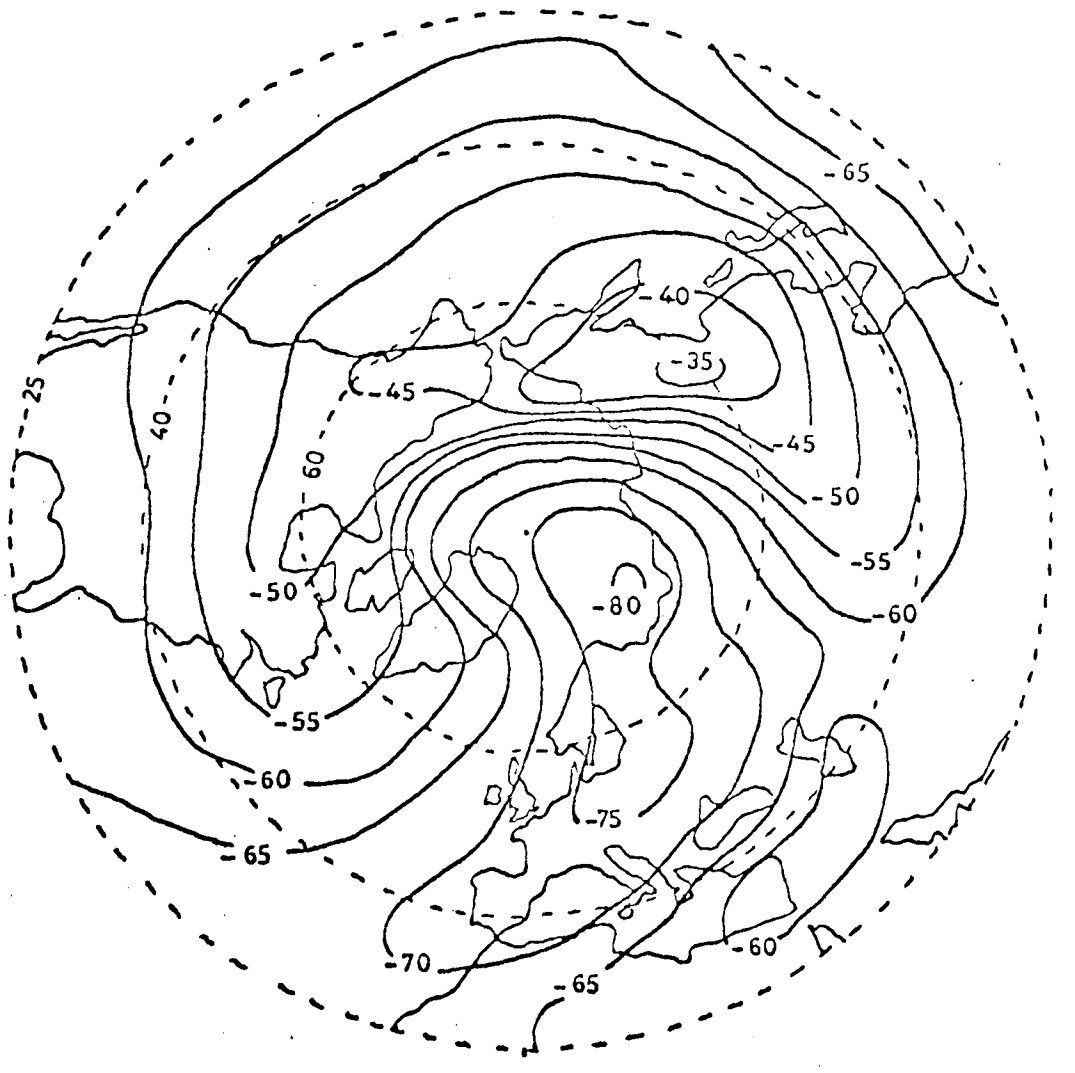
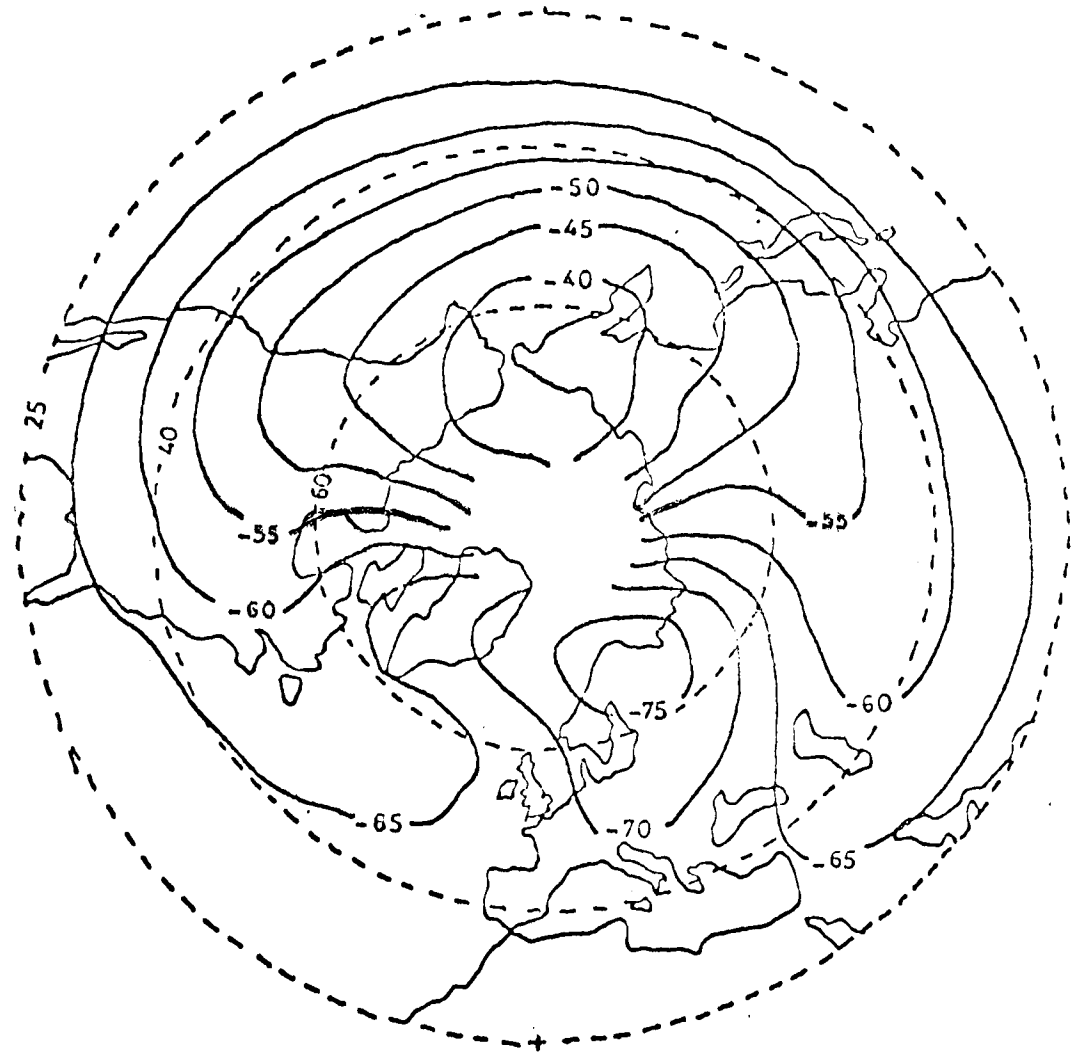


Figure 3.10b. As Figure 3.10a but for 28.1.73



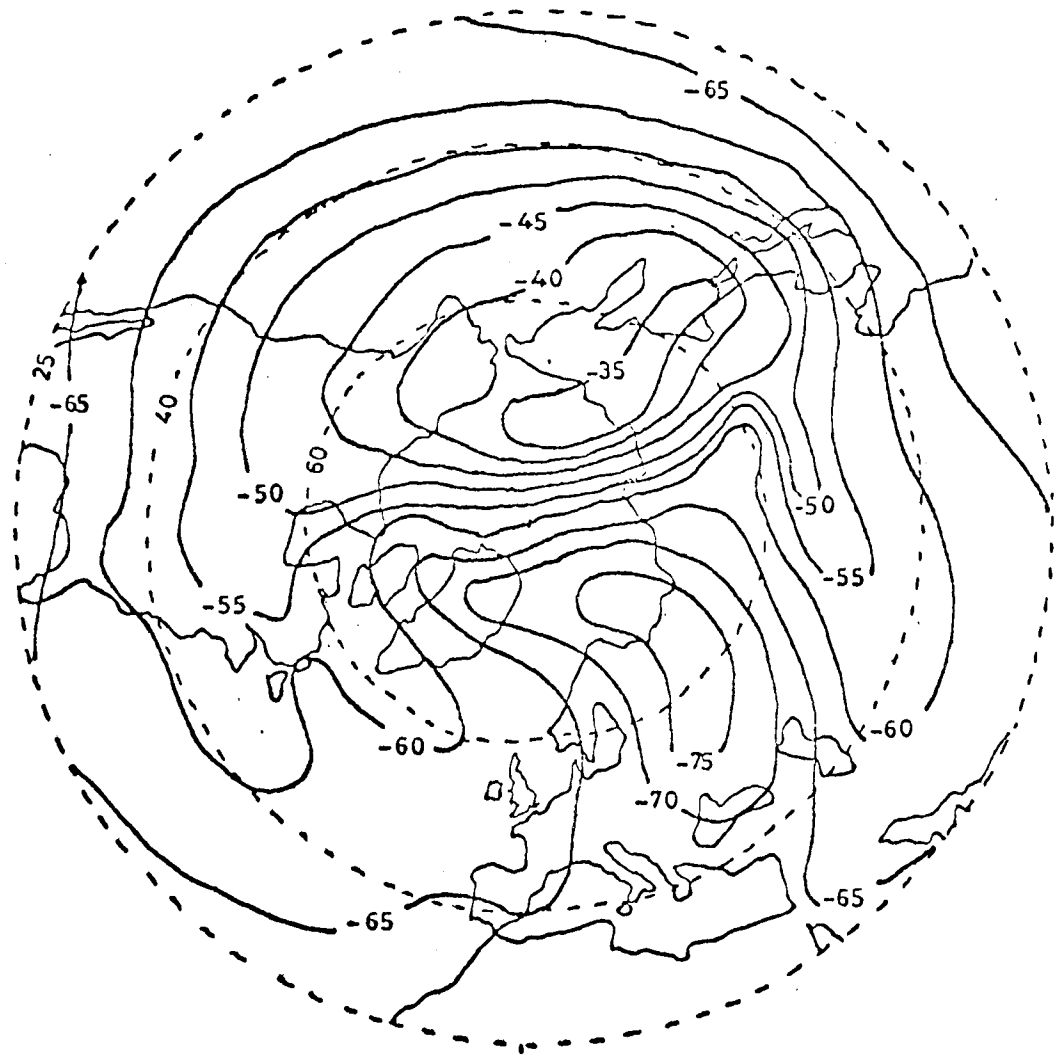


Figure 3.10c. As Figure 3.10a but for 30.1.73

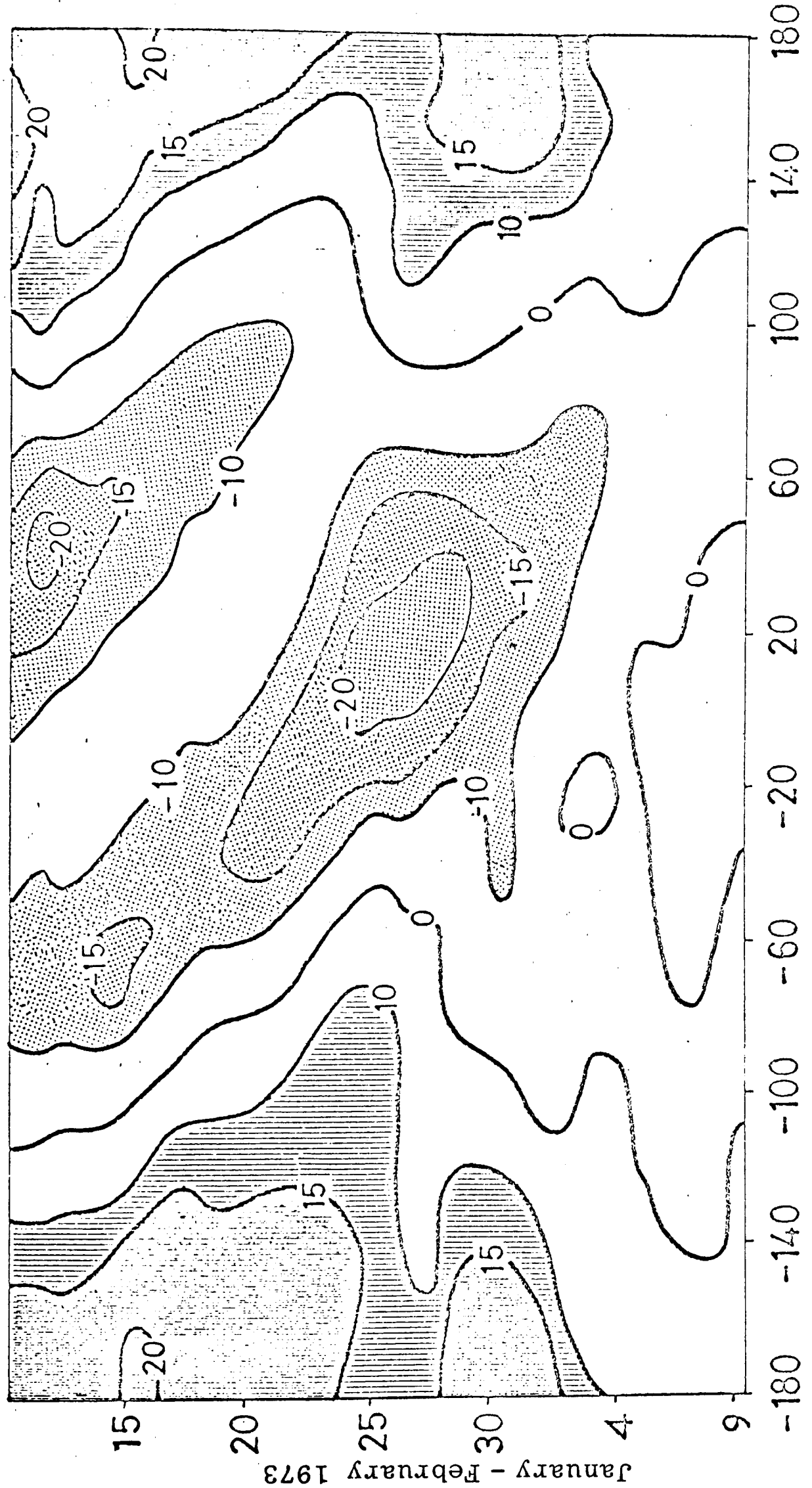


Figure 3.11a. Longitude-time section of the departure of temperature from the zonal mean around 60°N at 50 mb derived from Fourier component retrieval for the period 11.1.73 to 10.2.73 (K)

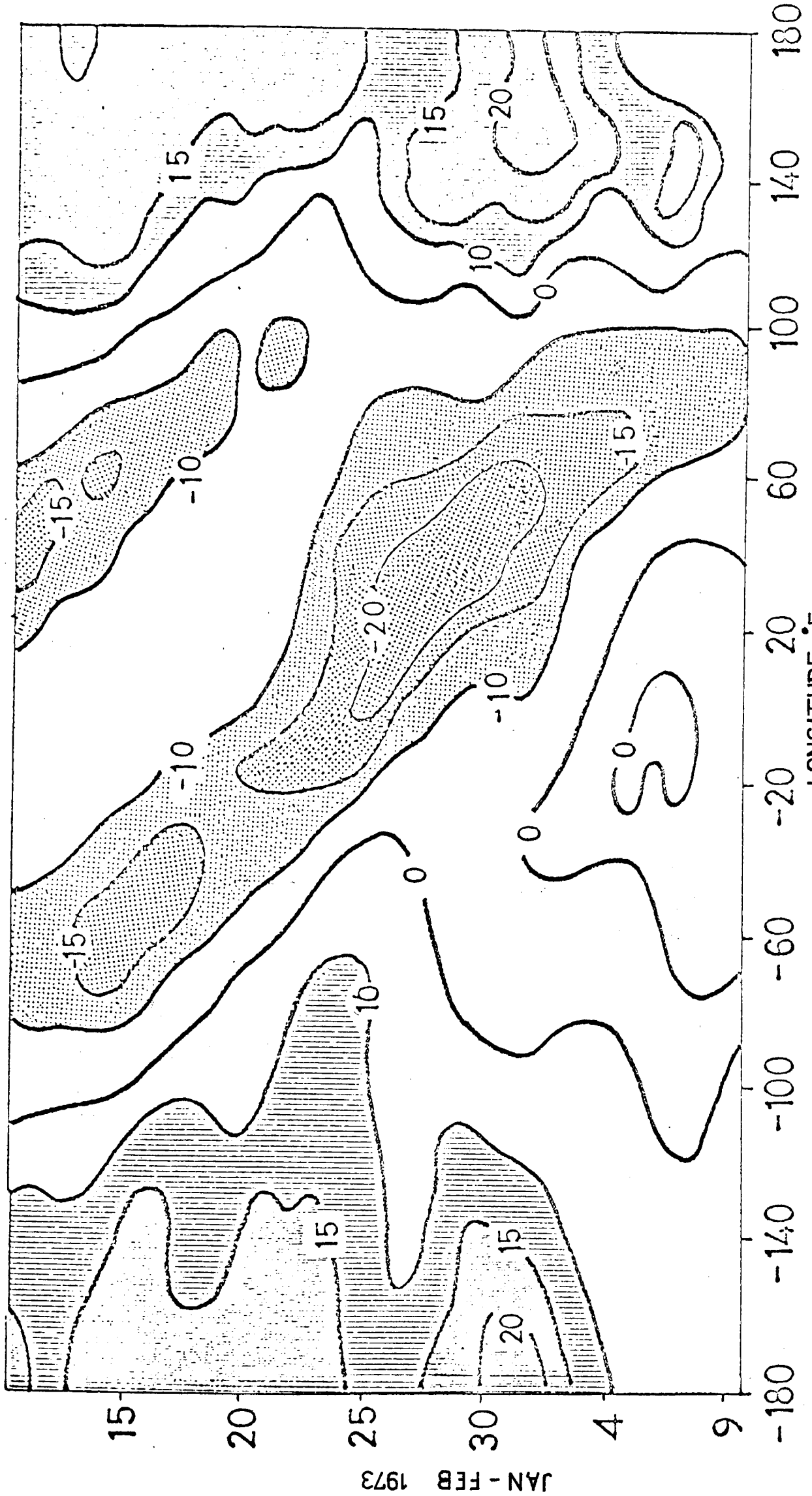


Figure 3.11b. As Figure 3.11a but derived from the 50 mb temperature analysis of the Free University of Berlin

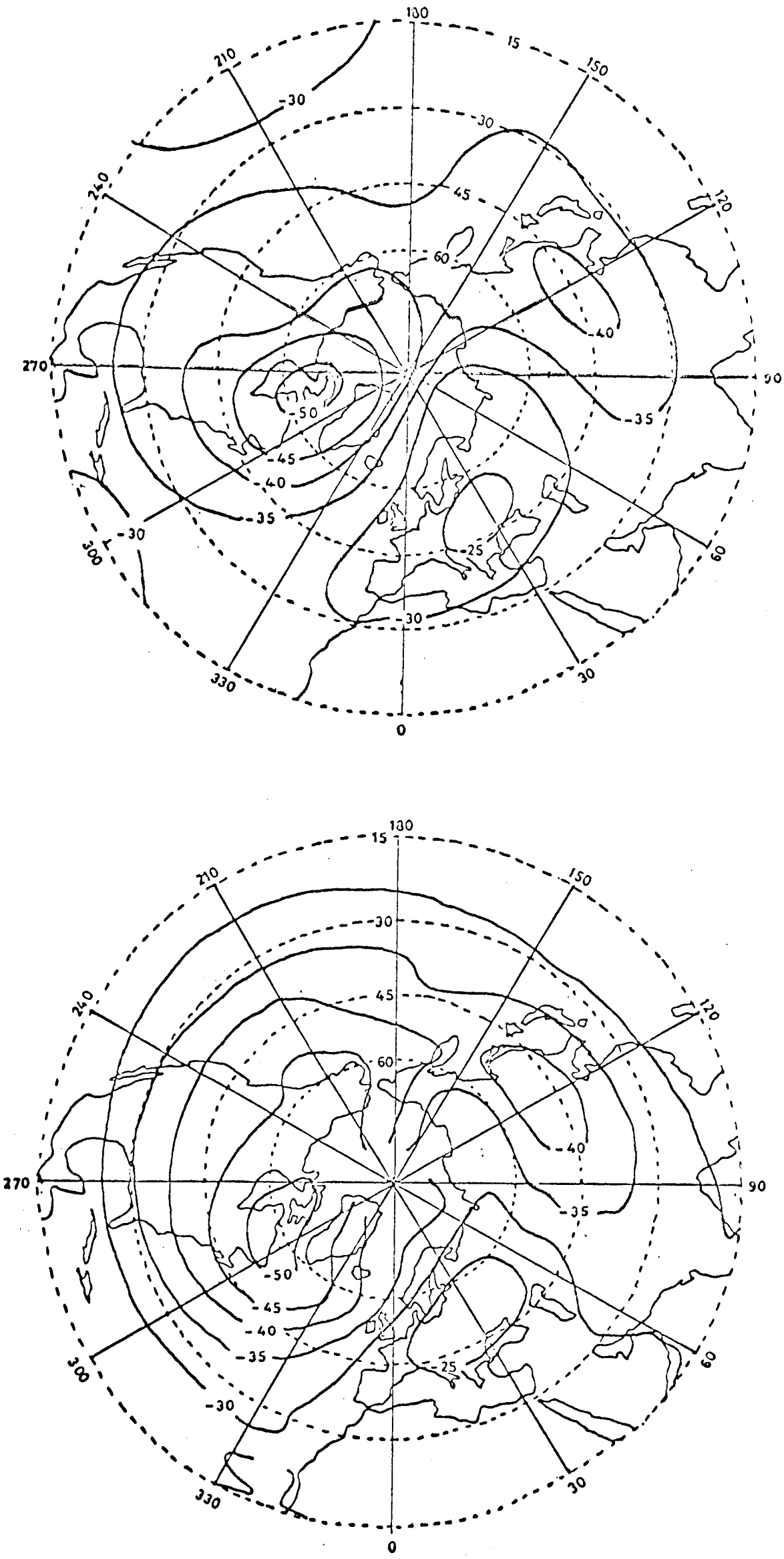
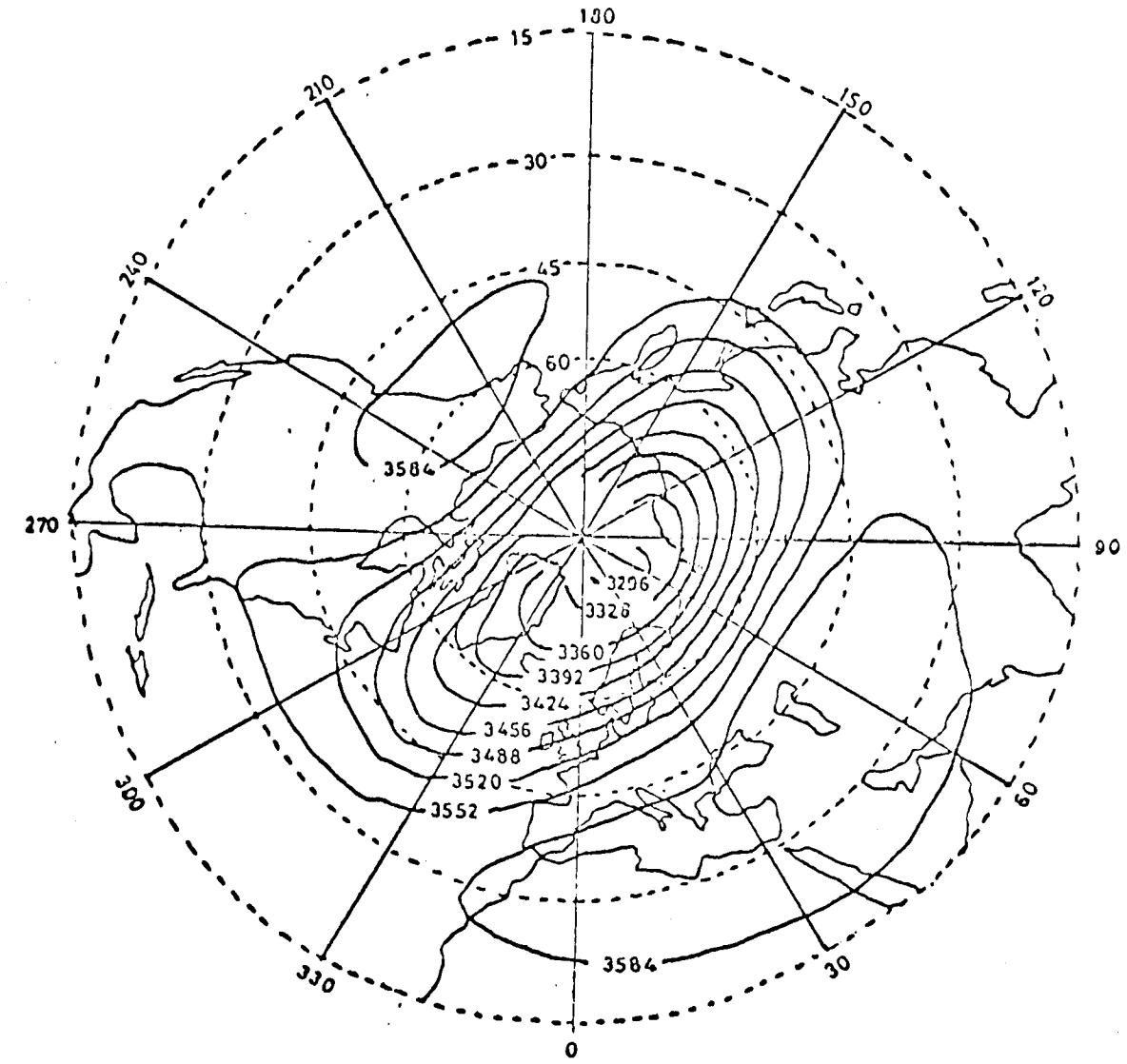


Figure 3.12. Comparison between the 5 mb temperature field derived from Fourier component retrieval and that produced by NOAA (1975) for 24.1.73 (^{o}C)



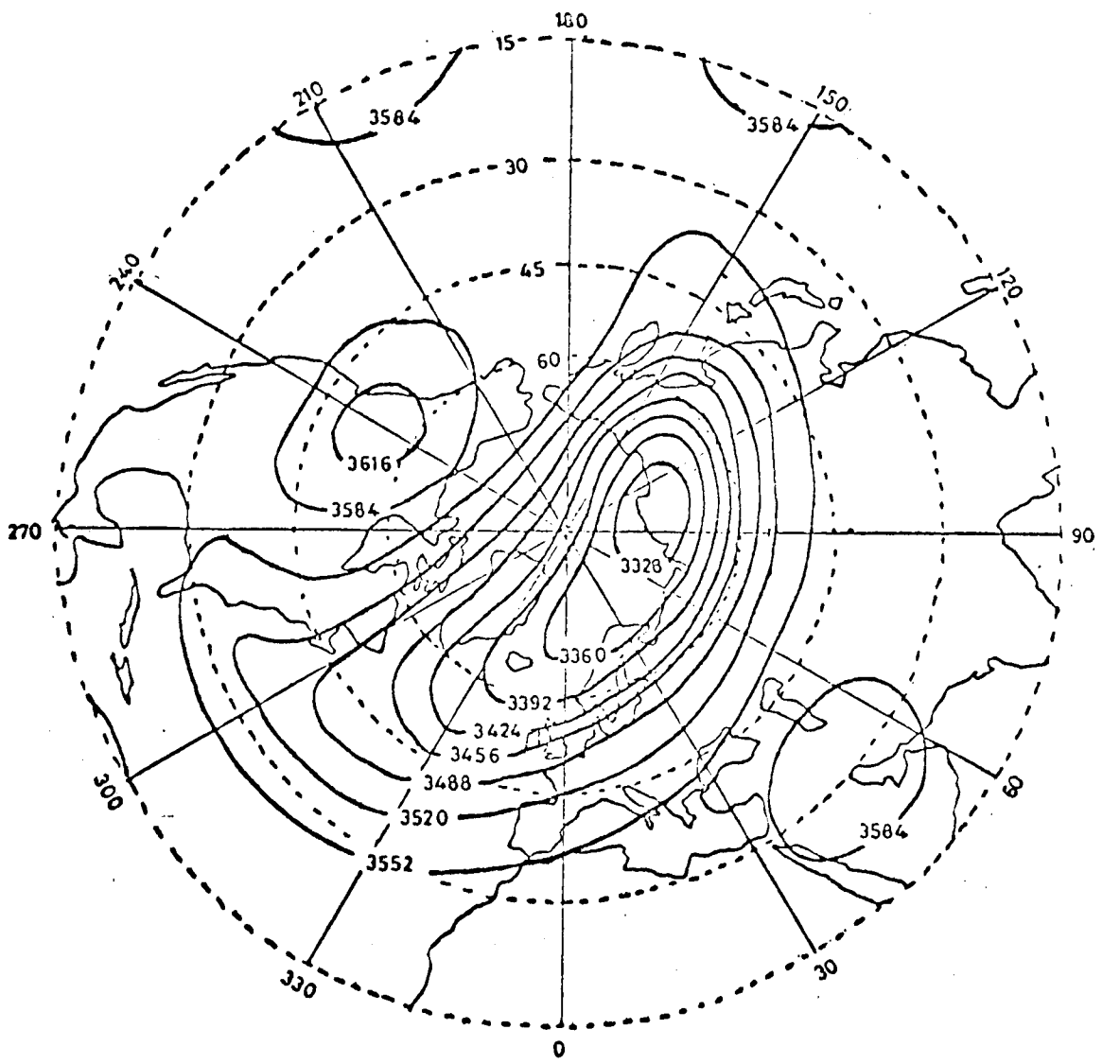


Figure 3.13. Comparison between the 5 mb geopotential height field derived from Fourier component retrieval and that produced by NOAA (1975) for 24.1.73 (Dm)

Barnett (1973) showed a tendency for the retrieved 10 mb geopotential heights to be slightly underestimated in high latitudes and slightly overestimated in low latitudes compared with conventional analyses, the same tendency as shown in Figure 3.13. In addition he showed that the differences between different analyses of approximately the same set of conventional data were as large as those between the retrieved and conventional analyses.

3.6. SUMMARY AND CONCLUSIONS

The effect of the constraints applied in any retrieval method is to bias the solution profile towards a prejudiced idea of what the profile should be. If the actual profile is very different from this the retrieval will generally be rather poor. The success of a retrieval method is thus dependent on how well the actual profiles are compatible with the a priori information used. The chance of success is improved by careful choice of constraints but any method will fail to some extent in dealing with profiles which are unusual relative to the constraints. It must be remembered however that in the absence of constraints the radiance measurements tell us absolutely nothing about the temperature at any given level.

We have described an attempt to develop simple constraints of a statistical nature, based on observations of stratospheric waves, to be compatible with most of the profiles to be retrieved. To this end some account was taken of latitudinal and wavenumber variations, although a more detailed study may have enabled more realistic differentiation in this respect. The comparisons with conventional measurement show that a measure of success has been achieved, but the very anomalous temperature structure of parts of the atmosphere during the sudden warming, as indicated, for example, by Figure 3.9(c), was not compatible with the a priori

information and consequently not well retrieved. In view of the very large temperature variations in space and time it was inevitable that the relatively simple constraints used would to some extent fall short in this respect. It would appear that constraints for sudden warming retrievals will only be substantially improved if regional climatologies for the various stages of these events can be constructed.

Introduction

We present in this chapter the mathematical basis of our sudden warming study. Equations describing the various energy conversion processes are developed. The concept of Available Potential Energy and the resolution of the motion and temperature fields into zonal mean and eddy components leads to a modified set of energy equations. Their validity for limited regions of the atmosphere, such as the near-hemispheric slabs used in this investigation, is discussed. Further modifications are presented in order that local energy conversions and latitudinal distributions of these processes may be described.

4.1. ENERGY CONVERSIONS IN THE ATMOSPHERE

The equation of atmospheric motion in x,y,z,t coordinates may be written, neglecting friction, as

$$\frac{d\underline{V}}{dt} + 2\underline{\Omega}\wedge\underline{V} = -\frac{1}{\rho}\underline{\nabla}p - \underline{\nabla}\Phi \qquad 4.1$$

where $\underline{V}$ (u,v,w) is the velocity, Ω the Earth's angular velocity, ρ the density, p the pressure and Φ the geopotential given by

$$\Phi = \int g\, dz$$

where z is height and g the acceleration due to gravity. $\underline{\nabla}$ is the three-dimensional gradient operator. Multiplication by $\underline{V}$ gives

$$\rho\frac{d}{dt}(K + \Phi) = -\underline{V}\cdot\underline{\nabla}p \qquad 4.2$$

where $K = \frac{1}{2}\underline{V}.\underline{V}$ and is the kinetic energy per unit mass.

If we consider a unit volume of air the work done by the surrounding air on the volume is $-\underline{\nabla}.p\underline{V}$ which may be expanded to give

$$-\underline{\nabla}\cdot p\underline{V} = -\underline{V}\cdot\underline{\nabla}p - p\underline{\nabla}\cdot\underline{V} \qquad 4.3$$

From equation 4.2 we see that part of the energy gain due to this work appears as kinetic energy, namely the work performed by the pressure force $-\underline{\nabla}p$. The remainder is work done on the volume by expansion or contraction. That this is conversion into internal energy of the volume may be shown as follows. Using equation 4.3 in 4.2 we have for the unit volume

$$\rho\frac{d}{dt}(K + \Phi) + \underline{\nabla}\cdot p\underline{V} = p\underline{\nabla}\cdot\underline{V} \qquad 4.4$$

The left hand side is the difference between the rate of change of mechanical energy and the work done on the volume which, by the first law of thermodynamics, is the difference between the rate of external heating, Q, and the rate of change of the internal energy, I. Thus the internal energy-balance equation is

$$\rho\underline{\nabla}\cdot\underline{V} = \rho Q - \rho\frac{dI}{dt} \qquad 4.5$$

The balance equation of gravitational potential energy for the unit volume is

$$\rho\frac{d\Phi}{dt} = \rho g\frac{dz}{dt} = \rho g w \qquad 4.6$$

which indicates that an increase in potential energy occurs when the state becomes more unstable through correlations of ρ and w. Subtracting equation 4.6 from 4.4 yields the kinetic energy balance equation

$$\rho\frac{dK}{dt} + \underline{\nabla}\cdot p\underline{V} = p\underline{\nabla}\cdot\underline{V} - g\rho w \qquad 4.7$$

Finally writing equations 4.5, 4.6, 4.7 in flux form we obtain

$$\frac{\partial}{\partial t}(\rho I) + \underline{\nabla}\cdot(\rho I\underline{V} + \underline{H}) = -p\underline{\nabla}\cdot\underline{V} \qquad 4.5'$$

$$\frac{\partial}{\partial t}(\rho\Phi) + \underline{\nabla}\cdot(\rho\Phi\underline{V}) = g\rho w \qquad 4.6'$$

$$\frac{\partial}{\partial t}(\rho K) + \underline{\nabla}\cdot(\rho K\underline{V} + p\underline{V}) = -g\rho w + p\underline{\nabla}\cdot\underline{V} \qquad 4.7'$$

where we have written $Q = -\underline{\nabla}.\underline{H}$, $\underline{H}$ being the heat flux across the unit volume boundary. Comparing the right hand sides of 4.5', 4.6' and 4.7' we see that each term appears twice with different sign. Since each represents a well defined process we may regard them as conversion rates between energy types. The left hand sides contain the local rates of change of the energy types and their flux divergences. In addition there is change of internal energy due to heat flux divergence $\underline{\nabla}.\underline{H}$ and change in kinetic energy due to work performed at the volume boundary $\underline{\nabla}.p\underline{V}$. There is no conversion between internal and potential energy. Furthermore the conversions $-p\underline{\nabla}.\underline{V}$ between internal and kinetic energy and $-g\rho w$ between potential and kinetic energy occur simultaneously and, for quasi-static motion, are proportional to one another*. Thus only the net result affects the kinetic energy and it is therefore convenient to combine internal and potential energy and refer to their sum as <u>total potential energy</u>.

4.2. UNAVAILABLE AND AVAILABLE POTENTIAL ENERGY

If we consider a column in the atmosphere of cross-sectional area δA it is easy to show that the potential energy of the column is

$$PE = \delta A \int_0^\infty \rho RT \, dz$$

where R is the gas constant and T is temperature.
The internal energy in the column is

* if integrated over an atmospheric column

$$IE = \delta A \int_0^\infty \rho c_v T \, dz$$

where c_v is the specific heat of air at constant volume. The total potential energy is therefore given by

$$TPE = \delta A \int_0^\infty \rho c_p T \, dz$$

where c_p is the specific heat of air at constant pressure since $R = c_p - c_v$. Rewriting the total potential energy in terms of the square of the speed of sound, given by

$$c^2 = \frac{c_p}{c_v} RT = \gamma RT$$

we obtain

$$TPE = \delta A \frac{c_p}{R\gamma} \int_0^\infty c^2 \rho \, dz$$

The kinetic energy of the column is

$$KE = \delta A \tfrac{1}{2} \int_0^\infty v^2 \rho \, dz$$

Therefore, representing an average over pressure by a bar, we have

$$\frac{KE}{TPE} \sim \frac{1}{2(\gamma-1)} \frac{\bar{v}^2}{\bar{c}^2} \sim \frac{1}{320}$$

if $\bar{v}$ is taken as $15 ms^{-1}$ and $\bar{c}$ as $300 ms^{-1}$. The kinetic energy, KE, is thus only a small fraction of the total potential energy, TPE, in the atmosphere. As far as energy conversions are concerned TPE is not a useful measure of the atmosphere's potential energy since its zero corresponds to the whole atmosphere being at OK, a state which is never approached. The zero of convertible potential energy occurs when the atmosphere is stably stratified with no horizontal density gradients. Although TPE is then large no conversion to KE can occur. Addition of heat in a manner which does not create density gradients increases TPE but does not release any energy for conversion. However

addition or subtraction of heat from a part of the atmosphere may create such gradients and make available energy for conversion. We may therefore regard the potential energy of the stably stratified state in which no horizontal density gradients exist as unavailable potential energy (UPE) and the difference between the actual TPE and UPE as the available potential energy (APE). APE is thus the amount of TPE that can be released from a given atmospheric state by bringing it adiabatically and reversibly to a horizontally uniform state. If only adiabatic, reversible processes take place then KE and APE suffer equal and opposite changes, maximum KE being attained for zero APE. In the real atmosphere non-adiabatic processes occur: increased production of APE by differential heating leads to larger pressure gradients and increased KE.

Lorenz (1955) introduced the concept of APE and showed that

$$\mathrm{APE} = \frac{c_p}{g p_0^{\kappa}(1+\kappa)} \int_S \int_{\theta=0}^{\infty} (p^{1+\kappa} - \bar{p}^{1+\kappa})\, d\theta ds \qquad 4.8$$

where p_0 is the surface pressure, $\bar{p}$ is the mean pressure on an isentropic surface ($p = \bar{p}$ everywhere in the horizontally uniform state), θ is potential temperature, $\kappa = R/c_p$ and S is the Earth's surface area. Since θ is not a convenient vertical coordinate Lorenz derived a useful approximation in which means are taken over isobaric surfaces:

$$\mathrm{APE} \simeq \tfrac{1}{2} \frac{c_p}{g} \int_S \int_{p=0}^{\infty} \bar{T} \left(1 - \frac{\Gamma}{\Gamma_d}\right)^{-1} \overline{\left(\frac{T''}{\bar{T}}\right)^2}\, dp dS \qquad 4.8'$$

where $\Gamma = -\partial T/\partial z$, $\Gamma_d = g/c_p$ (the dry adiabatic lapse rate) and T'' represents a departure from the isobaric mean temperature $\bar{T}$. To obtain equation 4.8' the integral in equation 4.8 is approximated by an expression involving the variance of pressure on isentropic surfaces, which itself may be approximated by the temperature variance on isobaric surfaces, provided Γ does not

approach Γ_d. Dutton and Johnson (1967) have shown that the isobaric approximation may yield higher values during winter than the exact expression, but is reasonable provided Γ is calculated as a function of space and time.

4.3. GENERATION AND CONVERSION OF AVAILABLE POTENTIAL ENERGY

It may be shown (e.g. Dutton and Johnson, 1967) that the proportion of diabatic heating, Q, which serves to increase APE is $1 - (\bar{p}/p)^\kappa$, the remainder increasing the unavailable potential energy. We may therefore write the generation of APE as

$$G = \frac{1}{g} \int_S \int_0^\infty Q(1 - (\bar{p}/p)^\kappa) \, d\theta dS \qquad 4.9$$

where the bar represents an isentropic average, which, by the same approximation as used to derive equation 4.8', gives

$$G \simeq \frac{1}{g} \int_S \int_0^\infty \frac{1}{\bar{T}} \left(1 - \frac{\Gamma}{\Gamma_d}\right)^{-1} Q''T'' \, dpdS \qquad 4.9'$$

where the reference surfaces are now isobaric.

Newell et al (1970) were able to calculate G (in fact G_z - see equation 4.19) for the troposphere using both the exact and the approximate equations 4.9 and 4.9'. They found large differences, presumably on account of the large relative slopes which may occur between isentropic and isobaric surfaces. Our results suggest that the approximation may be better in the stratosphere (see Chapter 7).

Because of its ubiquitous nature nearly all frictional heating is a source of UPE. The generation G is thus accomplished almost completely by differential radiative heating and cooling.

The conversion $g\rho w - p\underline{\nabla}\cdot\underline{V}$ per unit volume between KE and TPE (equations 4.5' - 4.7') may be integrated over the whole atmosphere to give

$$\int_\tau (g\rho w - p\underline{\nabla}\cdot\underline{V}) \, d\tau = \frac{1}{g} \int_S \int_0^\infty \alpha\omega \, dpdS \qquad 4.10$$

for the KE to APE conversion, where $\alpha=\rho^{-1}$ and $\omega=dp/dt$, the vertical 'pressure velocity'.

Using these relations we can now write down the balance equations for the APE and KE of the atmosphere:

$$\frac{\partial}{\partial t}APE = \frac{1}{g}\int_S\int_0^\infty \alpha\,\omega\,dpdS + \frac{1}{g}\int_S\int_0^\infty \frac{1}{\bar{T}}\left(1-\frac{\Gamma}{\Gamma_d}\right)^{-1} Q''T''\,dpdS \qquad 4.11$$

$$\frac{\partial}{\partial t}KE = \frac{1}{g}\int_S\int_0^\infty -\alpha\,\omega\,dpdS - F \qquad 4.12$$

The atmospheric energy cycle thus consists of three phases: generation of APE, conversion of APE to KE and dissipation of KE by friction, F, which over the long term must balance the other two processes. Since most of the frictional loss occurs at the surface it is unlikely to be important in the stratospheric energy budget. Moreover it would be difficult to estimate and has thus been neglected. This point is discussed further in Chapter 5 in connection with the calculation of the mean meridional circulation.

4.4. ZONAL MEAN AND EDDY COMPONENTS OF THE ENERGY EQUATIONS

We saw in Chapter 1 that in middle and high latitudes the large scale waves, or eddies, play a dominant role in the atmospheric circulation by transporting heat and momentum, while in low latitudes the zonal mean meridional circulation is the more important mechanism for such transports. To examine these separate roles it is convenient to separate the motion and temperature fields into zonal mean components and eddy components denoted by subscripts Z and E respectively. The following averaging notation will be used:

zonal mean : $[x] = \frac{1}{2\pi}\int_0^{2\pi} x\,d\lambda$, where λ is longitude;

eddy component : $x' = x - [x]$;

meridional mean : $\{x\} = (\sin\phi_1 - \sin\phi_2)^{-1} \int_{\phi_1}^{\phi_2} x \cos\phi \, d\phi$

where ϕ_1 and ϕ_2 are the bounding latitudes;

spatial mean : $\bar{x} = \{[x]\}$;

departure of zonal mean from spatial mean:

$$x^* = [x] - \bar{x};$$

departure of point value from spatial mean:

$$x'' = x - \bar{x}.$$

Parentheses ( ) will be used only for grouping terms. Using this notation we may write down the set of equations giving the average energy amounts and conversions for an atmospheric slab between pressure levels p_1 and p_2 separated by Δp mb and latitudes ϕ_1 and ϕ_2. APE and KE are written as A and K.

$$K_Z = \frac{1}{\Delta p} \int_{p_1}^{p_2} \tfrac{1}{2}\{[u]^2 + [v]^2\} \frac{dp}{g} \simeq \frac{1}{\Delta p} \int_{p_1}^{p_2} \tfrac{1}{2}\{[u]^2\} \frac{dp}{g} \qquad 4.13$$

since $[v]^2 \ll [u]^2$;

$$A_Z = \frac{1}{\Delta p} \int_{p_1}^{p_2} \tfrac{1}{2} c_p \xi \{T^{*2}\} \frac{dp}{g} \qquad 4.14$$

where $\xi = \Gamma_d(\Gamma_d - \bar{\Gamma})^{-1}\bar{T}^{-1}$;

$$K_E = \frac{1}{\Delta p} \int_{p_1}^{p_2} \tfrac{1}{2}\overline{u'^2 + v'^2} \frac{dp}{g} \qquad 4.15$$

$$A_E = \frac{1}{\Delta p} \int_{p_1}^{p_2} \tfrac{1}{2} c_p \xi \overline{T'^2} \frac{dp}{g} \qquad 4.16$$

$$A_Z \rightarrow K_Z = \frac{1}{\Delta p} \int_{p_1}^{p_2} -\{\omega^*\alpha^*\} \frac{dp}{g} \qquad 4.17$$

$$A_E \rightarrow K_E = \frac{1}{\Delta p} \int_{p_1}^{p_2} -\overline{\omega'\alpha'} \frac{dp}{g} \qquad 4.18$$

$$G_Z = \frac{1}{\Delta p} \int_{p_1}^{p_2} \xi\{Q^*T^*\} \frac{dp}{g} \qquad 4.19$$

$$G_E = \frac{1}{\Delta p} \int_{p_1}^{p_2} \xi \overline{Q'T'} \frac{dp}{g} \qquad 4.20$$

As a result of partitioning, two terms representing conversion between the eddy and zonal mean components of A and K occur. Before discussing these conversions it is necessary to consider the use of the set of equations 4.13 - 4.20 in studying the energetics of a limited region of the atmosphere and to introduce equations to describe additional processes which arise as a consequence of not treating the whole atmosphere.

4.5. AVAILABLE POTENTIAL ENERGY AND THE LIMITED ATMOSPHERIC REGION

Available potential energy is strictly only defined for the whole atmosphere, although, since interaction between the hemispheres is usually small, the concept may be applied to a single hemisphere. In practice lack of data usually requires that the spatial averages in equations 4.13 - 4.20 be taken over a smaller latitude range. In this case, as long as all the important scales of motion are included, it is acceptable to regard the values calculated as approximations to those which would be obtained if all latitudes were sampled.

A common misconception lies in attributing local meaning to the integrands of those of equations 4.13 - 4.20 which relate to the available potential energy. While it is possible to define by equations 4.13 and 4.15 the zonal and eddy kinetic energy at a given latitude, available potential energy cannot have a local meaning. It is possible to define the contribution of a given region to the global available potential energy (Smith, 1969) but it is easy to see that a local value of the integrand of equation 4.8' does not give this contribution since, given static stability, it is always positive and can never express the negative contributions which occur when the local pressure is less than the isentropic mean pressure. It is therefore correct to consider only total potential energy if, for example, we wish to calculate

the latitudinal distribution of the various energy exchanges. To show how the local value of the integrand of the $A_z \rightarrow K_z$ conversion (equation 4.17) differs from the actual zonal mean conversion between total potential and kinetic energy we first rewrite the kinetic energy balance equation (4.7) using pressure as a vertical coordinate, as given by Van Mieghem (1972):

$$\frac{\partial K}{\partial t} + \underline{\nabla}\cdot(K + \Phi)\underline{V} + \frac{\partial}{\partial p}(K + \Phi)\omega = -\omega\alpha \qquad 4.21$$

where $\underline{V}$ is the "horizontal" velocity on an isobaric surface. The right hand side is minus the conversion of kinetic to total potential energy. We may expand this to give

$$[\omega\alpha] = [([\omega] + \omega')([\alpha] + \alpha')] = [\omega][\alpha] + [\omega'\alpha'] \qquad 4.22$$

Averaging meridionally we obtain

$$\overline{\omega\alpha} = \{[\omega][\alpha]\} + \overline{\omega'\alpha'} = \{\omega^*\alpha^*\} + \bar{\omega}\bar{\alpha} + \overline{\omega'\alpha'} \qquad 4.23$$

If the meridional average extends over all latitudes, $\bar{\omega}$ is identically zero and equation 4.23 contains the integrands of the $K_z \rightarrow A_z$ and $K_E \rightarrow A_E$ conversions, equations 4.17 and 4.18. For a limited latitudinal average $\{\omega^*\alpha^*\}$ represents only a part of the zonal kinetic energy loss since $\bar{\omega}$ is usually non-zero. For a local conversion we must therefore retain the term involving $\bar{\omega}$ and use the first term on the right hand side of expression 4.22 as the conversion between zonal mean total potential energy and zonal kinetic energy. The local conversion between eddy kinetic and eddy total potential energy may however be interpreted as a contribution to the global KE-AE conversion.

If we look again at equation 4.21 we see that if the atmospheric region under investigation does not extend to all latitudes and/or does not contain the whole atmospheric depth we must include in the kinetic energy budget the boundary pressure-work

terms $\underline{\nabla}.\Phi\underline{V}$ and/or $(\partial/\partial p)\Phi\omega$ respectively. Using the same notation as in equations 4.13 - 4.20 their zonal and eddy components may be written (following Muench, 1965) as

$$B_{\Phi Z} = \frac{1}{\Delta p}\frac{1}{S}\int_{p_1}^{p_2}\Big(([v]\Phi^*)_{\phi_1} 2\pi a \cos\phi_1 - ([v]\Phi^*)_{\phi_2} 2\pi a \cos\phi_2\Big)\frac{dp}{g} + \frac{1}{\Delta p}\Big(\frac{1}{g}\{\omega^*\Phi^*\}_{p_1} - \frac{1}{g}\{\omega^*\Phi^*\}_{p_2}\Big) \qquad 4.24$$

$$B_{\Phi E} = \frac{1}{\Delta p}\frac{1}{S}\int_{p_1}^{p_2}\Big([v'\Phi']_{\phi_1} 2\pi a \cos\phi_1 - [v'\Phi']_{\phi_2} 2\pi a \cos\phi_2\Big)\frac{dp}{g} + \frac{1}{\Delta p}\Big(\frac{1}{g}(\overline{\omega'\Phi'})_{p_1} - \frac{1}{g}(\overline{\omega'\Phi'})_{p_2}\Big) \qquad 4.25$$

where $S = 2\pi a^2(\sin\phi_1 - \sin\phi_2)$, the area of the isobaric surface. The two terms on the right hand side of each equation represent the net work done on the volume by work at the vertical and isobaric bounding surfaces respectively. Calculated over the near hemispheric region these expressions, like 4.13 - 4.20, give values which are usually close to those that would obtain for the whole hemisphere. To calculate local kinetic energy change as a result of work done by pressure forces, an expression analagous to 4.22 must be derived from equation 4.21.

The drawback of neglecting terms involving isobaric means in order to obtain a more representative picture of the hemispheric energy budget is that imbalance is thereby introduced into the energy budgets of the zonal mean components. We can determine the error so caused in our calculated zonal kinetic energy budget by comparing the sum of the $A_Z \rightarrow K_Z$ conversion and zonal mean pressure-work terms calculated excluding the isobaric mean components (i.e. using equations 4.17 and 4.24) with the sum obtained when these components are included, namely

$$\frac{1}{\Delta p}\int_{p_1}^{p_2}\left(-\left\{\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}([v][\Phi]\cos\phi)\right\}-\left\{\frac{\partial}{\partial p}([\omega][\Phi])\right\}-\left\{[\omega][\alpha]\right\}\right)\frac{dp}{g} \qquad 4.26$$

The terms of expression 4.26 are difficult to calculate accurately but together they represent the total gain of zonal kinetic energy through the action of the mean meridional circulation and, as such, can be expressed by the integral

$$\frac{1}{\Delta p}\int_{p_1}^{p_2}\left\{f[u][v]\right\}\frac{dp}{g} \qquad 4.27$$

which is much easier to calculate. (This relationship is derived in Appendix 1). Figure 4.1 compares the two sums for three stratospheric layers during the stratospheric warming investigated. Prior to the peak of the event on 28.1.73 the upper and lower layers lost zonal kinetic energy while the middle stratosphere gained zonal kinetic energy as a result of vertical motion systems only partially sampled within the regions. After the warming peak actual conversions are very close to hemispheric estimates given by equations 4.17 and 4.24.

Finally in this section we give expressions for the gain in energy of the region due to advection of zonal and eddy kinetic energy across the boundaries. These have a form similar to that of the pressure-work terms and are

$$B_{KZ} = \frac{1}{\Delta p}\frac{1}{S}\int_{p_1}^{p_2}\left(([v]\tfrac{1}{2}[u]^2)_{\phi_1}2\pi a\cos\phi_1 - ([v]\tfrac{1}{2}[u]^2)_{\phi_2}2\pi a\cos\phi_2\right)\frac{dp}{g}$$
$$+\frac{1}{\Delta p}\left(\frac{1}{g}\{[\omega]\tfrac{1}{2}[u]^2\}_{p_1}-\frac{1}{g}\{[\omega]\tfrac{1}{2}[u]^2\}_{p_2}\right) \qquad 4.28$$

$$B_{KE} = \frac{1}{\Delta p}\frac{1}{S}\int_{p_1}^{p_2}\left(([v]\tfrac{1}{2}[u'^2+v'^2])_{\phi_1}2\pi a\cos\phi_1 - ([v]\tfrac{1}{2}[u'^2+v'^2])_{\phi_2}\right.$$
$$\left.2\pi a\cos\phi_2\right)\frac{dp}{g} + \frac{1}{\Delta p}\left(\frac{1}{g}\{[\omega]\tfrac{1}{2}[u'^2+v'^2]\}_{p_1}\right.$$
$$\left.-\frac{1}{g}\{[\omega]\tfrac{1}{2}[u'^2+v'^2]\}_{p_2}\right) \qquad 4.29$$

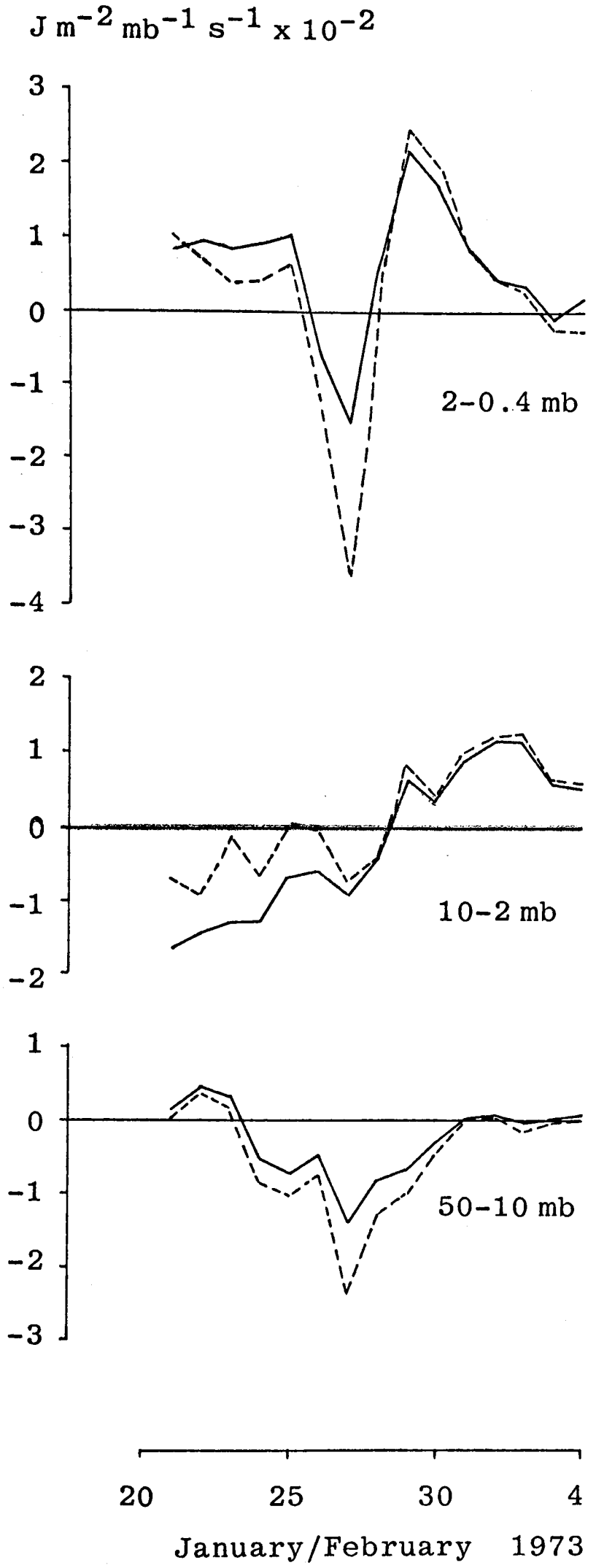


Figure 4.1. Comparison for three stratospheric layers of the sum of the $A_Z \rightarrow K_Z$ conversion and the zonal mean pressure-work terms as calculated when isobaric mean components are included (----) and when excluded (——) for the period 20.1 - 4.2.73 (see text for details)

Muench (1965) presents two similar boundary terms often referred to as 'boundary fluxes of available potential energy'. They should perhaps be considered as boundary fluxes which increase the temperature variance within the region, and as such should be included in the regional budget of the quantities defined by equations 4.14 and 4.16. Whether or not they serve to provide a better estimate of the hemispheric available potential energy is unclear. Their form below is essentially that given by Muench (1965) except that a term involving the eddy heat transport has been omitted since we have chosen to include it in the A_E-A_Z conversion discussed in section 4.6.

$$B_{AZ} = \frac{1}{\Delta p}\frac{1}{S}\int_{p_1}^{p_2} \tfrac{1}{2}\xi c_p \Big(([v]T^{*2})_{\phi_1} 2\pi a \cos\phi_1 - ([v]T^{*2})_{\phi_2} 2\pi a \cos\phi_2 \Big)\frac{dp}{g} + \frac{1}{\Delta p}\tfrac{1}{2}\xi c_p \Big(\frac{1}{g}\{[\omega]T^{*2}\}_{p_1} - \frac{1}{g}\{[\omega]T^{*2}\}_{p_2} \Big) \quad 4.30$$

$$B_{AE} = \frac{1}{\Delta p}\frac{1}{S}\int_{p_1}^{p_2} \tfrac{1}{2}\xi c_p \Big([vT'^2]_{\phi_1} 2\pi a \cos\phi_1 - [vT'^2]_{\phi_2} 2\pi a \cos\phi_2 \Big)\frac{dp}{g} + \frac{1}{\Delta p}\tfrac{1}{2}\xi c_p \Big(\frac{1}{g}(\overline{\omega T'^2})_{p_1} - \frac{1}{g}(\overline{\omega T'^2})_{p_2} \Big) \quad 4.31$$

4.6. CONVERSION BETWEEN EDDY AND ZONAL MEAN ENERGY

As a result of partitioning the energy into zonal mean and eddy components expressions describing the conversions between them naturally arise. We shall see in Chapter 5 that the balance equation of zonal mean kinetic energy (equation 5.10) contains a term involving the convergence of eddy momentum flux which represents eddy to zonal kinetic energy conversion, namely

$$K_E \rightarrow K_Z = \frac{1}{\Delta p}\int_{p_1}^{p_2}\left\{\frac{-[u]}{a\cos^2\phi}\frac{\partial}{\partial\phi}([u'v']\cos^2\phi) - [u]\frac{\partial}{\partial p}([u'\omega'])\right\}\frac{dp}{g} \quad 4.32$$

Many authors have represented this conversion by the form

$$K_E \rightarrow K_Z = \frac{1}{\Delta p}\int_{p_1}^{p_2}\left\{[u'v']\frac{\cos\phi}{a}\frac{\partial}{\partial\phi}\left(\frac{[u]}{\cos\phi}\right) + [u'\omega']\frac{\partial}{\partial p}[u]\right\}\frac{dp}{g} \qquad 4.33$$

The first terms of the integrands of equations 4.32 and 4.33 are related by the identity

$$\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}([u][u'v']\cos\phi) = \frac{[u]}{a\cos^2\phi}\frac{\partial}{\partial\phi}([u'v']\cos^2\phi) + [u'v']\frac{\cos\phi}{a}\frac{\partial}{\partial\phi}\left(\frac{[u]}{\cos\phi}\right) \qquad 4.34$$

A similar identity relates the vertical momentum flux terms. If the conversion is integrated over the globe the left hand side of expression 4.34 is zero and 4.32 and 4.33 give identical results. Momentum flux divergence is difficult to calculate accurately (see Chapter 5). Therefore, when average conversion over a limited region is required, it is convenient to employ expression 4.33 and to evaluate the left hand side of 4.34 as boundary integrals. As a result some authors have regarded the integrand of expression 4.33 as giving the local conversion, but this is incorrect since, by neglecting the local contribution of the left hand side of 4.34, it gives only a portion of the total conversion 4.32. Physical interpretations of the two expressions clarify this difference. Expression 4.32 describes convergence of eastward momentum flux in regions of eastward flow giving local increase of the zonal kinetic energy. Expression 4.33 gives a positive conversion if momentum flux is up the westerly wind gradient. This is only valid for a region with no flow through the boundaries since, given zero boundary fluxes, this process produces gain of K_Z where the zonal flow is strong but loss of K_Z where it is weak, giving a net gain only for the whole region. We compare in Figure 4.2 height-latitude cross-sections of the integrands of expressions 4.32 and 4.33 for 29.1.73 to illustrate how qualitatively different they may be in practice.

The conversion between zonal and eddy available potential

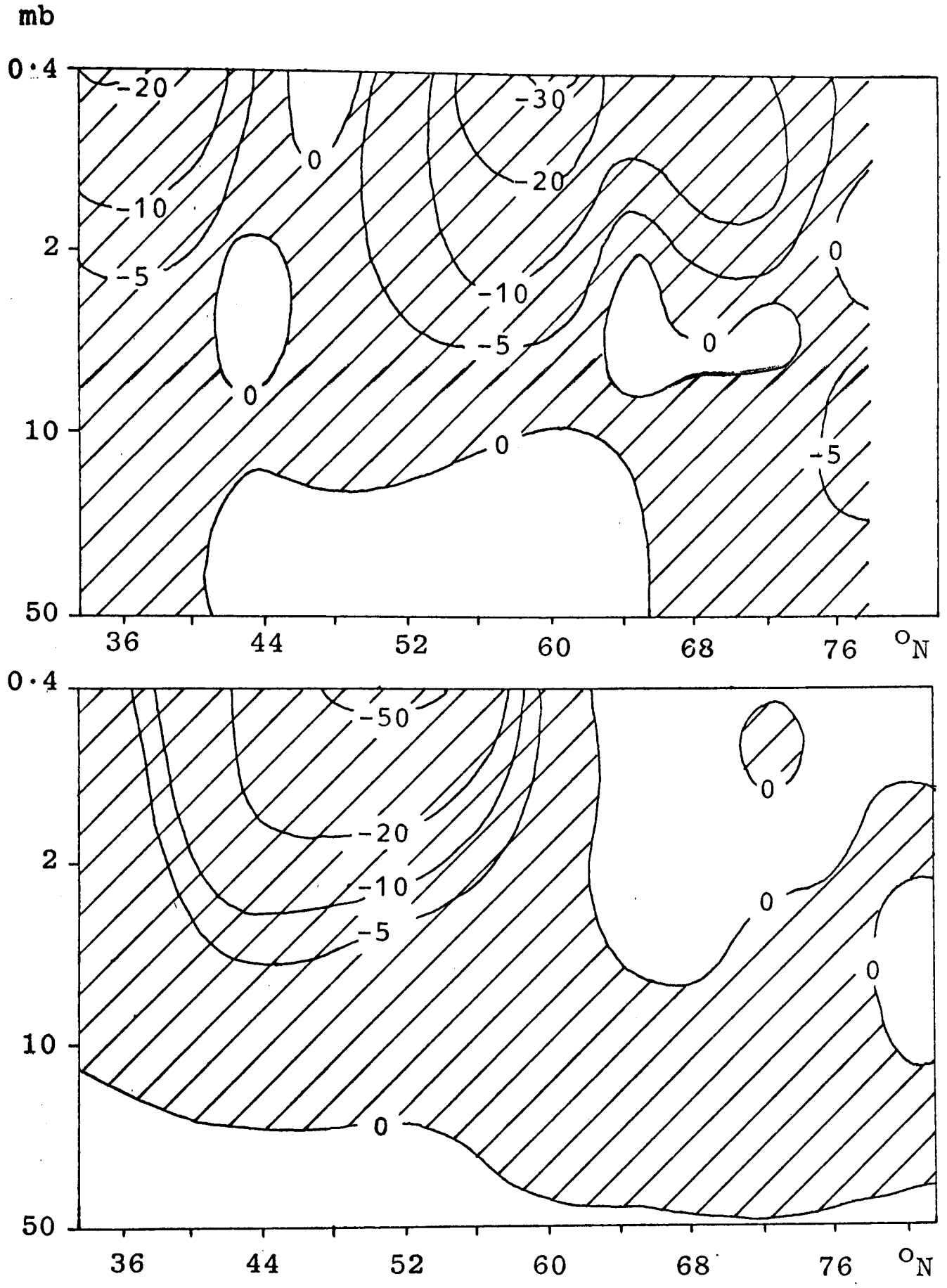


Figure 4.2. A comparison between height-latitude sections of the integrands of expressions 4.32 (upper) and 4.33 (lower). The latter is often incorrectly used to represent $K_E \rightarrow K_Z$ conversion at a given latitude, shown correctly in the upper diagram. The data used are for 29.1.73

energy may be expressed in two ways analagous to the K_E-K_Z conversions 4.32 and 4.33. We prefer to use the form involving the eddy heat flux[1] divergence, which does not require the evaluation of boundary integrals:

$$A_E \rightarrow A_Z = \frac{1}{\Delta p}\int_{p_1}^{p_2}\left\{-c_p \xi T^* \left(\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}([v'T']\cos\phi) - \frac{\partial}{\partial p}[\omega'T']\right\}\frac{dp}{g} \qquad 4.35$$

We may assume that where the integrand is locally positive (negative) the eddy (zonal) potential energy field is being depleted at the expense of the zonal (eddy) component.

The methods of calculating the eddy correlations of the form [x'y'] which appear in the energy equations of this chapter are outlined in Appendix 2.

---

[1] Throughout this thesis the term 'eddy heat flux' will refer to the quantity [v'T']. Strictly this is an eddy flux of temperature but it is, of course, proportional to the heat flux per unit mass $c_p[v'T']$.

Introduction

In this chapter we discuss possible methods of estimating vertical motion in the atmosphere. Solution of the thermodynamic equation is most suited to our case and we describe how non-adiabatic heating was specified. It is possible to obtain an independent estimate of the zonal mean vertical and meridional motions via the zonal mean momentum equation and we compare the mean meridional circulations implied by heat and momentum balances. We suggest reasons for the discrepancies found and discuss their implication for the zonal kinetic and available potential energy budgets.

5.1. METHODS OF CALCULATING VERTICAL MOTIONS

We give below four equations which may be solved under various conditions for the vertical motion and discuss briefly their suitability or otherwise for our study.

(1) The continuity equation:

$$\frac{\partial \omega}{\partial p} = -\underline{\nabla}_h \cdot \underline{V}_h \qquad 5.1$$

(2) The thermodynamic equation:

$$\frac{\partial T}{\partial t} = -\underline{V}_h \cdot \underline{\nabla}_h T - \omega\frac{\partial T}{\partial p} + \frac{\kappa T \omega}{p} + \frac{Q}{c_p} \qquad 5.2$$

(3) The geostrophic vorticity (ζ_g) equation:

$$\frac{\partial \omega}{\partial p} = \left(\frac{\partial}{\partial t} + \underline{V}_g \cdot \underline{\nabla}_h\right) \log_e (\zeta_g + f) \qquad 5.3$$

(4) The "ω"-equation:

$$\nabla^2 \sigma\omega + f^2\frac{\partial^2 \omega}{\partial p^2} = -f\frac{\partial}{\partial p}(\underline{V}_h \cdot \underline{\nabla}_h \zeta) + g\nabla^2(\underline{V}_h \cdot \nabla_h \frac{\partial z}{\partial p}) + \frac{R}{pc_p}\nabla^2 Q \qquad 5.4$$

Method (1) is generally not suitable since $\underline{\nabla}_h \cdot \underline{V}_h$ is an order of magnitude smaller than its components. Thus $\underline{V}_h$ must be known to 1% accuracy in order to calculate $\partial\omega/\partial p$

to 10% accuracy. Methods 2-4 may be used with geostrophic winds and will generally be correct to within 10%. Since the zonal mean ω is an order of magnitude smaller than grid point values of ω its accuracy may be rather worse than this. Both methods 3 and 4 require a specification of ω at the lower boundary or some other level. Such a field at 50mb was not obtainable in our case so method 2 was employed with a three-dimensional radiation scheme to specify Q. This scheme is described in the following section.

5.2. RADIATIVE HEATING AND COOLING

Vertical motions were evaluated on a three-dimensional grid between z = 3.0 and z = 7.8 where $z = -ln(p/p_s)$, p is pressure and p_s is surface pressure. The radiative heating scheme, used in both the calculation of vertical velocities and the generation of available potential energy, incorporated heating by absorption of solar radiation by ozone and long wave cooling by carbon dioxide and ozone. We describe each below:

(a) Solar heating

Heating by absorption of ultra-violet radiation by ozone for each day was calculated as a function of height, latitude and longitude at levels above z = 5.8, where photochemical equilibrium was assumed, and as a function of height and latitude only in the region z = 3.8 - 5.6, where the ozone distribution is controlled predominantly by dynamics. Heating rates are small in the lower stratosphere and were not calculated below z = 3.8. The values were calculated by the method of Harwood and Pyle (1975). The amount of ozone in the path of the solar beam is evaluated for each height and latitude and the energy depletion of the solar beam in traversing each height-latitude rectangle is determined. From this depletion the heating rate is deduced.

Because the path length of the solar beam varies throughout the day heating rates were evaluated for five times during the day and a daily average calculated. Ozone mixing ratios from the Oxford two-dimensional general circulation model were used and are given in Appendix 3.

Above z = 5.8 the ozone mixing ratios were adjusted for each longitude according to a simple temperature dependence scheme developed by Pyle (private communication). The equilibrium mixing ratio, E', at temperature T' is given in terms of the equilibrium mixing ratio, E, at temperature T by the relation

$$E' = E \exp\left(\frac{1000}{T'}\right) \Big/ \exp\left(\frac{1000}{T}\right) \qquad 5.5$$

where T is taken from the model run which produced the equilibrium mixing ratios, E. The scheme is described more fully in Appendix 3.

(b) Infra-red cooling

Using the technique employed by Harwood and Pyle (1975), and suggested by Houghton (1968), radiative cooling due to CO_2 at 15μm and O_3 at 9.6μm was calculated, as a function of temperature and height only, from the relation

$$\frac{dT}{dt} = k_1(z,X)\, B(T) + k_2(z,X) \qquad 5.6$$

where X is the emitting constituent and B(T) the Planck function at temperature T for the emitting wavelength. A table of the constants k_1 and k_2 is given in Appendix 3. Daily cooling rates on a height-latitude-longitude grid were obtained above level z = 3.8, while below this level cooling rates were set to zero. The four lowest levels were thus assumed to be in radiative equilibrium, a reasonable assumption for the lower stratosphere.

All heating and cooling rates were calculated at model grid points and linearly interpolated on to the finer grid required

for our calculations.

5.3. COMPARISON OF INDEPENDENTLY CALCULATED MEAN MERIDIONAL CIRCULATIONS

By using the thermodynamic equation we obtain vertical motions which are just those required to balance the heat budget, but there is no guarantee that momentum balance is thereby achieved. Similarly vertical velocities determined from the geostrophic vorticity equation would approximately balance the momentum budget but not necessarily the heat budget. By comparing such independent estimates of vertical motion an assessment of the internal consistency of the data may be made. For the reasons given in section 5.1 it was not possible to calculate vertical motions on a global grid other than by the thermodynamic equation. However a useful comparison restricted to zonal mean motions can be made. Given the zonally averaged thermodynamic equation

$$\frac{\partial}{\partial t}[T] + \frac{[v]}{a}\frac{\partial}{\partial \phi}[T] - [\omega]\left(\frac{R[T]}{pc_p} - \frac{\partial}{\partial p}[T]\right) = \frac{[Q]}{c_p}$$
$$-\frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}([v'T']\cos\phi) - \frac{\partial}{\partial p}[\omega'T'] + \frac{R}{pc_p}[\omega'T'] \qquad 5.7$$

the zonally averaged zonal momentum equation,

$$\frac{\partial}{\partial t}[u] + [v]\left(-f + \frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}([u]\cos\phi)\right) + [\omega]\frac{\partial}{\partial p}[u]$$
$$= -\frac{1}{a\cos^2\phi}\frac{\partial}{\partial \phi}([u'v']\cos^2\phi) - \frac{\partial}{\partial p}[u'\omega'] \qquad 5.8$$

and the zonally averaged continuity equation

$$\frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}([v]\cos\phi) + \frac{\partial}{\partial p}[\omega] = 0 \qquad 5.9$$

and assuming that the diabatic heating, Q, is specified, then each pair of the set of equations 5.7 - 5.9 may be solved for an independent estimate of the mean meridional circulation [v] and [ω] . The extent to which the [v] and [ω] calculated from two of the set do not satisfy the third equation is a measure of the

internal inconsistencies in the data set.

Two estimates of the mean meridional circulation were calculated using the equation pairs 5.7 with 5.9 and 5.8 with 5.9 for a five-day period (26 - 30.1.73) during the peak of the sudden warming under investigation when the meridional circulation was most pronounced. We use subscripts T to denote results obtained from thermodynamic balance (equations 5.7 and 5.9) and M to denote the momentum balancing mean meridional circulation (from equations 5.8 and 5.9). In the figures values of $[\omega]$ have been converted into vertical velocities (in cms^{-1}). In our case a good approximation to the $[\omega]$ satisfying equations 5.7 and 5.9 was already available in the zonal mean of the grid point vertical motion estimates. Since geostrophic winds were used in calculating grid point ω's the $[\omega]$ derived from them does not satisfy equation 5.7 exactly. The error is small however as shown in Table 5.1 where the difference between the final two columns is the total error due to the use of the geostrophic approximation and finite difference techniques. We shall therefore refer to a value of $[\omega]$ so calculated as $[\omega]_T$. The full curve in Figure 5.1a shows the latitudinal distribution of the five-day averaged values of $[\omega]_T$ at 2mb. To obtain $[v]_T$ from equation 5.9 the boundary condition $[v] = 0$ at 90^oN was used and $[\omega]_T$ was linearly extrapolated from 80^oN to the pole. The five-day averaged values of $[v]_T$ at 2mb are shown by the full curve in Figure 5.1b. The broken curves in Figures 5.1a and 5.1b are approximate solutions for $[\omega]_M$ and $[v]_M$ of equations 5.8 and 5.9. Previous analyses of the zonal mean zonal momentum balance (Vincent, 1968; Hirota and Sato, 1969; Barnett, 1973) have shown that terms involving $[\omega]$ and ω' are unimportant. Equation 5.8 was thus solved for $[v]_M$ neglecting the second terms on both sides. Equation 5.9

TABLE 5.1

Terms of the zonal mean thermodynamic equation for 2 mb at 76°N for the period 21.1.73 to 4.2.73. Column 7 gives the balance of columns 1 - 6 while column 8 shows the observed rates of change of zonal mean temperature.

units : K day^{-1}

Date Jan/Feb	1	2	3	4	5	6	7	8
21	-6.28	3.62	-0.23	0.12	-0.35	0.05	-3.07	-2.15
22	-5.99	4.00	-0.18	0.22	-0.41	0.09	-2.27	-1.65
23	-5.91	8.63	-4.72	0.13	-0.83	0.29	-2.41	-0.55
24	-5.88	7.60	-2.14	0.17	-1.24	0.32	-1.17	-0.80
25	-5.75	5.98	-0.49	0.07	-1.78	0.35	-1.62	-0.60
26	-5.77	17.80	-11.85	-0.03	-2.91	0.34	-2.42	1.30
27	-6.13	40.36	-37.03	0.98	3.29	-1.35	0.12	4.40
28	-6.89	47.79	-41.96	2.76	0.51	-0.29	1.92	4.00
29	-7.09	32.95	-26.25	1.96	-5.10	0.83	-1.70	1.10
30	-7.11	24.19	-15.31	0.38	-5.59	0.47	0.27	-1.70
31	-6.61	7.02	-4.25	0.33	-2.14	0.17	-5.48	-4.60
1	-5.91	0.30	0.61	0.12	-0.69	0.03	-5.54	-6.05
2	-5.12	3.70	-3.77	0.30	-0.67	0.04	-5.52	-5.05
3	-4.75	0.94	0.81	0.03	-0.37	0.04	-3.30	-3.40
4	-4.40	4.22	-1.68	-0.01	-0.02	-0.02	-1.91	-2.15

1: $[Q]/c_p$ 2: $-\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}([v'T']\cos\phi)$

3: $[\omega]\left(\frac{R[T]}{pc_p} - \frac{\partial}{\partial p}[T]\right)$ 4: $-\frac{[v]}{a}\frac{\partial}{\partial\phi}[T]$

5: $-\frac{\partial}{\partial p}[\omega'T']$ 6: $\frac{R}{pc_p}[\omega'T']$

7: $\frac{\partial}{\partial t}[T]$ (implied by terms 1 - 6)

8: $\frac{\partial}{\partial t}[T]$ (observed)

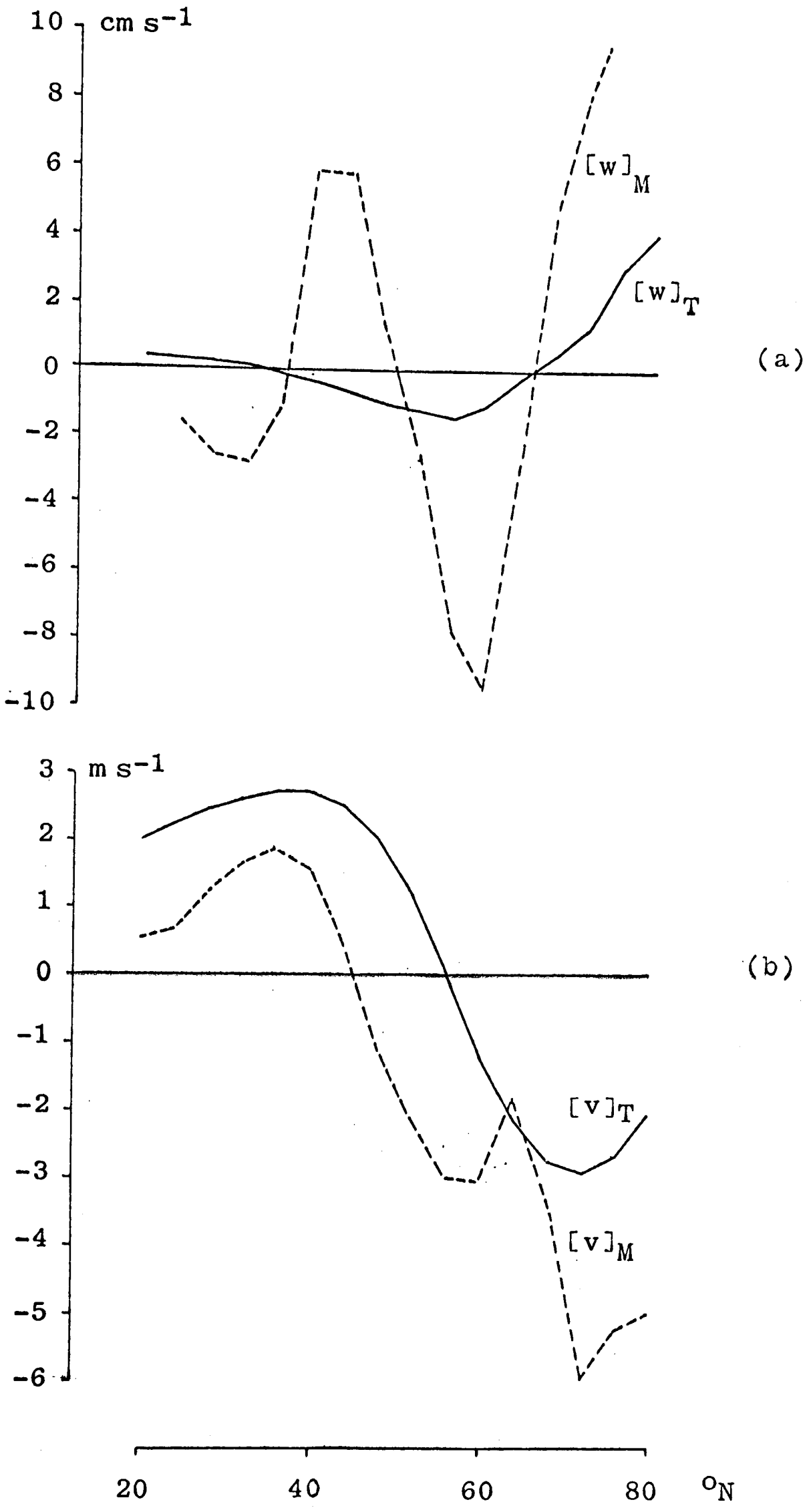


Figure 5.1a,b. Latitudinal distributions of zonal mean vertical velocity (a) and zonal mean meridional velocity (b) at 2mb as calculated from the zonal mean thermodynamic equation (T) and the zonal mean zonal momentum equation (M). Values plotted are means for 26 - 30.1.73

was solved for $[\omega]_M$ using $[\omega]_T$ at 50mb as a lower boundary condition, no better alternative being available. Estimates of the neglected terms using the calculated $[\omega]_M$ and $[u'\omega']$ evaluated from the grid point vertical motion calculations confirmed that they were one or two orders of magnitude smaller than the terms retained. Although there is qualitative agreement between $[v]_T$ and $[v]_M$ the vertical components are very different in both magnitude and latitudinal distribution.

The corresponding results for the comparison at the 10mb level are given in Figures 5.1c and 5.1d. Here there is closer agreement although momentum balance still requires upward motion at 40^oN in contrast to the sinking required by heat balance.

The smoother spatial variation of the thermodynamically derived velocity is matched by smoother variation in time also. This is shown in Figure 5.2 where we have plotted against time both vertical velocity estimates for 52 and 72^oN at 2 and 10mb.

The large fluctuations in the momentum-balanced velocities and the substantial differences in magnitude between the two sets of results warrant further discussion. The former are almost certainly a reflection of the difficulty in calculating the divergence of the horizontal eddy momentum flux, the major term in equation 5.8 on which $[v]_M$ depends. The momentum flux itself depends on the difference between two, usually large, but often similar products as indicated in the method of calculation given in Appendix 2. An example of its sensitivity is illustrated in Figure 5.3. The latitudinal distributions of wavenumber one eddy momentum flux at 2mb on 28.1.73 and the retrieved quantities from which it was calculated are plotted. The sensitivity of the momentum flux to slight changes in the latitudinal gradients of the height coefficients is very evident. At 64^oN, for example, a

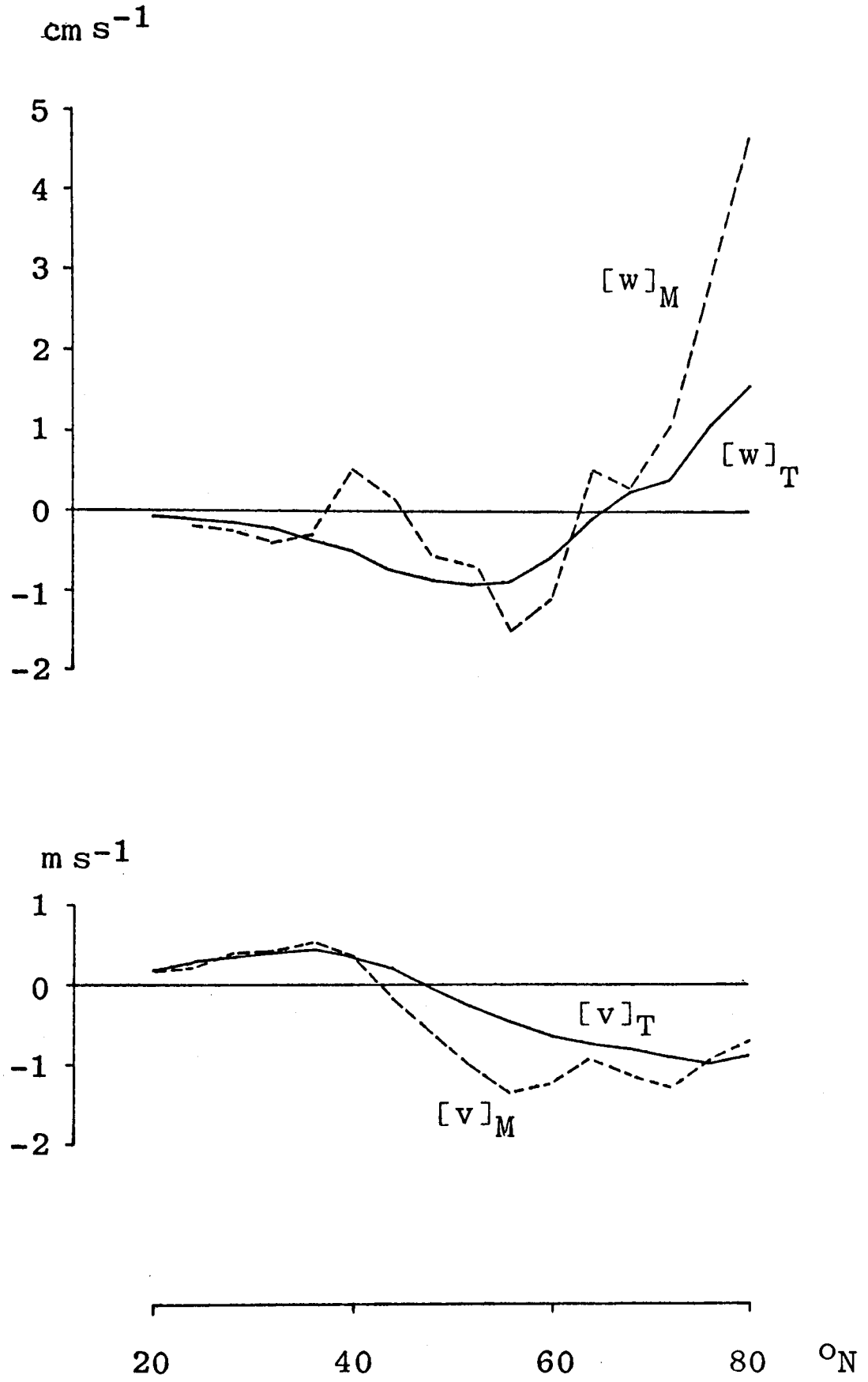


Figure 5.1c,d. As Figure 5.1a,b but at 10 mb

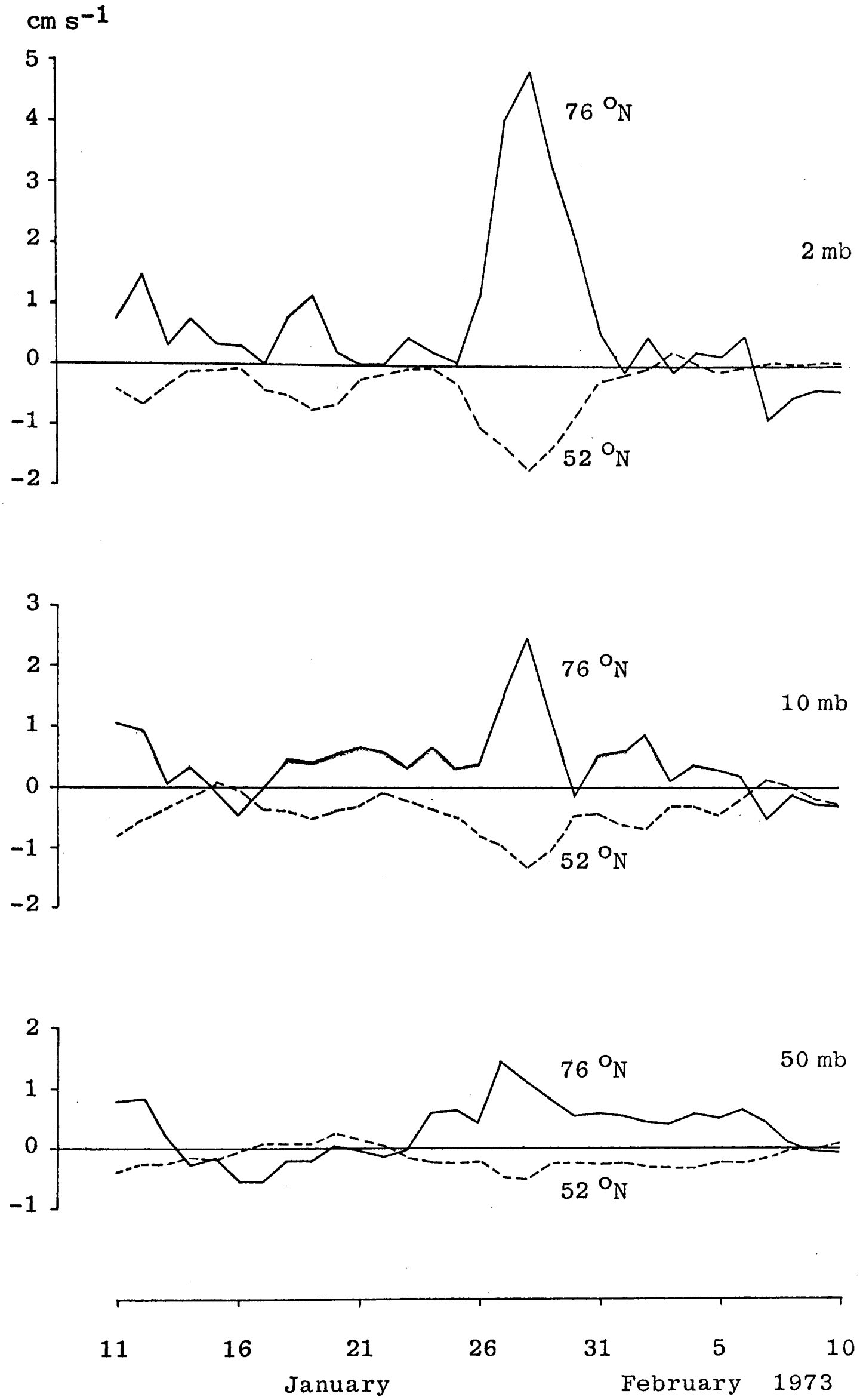


Figure 5.2b. Zonal mean vertical velocity at 76°N and 52°N calculated from zonal mean thermodynamic equation plotted against time for 2, 10 and 50 mb. (The 50 mb plot is included for later reference)

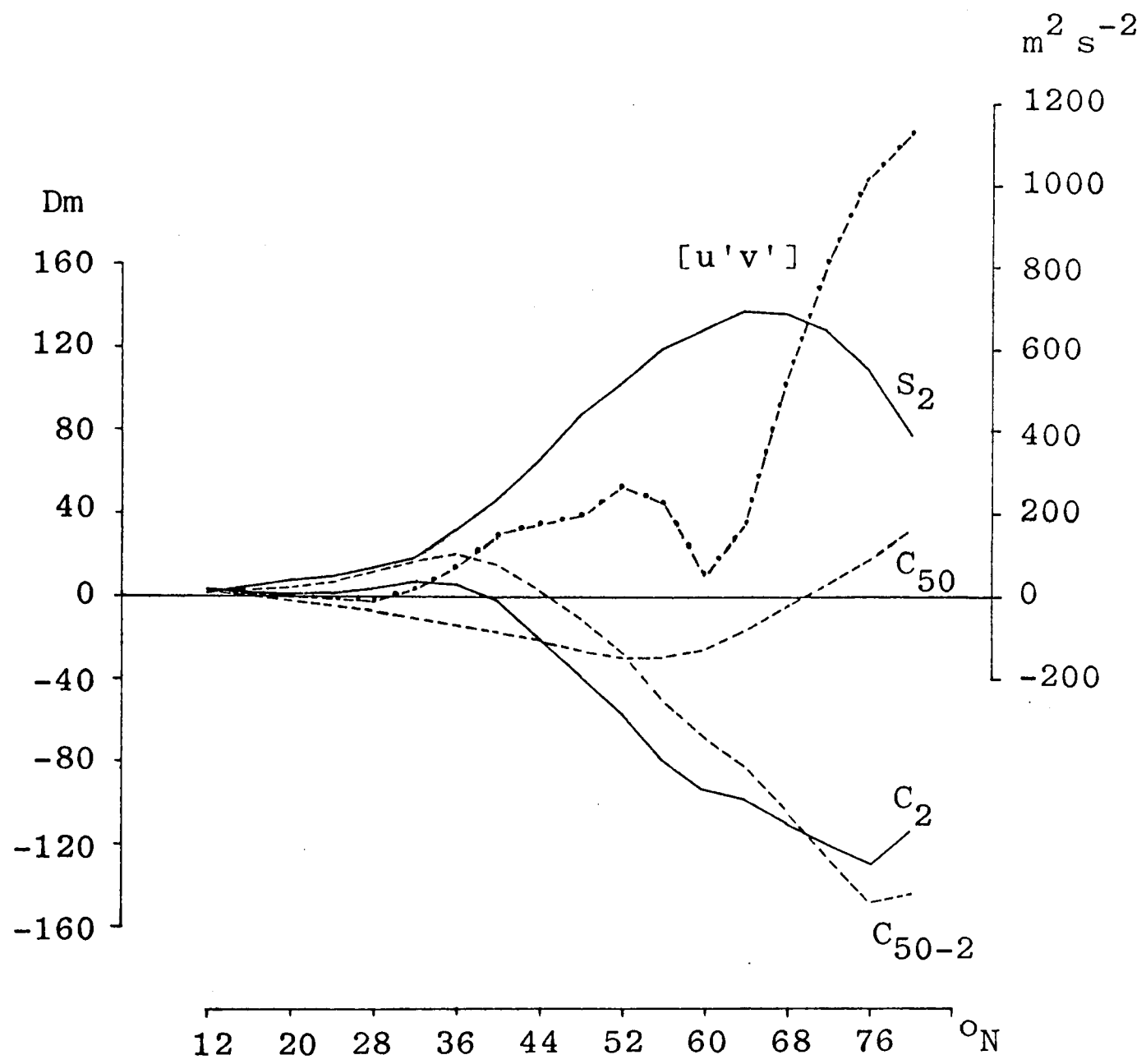


Figure 5.3. Latitudinal distributions for 28.1.73 at 2 mb of wavenumber 1 eddy momentum flux [u'v'] and retrieved wavenumber 1 sine and cosine Fourier coefficients of geopotential height (S_2, C_2), and the latitudinal distributions of the 50 mb cosine coefficient of geopotential height and the 50 - 2 mb thickness cosine coefficient (C_{50}, C_{50-2})

5% increase in the (negative) value of the 50-2mb thickness cosine coefficient (which added to the 50mb height coefficient gives the 2mb height coefficient) would have given a 100% larger value of $[u'v']$ at $60^{o}N$. This particular irregularity in the latitudinal gradient of the cosine coefficient is probably a consequence of the use of different covariance matrices for middle and high latitude retrieval as described in Chapter 3. From a more extensive study of this problem it might have been possible to eliminate systematic retrieval errors of this sort and produce a smoother set of momentum flux data. We shall see in Chapter 6 that there was much greater continuity in the eddy heat fluxes, in both space and time, than in the eddy momentum fluxes. Since the eddy heat flux convergence is largely balanced by the term involving $[\omega]$ in equation 5.7 (see Table 5.1) a smooth mean meridional circulation was obtained from equations 5.7 and 5.9.

Substantial differences between independent mean meridional circulations deduced from the equations 5.7 - 5.9 were also found by Vincent (1968) and Hartmann (1976). In the former study estimates of the mean vertical motion below 10mb in January 1964 were found to be twice as large in high latitudes when momentum balance was a requirement as when the heat budget was satisfied. Different choice of boundary values had little effect on $[v]_T$, which was also the case in our evaluations: $[\omega]_T$ linearly extrapolated to the pole from $80^{o}N$ and the value of $[\omega]_T$ at $80^{o}N$ used for all poleward latitudes gave essentially the same results.

Hartmann's southern hemisphere study employed satellite data with a minimum information retrieval method. He calculated three mean meridional circulations, averaged over two relatively inactive winter months, using the continuity equation with, in

turn, the momentum equation, the adiabatic thermodynamic equation and the thermodynamic equation including diabatic heating. The mean meridional velocities are reproduced in Figure 5.4. They are about an order of magnitude smaller than those we calculated during the sudden warming period. At 10 and 3mb the velocities derived from diabatic heat balance were more poleward at all latitudes than those determined from momentum balance, a result consistent with our own findings for 2mb and for 10mb north of 40°N. At 1mb the mean circulations were qualitatively different, heat balance requiring poleward motion at all latitudes with sinking in high latitudes and momentum balance being achieved by the two cell pattern, with high latitude rising, observed at lower levels. The neglect of friction is suggested by Hartmann as a possible contributory factor to the discrepancy between the circulations. For westerly flow the Coriolis acceleration acting on an equatorward mean meridional velocity, as found in high latitudes by Hartmann, decelerates the mean flow to counteract acceleration by eddy momentum flux convergence. Inclusion of friction provides an additional deceleration of the mean flow and thus a less strong equatorward meridional velocity is required. For easterly mean flow, however, inclusion of friction necessitates a stronger equatorward meridional velocity since the Coriolis torque accelerates the mean wind in this case. Therefore if friction were an important process we should expect [v] derived from the frictionless momentum equation to at least reduce, if not reverse, its poleward deficit relative to the thermally balanced [v] during a reversal of the zonal mean wind. Little evidence of such a reduction was found when the flow reversed during our period of investigation as is shown in Figure 5.5, where the two meridional velocities and the

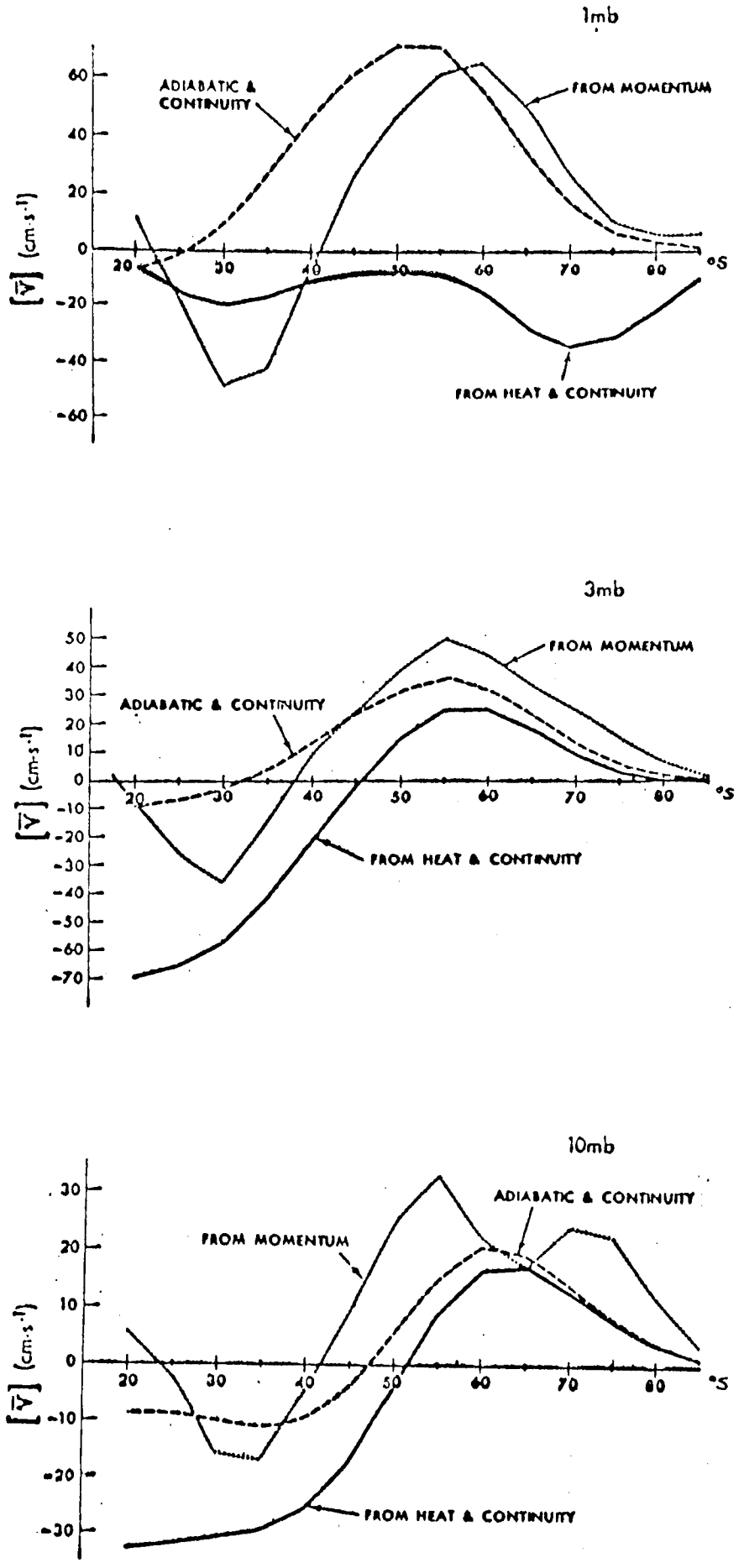


Figure 5.4. Latitudinal distributions of zonal mean meridional velocity at 1, 3 and 10 mb during the southern hemisphere winter of 1973 calculated using the zonal mean zonal momentum equation, the zonal mean adiabatic thermodynamic equation and the full zonal mean thermodynamic equation. (Taken from Hartmann, 1976)

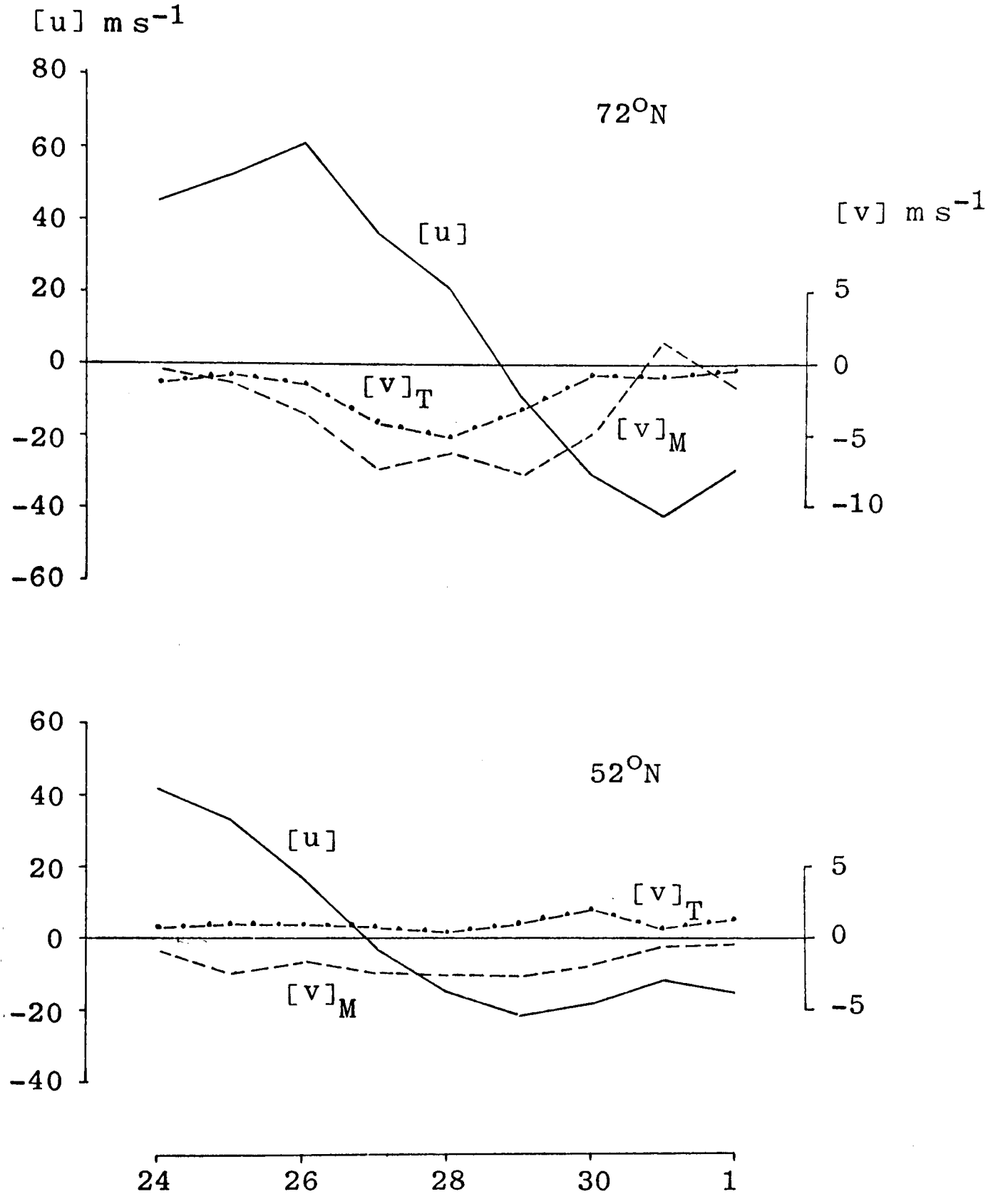


Figure 5.5. Plots against time of the zonal mean zonal wind [u], zonal mean meridional velocity calculated from heat balance ($[v]_T$) and that calculated from momentum balance ($[v]_M$) at 72°N (upper) and 52°N (lower) at 2 mb

zonal mean zonal wind are plotted against time for 72 and 52^{o}N at 2mb.

Since the discrepancy is consistently of the same sign a possible cause is systematic retrieval error. We saw in Chapter 3 that large temperature changes with height could not always be resolved. In such cases the levels at which temperature is likely to be best retrieved (near the peaks of the weighting functions) do not coincide with those where geopotential height is most accurately specified. Thus the mean meridional circulation from heat balance, which is largely controlled by eddy temperature and meridional wind correlation [v'T'] is not likely to agree with that derived from momentum balance, governed principally by eddy velocity correlation [u'v'].

Comparison of eddy fluxes calculated from retrieved data and from direct measurements will never be possible on account of the sparsity of rocket-sonde observations. However a useful alternative experiment would be to compute radiances implied by the temperature structure of a three-dimensional numerical circulation model simulating an active winter period and retrieve temperature and geopotential height profiles from these. Comparison of eddy fluxes calculated directly from the model and from retrieved data could then be made. Such an experiment could also indicate the sort of accuracy which can be obtained in heat, momentum and energy budget studies using retrieved data, which is otherwise difficult to determine.

5.4. IMPLICATIONS OF THE MEAN MERIDIONAL CIRCULATION COMPARISON ON THE KINETIC ENERGY BUDGET

Since the three-dimensional vertical motion field was calculated to satisfy the heat balance, within the small error resulting from use of the geostrophic approximation, the available

potential energy budget must also balance. The mean meridional circulation comparison has indicated that this vertical motion field is not compatible with momentum balance or, therefore, with kinetic energy balance. We show in this section the error which occurs in the zonal mean kinetic energy budget through use of the thermodynamically derived velocities.

If we multiply equation 5.8 by $[u]$ we obtain the balance equation of the east-west, or x-, component of the zonal kinetic energy, K_{zx}:

$$\frac{\partial}{\partial t}\tfrac{1}{2}[u]^2 + \frac{[v]}{a\cos\phi}\frac{\partial}{\partial\phi}(\tfrac{1}{2}[u]^2\cos\phi) + [\omega]\frac{\partial}{\partial p}\tfrac{1}{2}[u]^2 - f[u][v]$$
$$= -\frac{[u]}{a\cos^2\phi}\frac{\partial}{\partial\phi}([u'v']\cos^2\phi) - [u]\frac{\partial}{\partial p}[u'\omega'] \qquad 5.10$$

The terms on the right hand side are conversions from the x-component of eddy kinetic energy, K_{EX}, while the terms on the left are the local and advective changes in K_{zx} and the Coriolis torque which represents conversion between the meridional, or y-, component of the zonal kinetic energy, K_{zy}, and K_{zx}.

We show in Appendix 1 that the $K_{zy}\rightarrow K_{zx}$ conversion is balanced by the gain in K_{zy} via the horizontal and vertical components of the zonal mean pressure-work term and the conversion from zonal mean potential energy. We therefore see that the imbalance in the K_z budget caused by using $[v]_T$ and $[\omega]_T$ to calculate the pressure-work terms and the potential energy conversion will be the imbalance which occurs in equation 5.10 if $f[v]_M[u]$ is replaced by $f[v]_T[u]$. Figure 5.6 shows the observed rate of change of K_{zx} and that implied by using $[v]_T$ in equation 5.10 for two two-day periods at 2mb during the sudden warming. Although $[v]_M$ and $[v]_T$ were qualitatively similar, since the rate of change of K_{zx} is a small difference between two large terms it is very sensitive to the value of $[v]$. During the period of

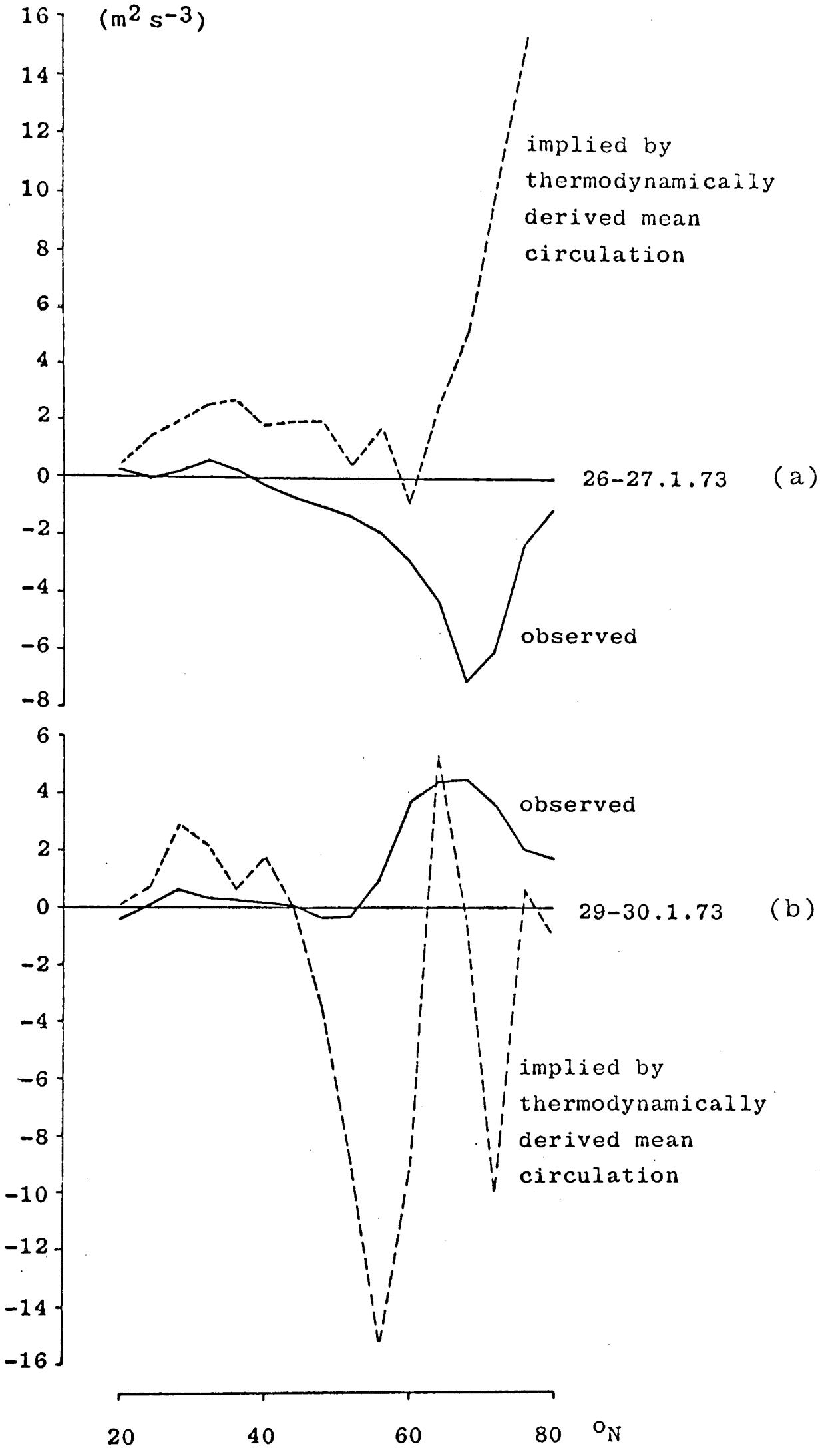


Figure 5.6. Comparison between the observed (i.e. directly calculated) rate of change of the zonal mean zonal kinetic energy at 2 mb and that derived from the balance equation (equation 5.10) using thermo - dynamically deduced mean meridional velocities, for two two - day periods during the 1973 sudden warming

weakening westerly flow prior to the circulation breakdown (Figure 5.6a) the conversion and generation processes imply a large gain in K_z while a loss of energy is shown at the time when the kinetic energy of the reversed, easterly, flow is increasing (Figure 5.6b). Note how the erratic behaviour of the momentum flux convergence near $60^{o}N$ affected the implied rate of change.

Similarly the rate of change of zonal available potential energy is a small balance between large terms and use of the momentum balanced mean velocities in its calculation would give results as unrealistic as those obtained here for zonal kinetic energy using the thermodynamically balanced mean meridional circulation. We see that the effect of this error on the area-averaged zonal kinetic energy budget will be more serious than that due to the neglect of isobaric mean processes discussed in Chapter 4. However the latter error would probably have been much larger had momentum balanced mean velocities been employed in that calculation.

5.5. SUMMARY

As a result of the inconsistencies in the data set we cannot obtain a balanced energy budget for the sudden warming event. The balance of any energy component is susceptible to relatively small errors in the conversions and generations since its rate of change is generally a small residual of these processes. Much can be learnt about the energy exchanges themselves even if they are subject to error which renders them unsuitable for rate of change calculations.

We shall employ thermodynamically derived mean vertical and meridional velocities in our budget calculations on account of their smooth spatial and temporal variation. In addition the

values we have obtained by this method are in good agreement with estimates of mean vertical and meridional velocities during a sudden warming made by Barnett (1973). However if the main source of error lies in incompatibility between the retrieved temperature and height fields there will be inherent error in both the heat and momentum budgets, and therefore in both the available potential and kinetic energy budgets. Thus it must be stressed that by adopting the heat-balanced mean meridional circulation estimate we are not implying that it is the actual circulation which occurred, nor does a balanced available potential energy budget imply that all its components are correct. We do believe, for the reasons given, that this circulation is closer to reality than that derived from momentum balance. It remains to be determined by numerical experiments of the type suggested whether one method of calculating the mean meridional circulation is generally more satisfactory than any other when using retrieved satellite data.

Chapter 6 A SYNOPTIC DESCRIPTION OF THE JANUARY - FEBRUARY 1973 SUDDEN WARMING

INTRODUCTION

Since the sudden warming phenomenon is characterised by a variety of dynamical processes its analysis may be approached from several standpoints. This first chapter on the results of the 1973 warming study is largely descriptive and deals mostly with the observed changes in stratospheric temperature and wind. After giving a brief account of the observable features common to most sudden warmings we view the 1973 warming as the intensification and movement of a thermal system. The direction of propagation of the warming is discussed and results are compared with similar studies. The changes in the vertical structure of the temperature and geopotential waves, particularly at the onset of the warming, illustrate the development of the disturbance. The evolution of the zonal mean temperature and wind fields, the eddy heat and momentum transports and the induced mean meridional circulation are examined.

The energy budget analysis, the importance of the rôle played by the eddy pressure - work term, and a comparison between the 1973 event and numerical simulations of sudden warmings are topics to be discussed in the following chapters.

6.1. THE OBSERVED SYNOPTICS OF SUDDEN WARMINGS

The first observation of a stratospheric warming was made by Scherhag in 1952. During the next ten years a number of observational studies of continental scale were made using rocket- and radio-sonde measurements, the availability of global analyses being very restricted at that time. However, the main features of sudden warmings were established and may be summarised as

follows.

(a) The initial stratospheric circulation is a strong westerly flow around a cold pole with highest temperatures often in middle latitudes. This is then perturbed and a predominantly wavenumber one or wavenumber two pattern evolves.

(b) Warming first occurs at high latitudes in the upper stratosphere and continues for a period of 5 - 10 days during which increase in temperature is observed at successively lower levels. The westerly flow weakens in accordance with the thermal wind equation (equation 1.5) as the latitudinal temperature gradient is reduced.

(c) If the increase in high latitude temperature is large enough to reverse the temperature gradient in a sufficiently thick layer an easterly circulation is induced at high levels which descends with time.

(d) Polar cooling begins in the easterly régime and descends with it. This eventually leads to a near isothermal polar atmosphere with a slack circulation. Except in the case of the final spring warming, when summer easterlies are established, westerly flow gradually increases with the return of the normal meridional temperature gradients.

The terms 'major' and minor' are often attributed to sudden warmings in which the circulation reversal does and does not extend below 10 mb respectively.

6.2. THE THERMAL STRUCTURE OF THE 1973 WARMING

It has become apparent that warmings may be divided into two classes on the basis of the way in which polar warming takes place (Quiroz et al, 1975): there are those in which a wavenumber one disturbance occurs in the height field, such as the warmings of 1966, 1969-70, 1970-71, 1973 and 1974, and others which are

predominantly wavenumber two in character of which those of 1957, 1958 and 1963 are examples. The former are associated with the merging of an eastward travelling thermal system with a standing wave, often over eastern Siberia, and poleward movement of the intensified warm region. In the wavenumber two warming wave interaction results in two warm centres converging on the pole from opposite sides. Interactions of quasi-stationary waves with travelling and transient waves do not always lead to sudden warmings however. Madden (1975) demonstrated that a 1 - 3 week cycle of eddy heat transport in eight years of 30 mb data was largely explained by quasi-stationary - transient wave interaction. Wave interactions must therefore be regarded as normal wintertime features which only occasionaly develop into sudden warmings.

The warming during late January 1973 was of the wavenumber one type. Quiroz et al (1975) gave a brief account of the development of this warming using satellite radiance data, but they were not able to draw conclusions regarding the vertical propagation of the warming. We shall endeavour to give a fuller description of this and other observable features of the event in this chapter.

The movement of thermal centres in the second half of January 1973 at 50 and 10 mb is illustrated in Figure 6.1. At 50 mb a warm centre moved east close to latitude 50^{o}N and merged with a quasi-stationary high temperature region near 150^{o}E. Slow poleward movement of the composite system followed. Figure 6.2, which shows the maximum temperatures in the travelling and quasi-stationary warm cells, indicates that significant rise in temperature at 50 mb did not occur until just before the centres merged. Although both the travelling and quasi-stationary warm regions were initially observable at 10 mb they did not merge at

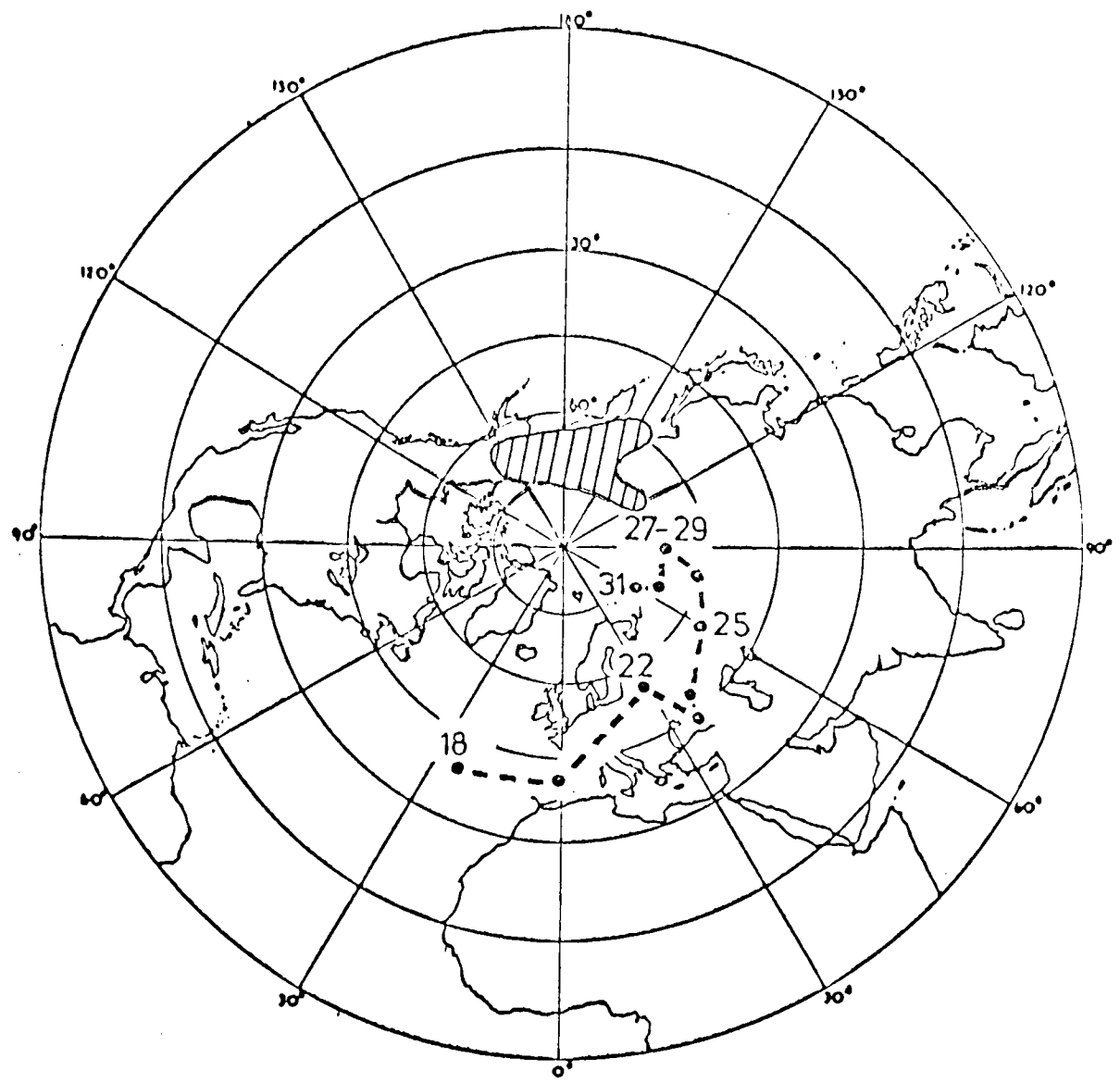
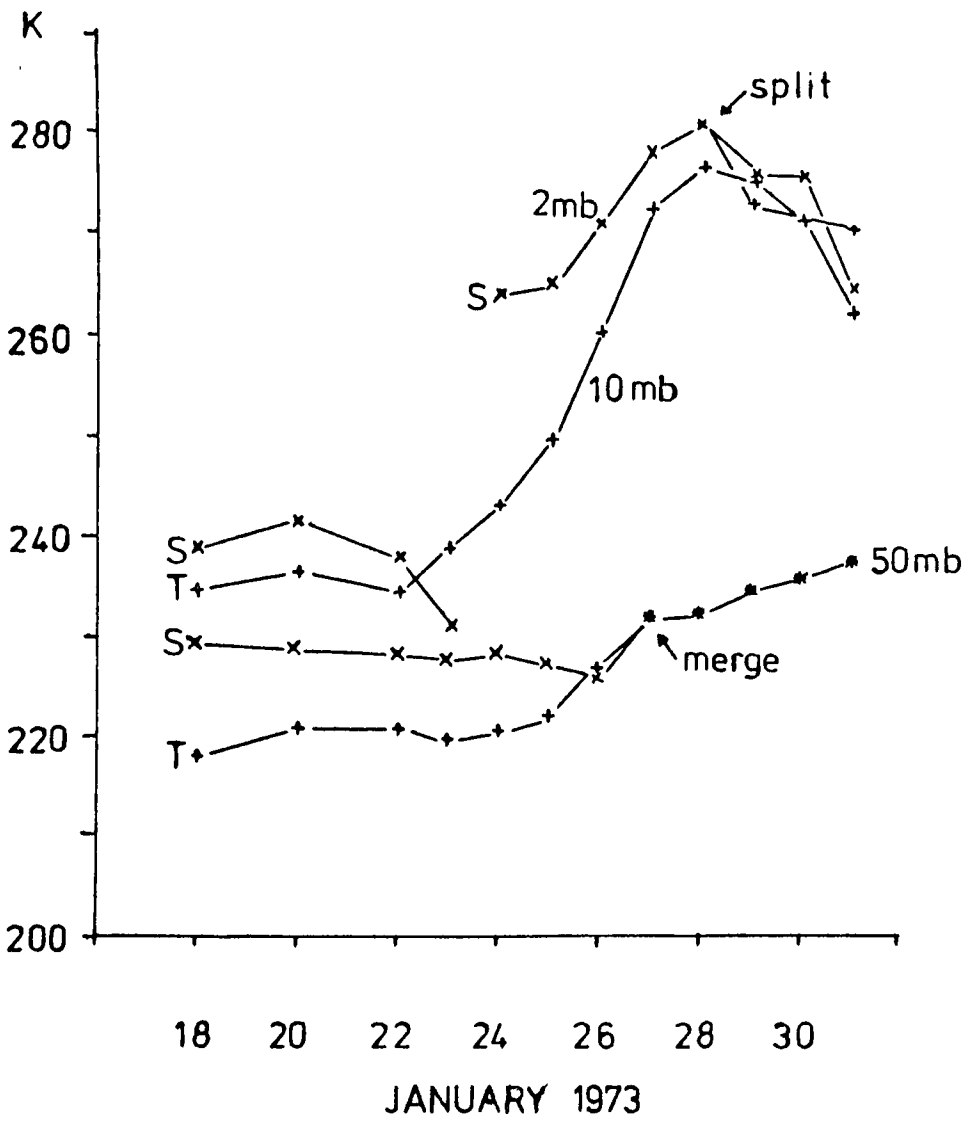


Figure 6.1. Movement of regions of high temperature at 10 mb (upper) and 50 mb (lower). Hot centres are denoted by circles plotted every two days prior to 22.1.73 and every day thereafter. The shaded region indicates the position of a quasi-stationary warm cell



JANUARY 1973

Figure 6.2. Plots against time of the maximum temperatures observed in the travelling (T) and quasi-stationary (S) warm cells at 2, 10 and 50 mb during the sudden warming. There was no marked travelling thermal system at 2 mb

this level. The latter cell relaxed while the eastward moving system intensified steadily from 22.1.73 onwards and became quasi-stationary near 90°E by 27.1.73. At 2 mb (not illustrated) one cell of an asymmetric wavenumber two thermal wave moving only slowly eastwards declined and amplification of the remaining warm region near 40°E led to a marked quasi-stationary wavenumber one pattern by 25.1.73. The development of the wavenumber one pattern at all three levels is illustrated more clearly in Figure 6.3 in which the departures of temperature from the zonal mean around 52°N, on 20, 22 and 25.1.73, and around 72°N, on 27.1.73, are given for each level. The latitudes chosen were close to the regions of highest temperature on those days. It is clear from Figure 6.1 that the travelling warm centres at 10 and 50 mb were associated with the same wave disturbance and since rapid warming occurred earlier at the 10 mb level it is possible, contrary to the theory outlined above, that the quasi-stationary warm region near 180°E did not play a significant rôle in the development of the warming. It is also evident that the intensification of the warm cell near 40°E at 2 mb was linked with the amplification further to the east at 10 mb.

An interesting feature following the period of maximum temperature, not shown in the figures, was the re-establishment at 10 mb of a warm region near 180°E although in mid-, rather than high, latitudes. The continued warming below in this area may have extended upwards to affect the middle stratosphere after cooling had begun at higher levels.

6.3. THE PROPAGATION OF THE WARMING

The question of whether the warming propagates downwards or upwards has often given rise to confusion. Hirota (1967) argued that the apparent downward propagation of warming observed from

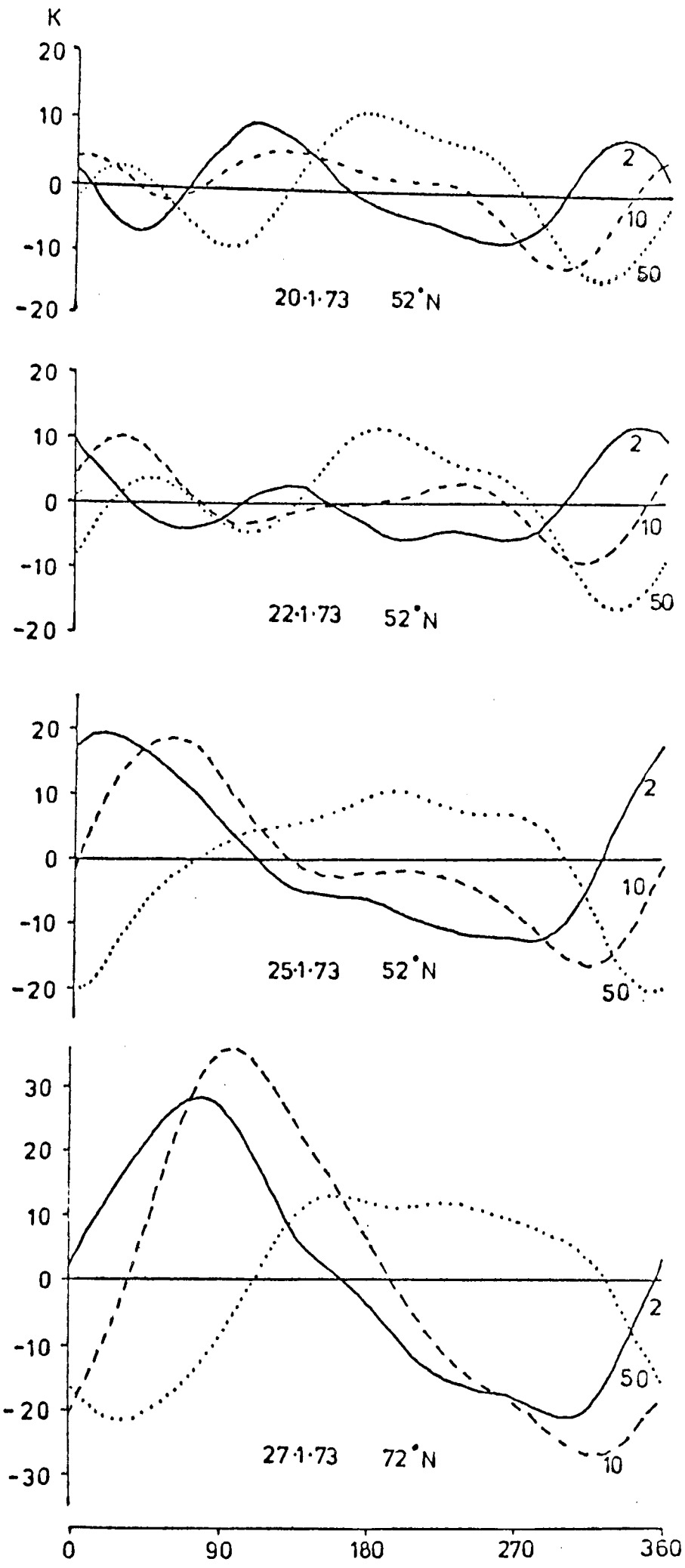


Figure 6.3. Departure of temperature from the zonal mean around 52°N on 20, 22 and 25.1.73 and around 72°N on 27.1.73 at 2, 10 and 50 mb

a fixed location results simply from the westward movement of a wave tilting westwards with height. By analysing the sudden warmings of January 1958, January 1963 and March 1965 he concluded that the warming propagates upwards and that the event occurs in two stages: an intensification stage in which the region of maximum rate of temperature increase in a westward tilting stationary wave moves upwards, and a migratory stage characterised by westward movement of the whole system. The distribution of temperature rise observed during the 1963 warming is reproduced in Figure 6.4a. A similar analysis of the 1973 warming, extending to the upper stratosphere but excluding the upper troposphere, has been made. Figure 6.4b shows the centres of maximum rate of temperature change at 60^{o}N in the period 22.1.73 - 2.2.73. The level of maximum warming moved steadily downwards from 22 - 26.1.73 giving a temperature rise of over 13K day^{-1} near 10 mb at 80^{o}E on 25.1.73. Simultaneous cooling occurred at similar longitudes in the high stratosphere and at similar levels between 280^{o} and 330^{o}E. Warming also occurred near 50 mb below the cooling in the middle stratosphere. During the declining stage of the warming event the same longitudes were associated with marked temperature change of the opposite sign, indicating the quasi-stationary nature of the disturbance. The pattern of temperature change at 72^{o}N was similar although much larger values were obtained, a maximum of over 20K day^{-1} being observed near 90^{o}E at 10 mb on 25.1.73. These results are in direct contrast with the conclusions drawn by Hirota (1967) given above. Of the three warmings Hirota studied two were wavenumber two disturbances while that of March 1965 was wavenumber one. The temperature change analysis of the latter warming (Figure 6.4c) does show the appearance of a new warming centre at 10 mb

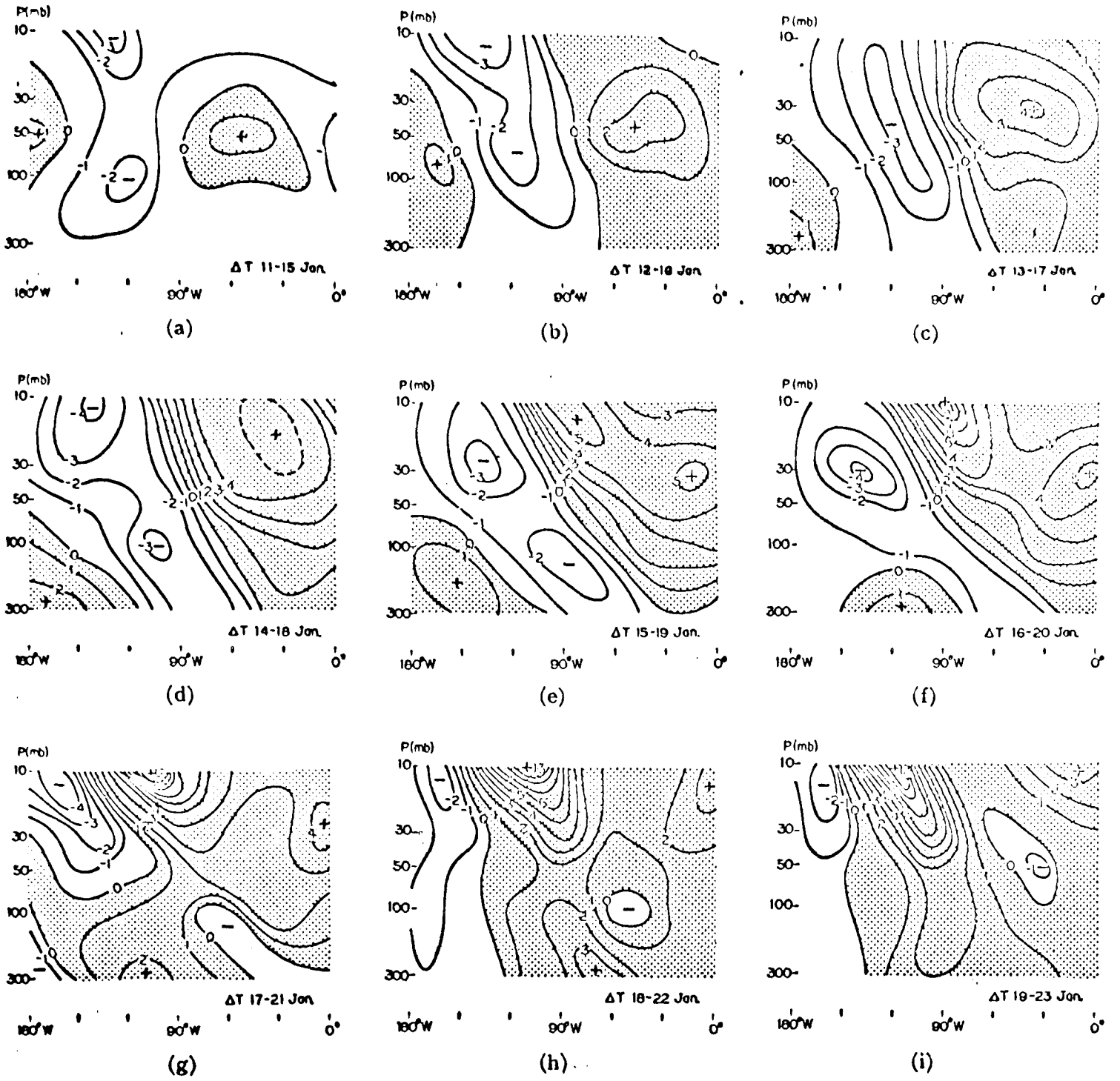


Figure 6.4a. Vertical cross sections of temperature change averaged over four days along 50°N during the period of the 1963 sudden warming ($^{o}C\,day^{-1}$). Shaded regions denote warming. (Taken from Hirota, 1967)

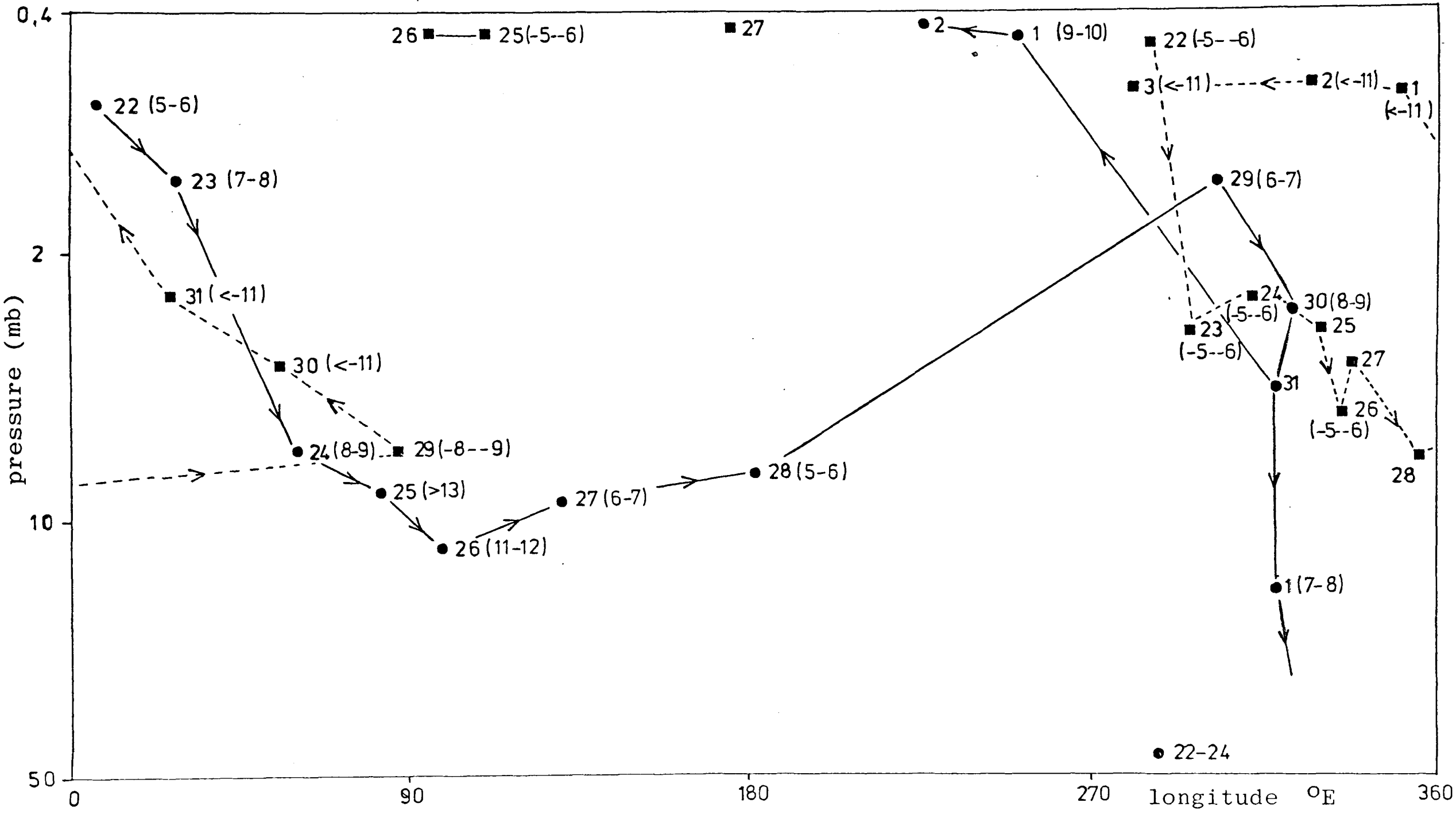


Figure 6.4b. Movement of centres of maximum rates of change of temperature at 60°N. Circles denote warming and squares cooling. The date in January and early February, and the rate of temperature change in K day $^{-1}$ (if greater than 5), are given alongside

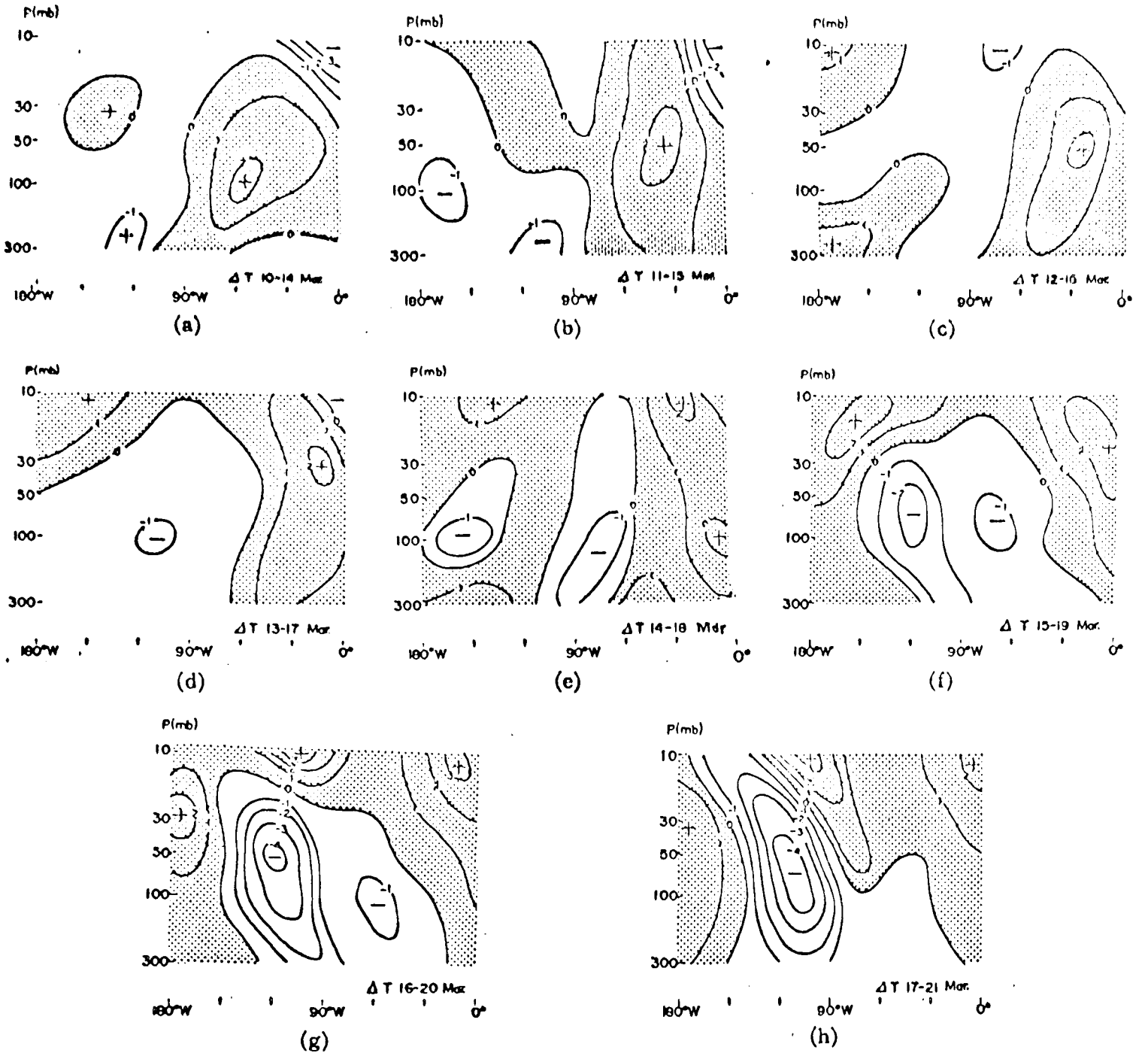


Figure 6.4c. As Figure 6.4a but for the warming period during March 1965. (Taken from Hirota, 1967)

around 18.3.65 which appears to have propagated downwards from the upper stratosphere. It is thus possible that downward propagation of warming is a normal feature of a wavenumber one warming. Alternatively, upward propagation in the troposphere and lower stratosphere and downward propagation from the upper stratosphere may be common to all warmings. It is clearly necessary to analyse the whole depth of the stratosphere to determine the direction of propagation of the major centre of warming.

Temperature change is, of course, only one aspect of the sudden warming event, and it is possible to describe the propagation of the disturbance in terms of other meteorological quantities. Figure 6.5a shows vertical time sections of the amplitude of wavenumbers one and two geopotential height during winter months, adapted from Muench (1965) and Hirota and Sato (1969). Clearly there is evidence of wave amplification occurring first in the troposphere and propagating into the stratosphere. Similar sections for the 50 - 0.4 mb region during the 1973 warming (Figure 6.5b) show no evidence of upward propagation; indeed maximum wavenumber 1 geopotential at $68^{o}N$ occurred first in the upper stratosphere and about one day later at 10 mb. A similar observation, during a warming in November 1975, of almost simultaneous wave amplitude maxima at 10 and 500 mb was made by Quiroz and Nagatani (1976), suggesting the possibility of downward propagation or stratospheric feedback. A closer look at the wavenumber one geopotential maximum at the end of January 1964 (Figure 6.5a) does show that, above 30 mb, there was a tendency for maximum amplitude to occur earlier at higher levels.

Finally we compare vertical time sections of kinetic energy

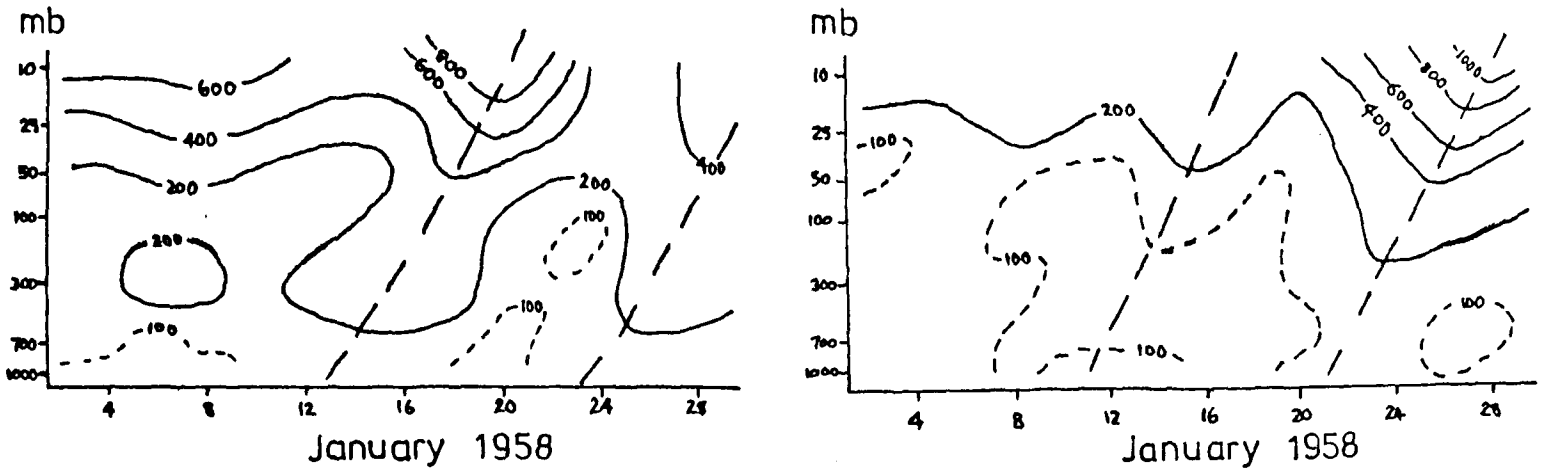
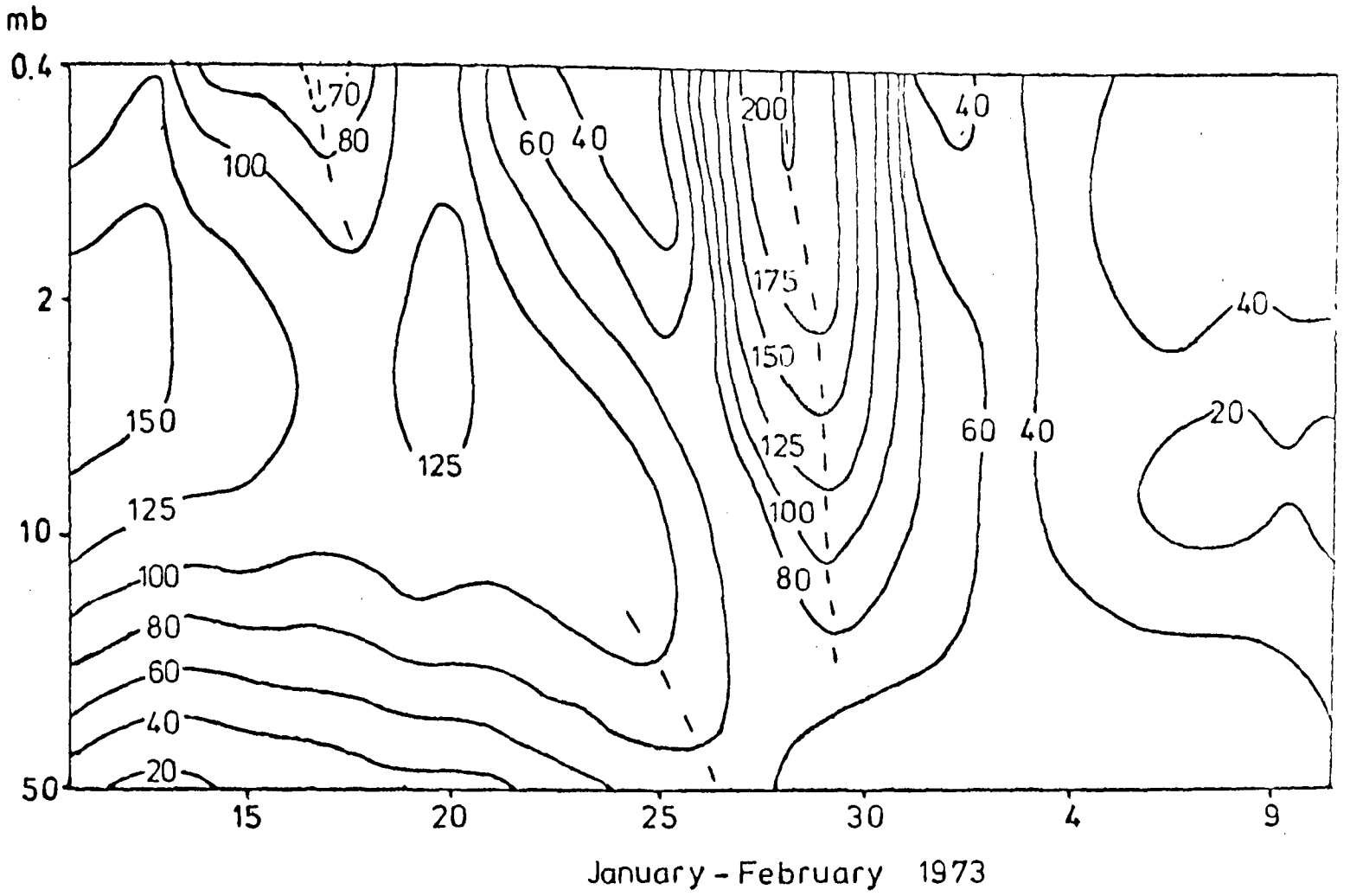


Figure 6.5a. Amplitude of wavenumber 1 (top left) and wavenumber 2 (top right) geopotential height (in metres) at 50°N during January 1958 (after Muench,1965), and similar sections for the amplitude of wavenumbers 1 (left) and 2 (right) during January 1963 (centre) and January - February 1964 (bottom) at 60°N (after Hirota and Sato, 1969)



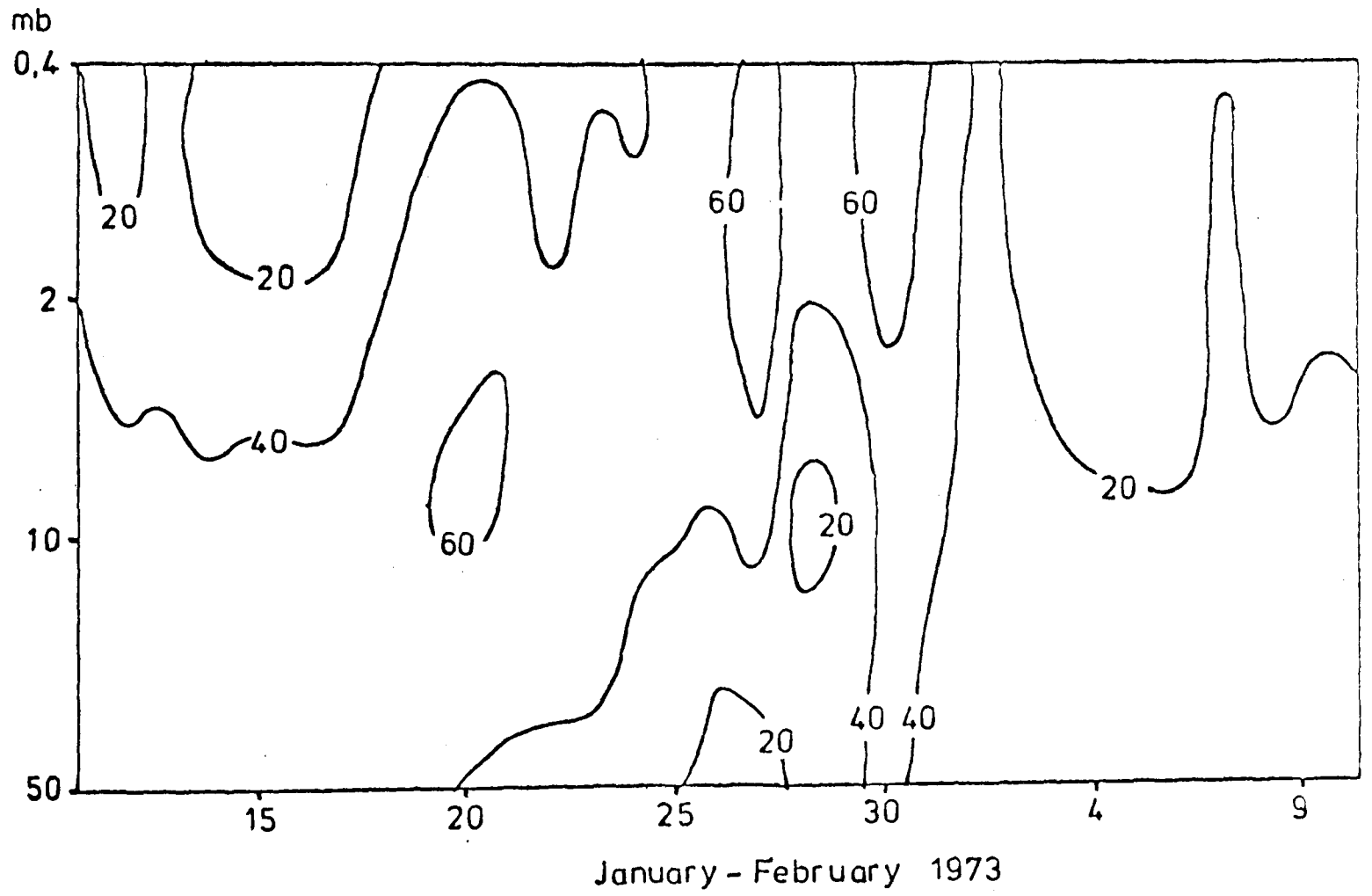


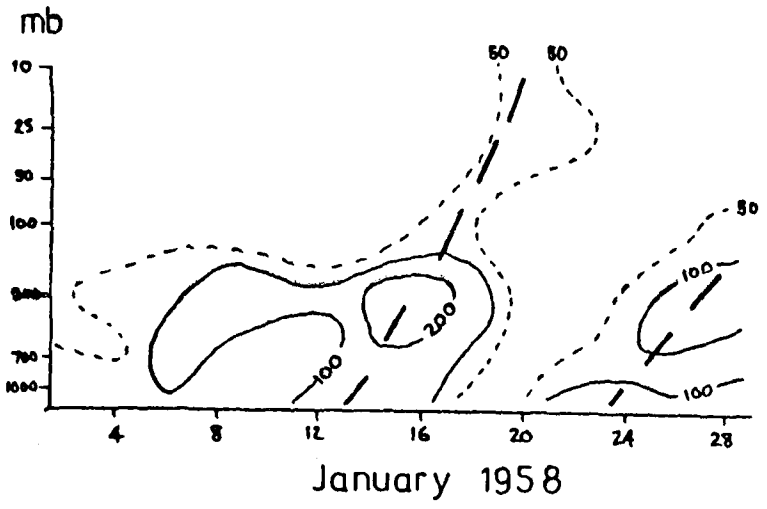
Figure 6.5b. Vertical - time sections of wavenumber 1 (upper) and wavenumber 2 (lower) geopotential height (in decametres) at 68°N during the period 11.1.73 - 10.2.73

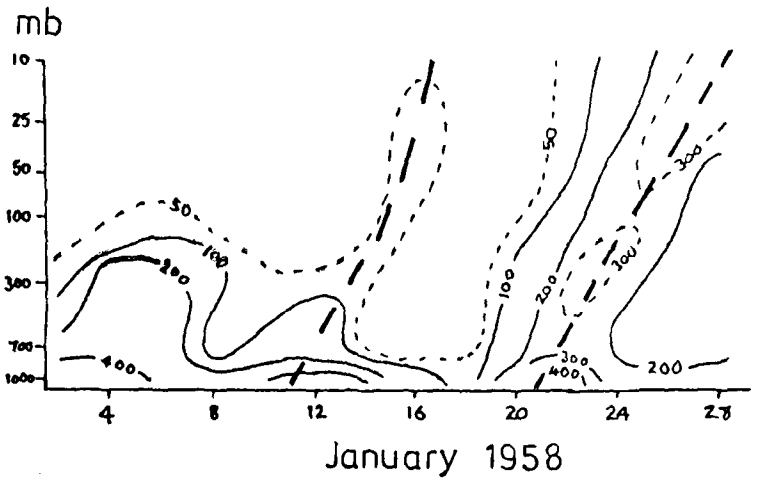
density. Once again there is good evidence in the literature for supposing that wave energy may propagate vertically from a tropospheric source with little attenuation (Figure 6.6a). However our results (Figure 6.6b) show that in the middle and upper stratosphere changes in kinetic energy occurred simultaneously at all levels prior to and after the 1973 warming event. Despite the fact that these observations do not reveal the existence of a tropospheric energy source during this warming there is no doubt that vertical propagation of energy via the pressure-work effect played a major rôle in the disturbance. Chapter 8 is devoted to a more detailed study of this important feature.

6.4. THE VERTICAL STRUCTURE OF WAVENUMBER ONE

The development of the 1973 warming is further illustrated by the changes in the vertical distribution of the amplitude and phase of the wavenumber one components of temperature and geopotential height. These are shown for 72°N in Figure 6.7, similar distributions occurring at neighbouring latitudes. Figure 6.8 demonstrates that these components describe the major changes in temperature and geopotential height in the middle and upper stratosphere. From Figure 6.7 it is seen that the temperature wave amplitude had a wavelike vertical structure with a maximum descending from near 1 mb on 23.1.73 to 10 mb on 29.1.73. Large amplification occurred between 25 and 27.1.73 giving an almost fivefold increase at 10 mb. The geopotential wave amplitude suffered a similar explosive amplification, increasing fourfold between 24 and 28.1.73 at 1 mb. The connection between the amplification of the temperature wave at 10 mb and that of geopotential at 1 mb is evident in Figure 6.8.

The distribution of wave phase shows that the geopotential





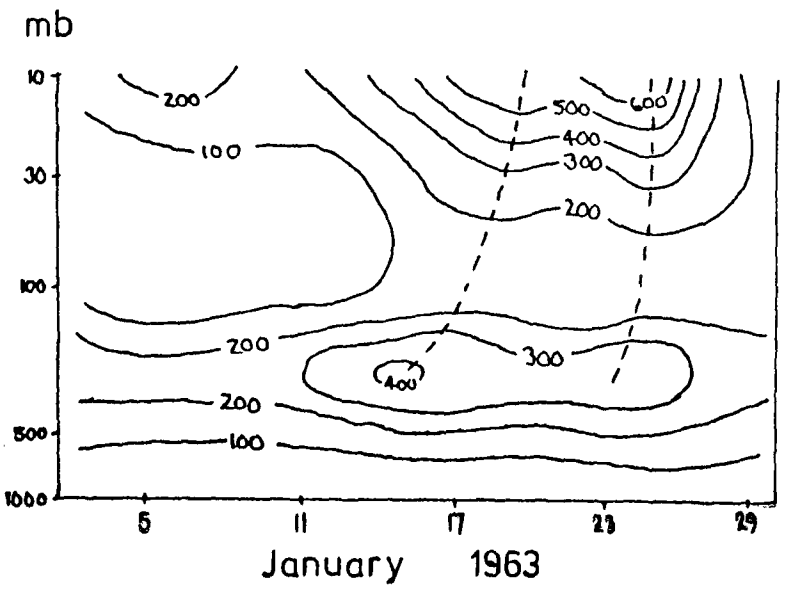
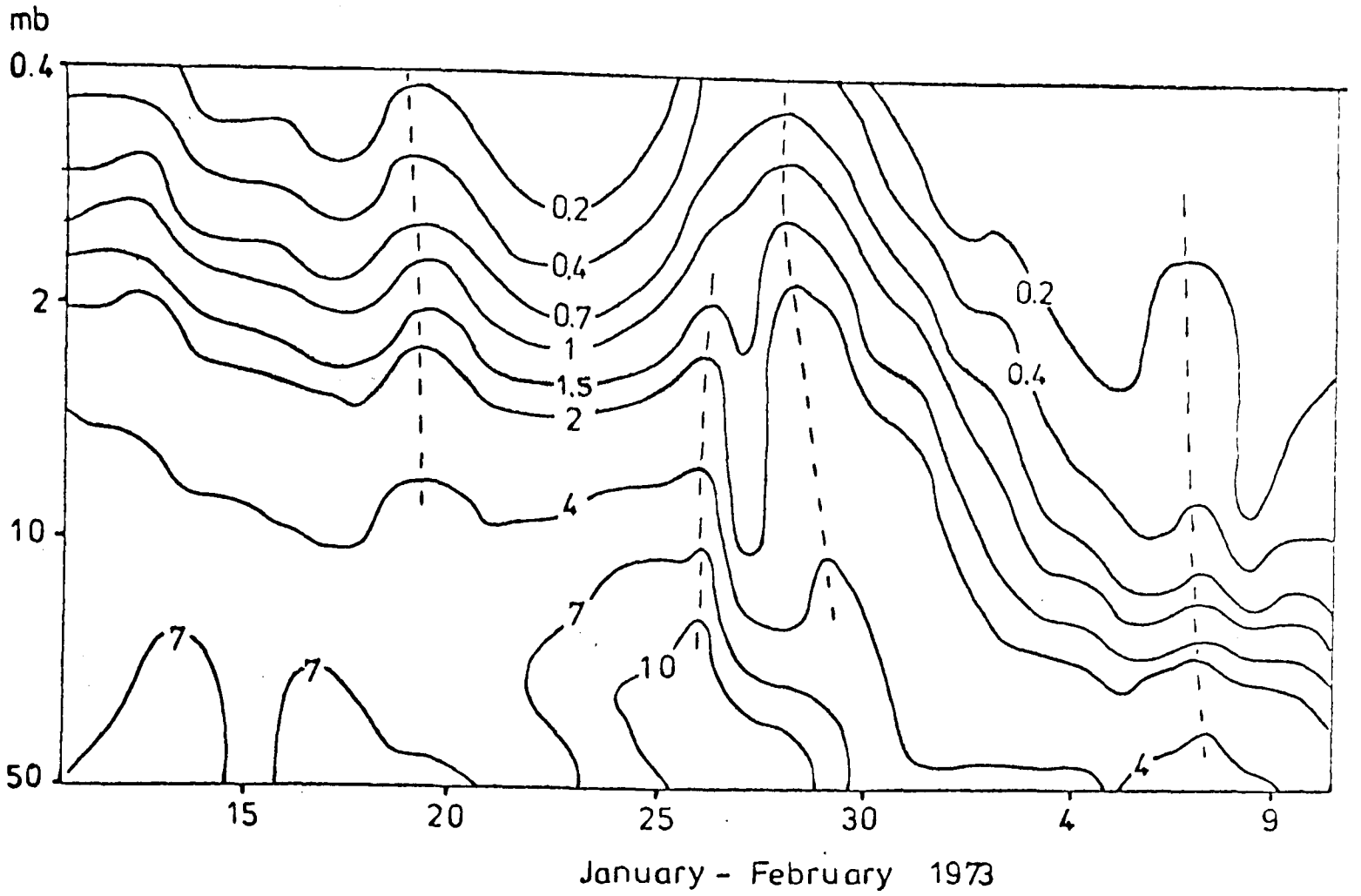


Figure 6.6a. Vertical - time sections of the kinetic energy density of the wavenumber 1 (top left) and wavenumber 2 (top right) meridional motion at 50^{o}N during January 1958 (10^{-1} erg cm^{-3}) (after Muench,1965), and the eddy kinetic energy averaged over 20 - 80^{o}N during January 1963 (10^{4}erg cm^{-2} mb^{-1}) (after Perry,1967)



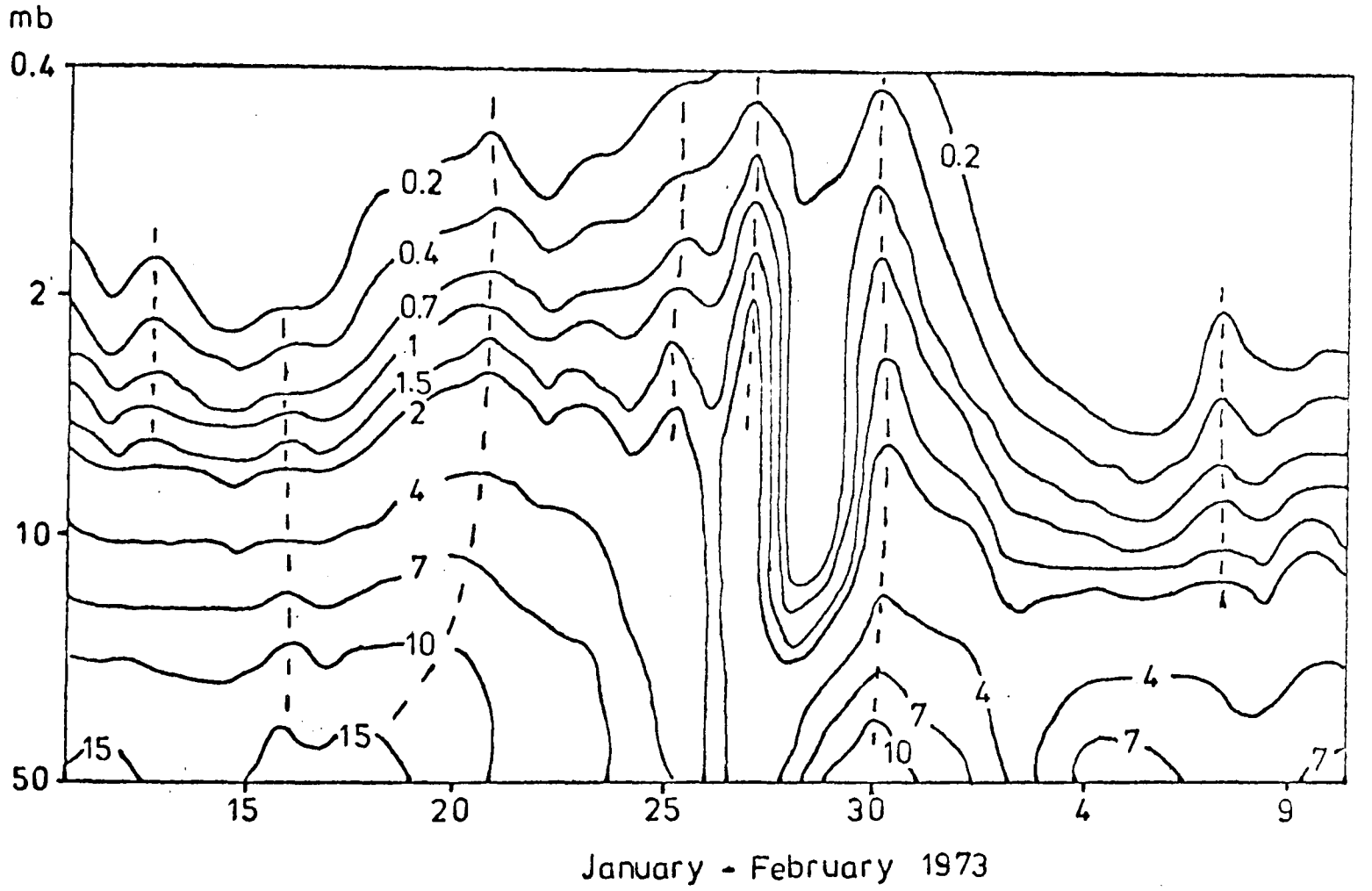


Figure 6.6b. Vertical - time sections of wavenumber 1 (upper) and wavenumber 2 (lower) eddy kinetic energy density at 68°N during the period 11.1.73 - 10.2.73 ($J\,m^{-3}$)

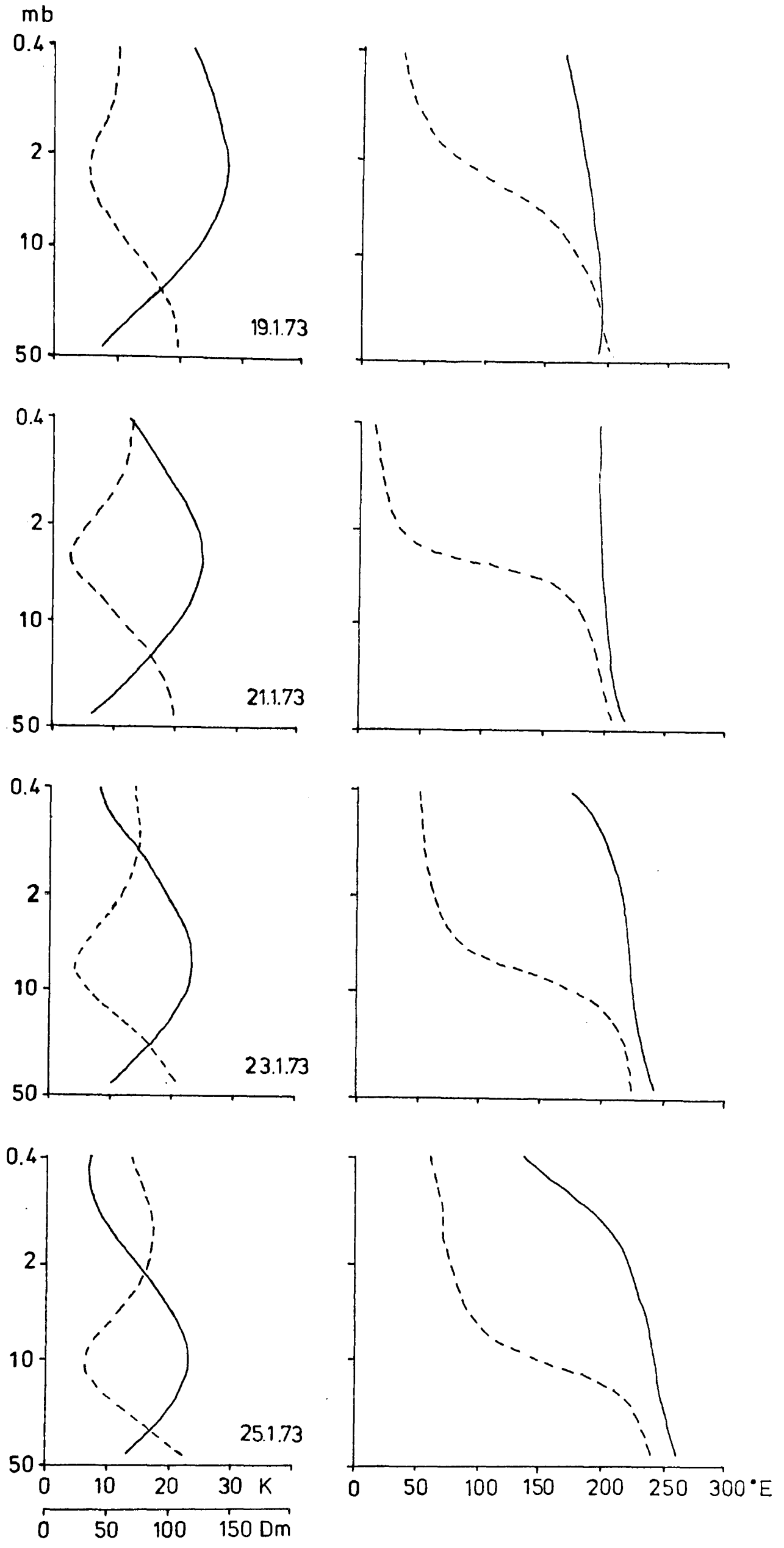


Figure 6.7. Vertical distributions of wavenumber 1 temperature (----) and geopotential height (——) amplitude (left) and phase (of maximum) (right) at 72°N during the period 19.1 - 3.2.73

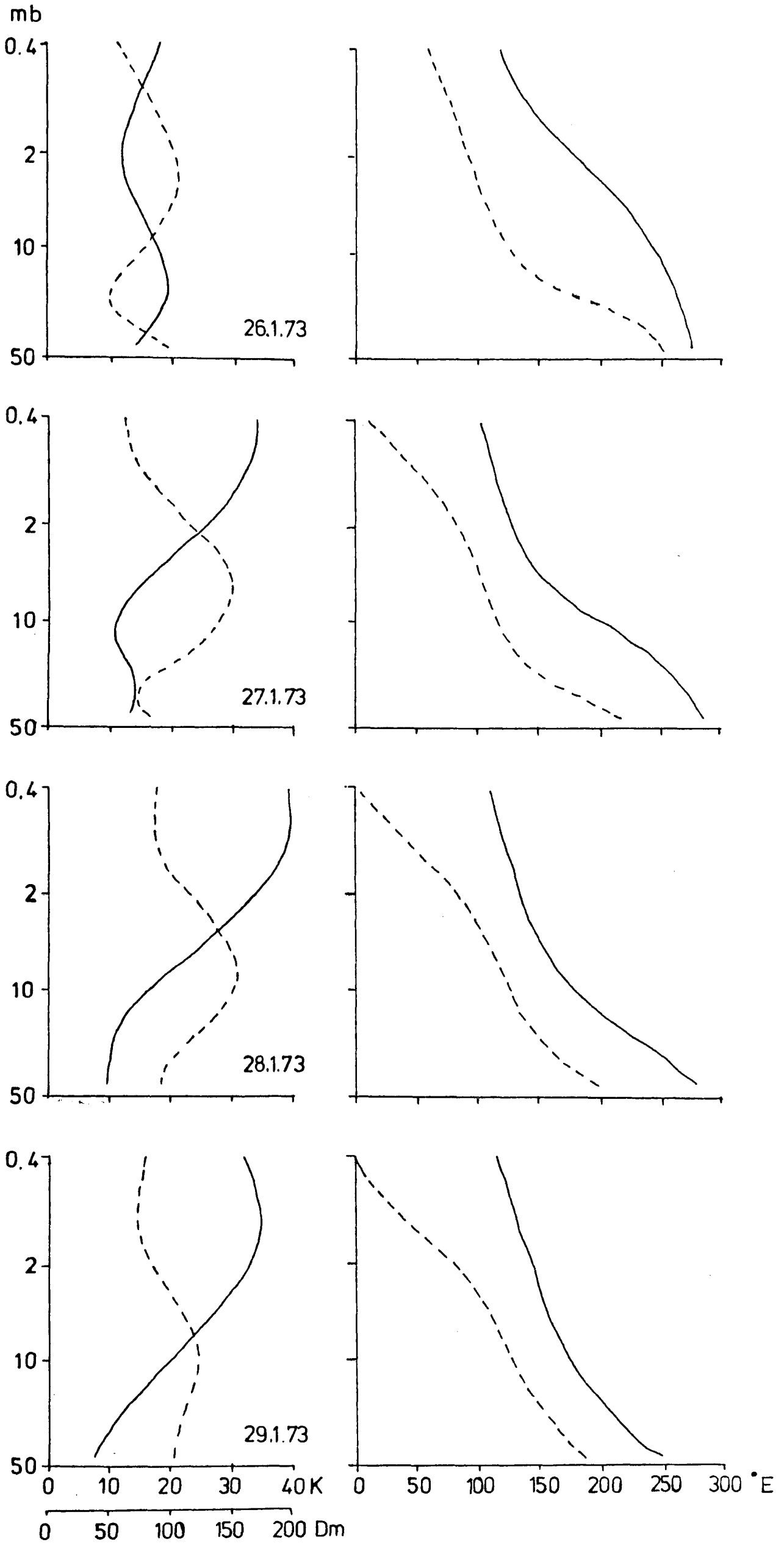


Figure 6.7 (continued)

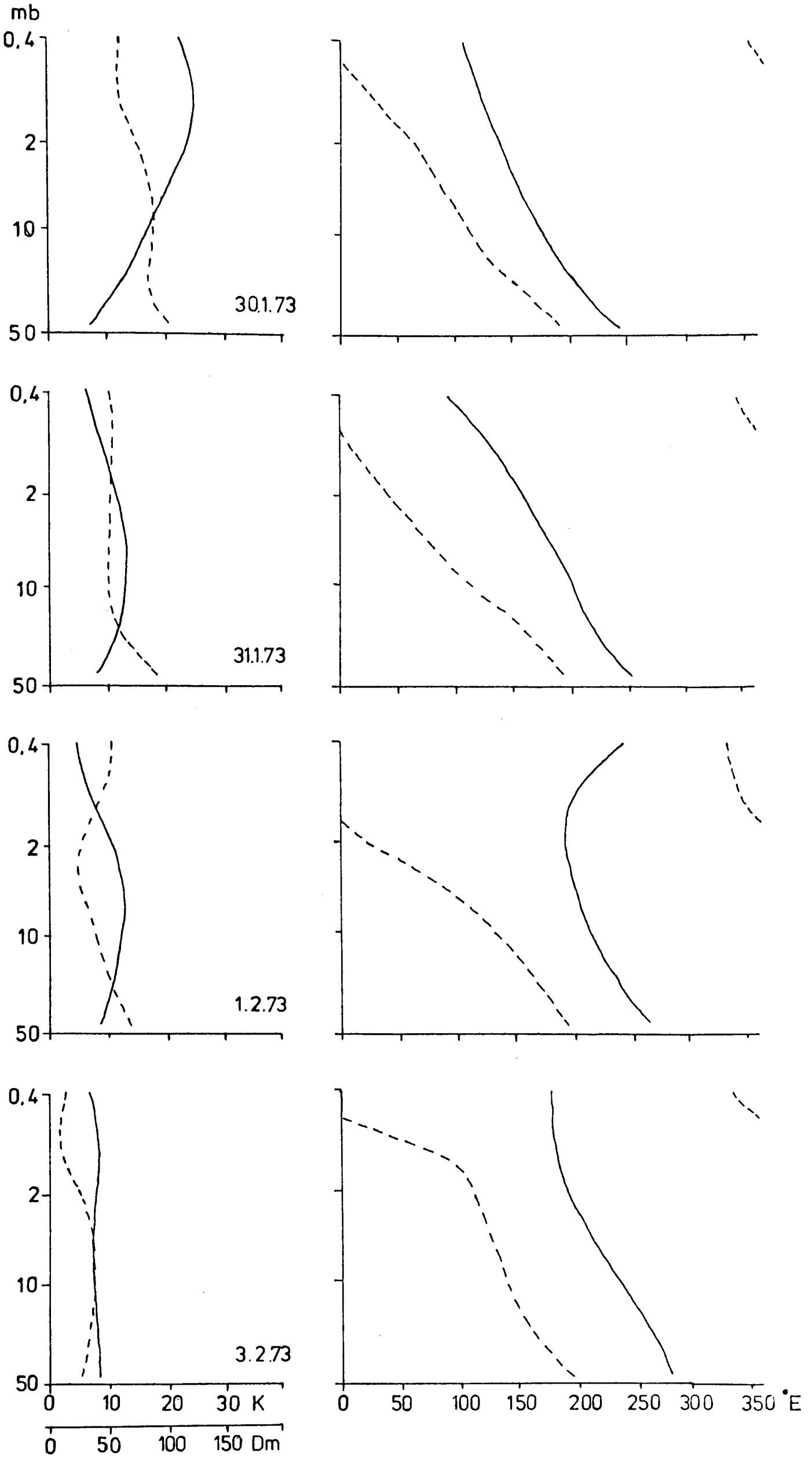


Figure 6.7 (continued)

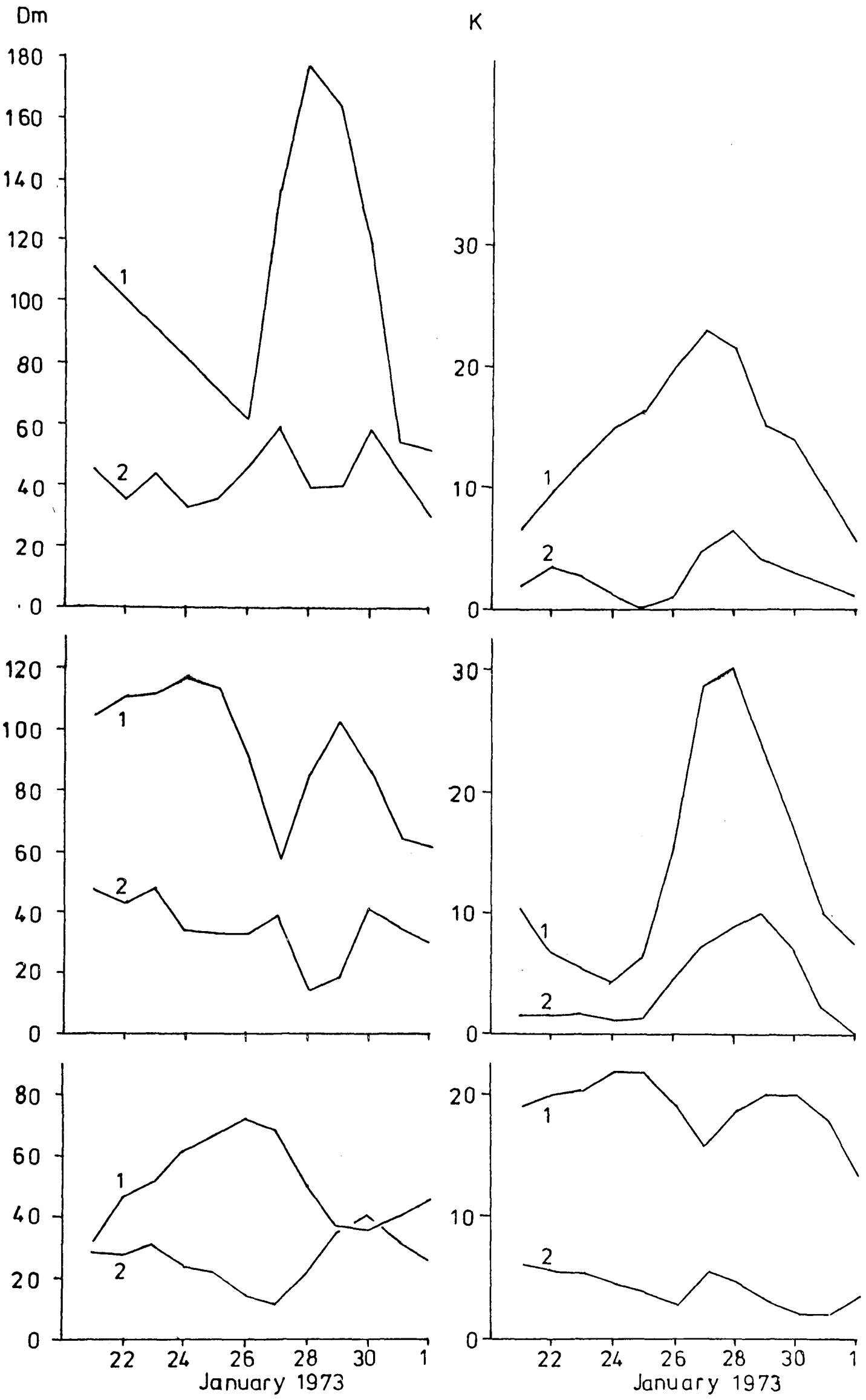


Figure 6.8. Wavenumbers 1 and 2 geopotential height (left) and temperature (right) at 72°N plotted against time for 2 mb (top), 10 mb (centre) and 50 mb (bottom)

ridge was to the east of the temperature ridge; this is typical of northern hemisphere planetary waves in winter and indicates westward tilt with height. Eliassen and Palm (1961) deduced that in a steady, quasi-geostrophic, adiabatic, stationary wave superimposed on a mean westerly flow the upward energy flux is proportional to the northward heat transport. This is a maximum when the geopotential ridge is $\frac{1}{4}$ wavelength east of the temperature ridge. Although the wave illustrated in Figure 6.7 was not steady we see that the optimum phase relation was approached on 26 and 27.1.73, and, as is described later, large vertical energy fluxes were observed at that time.

As previously noted by Muench (1965) maximum wave amplitudes were associated with phase slopes nearer to the vertical. This was particularly evident in the temperature wave. In addition there was a tendency for maximum geopotential amplitude to occur at levels where the temperature amplitude was a minimum, a consequence of the way in which layers of high temperature amplitude tend to be almost 180^{o} out of phase with similar over- or under-lying regions. The wave structure on 21.1.73 in Figure 6.8 is a good example. Hartmann (1976) showed that this is a feature of the southern hemisphere winter also. The author is not aware of any argument which has been put forward to explain this commonly observed structure.

6.5. THE ZONAL MEAN TEMPERATURE AND ZONAL WIND FIELDS

Although it has been shown that the rapid warming which occurred over Asia was associated with cooling on the opposite side of the globe, such cooling is usually not a very significant feature during the development of a sudden warming. The zonal mean temperatures therefore generally reflect well the warming process. Similarly the zonal mean zonal wind is a convenient

measure of the extent to which the warming destroys the normal polar vortex, even though near the pole the flow may be largely meridional in character with the geopotential "high" centred off the pole. In this section the 1973 sudden warming is described in terms of the evolution of the zonal mean temperature and zonal wind fields. The 2, 10 and 50 mb levels have been chosen to illustrate the different behaviour of the high, middle and lower-middle stratosphere. The rôles of the eddy heat and momentum fluxes are also discussed. In the accompanying figures the eddy fluxes have been plotted as $[v'T'] \cos\phi$ and $[u'v'] \cos^2\phi$, where ϕ is latitude, since these are the quantities whose divergence effects local changes in the zonal mean temperature and zonal wind respectively.

2 mb temperature and zonal wind

Figure 6.9 shows the latitudinal distribution of the zonal mean temperature and eddy heat flux for selected days during the warming event. Throughout the period 11.1.73 - 25.1.73 high latitudes were warmer than mid-latitudes, a feature which satellite observations have shown to be common in the winter upper stratosphere. From 26 - 29.1.73 a rise from 246K to 260K occurred at $80^{\circ}N$ in response to increased eddy heat flux convergence in high latitudes. An equally rapid fall in temperature through radiative loss and adiabatic cooling by vertical motion followed when the eddy heat flux weakened. Despite heat flux divergence south of $60^{\circ}N$ temperatures in mid-latitudes remained constant indicating compensatory adiabatic heating in the sinking branch of the mean meridional circulation.

Figure 6.10 shows the mean meridional circulations implied by heat balance on four days during the warming. By 2.2.73 the high latitude cell had reversed above 2 mb to offset radiative

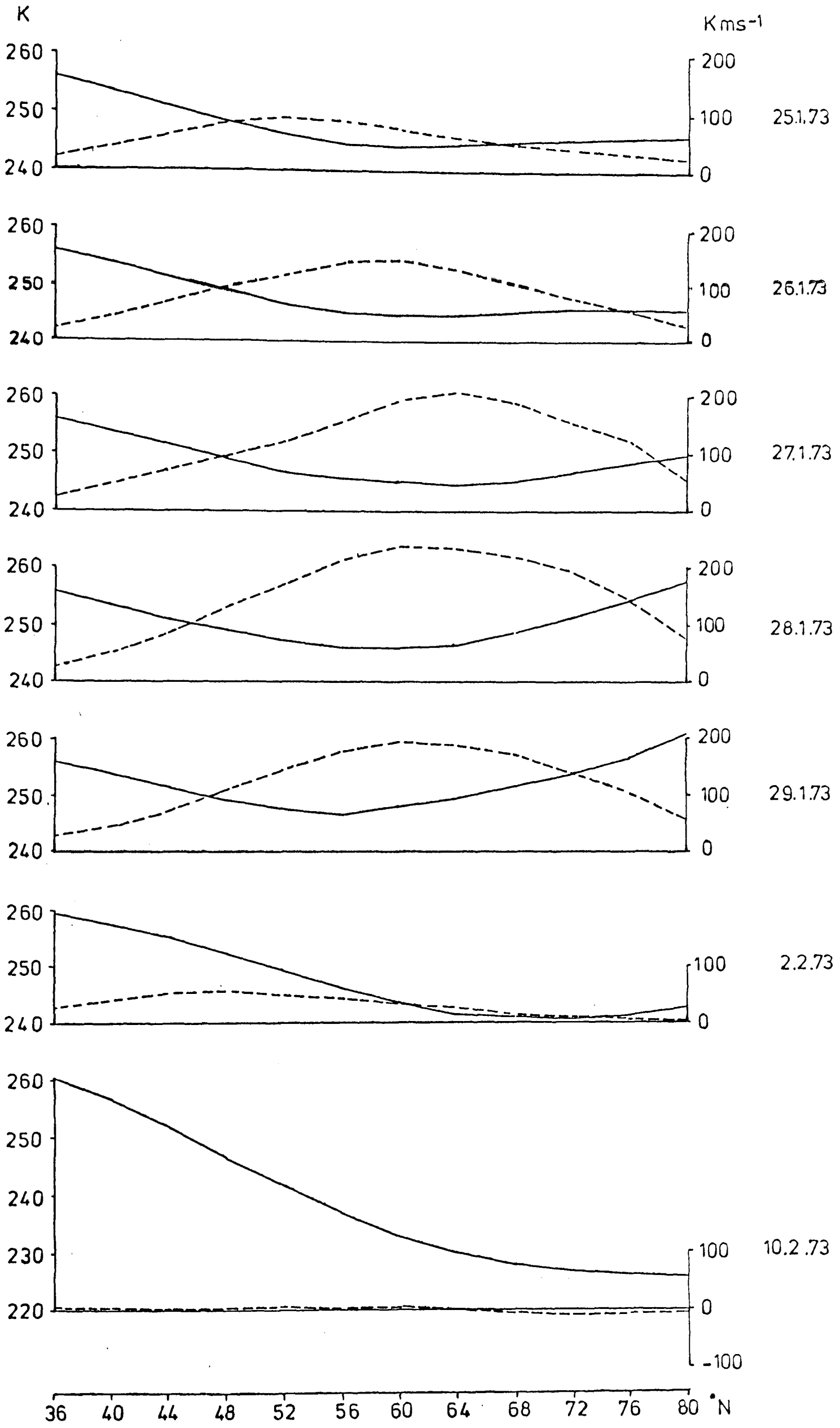
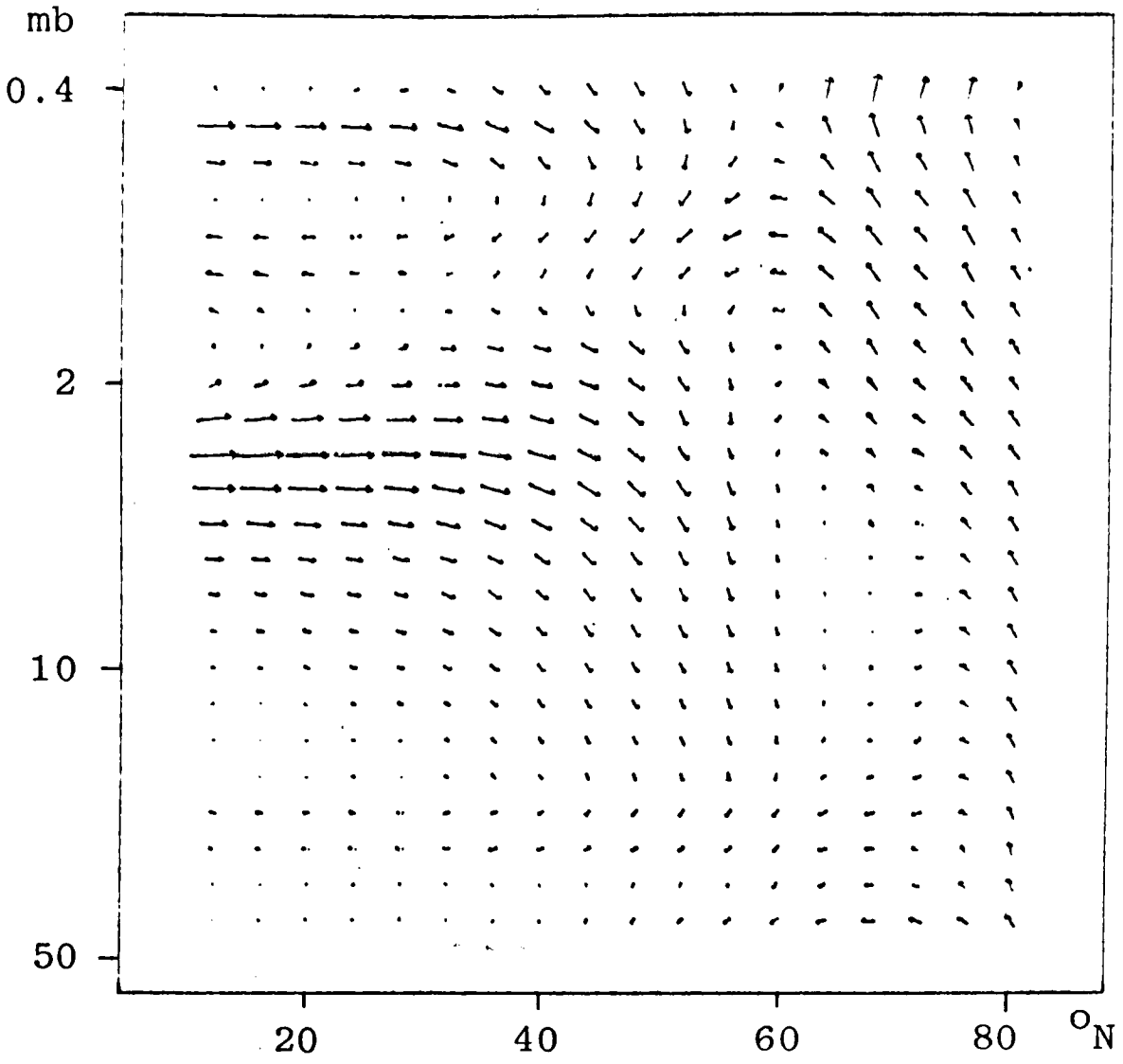
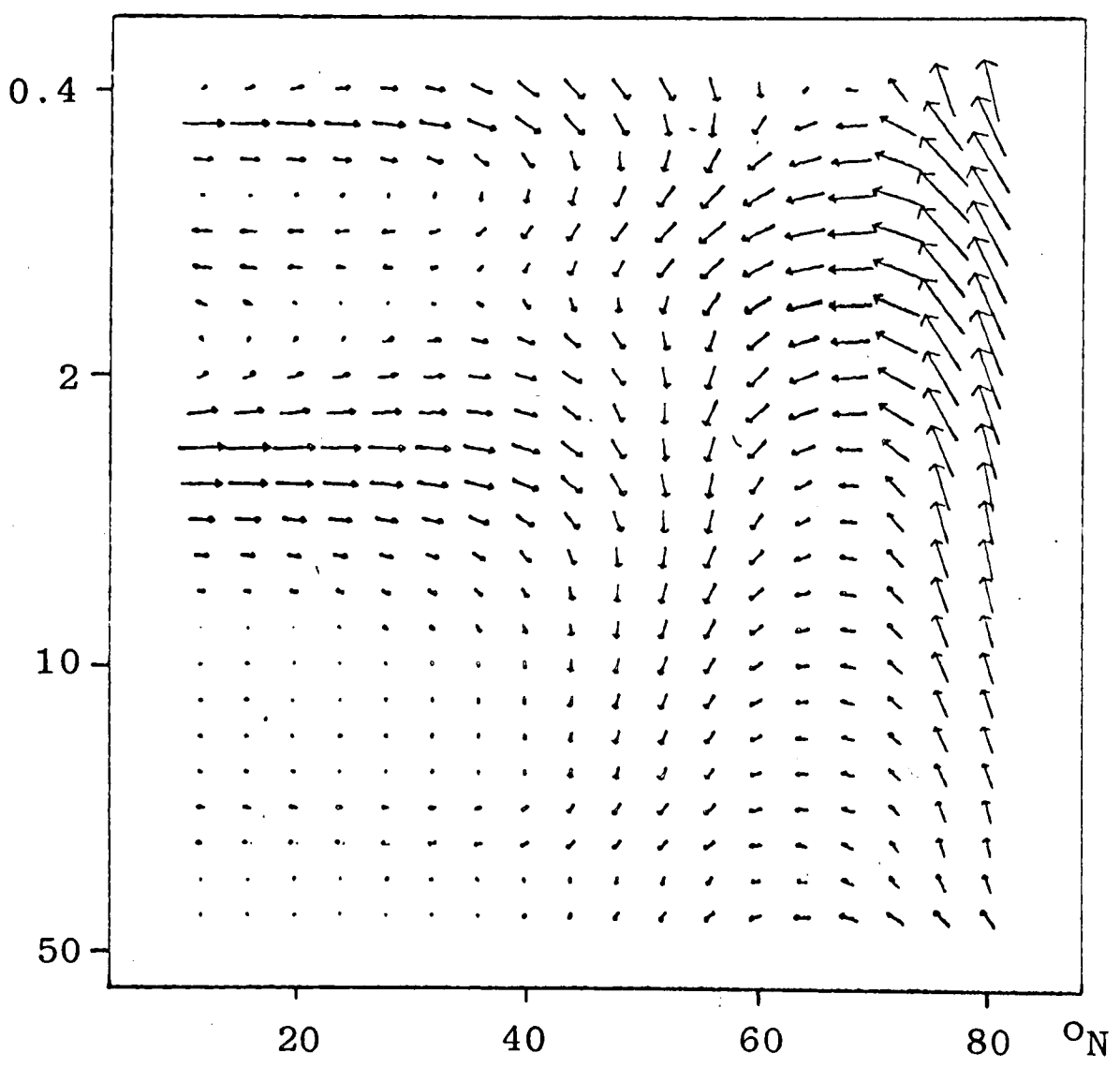


Figure 6.9. Latitudinal distributions of zonal mean temperature [T] (——) and eddy heat flux plotted as [v'T'] *cos* ϕ (- - -) at 2 mb for selected days during the sudden warming



26.1.73



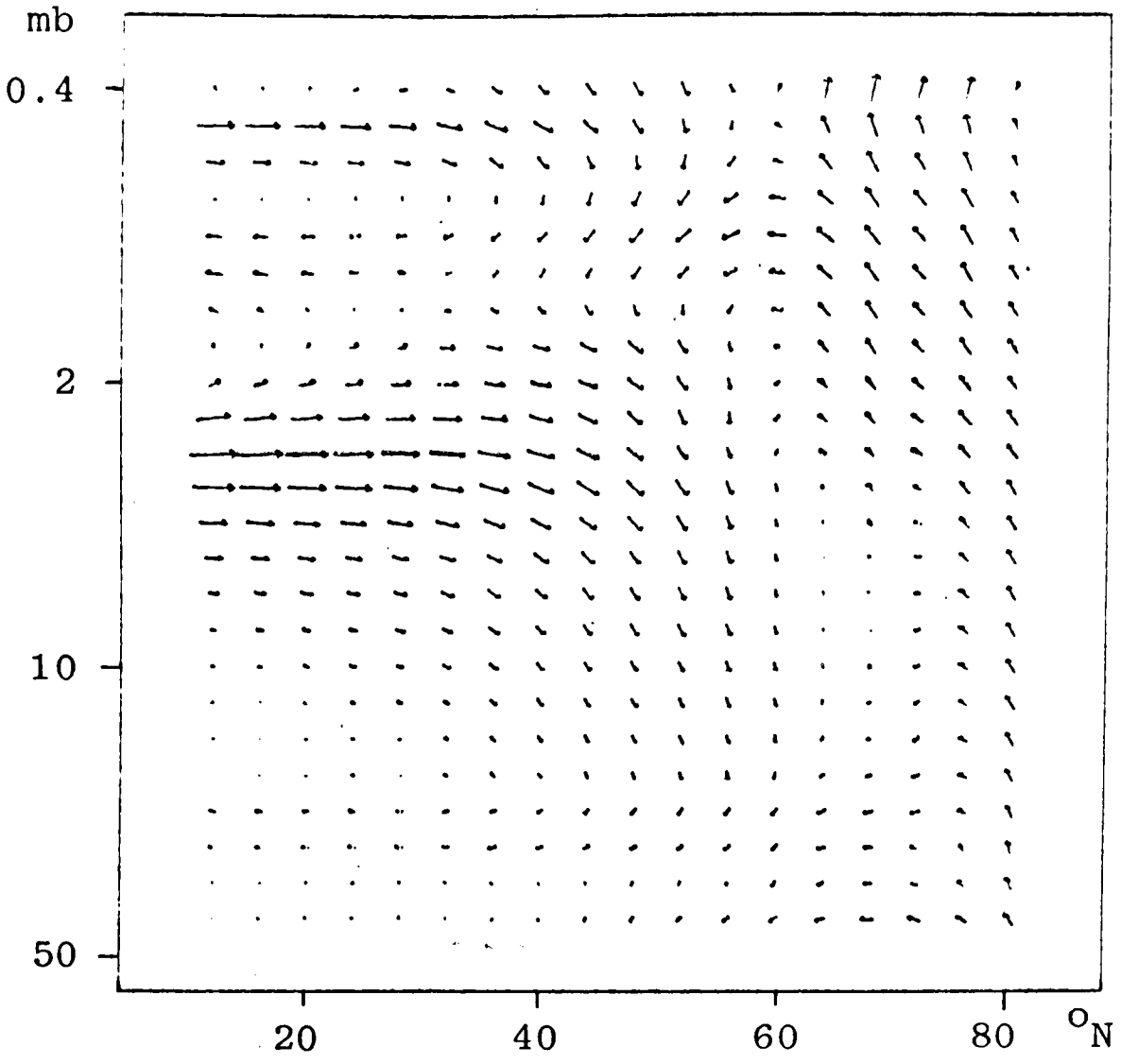
28.1.73

horizontal (isobaric) scale:

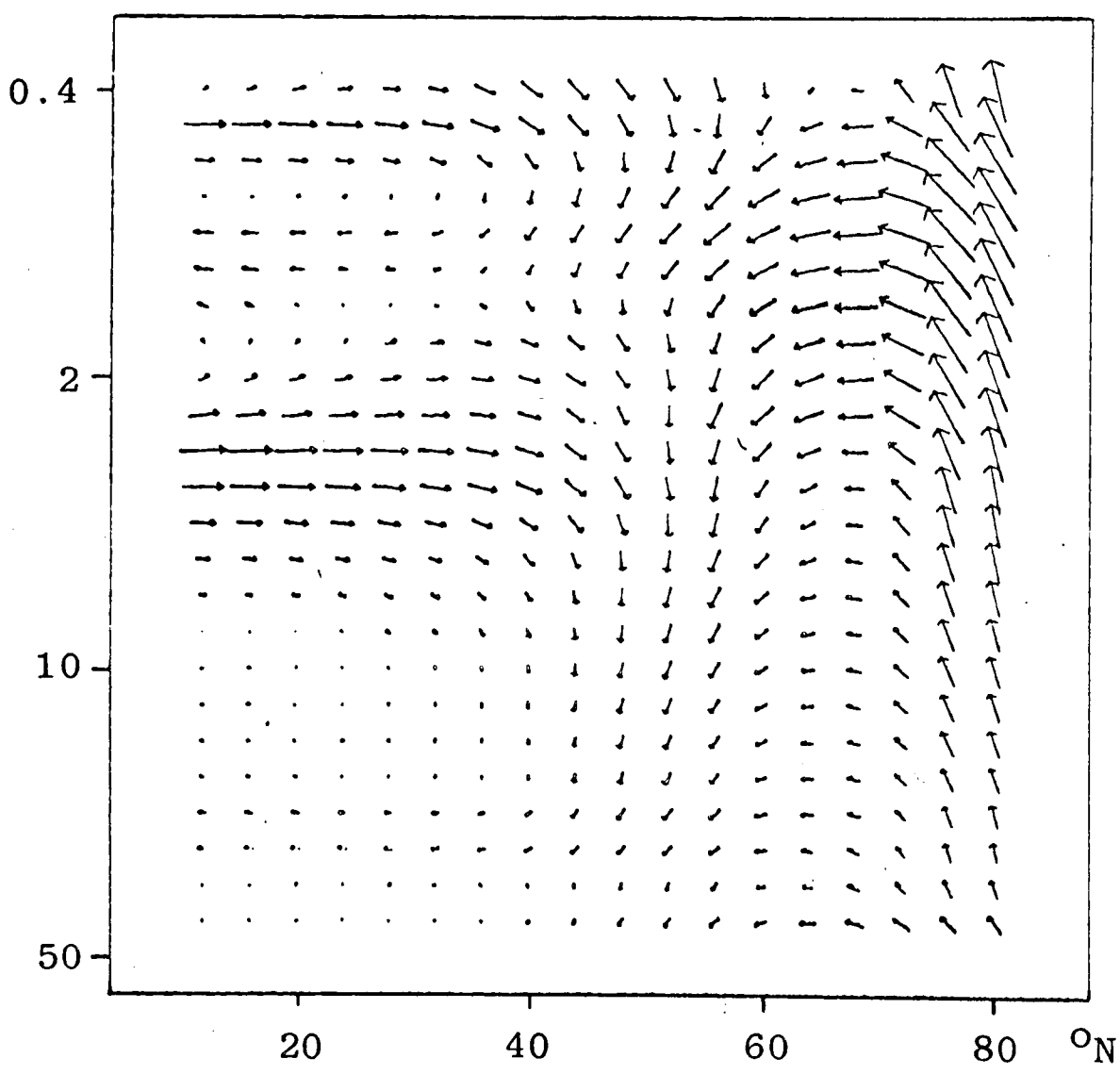
$\longmapsto$ 10 ms^{-1}

vertical scale:

I $\frac{10}{220}$ ms^{-1} = 4.5 $cm\,s^{-1}$



26.1.73



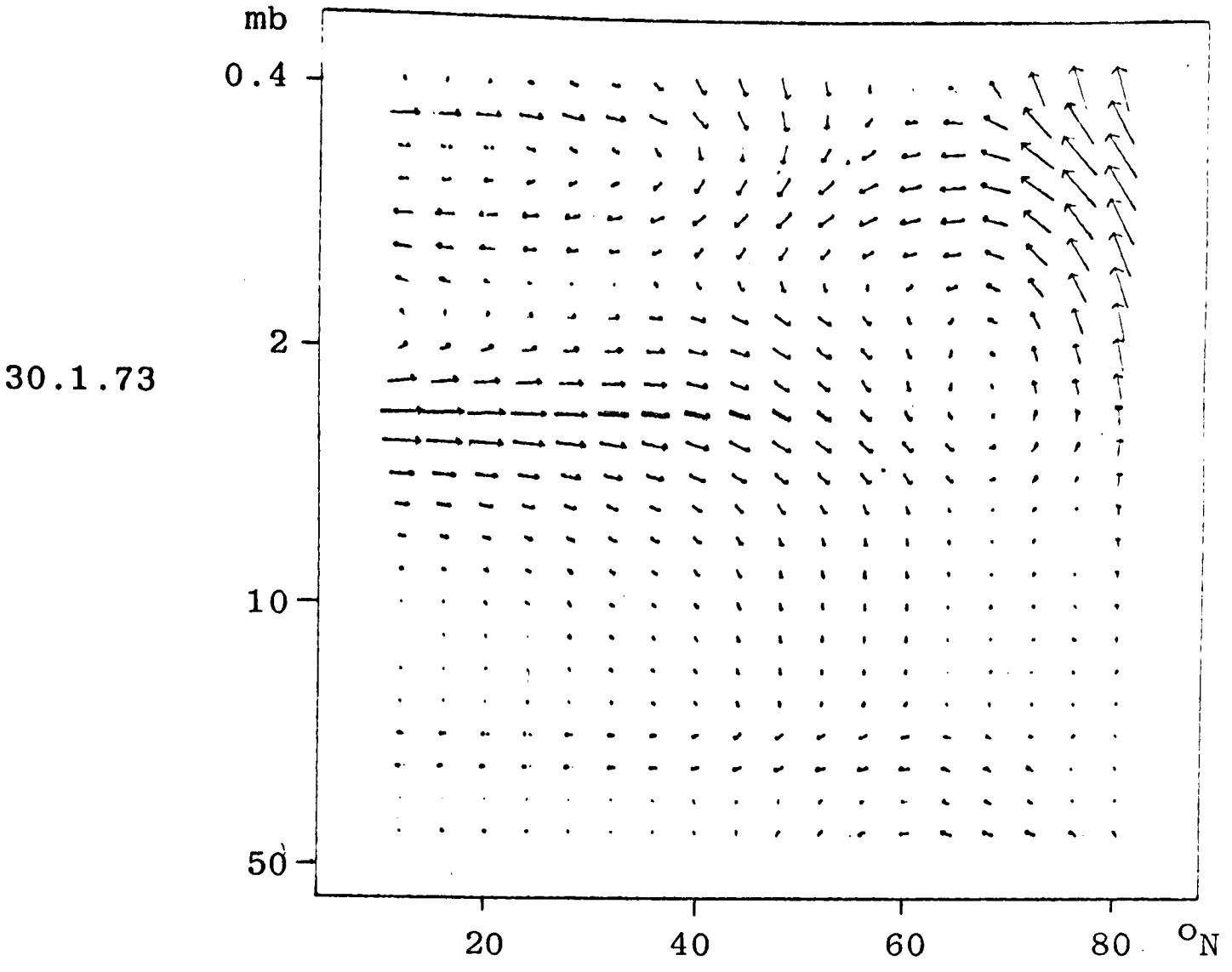
28.1.73

horizontal (isobaric) scale:

⟷ 10 ms^{-1}

vertical scale:

I $\frac{10}{220}$ ms^{-1} = 4.5 $cm\,s^{-1}$



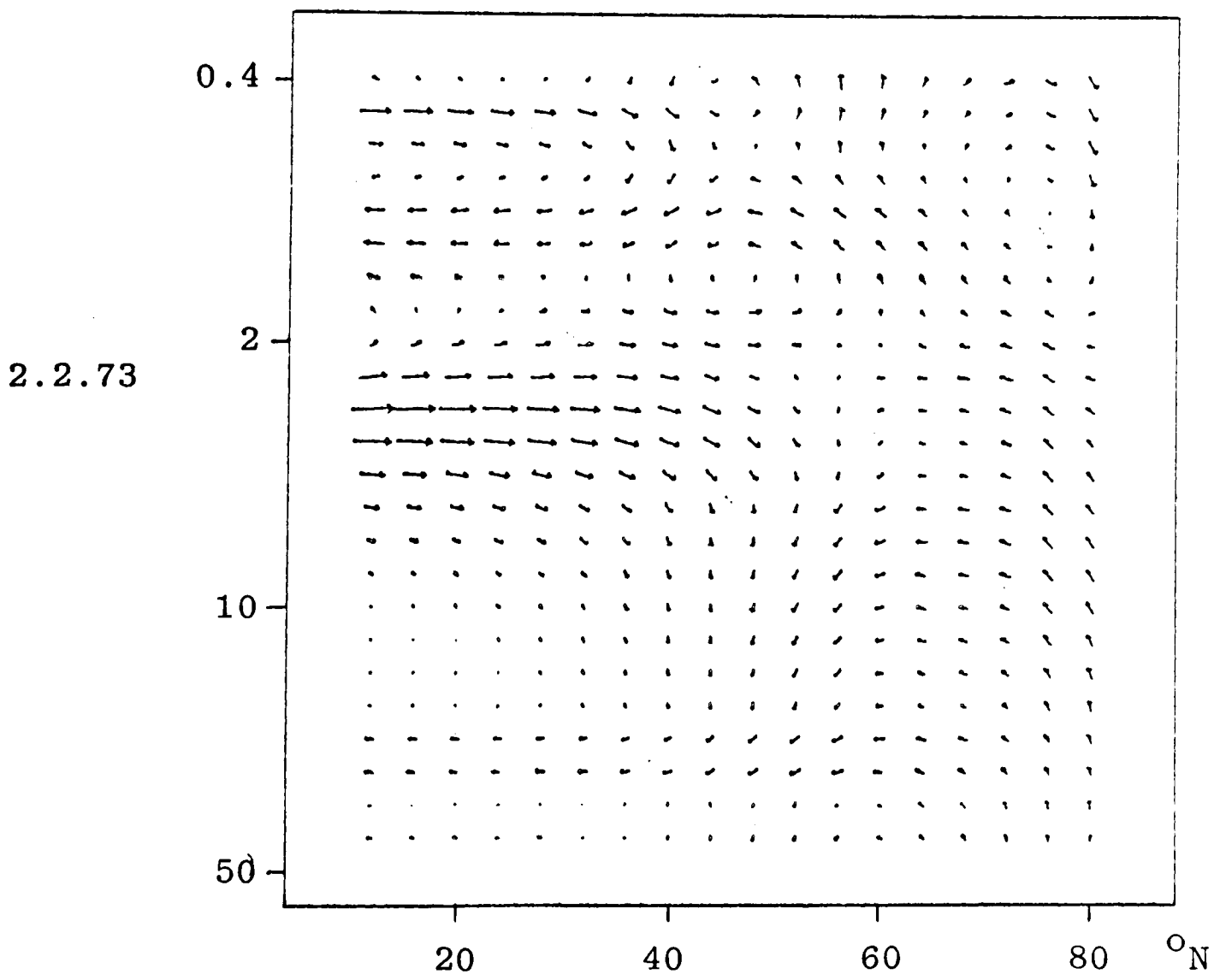
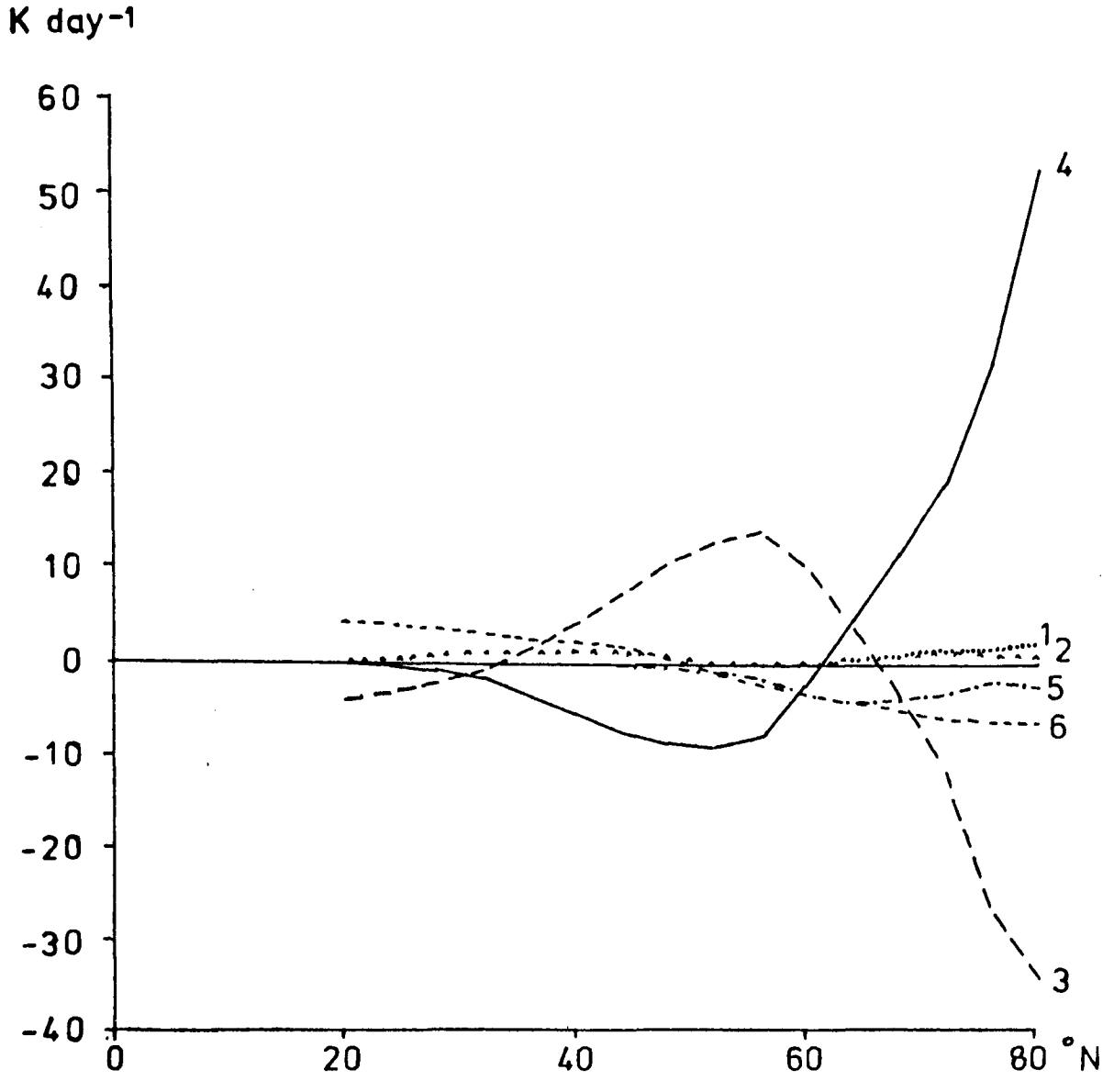


Figure 6.10. Mean meridional circulations on 26, 28, 30.1.73 and 2.2.73 as determined using the thermodynamic and continuity equations



$$1 = \frac{\partial [T]}{\partial t}$$

$$2 = -\frac{[v]}{a}\frac{\partial [T]}{\partial \phi}$$

$$3 = [\omega]\left(\frac{R[T]}{c_p} - \frac{\partial [T]}{\partial p}\right)$$

$$4 = -\frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}[v'T']\cos\phi$$

$$5 = -\frac{\partial}{\partial p}[\omega' T']$$

$$6 = \frac{[Q]}{c_p}$$

The term $\frac{R}{pc_p}[\omega'T']$ was negligible

Figure 6.11. Latitudinal distribution of the terms of the zonal mean thermodynamic equation at 2 mb averaged over the 5 - day period 26 - 30.1.73 in units of K day^{-1}

cooling. The strong poleward flow centred at 3 mb in low latitudes is unrealistic and is caused by too large a vertical gradient of radiative heating (see Appendix 3). It is interesting to compare the vertical motion estimates (shown more clearly in Figure 5.2a) with those of Barnett (1973) for the warming in January 1971. Eddy heat flux convergence in high latitudes at 2 mb was larger in that event, and with temperature changes similar to those of the 1973 warming stronger vertical velocities were implied. He calculated a value of 6.9 cms^{-1} averaged over the polar cap north of $70^{o}N$ at 1 mb on 5.1.73 compared with 4.8 cms^{-1} calculated at 2 mb and $76^{o}N$ on 28.1.73. Downward motion between 45^{o} and $70^{o}N$ of 1.9 cms^{-1} was very similar to the mid-latitude values at the peak of the 1973 event.

During early February a steady fall in temperature occurred at all but tropical latitudes and a strong equator to pole temperature gradient became established.

The heat balance during the warming is illustrated in the latitudinal distributions of the main components of the thermodynamic equation (equation 5.9) in Figure 6.11. The values plotted are averaged over the five-day period 26 - 30.1.73. Although the temperature changes were much larger than is normally observed, even during sudden warmings it appears that there is a near equilibrium between the effects of eddy heat flux divergence (terms 4 & 5), vertical motion (3) and radiative heating (6) in middle and high latitudes, and between vertical motion and radiative heating in low latitudes.

In a study of the stratospheric heat balance of the southern hemisphere during the 1973 winter (1.7.73 - 6.9.73) Hartmann (1976) found that strong sinking motion in high latitudes of the upper stratosphere was required to offset radiative loss, eddy

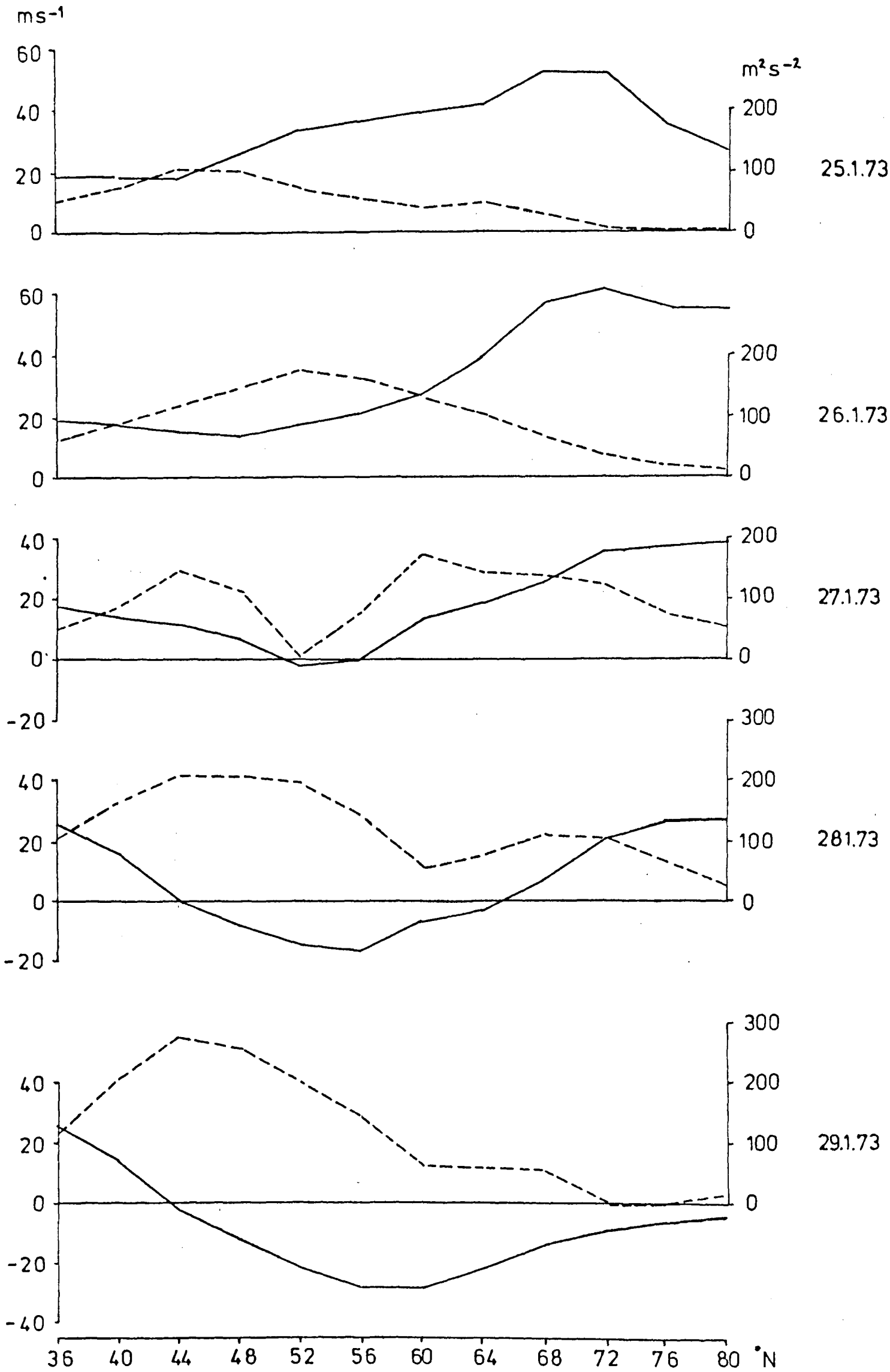
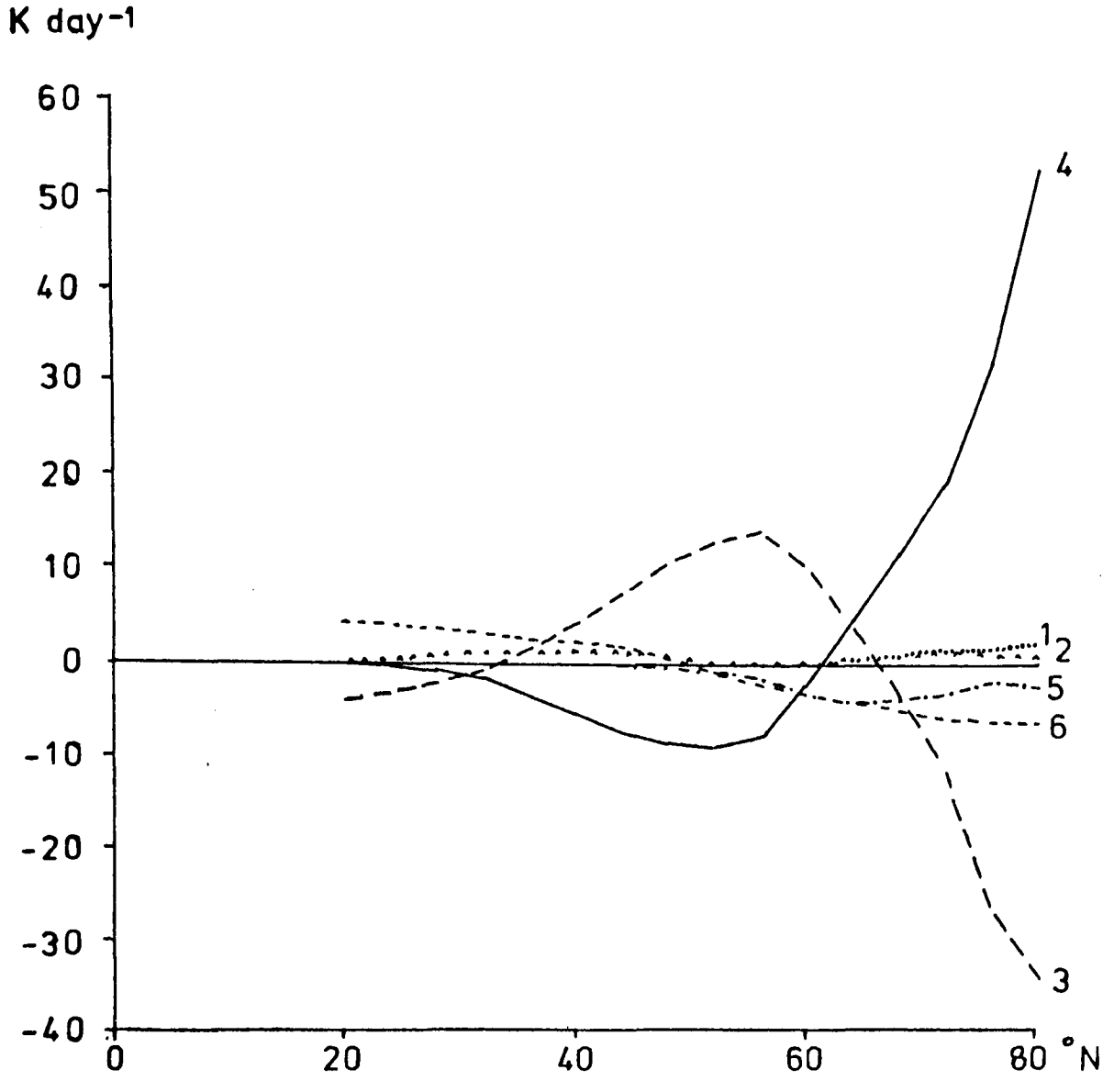


Figure 6.12. Latitudinal distributions of zonal mean zonal wind [u] (——) and eddy momentum flux plotted as $[u'v']\cos^2\phi$ (----) at 2 mb for selected days during the sudden warming



$$1 = \frac{\partial [T]}{\partial t}$$

$$2 = -\frac{[v]}{a}\frac{\partial [T]}{\partial \phi}$$

$$3 = [\omega]\left(\frac{R[T]}{c_p} - \frac{\partial [T]}{\partial p}\right)$$

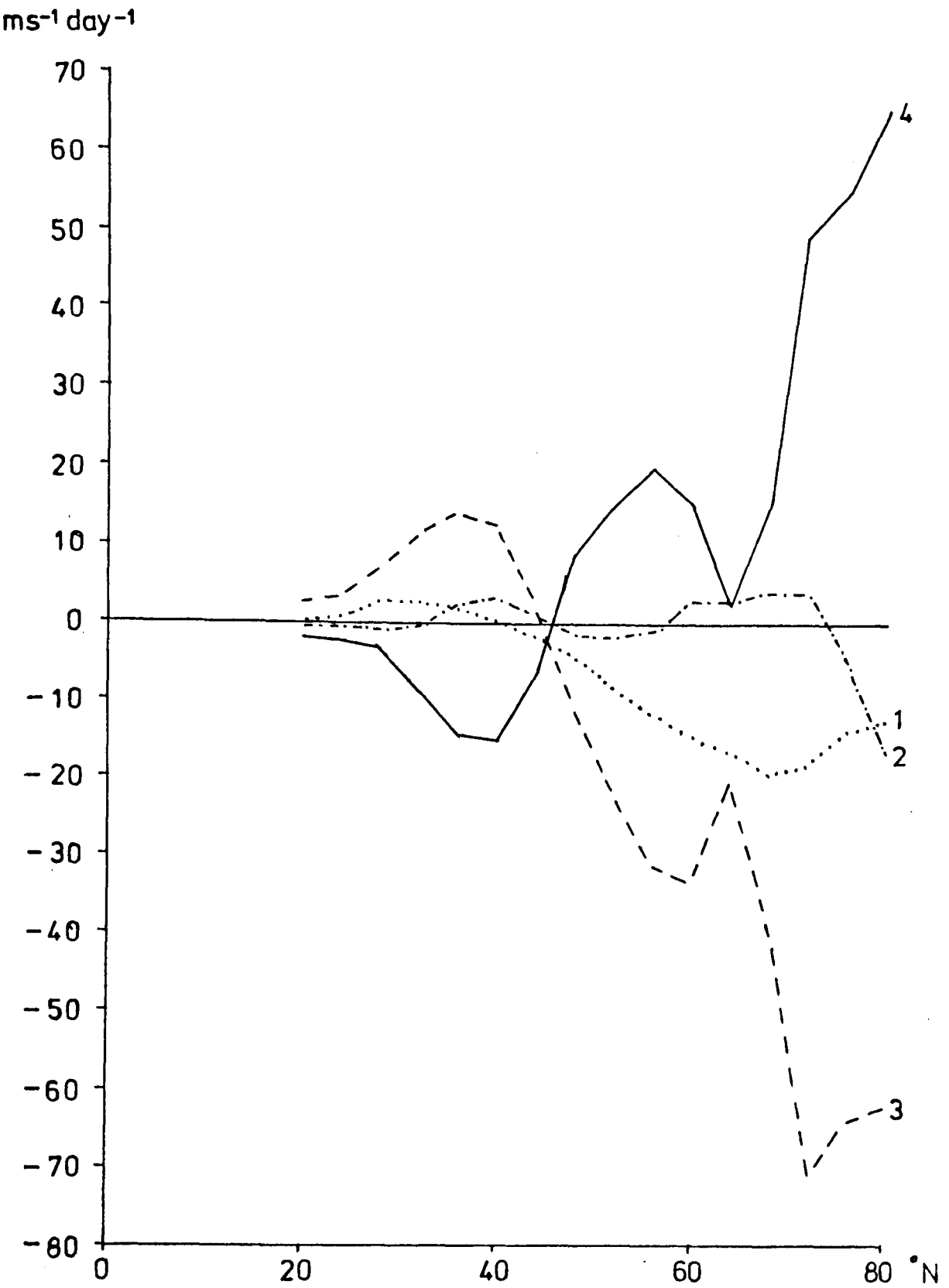
$$4 = -\frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}[v'T']\cos\phi$$

$$5 = -\frac{\partial}{\partial p}[\omega' T']$$

$$6 = \frac{[Q]}{c_p}$$

The term $\frac{R}{pc_p}[\omega'T']$ was negligible

Figure 6.11. Latitudinal distribution of the terms of the zonal mean thermodynamic equation at 2 mb averaged over the 5 - day period 26 - 30.1.73 in units of K day^{-1}



$$1 = \frac{\partial [u]}{\partial t}$$

$$2 = -[v] \frac{1}{a \cos \phi} \frac{\partial}{\partial \phi} [u] \cos \phi$$

$$3 = f[v]$$

$$4 = \frac{1}{a \cos^2 \phi} \frac{\partial}{\partial \phi} [u'v'] \cos^2 \phi$$

Figure 6.13. Latitudinal distribution of the major terms of the zonal mean zonal momentum equation at 2 mb averaged over the 5 - day period 26 - 30.1.73 in units of $m\,s^{-1}\,day^{-1}$

heat flux convergence being insufficient. In the winter period studied in this thesis vertical motion in high latitudes was predominantly upwards (Figure 5.2). If both these winters are typical it would appear that the winter upper stratospheric circulations are different in the two hemispheres on account of differences in eddy heat flux. Indeed Barnett (1975) suggested that the reason why intense polar heating does not occur in southern hemisphere warmings is that the westward tilts with height of the waves are much smaller than observed in the northern hemisphere and consequently eddy heat transport is weaker.

The changes with time of the latitudinal distribution of zonal mean zonal wind, given in Figure 6.12, can possibly be explained in the early stages of the warming by eddy momentum flux divergence although as indicated in Chapter 5 its accuracy is uncertain. The high latitude increase in zonal wind was short-lived, the easterly acceleration in middle latitudes extending quickly northward. Circulation reversal was first observed at 52^{o}N on 27.1.73 and the flow was easterly at all latitudes north of 44^{o}N by 29.1.73 in spite of continued convergence of momentum flux in high latitudes. Average values for the period 26 - 30.1.73 of the major terms in the zonal mean zonal momentum balance equation (equation 5.8) are plotted in Figure 6.13. The Coriolis torque (term 3) and the horizontal advection (2) plotted were calculated to give momentum balance but their distributions are, qualitatively, not too dissimilar to those calculated using thermally derived $[v]$, as was indicated in Figure 5.1b. It may therefore be concluded that, as in the case of the heat budget, the imbalance between the major processes was not large relative to their magnitude, but it was sufficient to ensure deceleration and reversal of the westerly flow at an

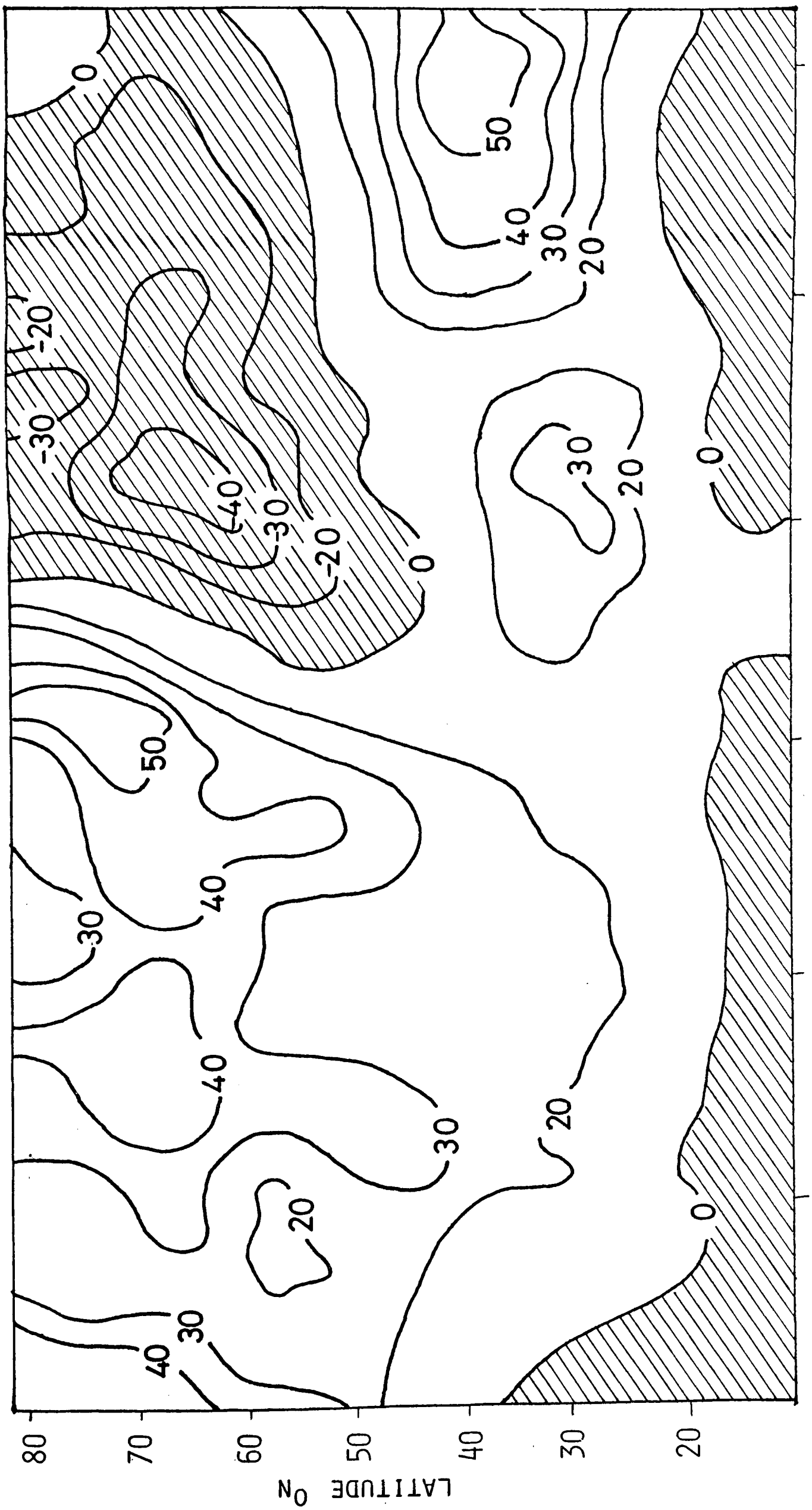


Figure 6.14. Latitude - time section of zonal mean zonal wind [u] (ms^{-1}) at 2 mb during the period 11.1.73 - 10.2.73. Easterlies are shaded

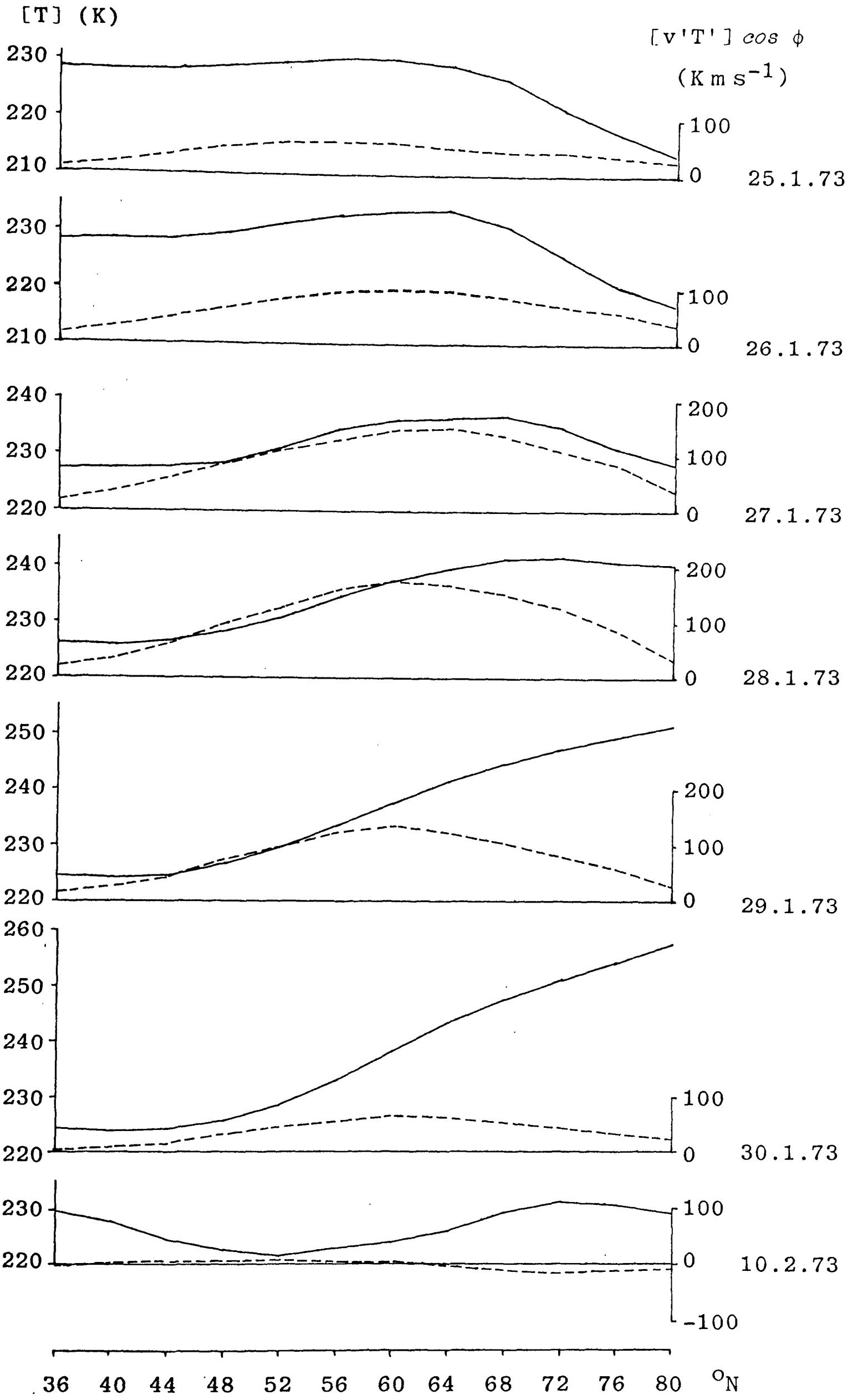


Figure 6.15. As Figure 6.9 but for 10 mb

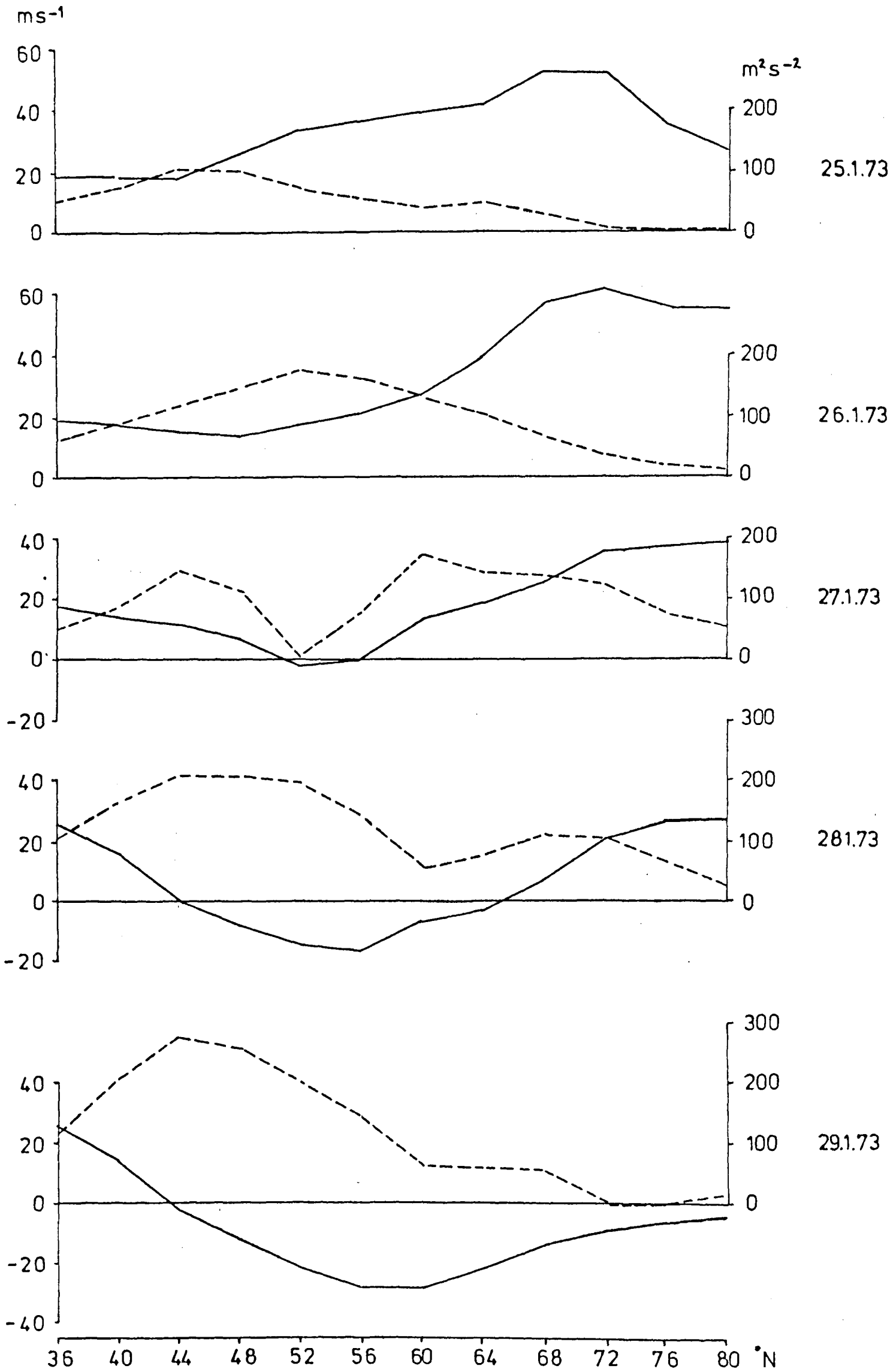
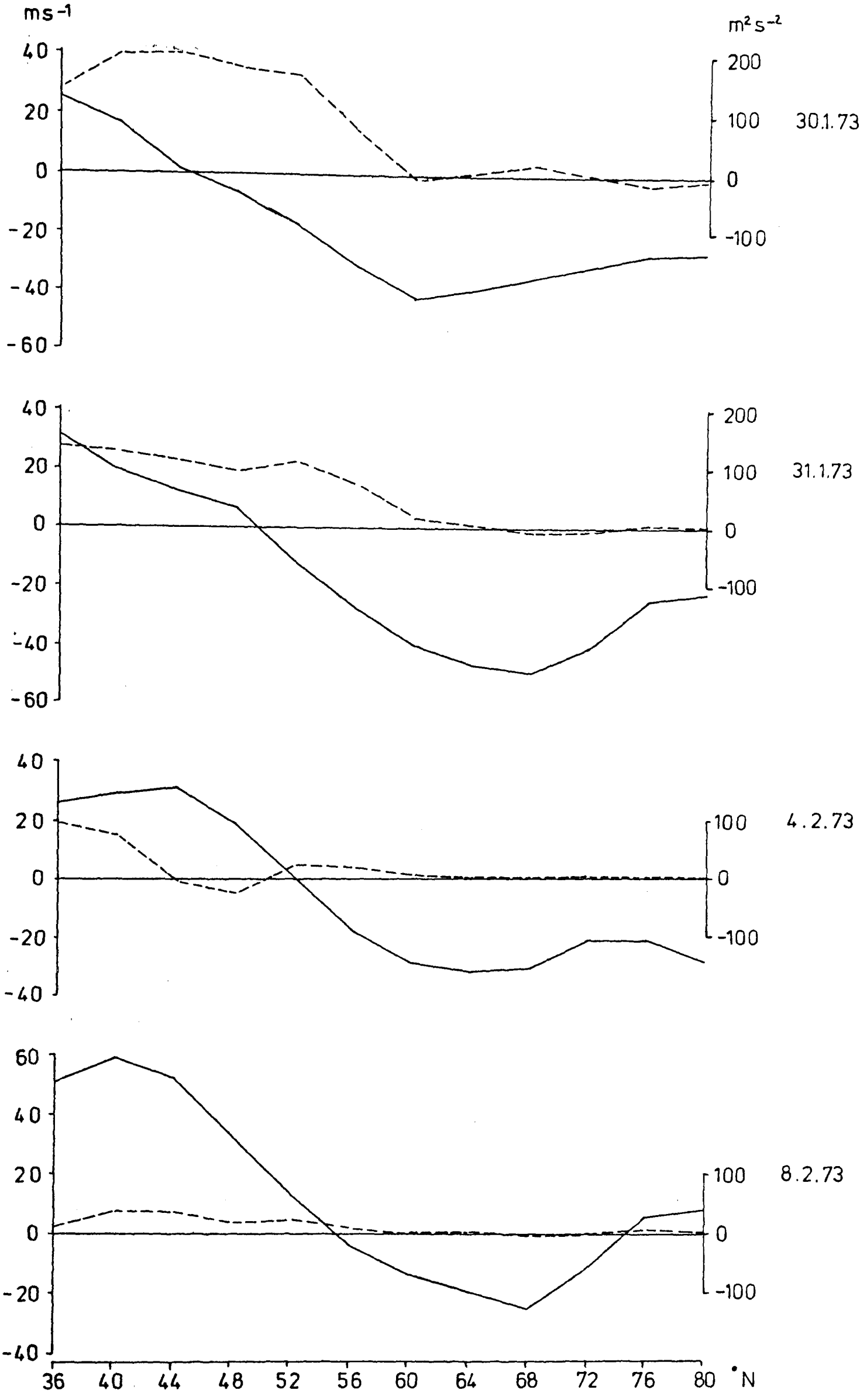


Figure 6.12. Latitudinal distributions of zonal mean zonal wind [u] (——) and eddy momentum flux plotted as $[u'v']\cos^2\phi$ (----) at 2 mb for selected days during the sudden warming

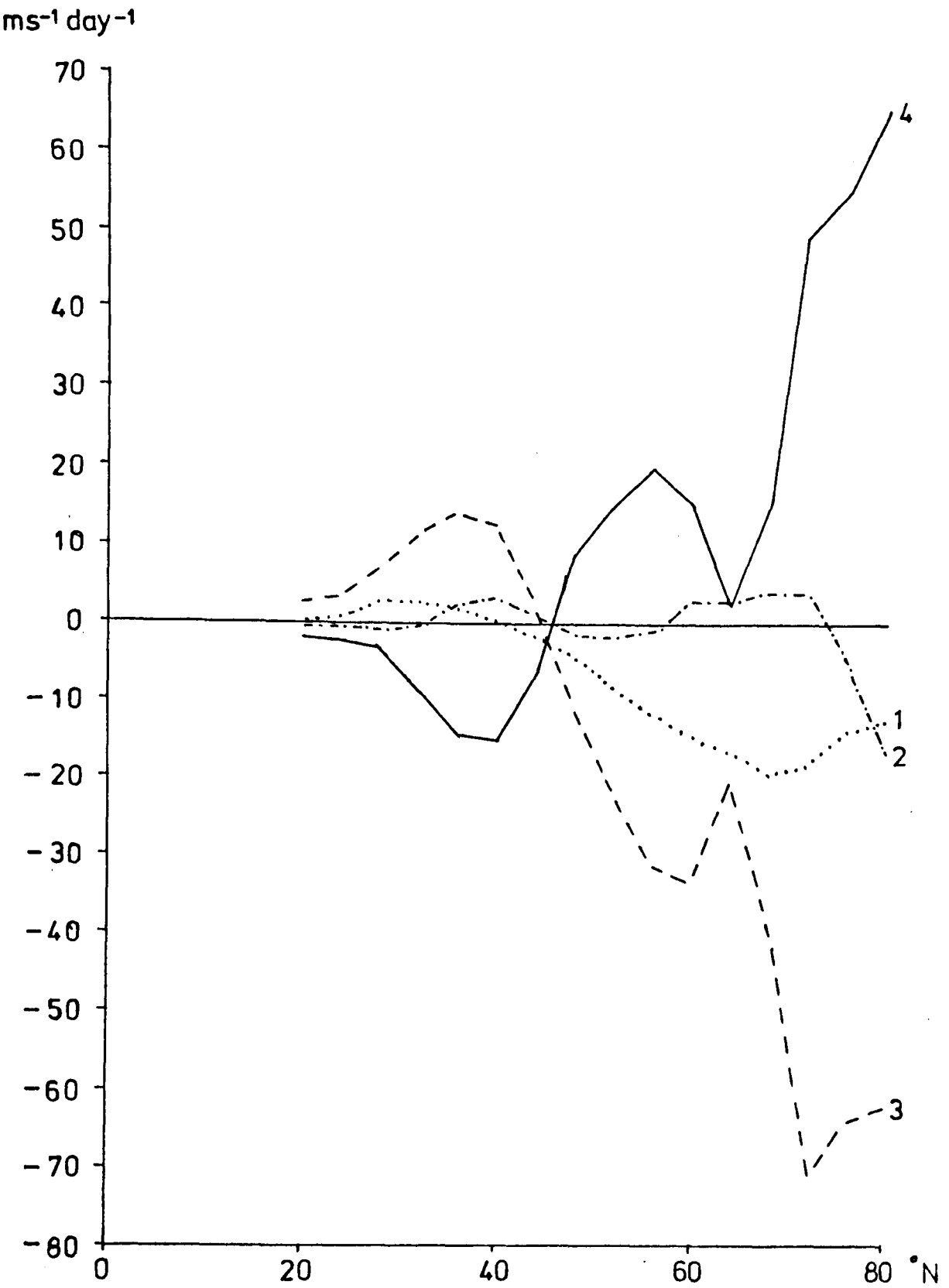

ms-1
m2 s-2
40
20
0
-20
-40
-60
200
100
0
-100
30.1.73
31.1.73
4.2.73
8.2.73
60
36
40
44
48
52
56
60
64
68
72
76
80
°N

were generally very much smaller than at 2 mb, however, and the westerlies weakened as a result of a large imbalance between the effects of the momentum flux divergence and the Coriolis torque as shown in Figure 6.18. The latitude-time section of the zonal mean zonal wind at 10 mb (Figure 6.19) shows that the westerly jet was not re-established in middle latitudes after the warming as it was at higher levels.

Summary of the differences in the zonal mean fields between 2 and 10 mb

The major differences in the evolution of the zonal mean temperature and zonal wind fields between 2 and 10 mb were clearly related to the wave amplification shown in Figure 6.7. The increase in eddy geopotential at 2 mb was associated with large momentum and heat fluxes, while at 10 mb, where only the temperature wave amplitude increased, heat fluxes were large but eddy momentum fluxes remained small. Since the mean meridional circulation strengthened with height, somewhat smaller eddy heat flux convergence in high latitudes at 10 mb gave substantially larger temperature increases than at 2 mb. In addition the vertical eddy heat flux led to warming at 10 mb but cooling at 2 mb. To emphasise the different relationship between eddy heat flux and polar warming at different levels, Figure 6.20 compares the northward eddy heat flux (in Kms^{-1}) at $68^{o}N$, which gives an indication of the heat flux convergence in high latitudes, with the zonal mean temperature at $80^{o}N$.

At 2 mb eddy momentum flux convergence offset to a considerable extent the zonal flow deceleration by the mean meridional circulation, but at 10 mb, although the net deceleration of the flow was smaller than at 2 mb, there was a large imbalance between the competing processes relative to their size.



$$1 = \frac{\partial [u]}{\partial t}$$

$$2 = -[v] \frac{1}{a \cos \phi} \frac{\partial}{\partial \phi} [u] \cos \phi$$

$$3 = f[v]$$

$$4 = \frac{1}{a \cos^2 \phi} \frac{\partial}{\partial \phi} [u'v'] \cos^2 \phi$$

Figure 6.13. Latitudinal distribution of the major terms of the zonal mean zonal momentum equation at 2 mb averaged over the 5 - day period 26 - 30.1.73 in units of $m\,s^{-1}\,day^{-1}$

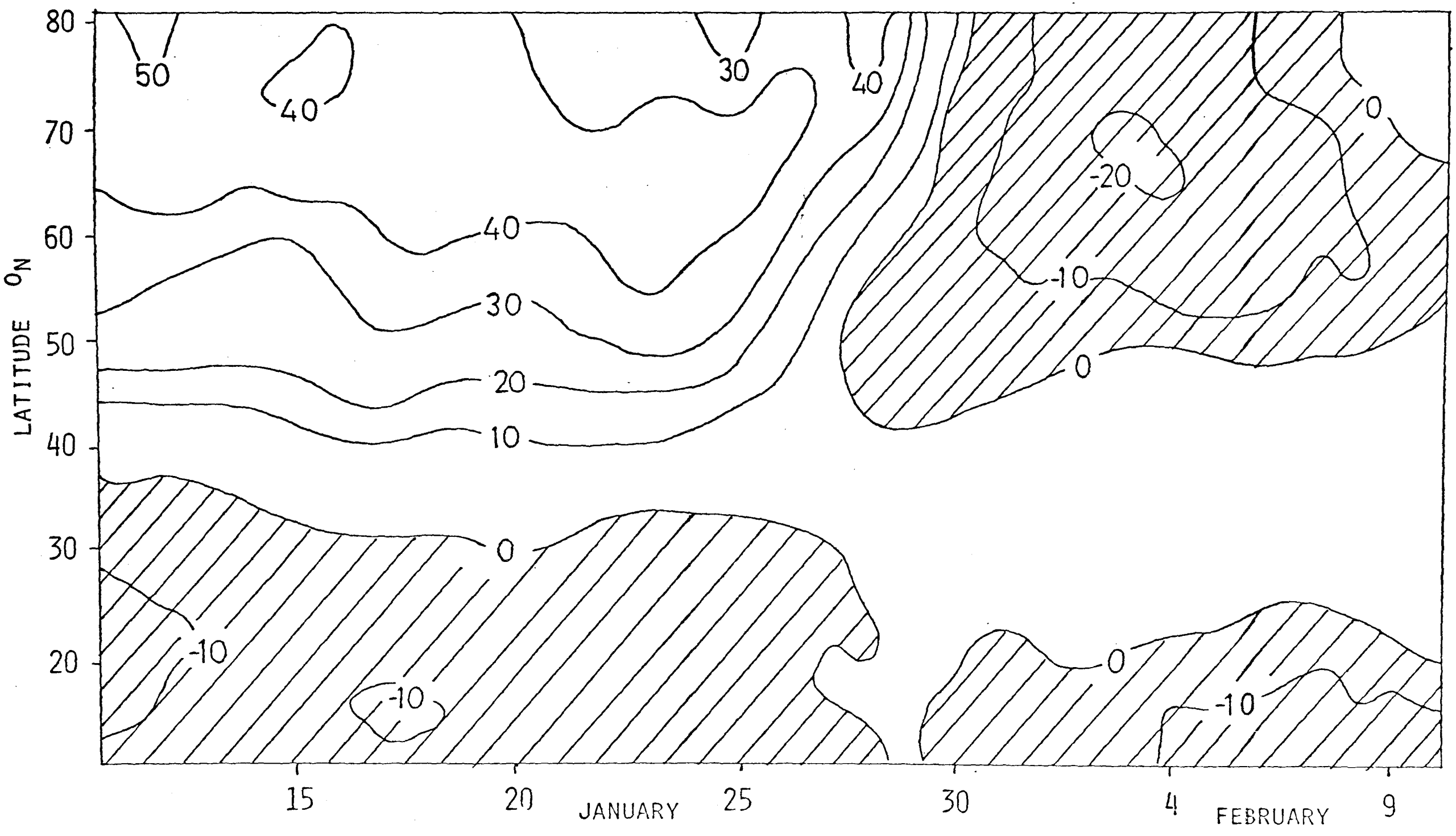


Figure 6.19. Latitude - time section of zonal mean zonal wind [u] (ms^{-1}) at 10 mb during the period 11.1.73 - 10.2.73. Easterlies are shaded

average rate of 20 ms^{-1} day^{-1} at 68°N. South of 40°N a net westerly acceleration occurred in this period and again after 3.2.73 as illustrated more clearly in the latitude-time section of Figure 6.14. A similar re-establishment of the westerly jet in mid-latitudes of the upper stratosphere after the January 1971 warming and after the minor warming of January 1975 was observed by Klinker (1976). The momentum fluxes at 2 mb during the warming peak were of similar magnitude to those calculated by Barnett (1973) for 5.1.71. An interesting feature of the latitudinal distribution of momentum flux on both these occasions was the double-peaked structure with a minimum flux value occurring near 60°N at the peak of the 1973 event (see Figure 5.3) and at 70°N on 5.1.71.

10 mb temperature and zonal wind

At 10 mb convergence of eddy heat flux in high latitudes led to very much larger polar warming than was observed at 2 mb although the heat flux values were a little smaller (Figure 6.15). The temperature increase at 80°N from 214-258K during the period 26 - 30.1.73 was three times larger than that at 2 mb, and compares well with observations given by Arpe (1976) of temperature rises of 45K at 10 mb during the warming periods 28.12.67 - 2.1.68 and 5 - 11.1.71. The mean values of the terms of the zonal mean thermodynamic equation for this period are given in Figure 6.16. The relative magnitudes of the eddy heat flux and vertical motion terms were similar in middle latitudes but very different north of 70°N to those observed at 2 mb.

The pattern of change in the zonal flow (Figure 6.17) was the same as in the upper stratosphere although in high latitudes circulation reversal occurred one day later and the easterlies established were only about half as strong. The momentum fluxes

50 mb temperature and zonal wind

The polar warming at 50 mb (Figure 6.21) was later and less rapid than at higher levels and was maintained for a longer period. The temperature at 80°N increased by 32K between 27.1.73 and 2.2.73 and remained almost constant for six days. An equivalent increase in temperature of 40K in 10 days was observed by Reed et al (1963) during the 1957 warming. The eddy heat flux was substantially smaller than at higher levels but reached its maximum value on the same day (28.1.73). It was able to effect large polar warming since cooling by vertical motion and radiative loss were less marked. A maximum zonal mean vertical velocity of 1.4 cms^{-1} at 76°N (Figure 5.2a) compares with a value at 70°N of 1.7 cms^{-1} at the peak of the 1957 event (Reed et al, 1963).

Figure 6.22 shows that the convergence of eddy momentum flux in high latitudes increased the westerly flow between 25 and 29.1.73. In the absence of this mechanism from 30.1.73 onwards the flow weakened by the action of the Coriolis torque and easterlies occurred for a short period (2 - 5.2.73) in high latitudes. Easterly flow was not established in middle latitudes as it was at higher levels.

Summary

The vertical distribution of the high latitude temperature and zonal wind changes which have been described are conveniently summarised in Figures 6.23 and 6.24, which are pressure-time sections of the zonal mean zonal wind at 68°N, the latitude at which the strongest easterly circulation occurred, and the change of zonal mean temperature at 80°N from its value on 11.1.73. The downward progression of these features of the warming event is well illustrated. The high level cooling and low level warming

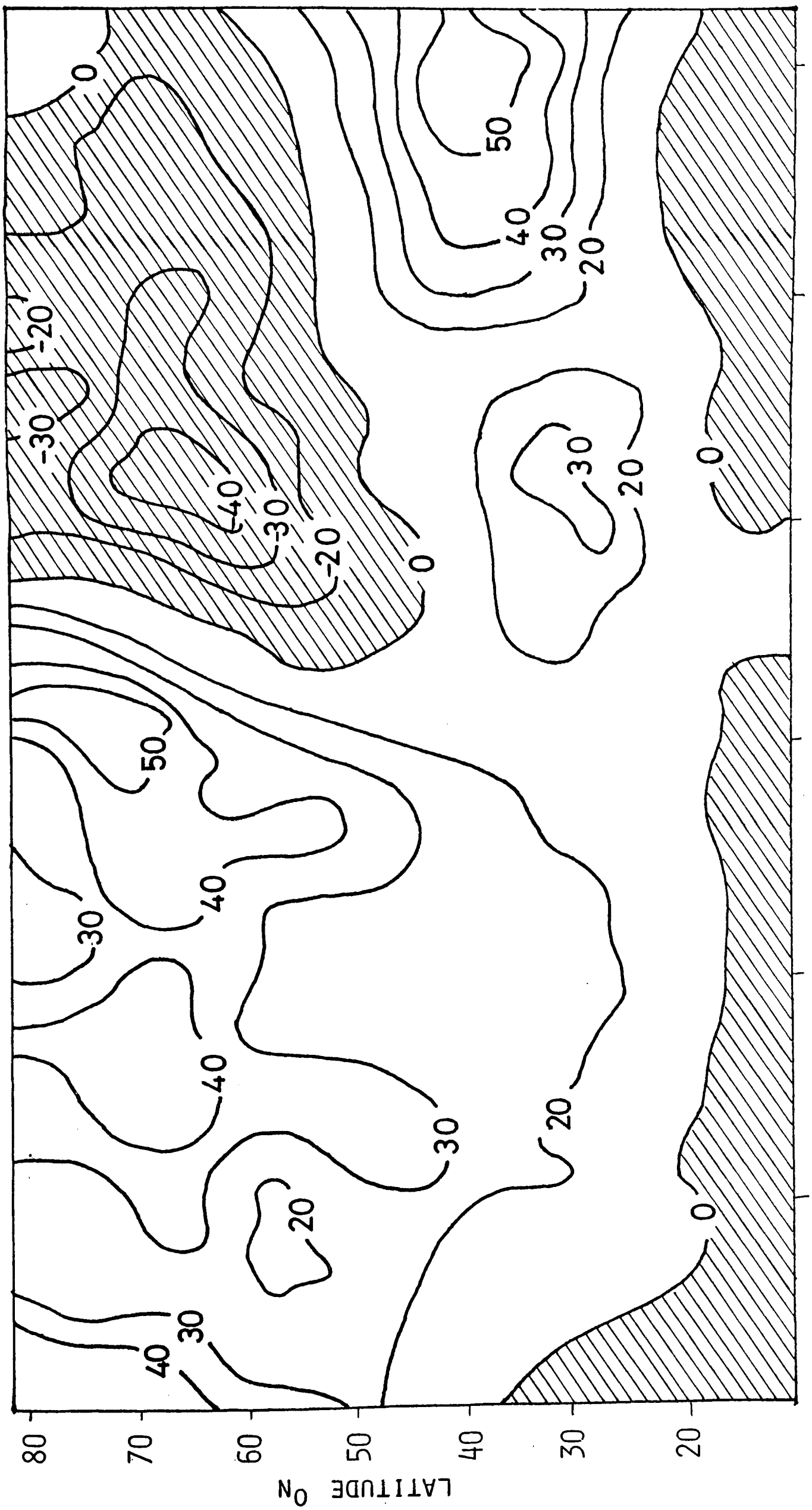


Figure 6.14. Latitude - time section of zonal mean zonal wind [u] (ms^{-1}) at 2 mb during the period 11.1.73 - 10.2.73. Easterlies are shaded

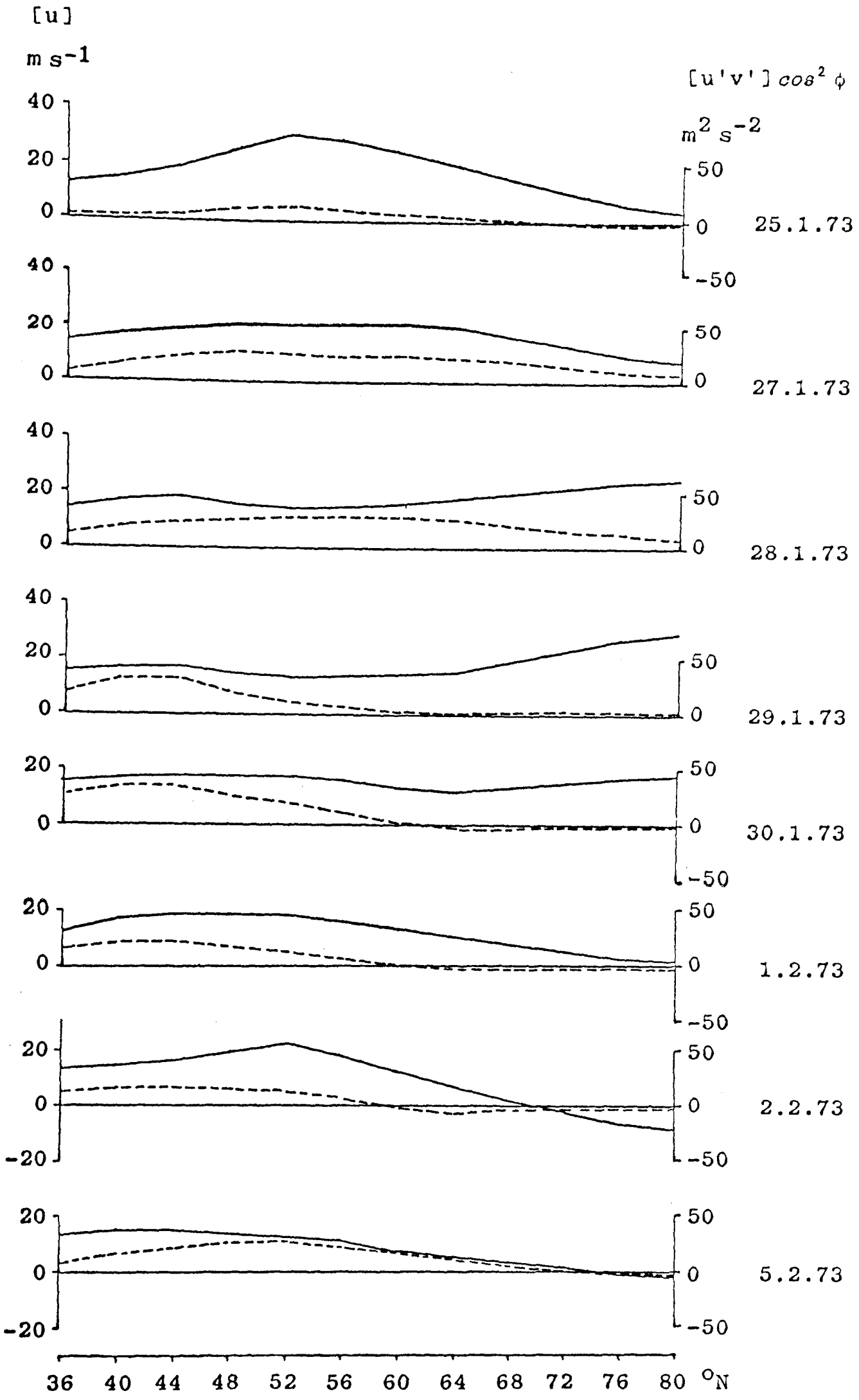


Figure 6.22. As Figure 6.12 but for 50 mb

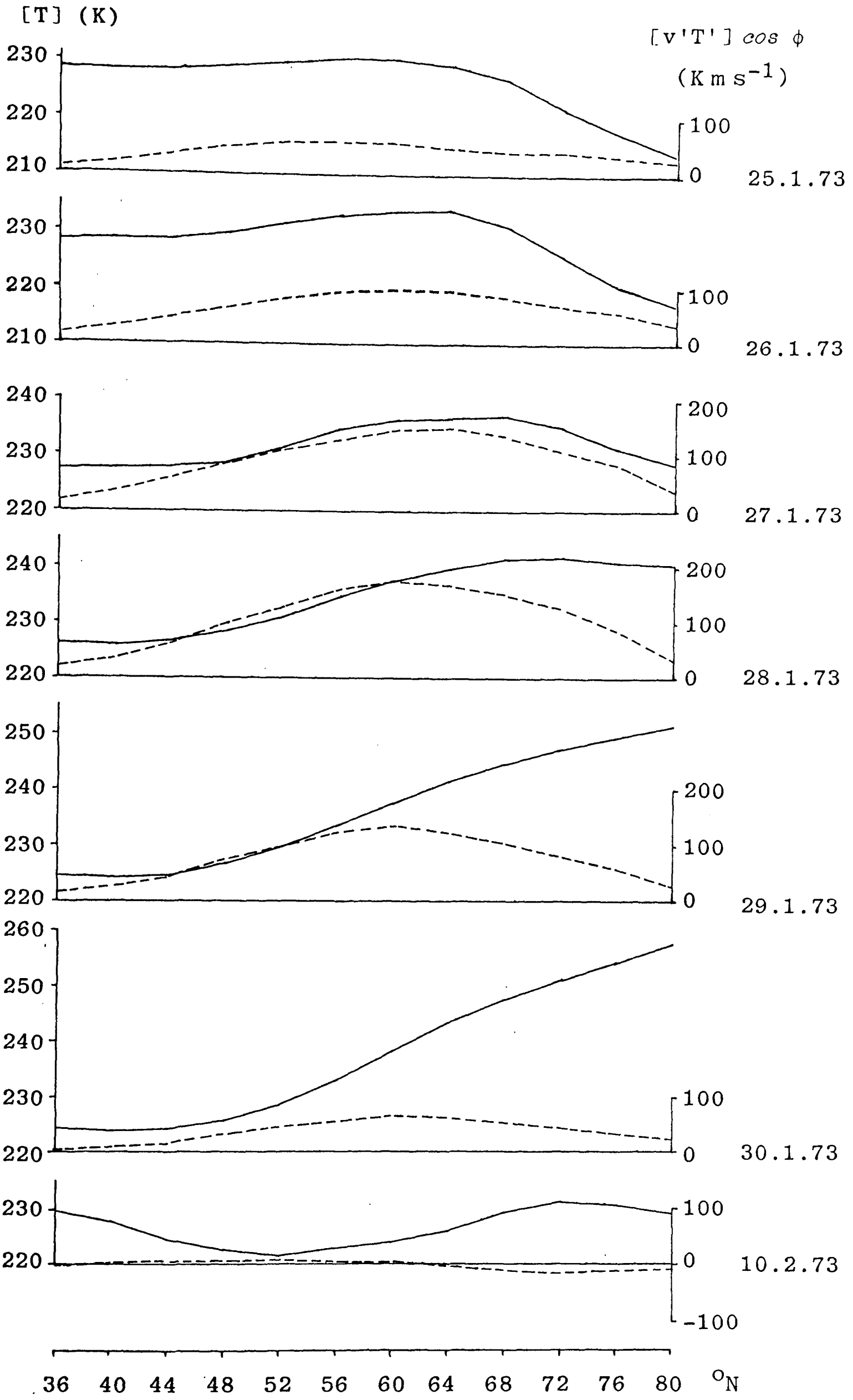
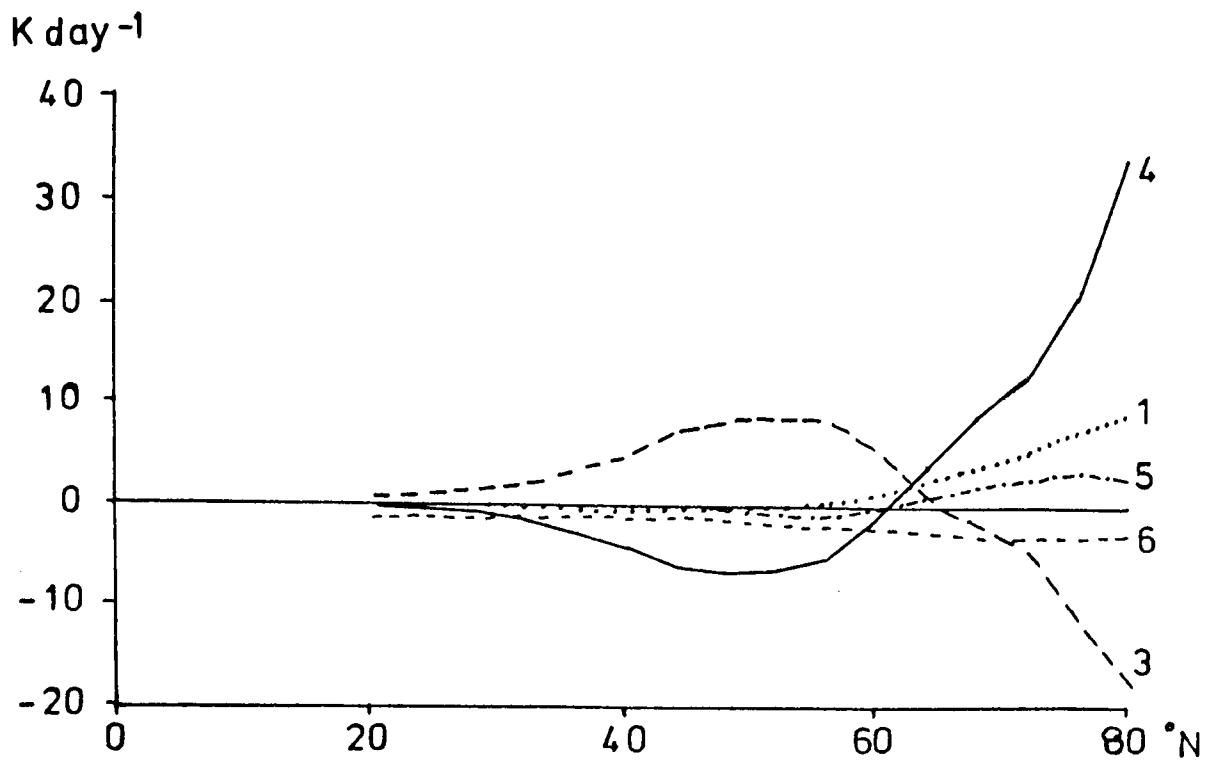


Figure 6.15. As Figure 6.9 but for 10 mb

resulted in a reduction of the temperature difference between 2 and 50 mb from 72K on 13.1.73 to -1K on 7.2.73, and is illustrated more graphically in Figure 6.25. The trend of the temperature at 12^{o}N is shown to be opposite that at 80^{o}N at all levels for most of the period. This common feature of the winter stratosphere was first discovered by Fritz and Soules (1970) but is not yet understood.



$$1 = \frac{\partial [T]}{\partial \phi}$$

$$3 = [\omega] \left(\frac{R[T]}{c_p} - \frac{\partial [T]}{\partial p} \right)$$

$$4 = -\frac{1}{a \cos\phi} \frac{\partial}{\partial \phi} [v'T'] \cos\phi$$

$$5 = -\frac{\partial}{\partial p} [\omega' T']$$

$$6 = \frac{[Q]}{c_p}$$

The terms $-\frac{[V]\partial[T]}{a\,\partial\phi}$ and $\frac{R}{pc_p}[\omega T]$ were negligible

Figure 6.16. As Figure 6.11 but for 10 mb

Introduction

The results of the previous chapter have demonstrated the very large scale changes which took place in the middle and upper stratosphere following the rapid amplification of a wave-number one planetary wave. In general wave energy may increase at the expense of the potential or kinetic energy of the mean state or may be altered through the action of pressure-work terms (see Chapter 4) and advective boundary fluxes. Analysis of the energetics of sudden warmings can thus indicate the mechanisms responsible for the rapid wave amplification which characterises them. Most studies of sudden warming energetics have, on account of the availability of data, been restricted to the troposphere and the low and middle stratosphere. However it has been shown that the warming process may begin in the upper stratosphere and propagate downwards, so a knowledge of the energy processes occurring there is of great importance. In this chapter the energetics of the 1973 warming in three stratospheric layers - the upper, upper-middle and lower-middle stratosphere - are examined and compared. By way of introduction a brief review of the findings of previous energy budget studies is given.

7.1. REVIEW OF SUDDEN WARMING ENERGY BUDGET STUDIES

Table 7.1 lists most of the important, published studies on the energetics of sudden warmings. All employed near-hemispheric data sets and described the energy processes in terms of the Lorenz cycle of energy conversions discussed in Chapter 4. In most results were averaged with respect to time and averaged over pressure and horizontal area. The much larger energy

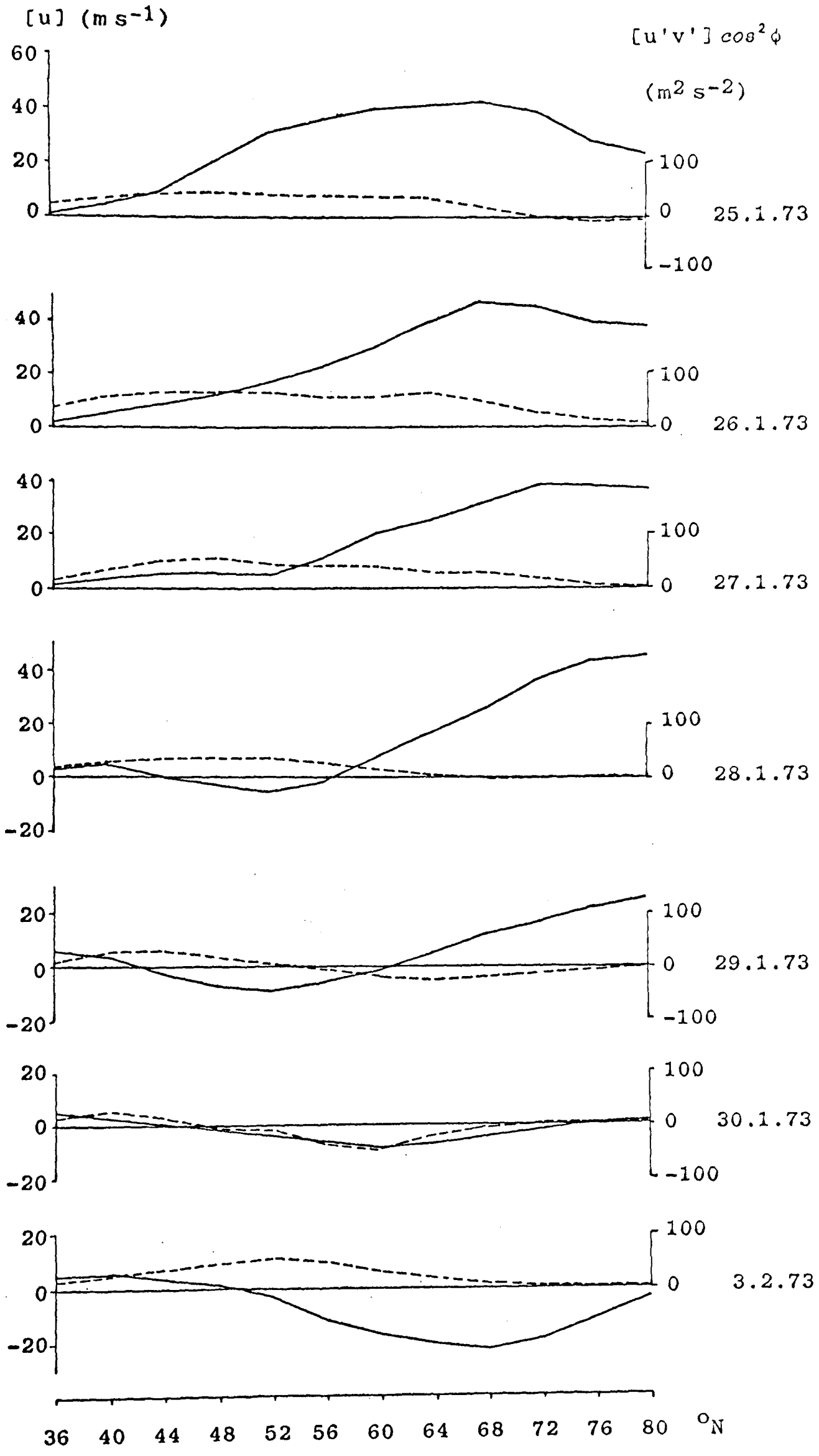


Figure 6.17. As Figure 6.12 but for 10 mb

TABLE 7.1 (continued)

Authors	Period studied & data frequency	Vertical extent of analysis	Vertical resolution	Other points
Klinker (1976)	1.12.70-31.1.71 Daily up to 10 mb weekly above 10 mb	50 - 0.4 mb	5 levels: 50,30,10,2 and 0.4 mb	
	1.12.74-24.1.75 As above	As above	As above	Minor warming
Arpe (1976)	1.11.67-30.1.68 Daily	75 - 7.5 mb	5 levels: 500,100,50, 30 and 10 mb	Calculated non linear interactions between wavenumbers
	1.11.70-30.1.71 Daily	As above	As above	" "

conversion rates of high latitudes tend to dominate the area-averaged energetics despite the relatively small area they occupy. Comparison between different studies is often complicated by the different choice of boundaries, but they present a fairly consistent picture of the energy cycle during sudden warmings in the stratosphere below 10 mb.

The energetics below 10 mb during sudden warmings

The sudden warming may be considered as consisting of two stages: the first is the amplifying stage in which eddy energy increases and zonal mean energy declines; the second, declining, stage begins when the zonal flow reverses. Figure 7.1 shows the energy cycle which normally occurs below 10 mb during the first stage. This direct baroclinic cycle ($A_Z \rightarrow A_E \rightarrow K_E$) was found in all the major sudden warmings listed in Table 7.1 except in the

were generally very much smaller than at 2 mb, however, and the westerlies weakened as a result of a large imbalance between the effects of the momentum flux divergence and the Coriolis torque as shown in Figure 6.18. The latitude-time section of the zonal mean zonal wind at 10 mb (Figure 6.19) shows that the westerly jet was not re-established in middle latitudes after the warming as it was at higher levels.

Summary of the differences in the zonal mean fields between 2 and 10 mb

The major differences in the evolution of the zonal mean temperature and zonal wind fields between 2 and 10 mb were clearly related to the wave amplification shown in Figure 6.7. The increase in eddy geopotential at 2 mb was associated with large momentum and heat fluxes, while at 10 mb, where only the temperature wave amplitude increased, heat fluxes were large but eddy momentum fluxes remained small. Since the mean meridional circulation strengthened with height, somewhat smaller eddy heat flux convergence in high latitudes at 10 mb gave substantially larger temperature increases than at 2 mb. In addition the vertical eddy heat flux led to warming at 10 mb but cooling at 2 mb. To emphasise the different relationship between eddy heat flux and polar warming at different levels, Figure 6.20 compares the northward eddy heat flux (in Kms^{-1}) at $68^{o}N$, which gives an indication of the heat flux convergence in high latitudes, with the zonal mean temperature at $80^{o}N$.

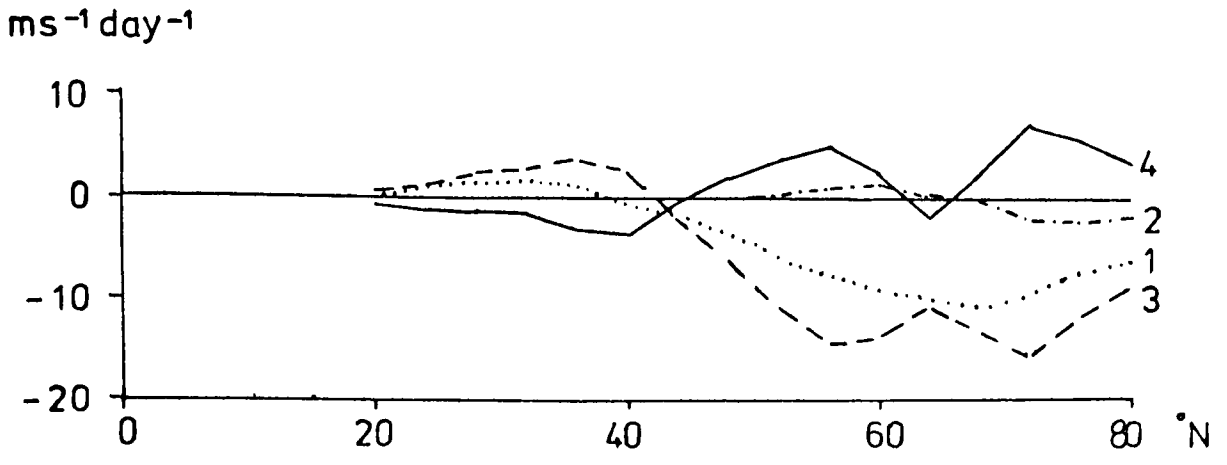
At 2 mb eddy momentum flux convergence offset to a considerable extent the zonal flow deceleration by the mean meridional circulation, but at 10 mb, although the net deceleration of the flow was smaller than at 2 mb, there was a large imbalance between the competing processes relative to their size.

1973 warming reported by Quiroz et al (1975), a point to be discussed later. Both Muench (1965) and Dopplick (1971) observed baroclinic cycles for non-warming periods in the 100 - 10 mb layer of the northern hemisphere and Hartmann (1976) calculated energetics in the same sense for July and August 1973 in the southern hemisphere 100 - 10 mb region. It would therefore appear that this is the normal wintertime energy cycle for the lower middle stratosphere, which, as noted by Muench (1965), becomes more intense at the onset of a sudden warming. Gain in K_E due to energy propagation from the troposphere via the eddy pressure-work effect was a significant feature in all the sudden warming events listed in Table 7.1. On account of its importance, Chapter 8 is devoted to a more detailed study of this process and its effects during the 1973 warming. Tropospheric forcing by this means is a normal wintertime feature (Miller, 1970) which is enhanced during sudden warmings (Miller and Johnson, 1970).

With the reversal of the zonal circulation all the internal conversions, with the exception of the barotropic (K_E-K_Z) conversion, reverse. The observed cycle during the declining phase is shown in Figure 7.2. The K_E-K_Z conversion, although it usually remains in the eddy to zonal sense, may reverse, in which case it tends to be rather weak. Pressure-work effects continue to supply K_E and deplete K_Z, although at reduced rates. Destruction of available potential energy by the interaction of the meridional distributions of temperature and solar heating and by radiative relaxation of temperature perturbations around the globe is usually observed throughout the warming event.

The energetics above 10 mb during sudden warmings

Until the advent of satellite measurements the energetics of the upper stratosphere were a subject for speculation only. It



$$1 = \frac{\partial [u]}{\partial t}$$

$$2 = -[v] \frac{1}{a \cos \phi} \frac{\partial}{\partial \phi} [u] \cos \phi$$

$$3 = f[v]$$

$$4 = \frac{1}{a \cos^2 \phi} \frac{\partial}{\partial \phi} [u'v'] \cos^2 \phi$$

Figure 6.18. As Figure 6.13 but for 10 mb

was assumed that the latitudinal distribution of solar radiation would be a large source of available potential energy and that a baroclinic energy cycle would prevail (Newell, 1964). The first attempt to investigate the energetics above 10 mb was made by Miller et al (1972). They made calculations for the 30 - 2 mb layer for seven separate days during the warming confined to the upper stratosphere in December - January 1969/70, using data from the satellite infra-red spectrometer (SIRS) experiment on Nimbus 3. An $A_Z \rightarrow A_E \rightarrow K_E$ cycle during the amplifying stage was accompanied by energy gain through the eddy pressure-work effect and a large $K_E \rightarrow K_Z$ conversion. Unlike the major warmings discussed above, this minor warming was characterised by a $K_E \rightarrow A_E$ conversion in the first stage in the 100 - 30 mb layer. The authors suggested that the absence of baroclinic conversion in the lower stratosphere may have been associated with the failure of this warming to affect the whole stratosphere. However the extent to which warming descends into the lower stratosphere is clearly dependent on other factors since the presence of baroclinic conversion in the lower stratosphere did not ensure the downward propagation of the minor warming of 1974-5 observed by Klinker (1976) or the upper stratospheric warming at the end of January, 1964 (Dopplick, 1971). In contrast with the results of Miller et al, Klinker (1976) calculated a $K_E \rightarrow A_E$ conversion between 30 and 2 mb during the first phase of both a major (1970-1) and a minor (1974-5) warming. Both these studies were based on few days' data and Klinker did not attempt to produce a complete budget above 10 mb. The results described below thus form the first detailed set of energetics calculations for the upper stratosphere during a major sudden warming.

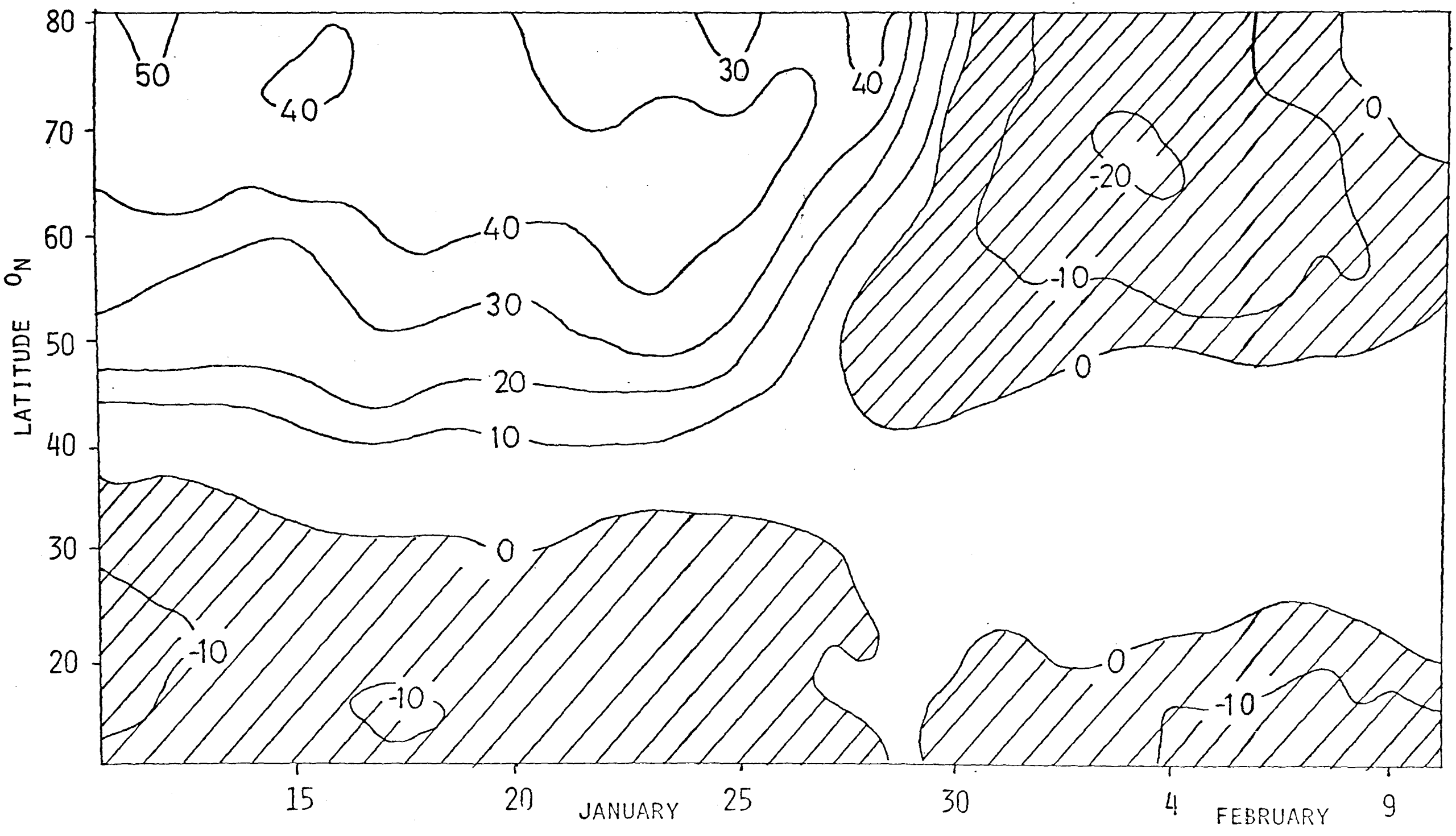
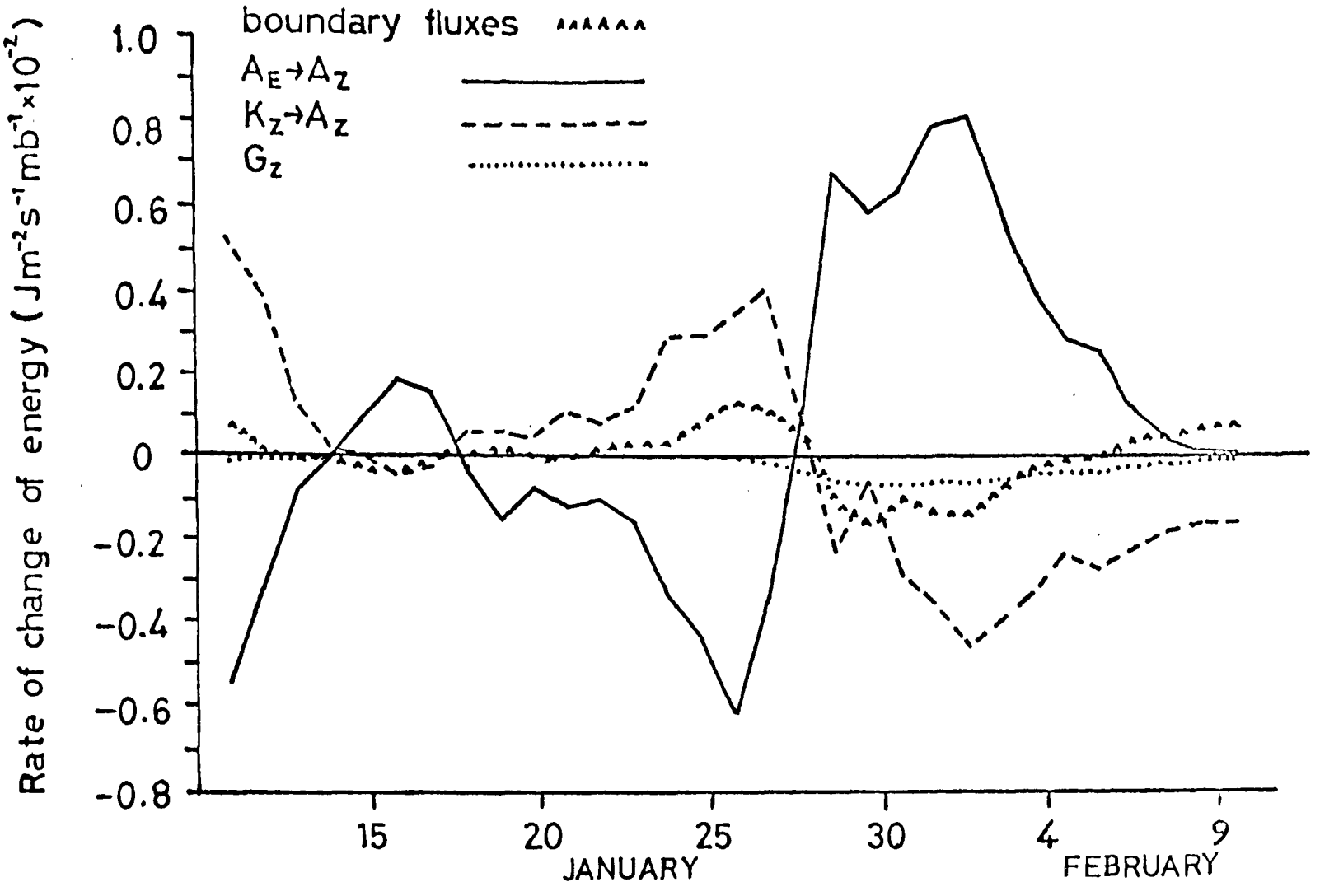
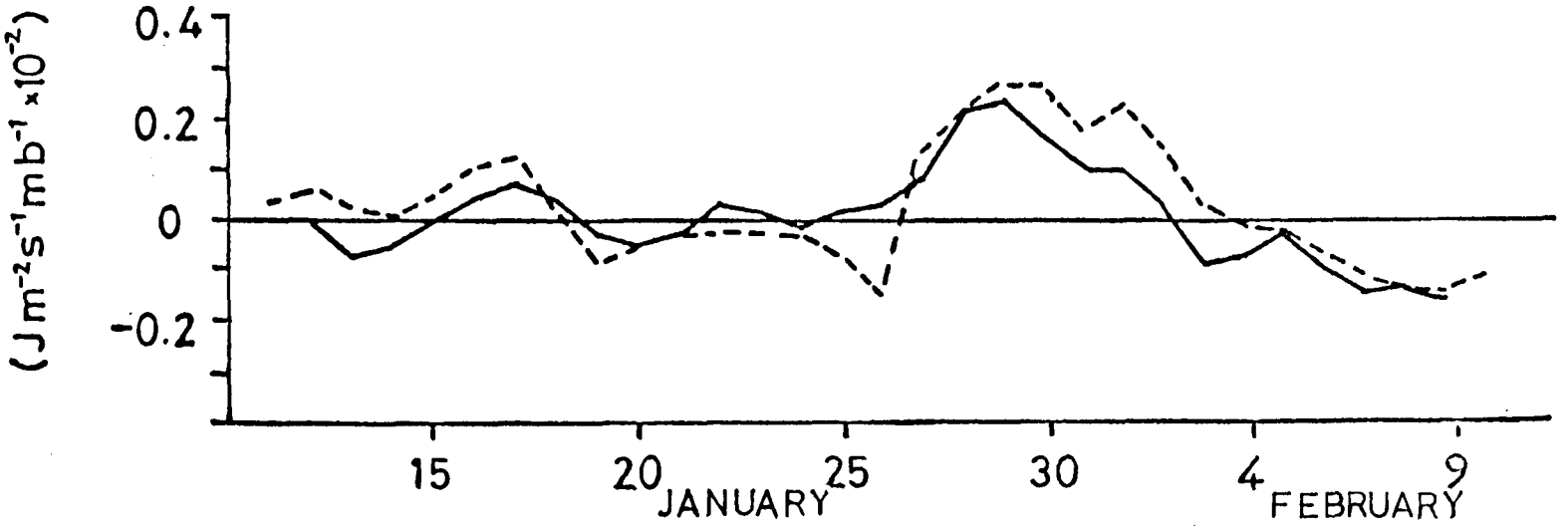


Figure 6.19. Latitude - time section of zonal mean zonal wind [u] (ms^{-1}) at 10 mb during the period 11.1.73 - 10.2.73. Easterlies are shaded



Rates of change of A_z due to energy generation, conversions and boundary fluxes 50–10 mb



Observed rate of change of A_z(——) compared with net rate of change due to above processes (-------)

Figure 7.3a. The zonal available potential energy budget for the region 50 - 10 mb, 14 - 82°N between 11.1.73 and 10.2.73

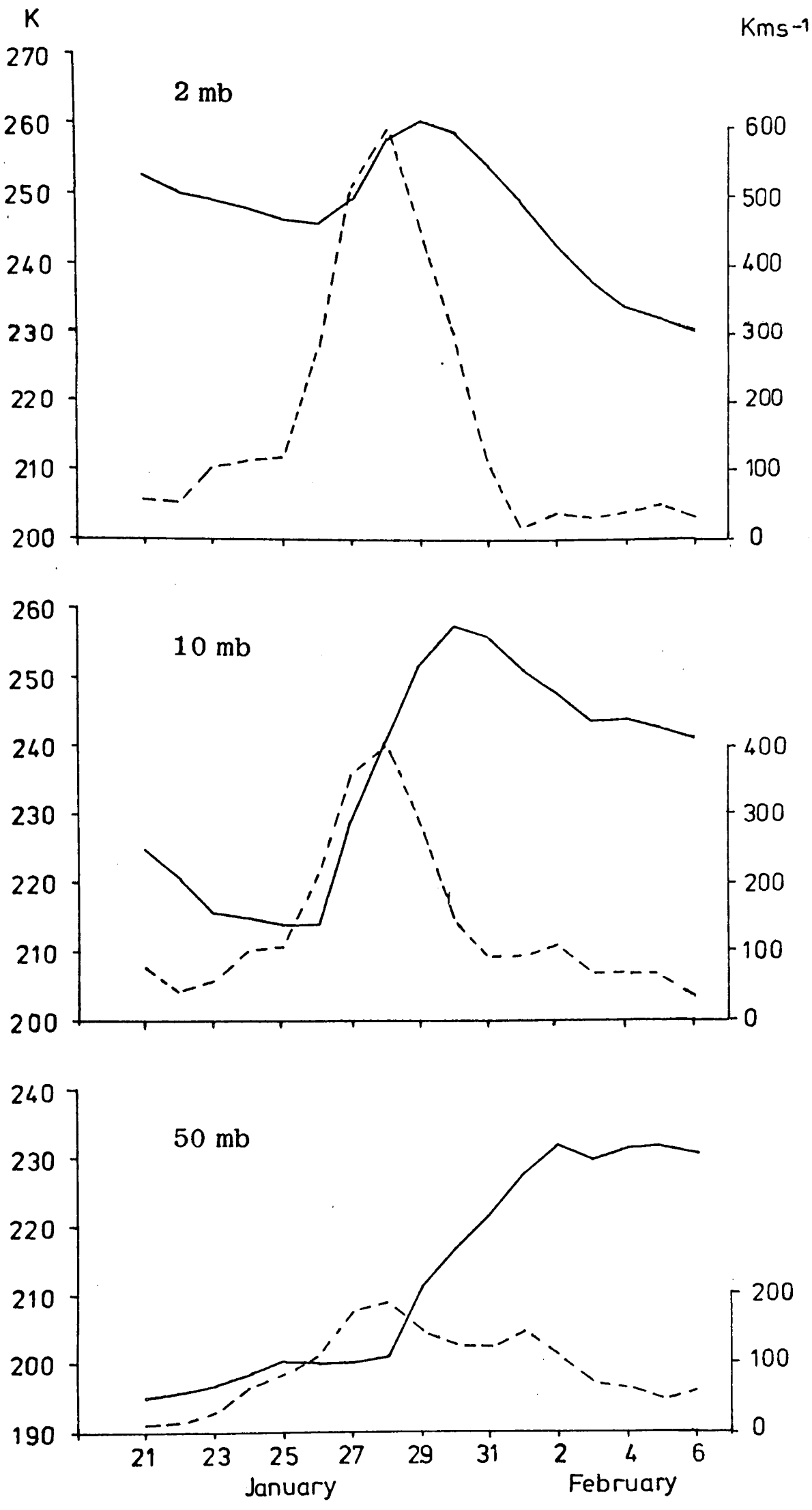
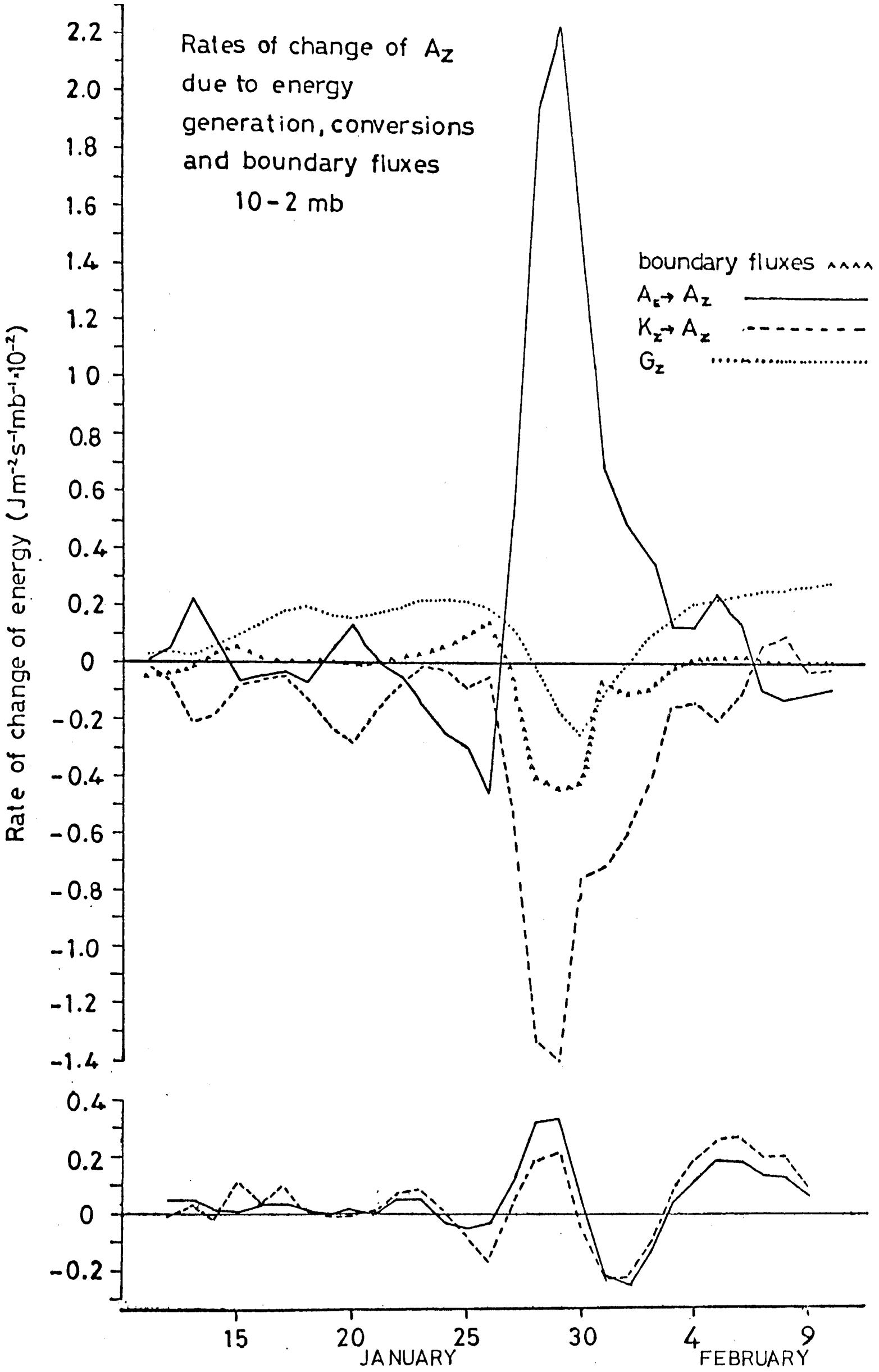


Figure 6.20. Zonal mean temperature [T] (K) (——) at 80^{o}N and eddy heat flux [v'T'] (K ms^{-1}) (----) at 68^{o}N plotted against time for 2 mb (top) 10 mb (centre) and 50 mb (bottom)



Observed rate of change of A_Z(——) compared with net rate of change due to above processes (- - - - - - -)

Figure 7.4a. The zonal available potential energy budget for the region 10-2 mb, 14-82 ^{O}N between 11.1.73 - 10.2.73

50 mb temperature and zonal wind

The polar warming at 50 mb (Figure 6.21) was later and less rapid than at higher levels and was maintained for a longer period. The temperature at 80°N increased by 32K between 27.1.73 and 2.2.73 and remained almost constant for six days. An equivalent increase in temperature of 40K in 10 days was observed by Reed et al (1963) during the 1957 warming. The eddy heat flux was substantially smaller than at higher levels but reached its maximum value on the same day (28.1.73). It was able to effect large polar warming since cooling by vertical motion and radiative loss were less marked. A maximum zonal mean vertical velocity of 1.4 cms^{-1} at 76°N (Figure 5.2a) compares with a value at 70°N of 1.7 cms^{-1} at the peak of the 1957 event (Reed et al, 1963).

Figure 6.22 shows that the convergence of eddy momentum flux in high latitudes increased the westerly flow between 25 and 29.1.73. In the absence of this mechanism from 30.1.73 onwards the flow weakened by the action of the Coriolis torque and easterlies occurred for a short period (2 - 5.2.73) in high latitudes. Easterly flow was not established in middle latitudes as it was at higher levels.

Summary

The vertical distribution of the high latitude temperature and zonal wind changes which have been described are conveniently summarised in Figures 6.23 and 6.24, which are pressure-time sections of the zonal mean zonal wind at 68°N, the latitude at which the strongest easterly circulation occurred, and the change of zonal mean temperature at 80°N from its value on 11.1.73. The downward progression of these features of the warming event is well illustrated. The high level cooling and low level warming

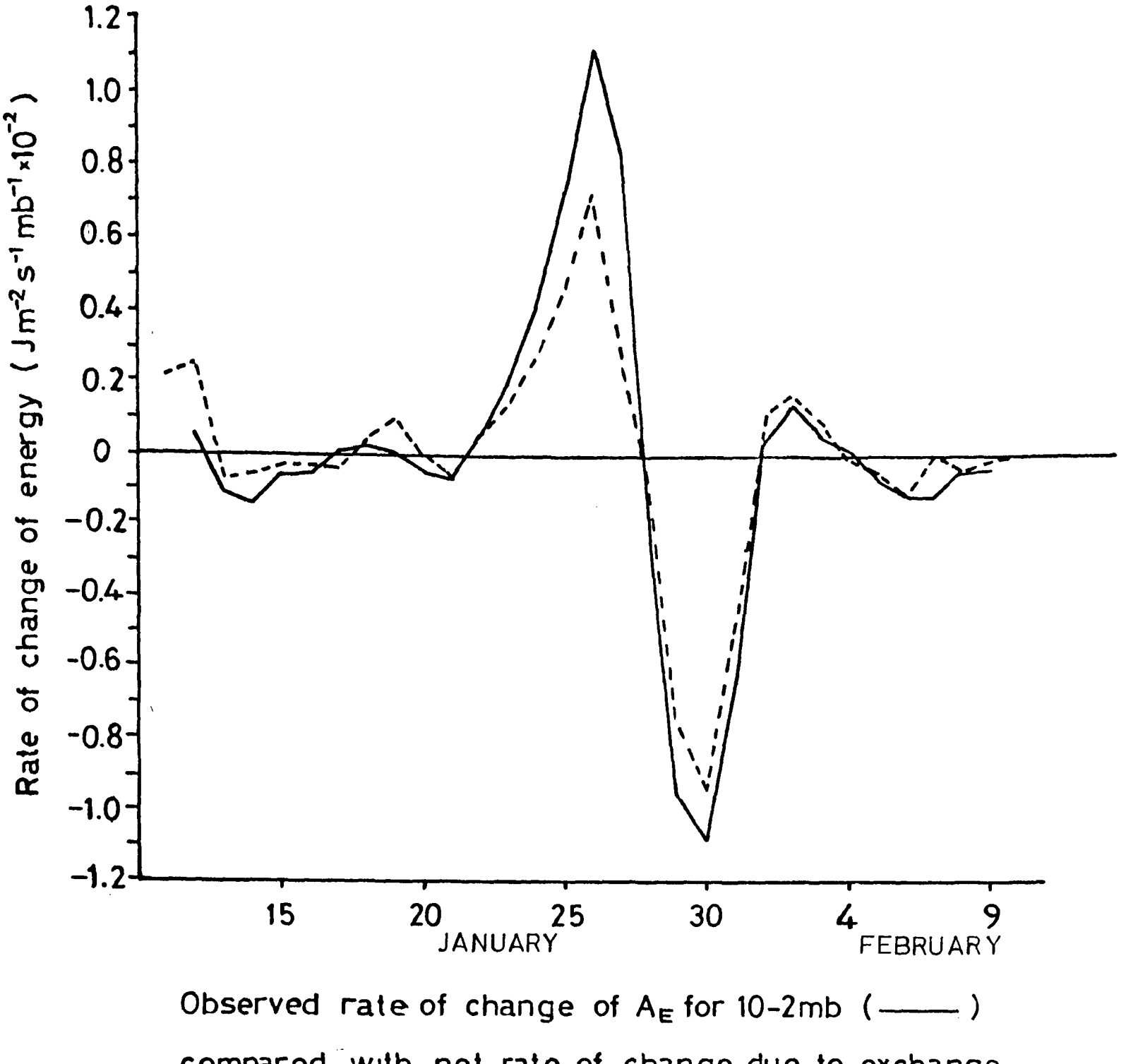


Figure 7.4b. The eddy available potential energy budget for the region 10 - 2 mb, 14 - 82°N between 11.1.73 - 10.2.73

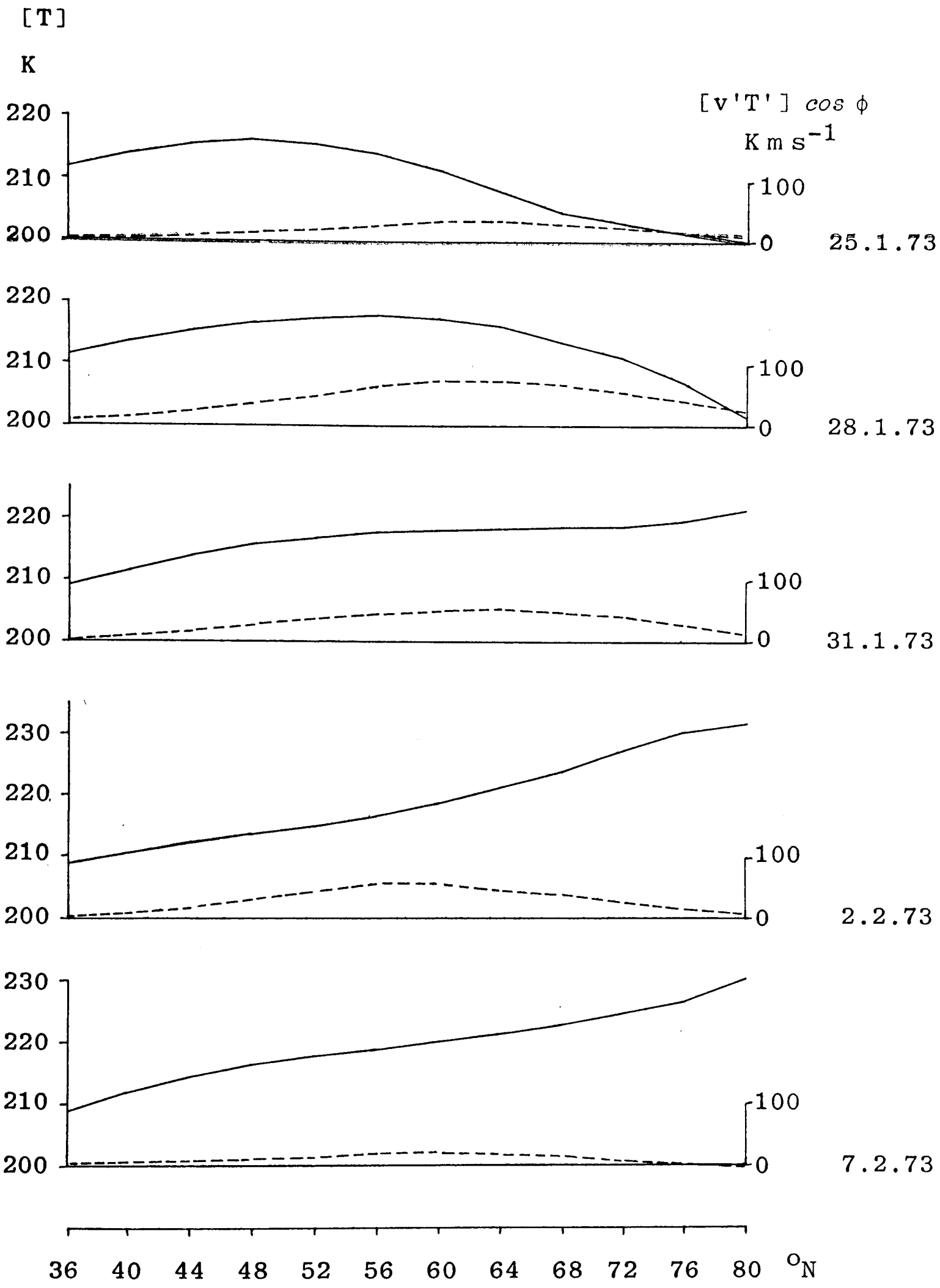
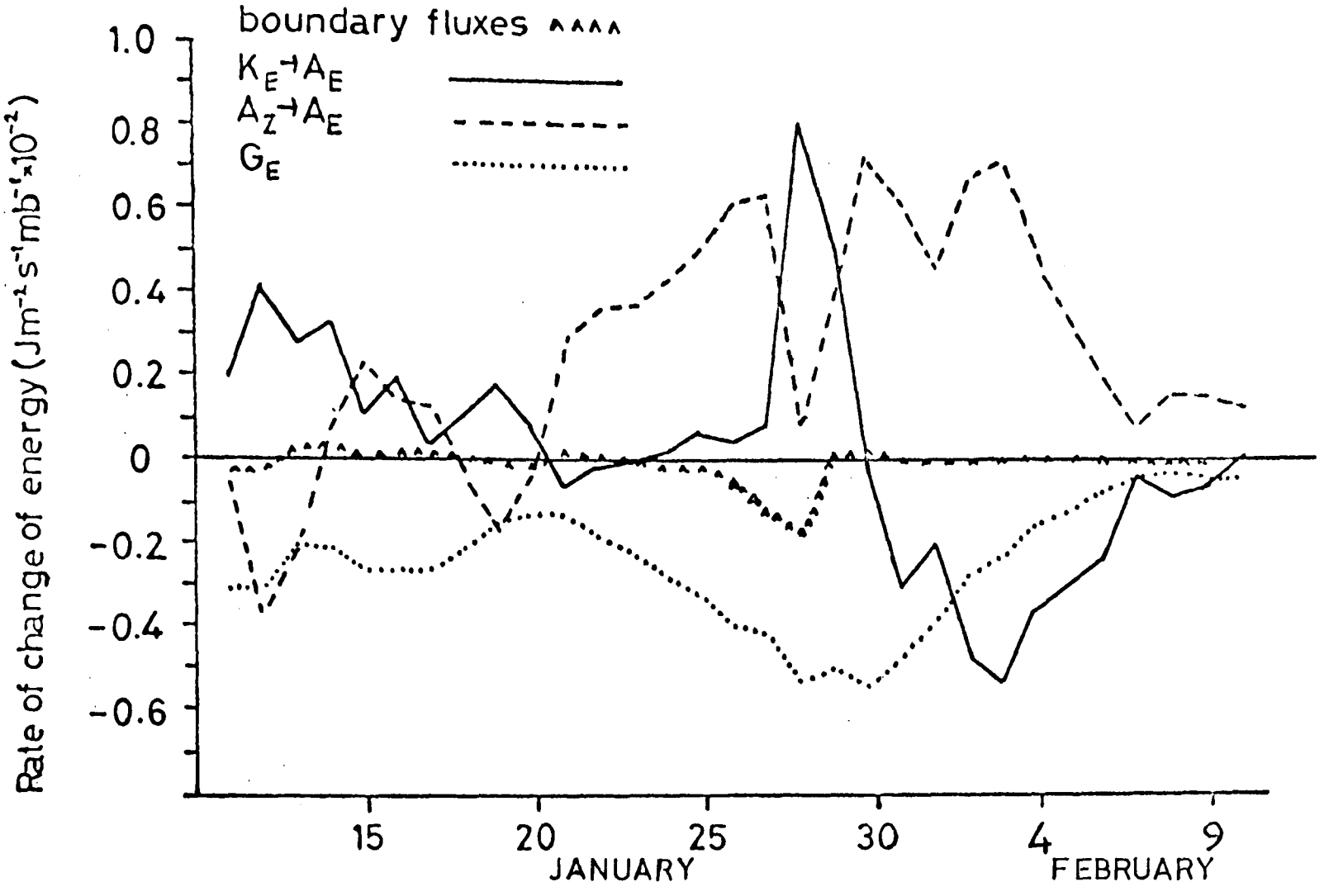
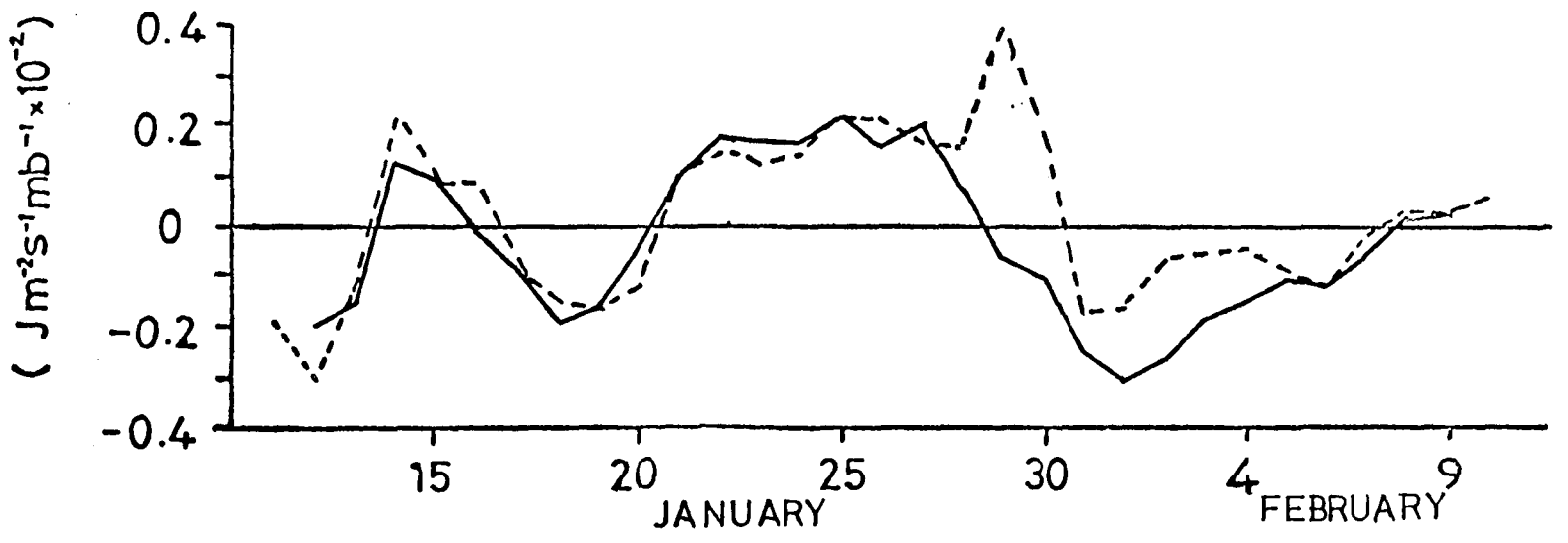


Figure 6.21. As Figure 6.9 but for 50 mb



Rates of change of A_E due to energy generation, conversions and boundary fluxes 2 - 0.4 mb



Observed rate of change of A_E (——) compared with net rate of change due to above processes (- - - - - -)

Figure 7.5b. The eddy available potential energy budget for the region 2 - 0.4 mb, 14 - 82°N between 11.1.73 - 10.2.73

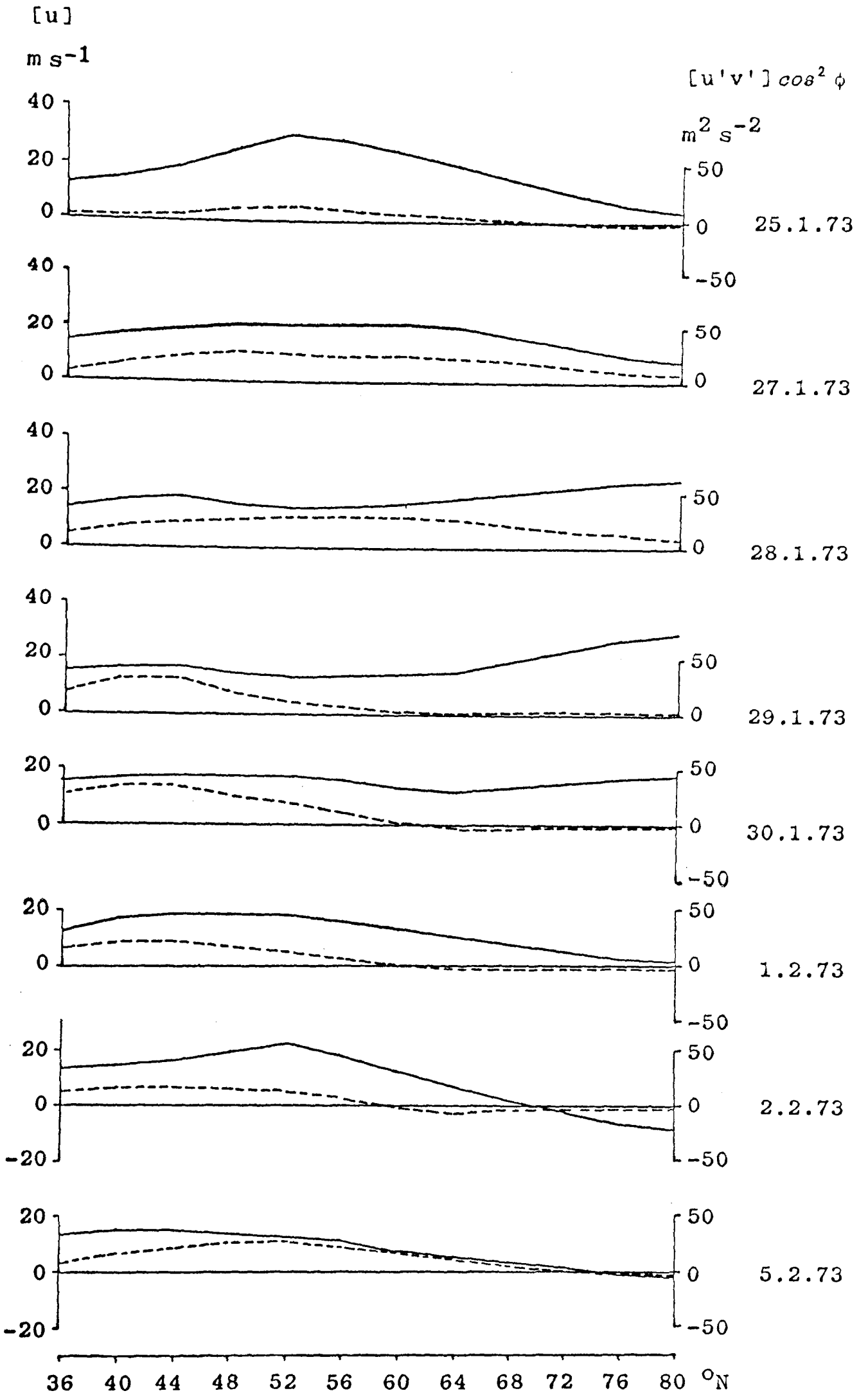


Figure 6.22. As Figure 6.12 but for 50 mb

as the pole cooled and a strong equator to pole temperature gradient developed (see Figure 6.25). This latter rate of increase in A_Z was the largest observed despite the fact that the conversion processes between 10 and 2 mb were two to three times larger than those in the top layer.

The reversal of the conversion processes below 2 mb at the peak of the warming occurred as a result of the poleward heat flux and the mean meridional circulation remaining in the same sense while the meridional temperature gradient reversed. Above 2 mb polar warming was short-lived and although there was a tendency for the A_E-A_Z conversion to reverse between 27 and 28.1.73 polar cooling quickly became established so that A_E remained a sink for A_Z and K_Z became a strong source via the mean meridional circulation. Radiative generation of A_Z was the major process throughout the warming in the top layer.

The eddy available potential energy budget

A similar tendency for the internal conversions to balance exists in the A_E budget where $A_E \rightarrow K_E$ conversion is usually observed to compensate the changes in A_E caused by conversion from A_Z to A_E.

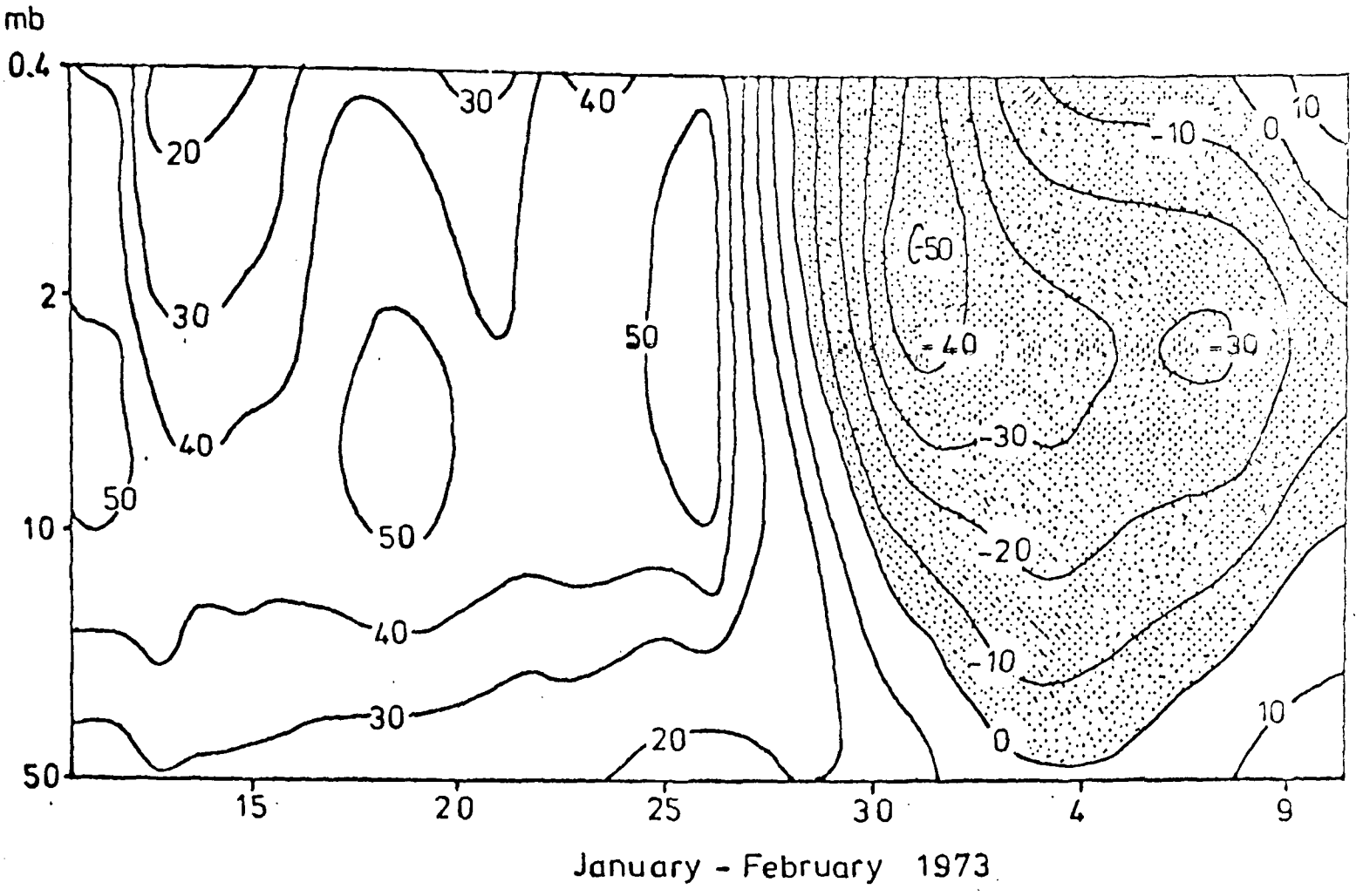


Figure 6.23. Vertical - time section of the zonal mean zonal wind [u] at 68°N during the period 11.1.73 to 10.2.73. Easterlies are shaded

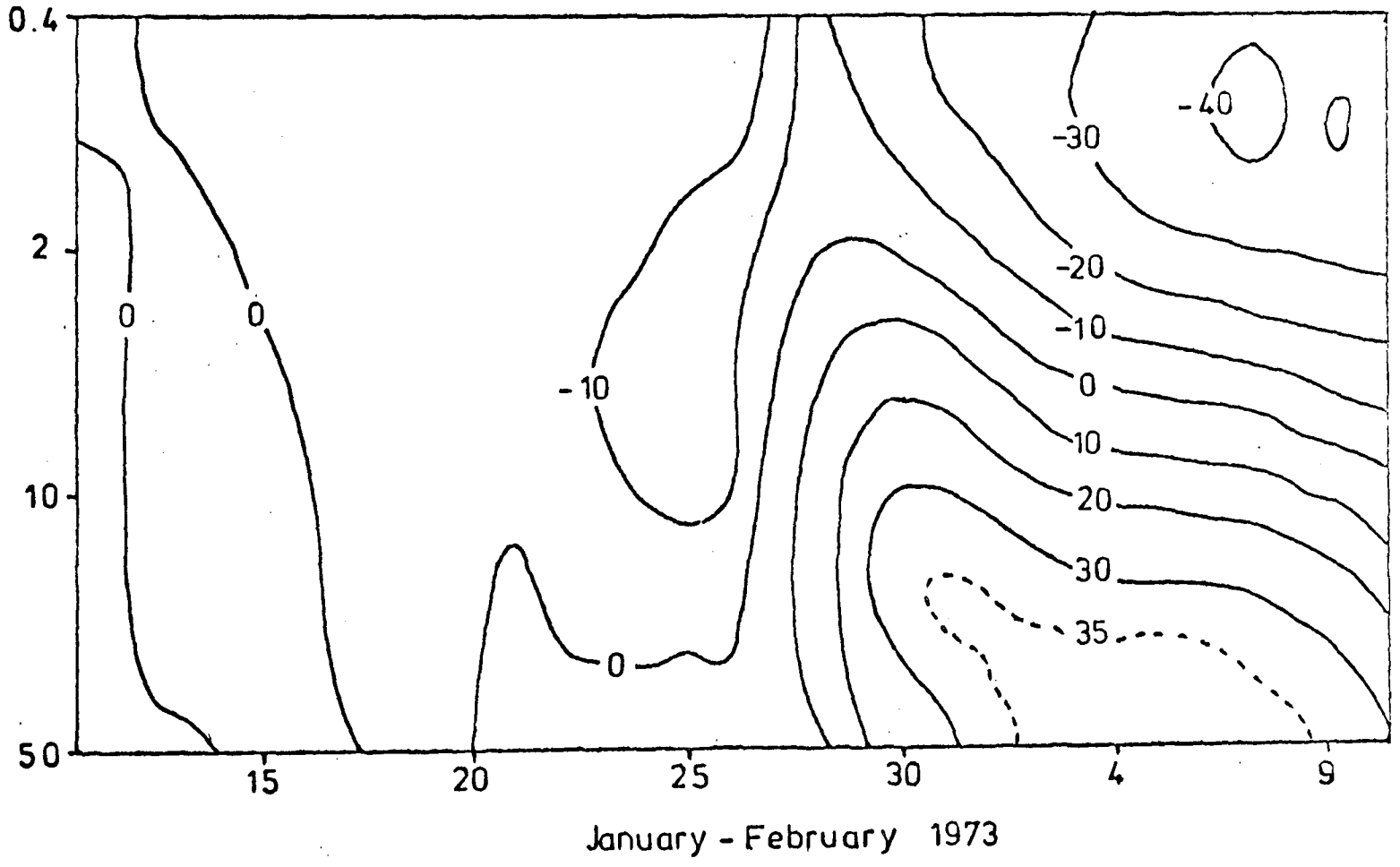


Figure 6.24. Vertical - time section of the change in zonal mean temperature at 80°N from its value on 11.1.73

rates of change of temperature shown in Figure 6.4b. It is clear that the anomaly of simultaneous increase of the $A_Z \rightarrow A_E$ and $K_E \rightarrow A_E$ conversions was a significant factor in the initiation of the warming. Since the A_Z budget was in near balance throughout the period it would appear that the eddy conversion was directed abnormally at this time. In Figure 7.6 time averaged height - latitude sections of the vertical shear of the zonal wind are compared with the distribution of $-[\omega'\alpha']$ ($A_E \rightarrow K_E$ conversion) in accordance with equation 7.1. Since $[v'T']$ was positive for the most part we should expect positive values of the wind shear to be associated with conversion from A_E to K_E. Although the correlation is far from exact there is broad agreement with this prediction. Note however that between 10 and 2 mb in the period 20 - 26.1.73 there is a significantly larger region contributing to the $K_E \rightarrow A_E$ conversion than would be expected from the zonal wind shear distribution. It therefore seems likely that this process was forced by external means. In the post-warming period very large $K_E \rightarrow A_E$ conversion throughout the 50 - 2 mb region was associated with the negative shear between westerlies at low levels and the strong easterly circulation which had been established in the upper stratosphere. As noted earlier eddy conversion in this sense was observed for a short period above 10 mb during the 1970-1 and 1974-5 warmings by Klinker (1976) with maximum values of -30×10^{-4} $JKg^{-1}s^{-1}$, but its significance was not clear since a detailed budget for this region was not calculated.

(ii) Kinetic Energy

It was shown in Chapter 5 that by using thermodynamically derived zonal mean vertical and meridional velocities a balanced zonal kinetic energy budget would not be obtained. It will be

resulted in a reduction of the temperature difference between 2 and 50 mb from 72K on 13.1.73 to -1K on 7.2.73, and is illustrated more graphically in Figure 6.25. The trend of the temperature at $12^{o}N$ is shown to be opposite that at $80^{o}N$ at all levels for most of the period. This common feature of the winter stratosphere was first discovered by Fritz and Soules (1970) but is not yet understood.

seen that there were similar imbalances in the eddy kinetic energy budget, almost certainly due to error in both the $K_E - K_Z$ conversion and in the eddy pressure-work term. In discussing the processes involved a qualitative approach must therefore be adopted. The evolutions of the processes contributing to the K_Z and K_E budgets of the 50 - 10 mb, 10 - 2 mb and 2 - 0.4 mb layers are given in Figures 7.7 - 7.9. On account of the imbalances only results for the period 21.1.73 - 4.2.73 are presented.

Zonal kinetic energy

Three distinct periods of change in the zonal flow are reflected in the observed changes of K_Z in the two upper layers: the increase of westerly winds prior to warming, the decline of these westerlies and the strengthening of the reversed, easterly circulation. Above 2 mb the replacement of polar easterlies with mid-latitude westerlies as the major jet is evident as a fourth phase. Below 10 mb the steady fall of K_Z throughout the period corresponded to steadily declining westerly flow. The changes in K_Z were generally larger than changes in A_Z.

It appears that during the amplifying stage energy was lost via the zonal mean pressure-work term, $B_{\phi z}$, and gained by the $K_E \rightarrow K_Z$ conversion at all levels. Above 2 mb these processes changed sign around 28.1.73. Interaction between K_Z and A_Z did not significantly affect the K_Z budget in the top layer, although between 10 and 2 mb $A_Z \rightarrow K_Z$ conversion was a major source of K_Z in the declining stage on account of the association between the mean meridional circulation and the very large pole to mid-latitude temperature gradient. The action of the mean meridional circulation largely controlled the K_Z budget below 10 mb from 31.1.73 onwards, A_Z being a source of K_Z and $B_{\phi z}$ a sink.

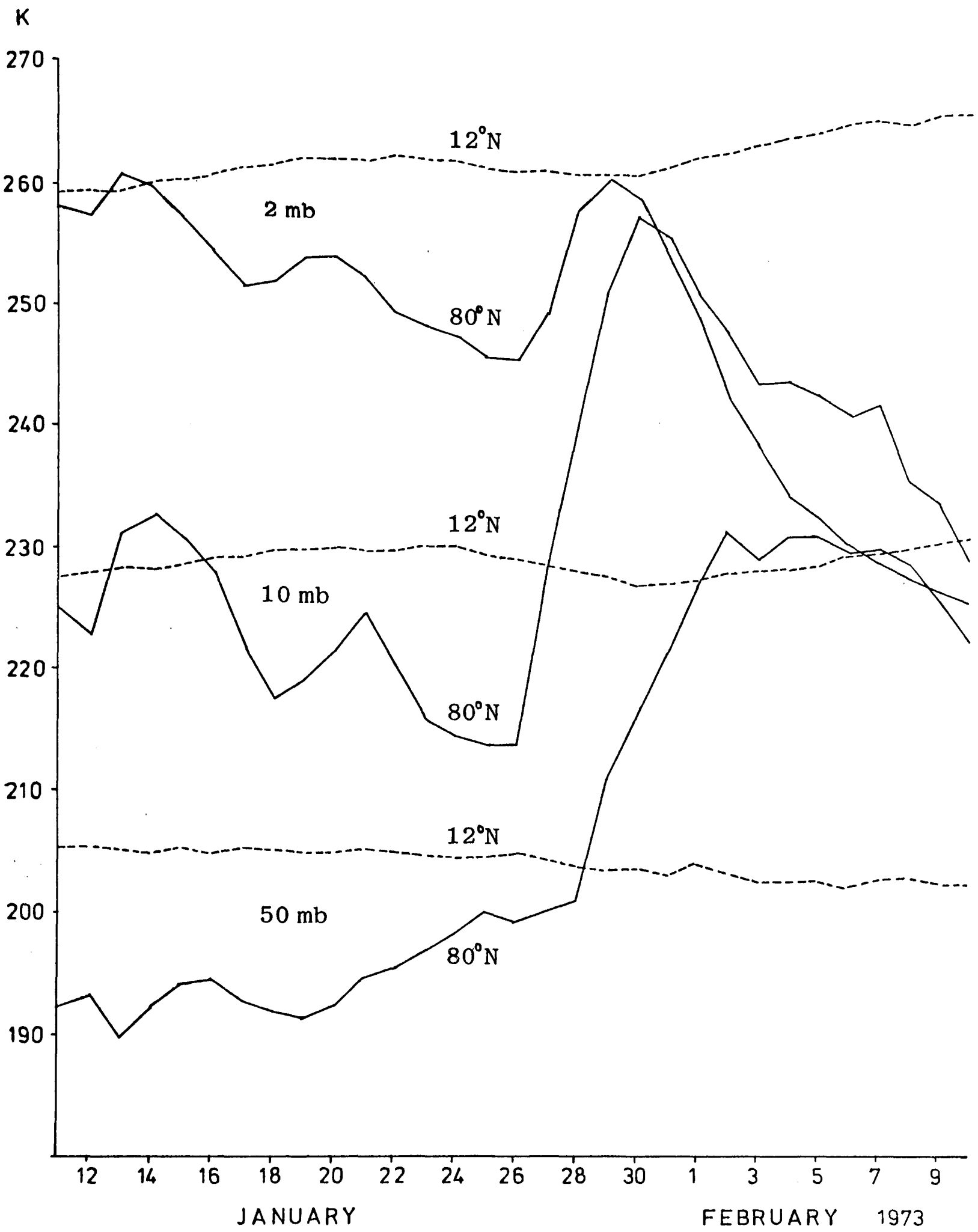


Figure 6.25. Zonal mean temperature [T] at 12^{o}N and 80^{o}N plotted against time for 2, 10 and 50 mb during the period 11.1.73 - 10.2.73

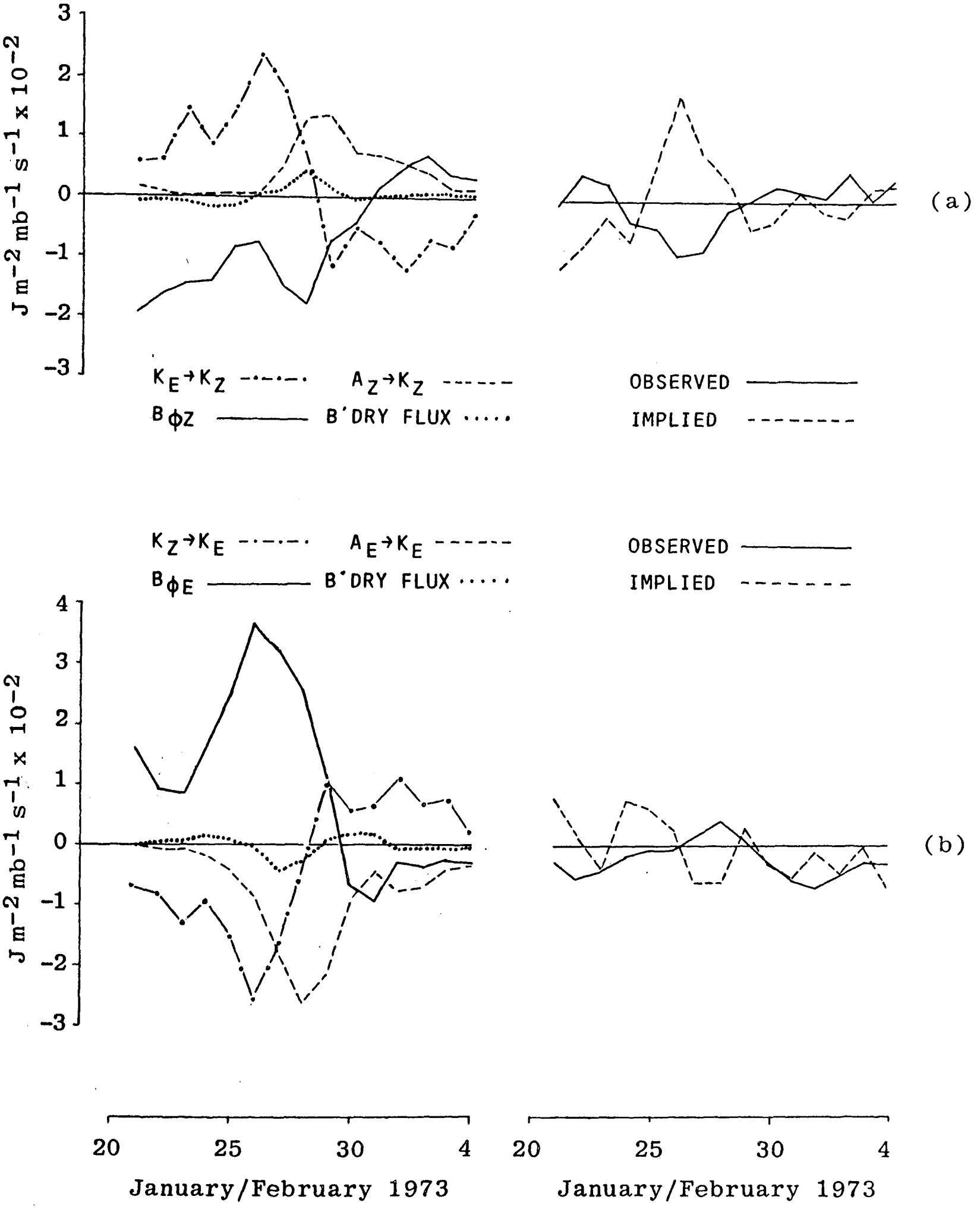


Figure 7.8. As Figure 7.7 but for the 10 - 2 mb layer

Introduction

The results of the previous chapter have demonstrated the very large scale changes which took place in the middle and upper stratosphere following the rapid amplification of a wave-number one planetary wave. In general wave energy may increase at the expense of the potential or kinetic energy of the mean state or may be altered through the action of pressure-work terms (see Chapter 4) and advective boundary fluxes. Analysis of the energetics of sudden warmings can thus indicate the mechanisms responsible for the rapid wave amplification which characterises them. Most studies of sudden warming energetics have, on account of the availability of data, been restricted to the troposphere and the low and middle stratosphere. However it has been shown that the warming process may begin in the upper stratosphere and propagate downwards, so a knowledge of the energy processes occurring there is of great importance. In this chapter the energetics of the 1973 warming in three stratospheric layers - the upper, upper-middle and lower-middle stratosphere - are examined and compared. By way of introduction a brief review of the findings of previous energy budget studies is given.

7.1. REVIEW OF SUDDEN WARMING ENERGY BUDGET STUDIES

Table 7.1 lists most of the important, published studies on the energetics of sudden warmings. All employed near-hemispheric data sets and described the energy processes in terms of the Lorenz cycle of energy conversions discussed in Chapter 4. In most results were averaged with respect to time and averaged over pressure and horizontal area. The much larger energy

Eddy kinetic energy

An unusual feature of the 1973 warming was the fact that very little change took place in the K_E content of the 50 - 10 mb layer. This observation contrasts not only with the results for the lower stratosphere during other warmings but also with those of Quiroz et al (1975) who studied the energetics of the 100 - 30 mb layer during this event. They observed a steady increase in K_E between 16 and 29.1.73. It appears that a rise in K_E occurred in the lower and the upper stratosphere but not at intermediate levels. The error in the conversion and generation processes is too large for definitive conclusions regarding their signs to be drawn from the K_E budget alone. However since a much closer balance was obtained for the K_Z budget of this layer it is probable that the $K_E \rightarrow K_Z$ conversion is a reasonable estimate and that the energy gain due to the eddy pressure-work term was overestimated.

In the middle layer energy gained via $B_{\phi E}$ was transferred to both A_E and K_Z, again with little net change in K_E. Above 2 mb, in contrast with lower levels, gain in K_E was very large. The kinetic energy field was largely decoupled from the available potential energy field with energy being transferred between K_E and K_Z and $B_{\phi E}$ acting as a source prior to 28.1.73 and a sink thereafter. Comparing Figures 7.9a and 7.9b it is clear that in the upper stratosphere the K_E-K_Z conversion was overestimated and the pressure-work terms underestimated throughout the warming period.

It should be noted that the horizontal components of the eddy pressure-work term $B_{\phi E}$ (equation 4.25) are identically zero if meridional velocities are calculated by the geostrophic approximation. The effect of these components on the kinetic

TABLE 7.1

SUMMARY OF IMPORTANT SUDDEN WARMING ENERGY BUDGET STUDIES OF HEMISPHERIC SCALE

Authors	Period studied & data frequency	Vertical extent of analysis	Vertical resolution	Other points
Reed, Wolfe & Nishimoto (1963)	25.1.57-9.2.57 5 days	50 mb level only		
Muench (1965)	1.1.58-31.1.58 2-4 days	100 - 10 mb	4 levels: 100,50,25 and 10 mb	Only the beginning of a warming at end of January sampled
Julian & Labitzke (1965)	1.1.63-31.1.63 2 days (stratospheric data every 5 days)	850 - 10 mb	8 levels: of which 100,30 and 10 mb in stratosphere	
Perry (1967)	1.1.63-31.1.63 Same data as J. & L. above + 1000 mb surface	1000 - 10 mb	9 levels: of which 100,30 and 10 mb in stratosphere	Calculated non-linear interactions between wavenumbers
Mahlman (1969)	15.11.58-15.12.58 daily	100 - 50 mb	2 levels: 100 and 50 mb	Warming of very limited latitudinal extent
Miller, Brown & Campana (1972)	16.12.69-15.1.70 Daily below 30 mb, between 3 and 7 day interval above	1000 - 2 mb	16 levels: 13 between 1000 and 30 mb + 10, 5 and 2 mb	Minor (or upper stratospheric) warming
Iwashima (1974)	23.12.67-8.1.68 daily	100 - 10 mb	4 levels	Analysed travelling and stationary waves separately
Quiroz, Miller & Nagatani (1975)	16.1.73-15.2.73	100 - 30 mb	Not stated	

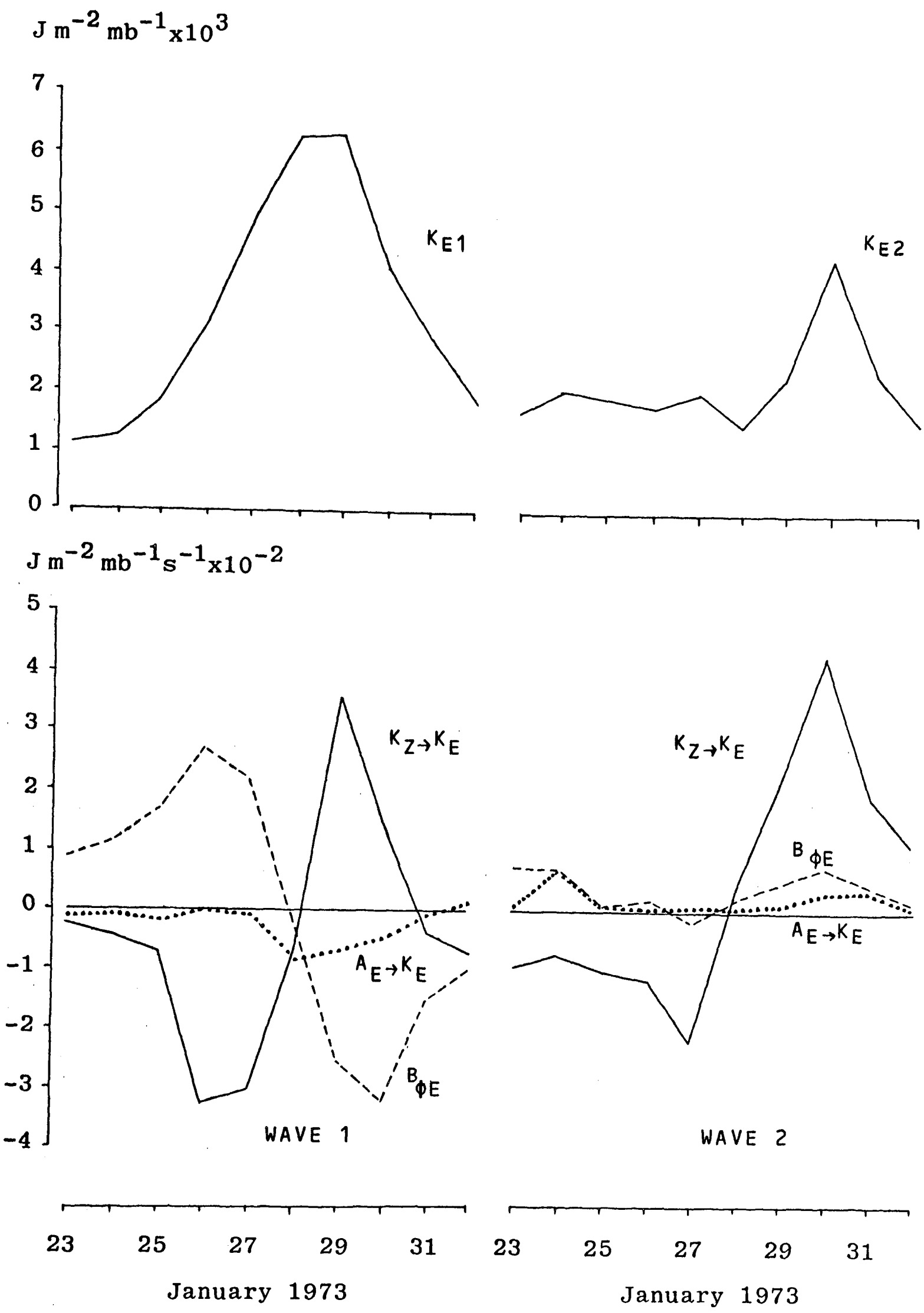


Figure 7.10. Mean content of wavenumber 1 (upper left) and wavenumber 2 (upper right) eddy kinetic energy in the 2 - 0.4 mb, 14 - 82°N region and the wavenumber 1 and 2 processes, respectively, contributing to the change in these energy components for the period 23.1.73 to 1.2.73. Energy exchanges between different wavenumbers are not shown

TABLE 7.1 (continued)

Authors	Period studied & data frequency	Vertical extent of analysis	Vertical resolution	Other points
Klinker (1976)	1.12.70-31.1.71 Daily up to 10 mb weekly above 10 mb	50 - 0.4 mb	5 levels: 50,30,10,2 and 0.4 mb	
	1.12.74-24.1.75 As above	As above	As above	Minor warming
Arpe (1976)	1.11.67-30.1.68 Daily	75 - 7.5 mb	5 levels: 500,100,50, 30 and 10 mb	Calculated non linear interactions between wavenumbers
	1.11.70-30.1.71 Daily	As above	As above	" "

conversion rates of high latitudes tend to dominate the area-averaged energetics despite the relatively small area they occupy. Comparison between different studies is often complicated by the different choice of boundaries, but they present a fairly consistent picture of the energy cycle during sudden warmings in the stratosphere below 10 mb.

The energetics below 10 mb during sudden warmings

The sudden warming may be considered as consisting of two stages: the first is the amplifying stage in which eddy energy increases and zonal mean energy declines; the second, declining, stage begins when the zonal flow reverses. Figure 7.1 shows the energy cycle which normally occurs below 10 mb during the first stage. This direct baroclinic cycle ($A_Z \rightarrow A_E \rightarrow K_E$) was found in all the major sudden warmings listed in Table 7.1 except in the

Service. During a sudden warming period the former values were at times larger by more than 50%. Imbalance in the kinetic energy budget is often attributed to frictional effects. Invariably however such imbalances are much larger than can be explained by this mechanism. Muench (1965), recognising that frictional loss could not explain the error, suggested that use of the adiabatic thermodynamic equation could have been the cause. Although this approximation may have contributed to the error, studies employing the full thermodynamic equation have suffered similar imbalances - for example that of Dopplick (1971) who, like Mahlman (1969), attributed a large residual in the K_Z budget to poor specification of $B_{\phi z}$.

Even if perfect data were available computational errors would remain, of which the use of finite difference techniques is probably the most significant and unpredictable, especially during periods of rapid temporal and spatial change. In any study of energetics it is therefore important to compare observed and implied rates of change of energy and to treat the calculated conversions and generations with an appropriate degree of caution.

(iv) The Lorenz energy cycles

To facilitate comparisons between the stratospheric layers and between pre- and post- warming energetics the conventional box diagrams have been compiled (Figures 7.11a, b). In the boxes are the observed (i.e. directly calculated) rate of change of energy type (top left), the rate of change implied by the conversions, generations and boundary fluxes (top right) and the amount of the energy type present. In order that the time scales of stratospheric energetics may be appreciated the units of conversion, generation, boundary flux and rate of change terms have been converted to $Jm^{-2}mb^{-1}day^{-1}$. The values shown are those

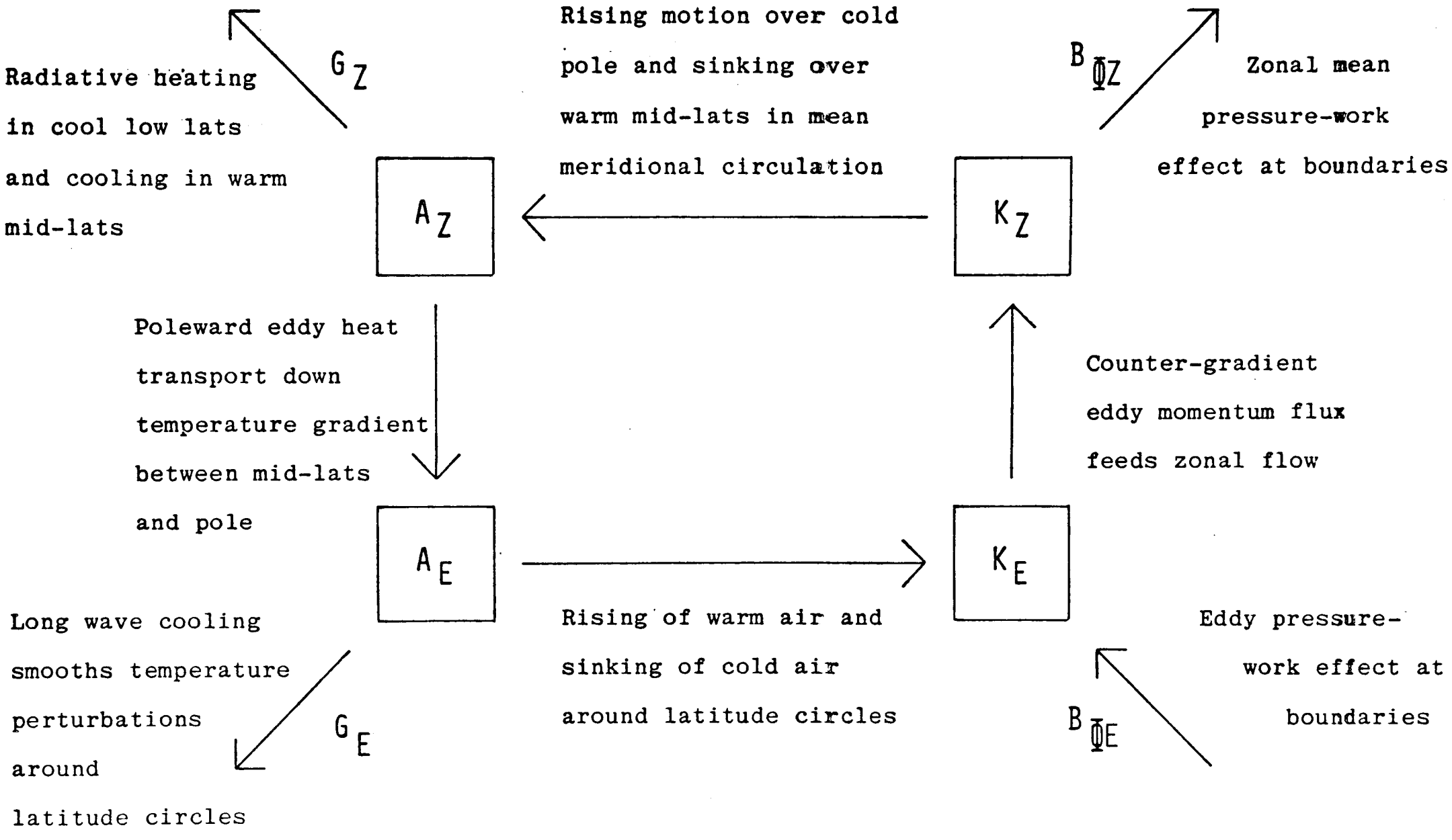


Figure 7.1. Commonly observed energy cycle below 10 mb during the amplifying stage of a sudden warming

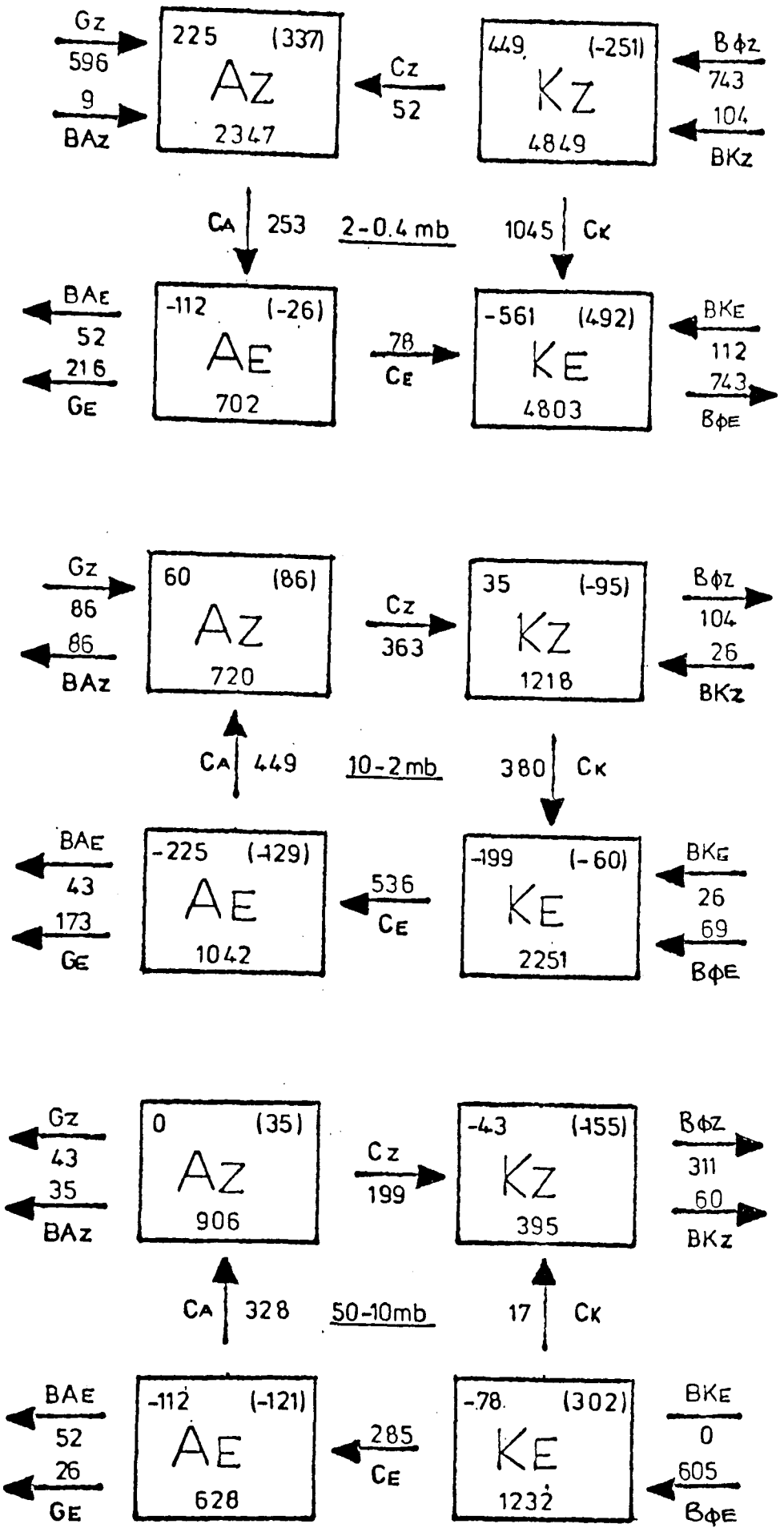


Figure 7.11b. Stratospheric energetics during the period 28.1.73 to 10.2.73. See caption to Figure 7.11a for explanatory notes

1973 warming reported by Quiroz et al (1975), a point to be discussed later. Both Muench (1965) and Dopplick (1971) observed baroclinic cycles for non-warming periods in the 100 - 10 mb layer of the northern hemisphere and Hartmann (1976) calculated energetics in the same sense for July and August 1973 in the southern hemisphere 100 - 10 mb region. It would therefore appear that this is the normal wintertime energy cycle for the lower middle stratosphere, which, as noted by Muench (1965), becomes more intense at the onset of a sudden warming. Gain in K_E due to energy propagation from the troposphere via the eddy pressure-work effect was a significant feature in all the sudden warming events listed in Table 7.1. On account of its importance, Chapter 8 is devoted to a more detailed study of this process and its effects during the 1973 warming. Tropospheric forcing by this means is a normal wintertime feature (Miller, 1970) which is enhanced during sudden warmings (Miller and Johnson, 1970).

With the reversal of the zonal circulation all the internal conversions, with the exception of the barotropic (K_E-K_Z) conversion, reverse. The observed cycle during the declining phase is shown in Figure 7.2. The K_E-K_Z conversion, although it usually remains in the eddy to zonal sense, may reverse, in which case it tends to be rather weak. Pressure-work effects continue to supply K_E and deplete K_Z, although at reduced rates. Destruction of available potential energy by the interaction of the meridional distributions of temperature and solar heating and by radiative relaxation of temperature perturbations around the globe is usually observed throughout the warming event.

The energetics above 10 mb during sudden warmings

Until the advent of satellite measurements the energetics of the upper stratosphere were a subject for speculation only. It

cycles prior to the warming peak although differences in generation terms and magnitudes of the conversions led to a large increase in A_E between 10 and 2 mb and in K_E between 2 and 0.4 mb. An important feature to note is that the pressure work terms and the barotropic conversion were about four times larger than other processes, excepting the radiative generation of A_Z above 2 mb which was a large energy source throughout. In the declining phase of the event strong reversed cycles below 2 mb maintained the warm pole much longer than in the upper stratosphere (Figure 6.25). A large increase in A_E in the middle stratosphere was observed during the 1969-70 warming by Miller et al (1972). However in that event the energy gain via the eddy pressure-work effect was much smaller than the internal conversions and the gain in A_E occurred via $A_Z \rightarrow A_E$ conversion rather than forced conversion from K_E. The dominance of the $B_{\phi E} \rightarrow K_E \rightarrow K_Z$ cycle above 10 mb was observed by Hartmann (1976) during the southern hemisphere winter of 1973. It is likely that this is a normal wintertime feature of the middle and upper stratosphere.

The overall effect of the warming was to transfer energy from the low and middle stratosphere to the upper stratosphere. This rearrangement of energy is seen clearly both in Figure 7.12, which shows the evolutions of the energy components in each layer, and in Table 7.2 where the sums of available and kinetic energy in each layer on 11.1.73 are compared with values for 10.2.73. The calculations show a 21% fall in energy of the 50 - 0.4 mb region over the period. In view of the changes in individual layers it is probable that this energy was lost to the mesosphere.

A comparison between the observed mean rate of change of A+K during the two phases of the period and the rates implied by conversion, generation and boundary terms is made in Table 7.3.

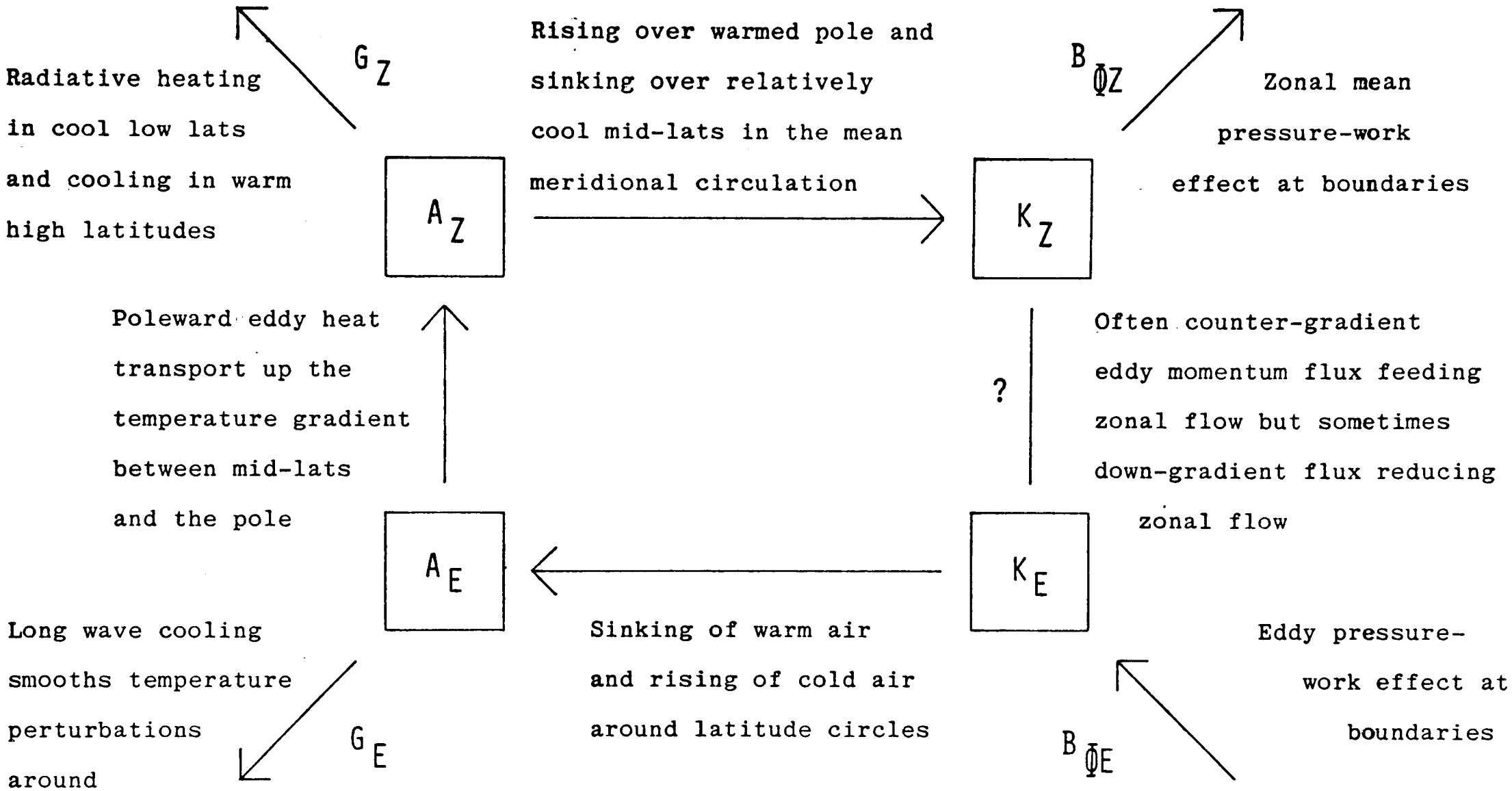


Figure 7.2. Commonly observed energy cycle below 10 mb during the declining stage of a sudden warming

TABLE 7.2

A+K CONTENT (MEAN OVER 14 - 82°N) ($Jm^{-2}mb^{-1} \times 10^3$)

mb	11.1.73	10.2.73	content on 10.2.73 as % of 11.1.73
2 - 0.4	8.9	13.5	152
10 - 2	9.4	3.9	41
50 - 10	5.8	1.7	29
50 - 0.4	24.1	19.1	79

TABLE 7.3

MEAN RATE OF CHANGE OF A+K ($Jm^{-2}mb^{-1}day^{-1} \times 10^3$)

	11 - 27.1.73		28.1 - 10.2.73	
	observed	implied	observed	implied
mb				
2 - 0.4	0.23	2.39	0.00	0.55
10 - 2	-0.12	0.04	-0.33	-0.20
50 - 10	-0.10	0.01	-0.23	0.06

It is seen that in each layer the implied energy gain was larger than observed in both periods. This suggests energy gains by the generation and/or pressure-work terms were consistently over-estimated, advective boundary fluxes being small.

The redistribution of energy between the different components within each layer and how these distributions compare with other observations is of interest. In Table 7.4 are presented the energy distributions in the stratosphere during a southern hemisphere winter (Hartmann, 1976) and a relatively quiet northern hemisphere winter (Dopplick, 1971), and during the pre- and post-

was assumed that the latitudinal distribution of solar radiation would be a large source of available potential energy and that a baroclinic energy cycle would prevail (Newell, 1964). The first attempt to investigate the energetics above 10 mb was made by Miller et al (1972). They made calculations for the 30 - 2 mb layer for seven separate days during the warming confined to the upper stratosphere in December - January 1969/70, using data from the satellite infra-red spectrometer (SIRS) experiment on Nimbus 3. An $A_Z \rightarrow A_E \rightarrow K_E$ cycle during the amplifying stage was accompanied by energy gain through the eddy pressure-work effect and a large $K_E \rightarrow K_Z$ conversion. Unlike the major warmings discussed above, this minor warming was characterised by a $K_E \rightarrow A_E$ conversion in the first stage in the 100 - 30 mb layer. The authors suggested that the absence of baroclinic conversion in the lower stratosphere may have been associated with the failure of this warming to affect the whole stratosphere. However the extent to which warming descends into the lower stratosphere is clearly dependent on other factors since the presence of baroclinic conversion in the lower stratosphere did not ensure the downward propagation of the minor warming of 1974-5 observed by Klinker (1976) or the upper stratospheric warming at the end of January, 1964 (Dopplick, 1971). In contrast with the results of Miller et al, Klinker (1976) calculated a $K_E \rightarrow A_E$ conversion between 30 and 2 mb during the first phase of both a major (1970-1) and a minor (1974-5) warming. Both these studies were based on few days' data and Klinker did not attempt to produce a complete budget above 10 mb. The results described below thus form the first detailed set of energetics calculations for the upper stratosphere during a major sudden warming.

afterwards, the region below 2 mb contributing more to this reduction than that above. These values compare with a ratio of 9:1 for the southern winter, the difference being due not so much to smaller amounts of available energy but to the very large K_Z component there. From Figure 7.12 it is seen that the ratio of eddy to zonal energy at the peak of the warming was a little less than 4:1 between 2 and 0.4 mb and nearly 6:1 between 10 and 2 mb. These ratios fell substantially after the event, but only from about 4.2.73 onwards and above 2 mb did the zonal components become substantially larger than the eddy components, with what was probably a short-lived maximum Z:E ratio of 7:1 occurring on 8.2.73. Over the 50 - 0.4 mb region the time averaged value of the E:Z ratio was 1.3:1 in the pre-warming period and 1:1 in the post-warming phase. In the southern hemisphere winter Z:E ratios of between 6:1 and 7:1 were observed for all three layers. Comparing the Z:E ratios for the lower stratosphere only, values of 6.7:1, 1.6:1 and 0.7:1 were obtained for the southern winter, the quiet northern winter and the whole of the sudden warming period respectively. If these results are typical it would appear that only occasionally does the stability of the northern hemisphere winter zonal circulation approach that of its southern counterpart. Moreover the relative stability of the circulation in a quiet northern winter compared with a disturbed northern winter is much less than that of a quiet southern winter compared with a quiet northern winter.

7.2. HEMISPHERIC ENERGETICS OF THE 1973 WARMING

Presented in this section are daily values of the energetics of the three stratospheric layers 50 - 10 mb, 10 - 2mb and 2 - 0.4 mb, averaged over latitudes 14 - 82°N. Values calculated at latitude $\emptyset^{o}$ were considered to be representative of the latitude band $\emptyset$ - 2° to $\emptyset$ + 2°. The rates of energy conversion, generation and advection contributing to each energy component are plotted against time and the sum of these terms are compared with the observed changes.[1] Each component is considered in turn, after which an overall view of the energetics is presented with the aid of the conventional energy cycle 'box' diagrams.

(i) Available potential energy

It was noted in Chapter 5 that use of vertical motion fields calculated to balance the thermodynamic equation in the evaluation of available potential energy conversions should ensure a balanced budget. In practice errors occur through use of the geostrophic approximation in calculating vertical motions, the use of finite difference techniques and the application of the energy relationships to a limited atmospheric region as discussed in Chapter 4. Figures 7.3a, 7.4a and 7.5a illustrate the variation with time of the processes which contributed to the changes in A_Z in the 50 - 10 mb, 10 - 2 mb and 2 - 0.4 mb layers respectively. The lower portions of each figure, which compare the calculated and the observed rate of energy change, indicate that in the two lower layers a near balanced budget was obtained and that errors of the type noted above were small. Above 2 mb errors were greater, although still not large relative to the magnitude of the

---

[1] Throughout the chapter 'observed' energy content refers to the value calculated from the definition of the energy type given by equations 4.13 - 4.16. Observed rates of energy change were evaluated using centred finite differences.

in geostrophic balance the wave energy equation

$$\frac{\partial}{\partial y}[v'\Phi'] + \frac{\partial}{\partial p}[\omega'\Phi'] = -[u'v']\frac{\partial}{\partial y}[u] - [u'\omega']\frac{\partial}{\partial p}[u] + f\sigma^{-1}\frac{\partial}{\partial p}[u]\ [v'\frac{\partial}{\partial p}\Phi'] \qquad 8.1$$

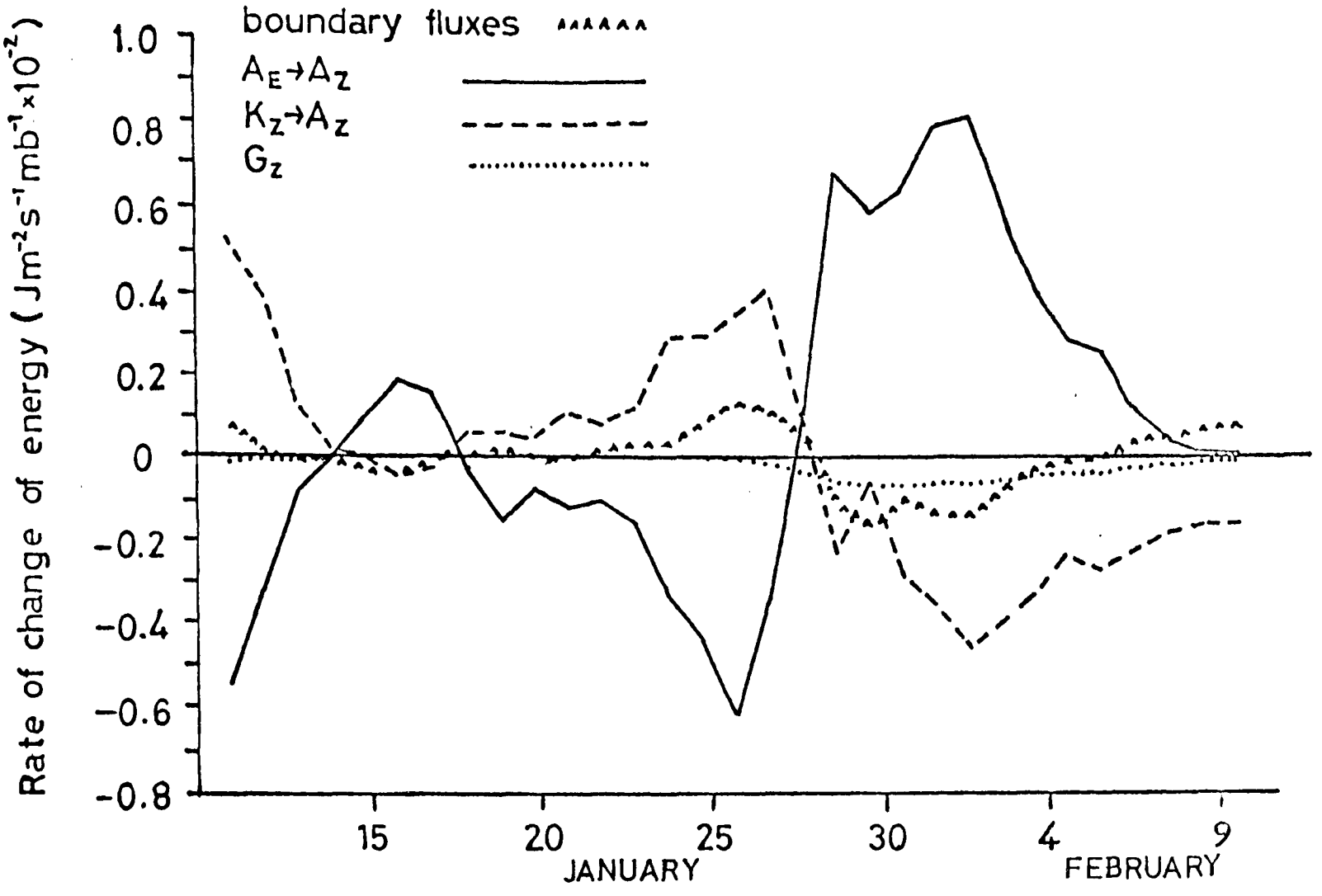
where Φ' is the perturbation geopotential, σ is a static stability parameter and other terms have their usual meanings. They regarded the terms on the right hand side as conversions from zonal kinetic and zonal potential energy to the corresponding eddy forms, and the left hand side as wave energy flux divergence. Using the thermal wind equation (equation 1.5) to replace $\partial[u]/\partial p$ in the last term it is seen that these conversion terms are those given in Chapter 4 as being appropriate only for a closed system, i.e. for a region with no flow through the boundaries. In order to present in a physically more meaningful way the conversions between the wave and the zonal flow Holton (1974) recast the equation in the form

$$\frac{\partial}{\partial y}([v'\Phi'] + [u][u'v']) + \frac{\partial}{\partial p}([\omega'\Phi'] - f\sigma^{-1}[u][v'\frac{\partial}{\partial p}\Phi'] + [u][u'\omega']) = [u]F \qquad 8.2$$

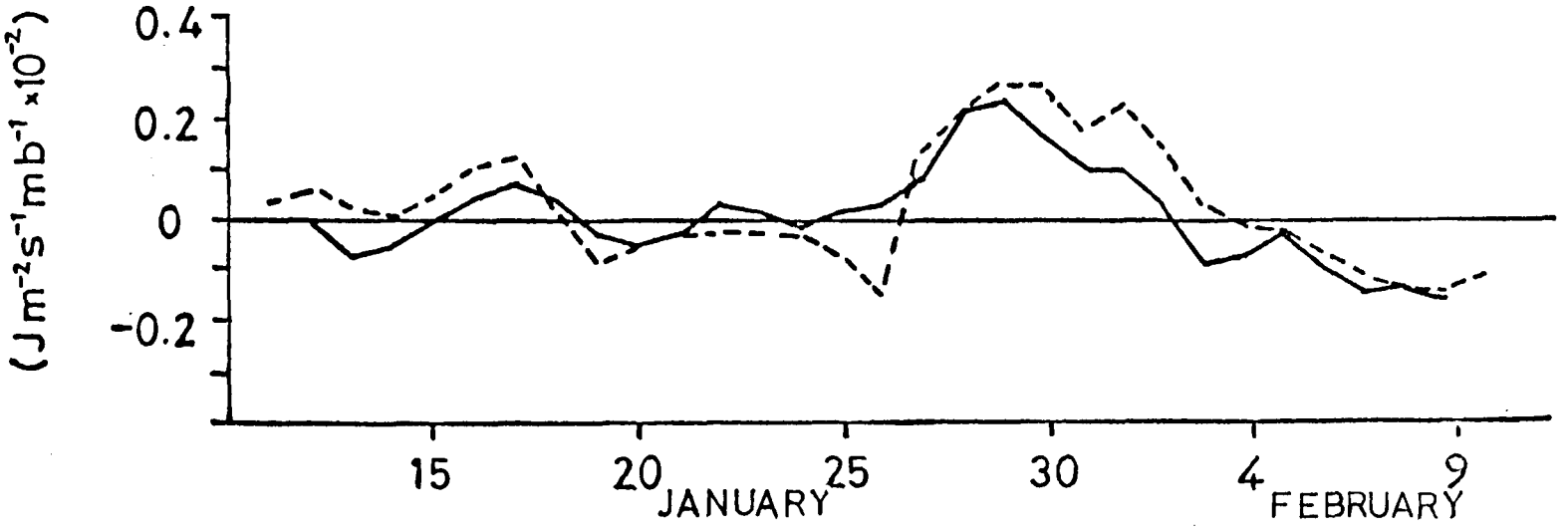
where F, the 'eddy forcing', is given by

$$\frac{\partial}{\partial y}[u'v'] + \frac{\partial}{\partial p}[u'\omega'] - \frac{\partial}{\partial p}(f\sigma^{-1}[v'\frac{\partial}{\partial p}\Phi'])$$

The third term in F is a rewritten form of the Coriolis torque term $f[u][v]$ which represents the total loss of zonal kinetic energy through the action of the mean meridional circulation. Holton thus regarded eddy forcing as all processes which alter the zonal kinetic energy while Eliassen and Palm considered wave - mean flow interaction as involving interchange between both eddy and both zonal mean energy components. As a result of the different conversion terms the energy flux divergences are different: the left hand side of equation 8.1 is the divergence of eddy kinetic energy flux due to pressure-work terms



Rates of change of A_z due to energy generation, conversions and boundary fluxes 50–10 mb



Observed rate of change of A_z(——) compared with net rate of change due to above processes (-------)

Figure 7.3a. The zonal available potential energy budget for the region 50 - 10 mb, 14 - 82°N between 11.1.73 and 10.2.73

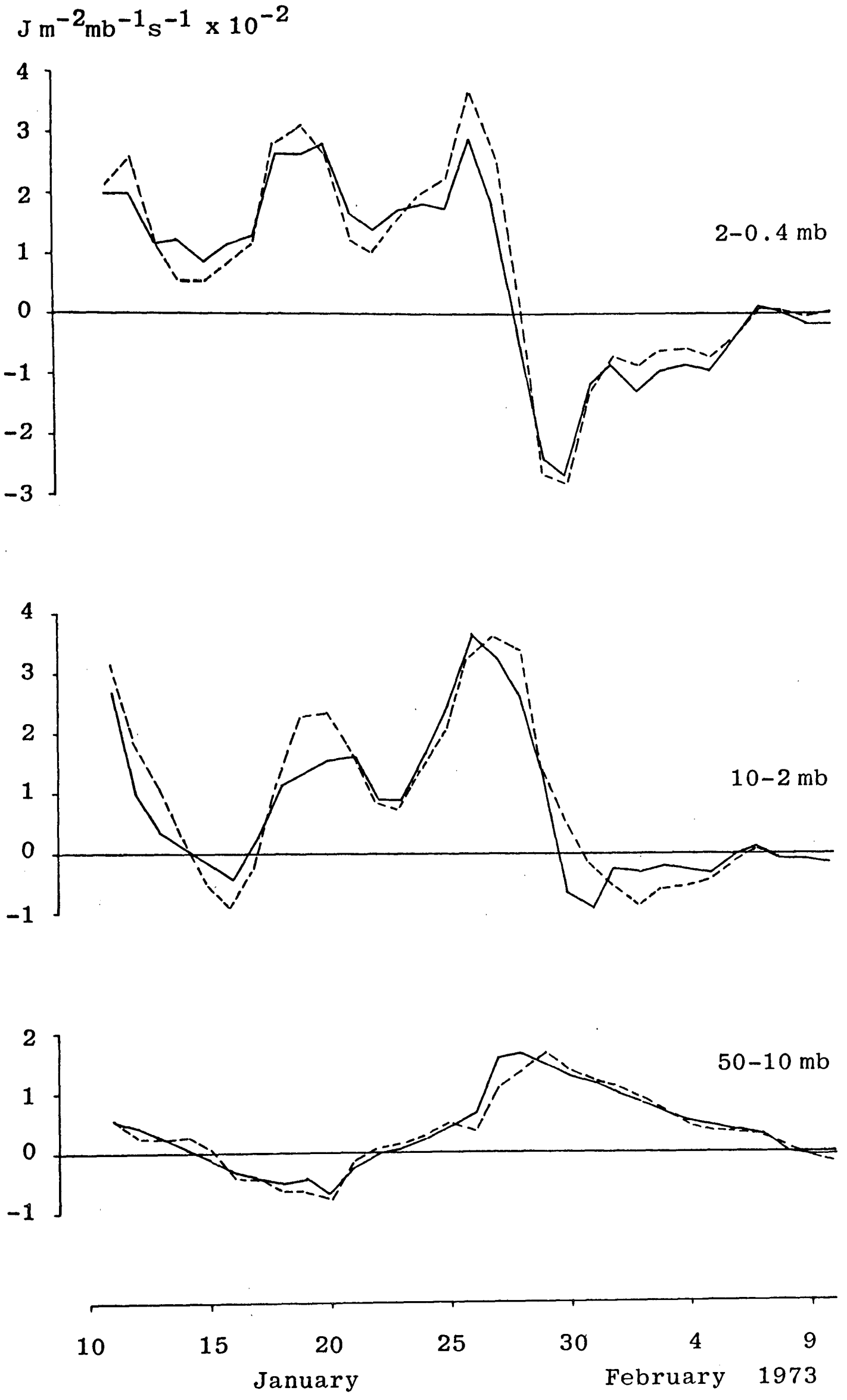
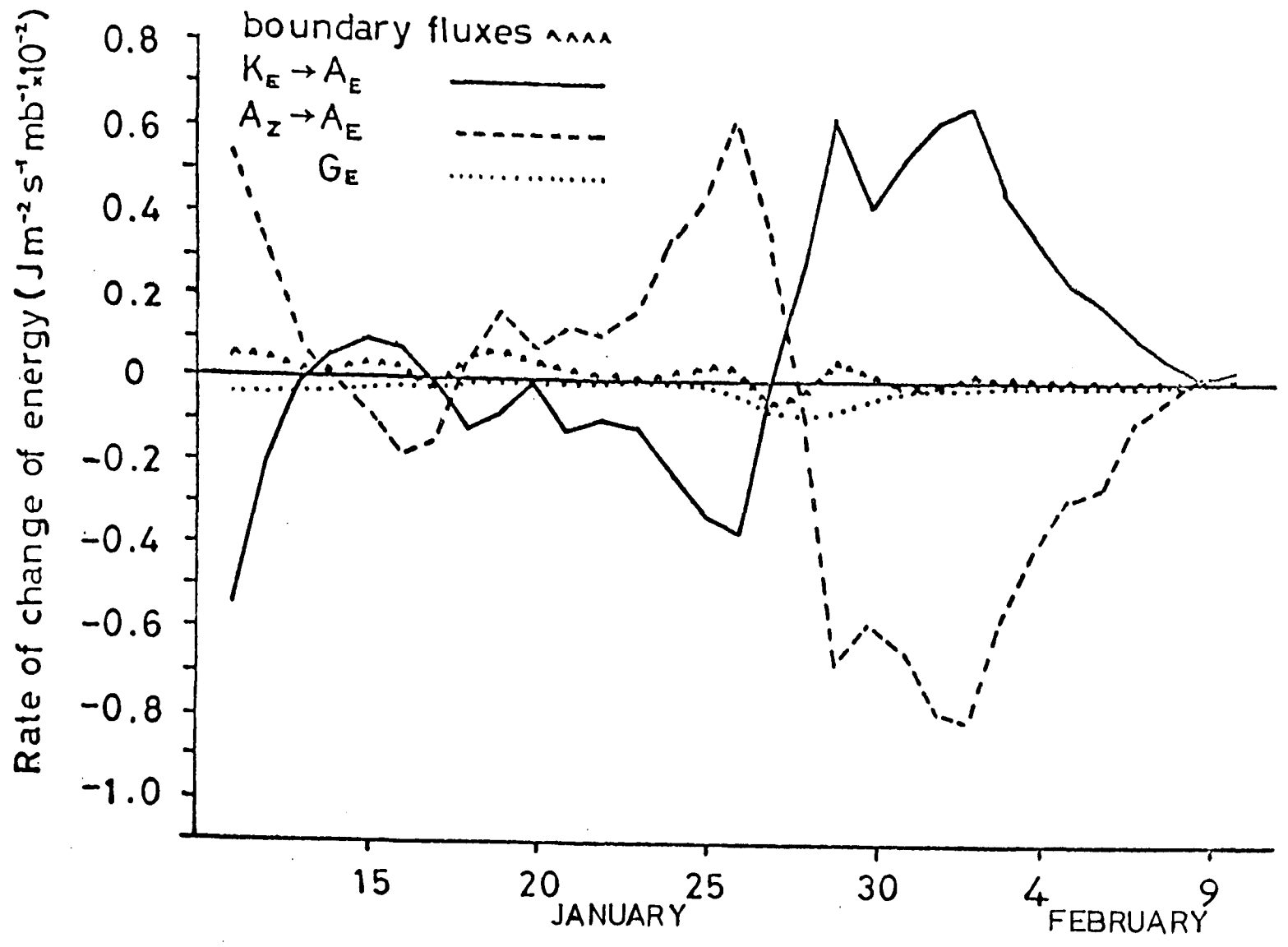
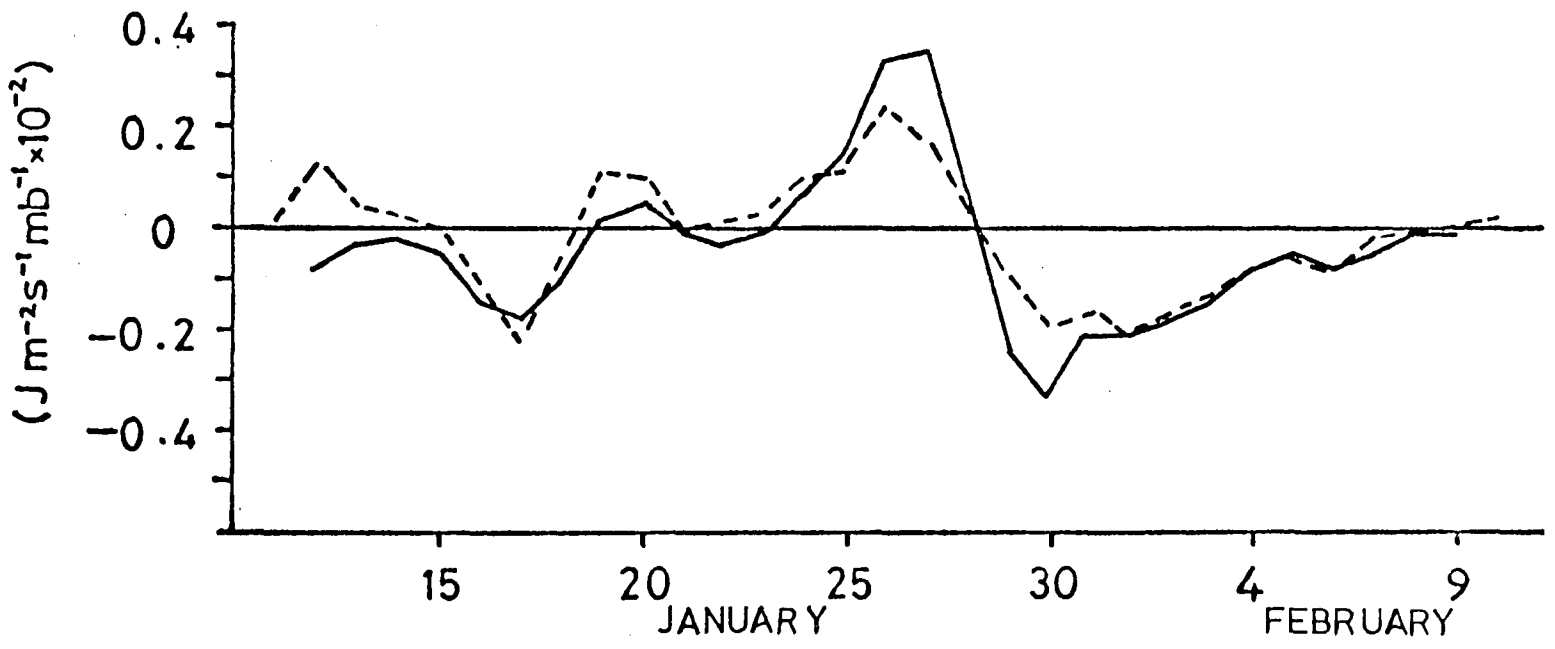


Figure 8.1. Comparisons of the convergence of the eddy vertical energy flux $B_{\Phi E}$ calculated directly (——) and calculated using the Eliassen and Palm (1961) approximation (----) for the 2 - 0.4, 10 - 2, and 50 - 10 mb layers during the period 11.1.73 to 10.2.73 (J m^{-2} mb^{-1} s^{-1} x 10^{-2})



Rates of change of A_E due to energy generation, conversions and boundary fluxes 50–10mb



Observed rate of change of A_E (——) compared with net rate of change due to above processes (- - - - -)

Figure 7.3b. The eddy available potential energy budget for the region 50 - 10 mb, 14 - 82°N between 11.1.73 and 10.2.73

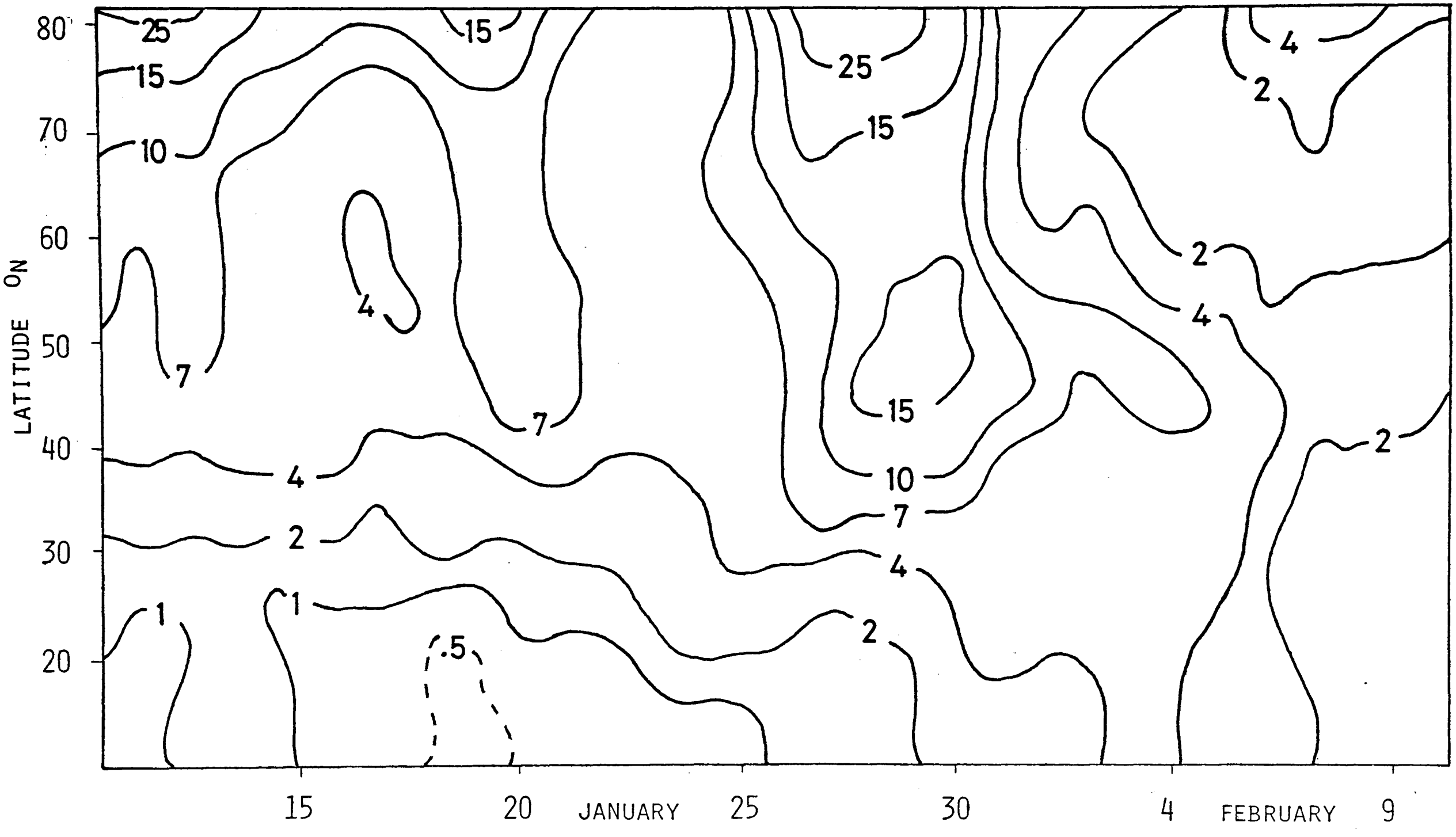
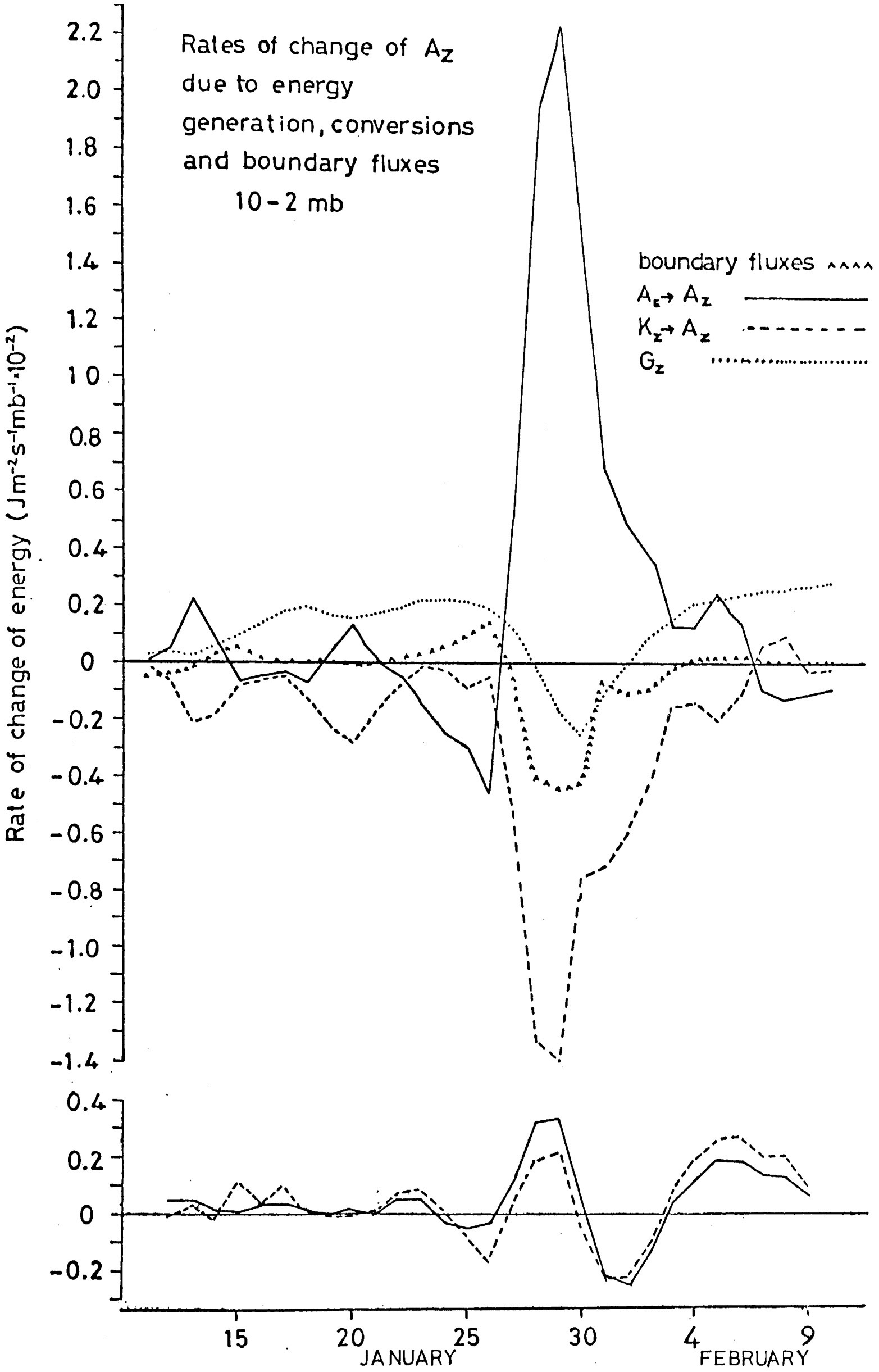


Figure 8.2. Latitude - time section of mean eddy kinetic energy content of the 2 - 0.4 mb layer for the period 11.1.73 to 10.2.73 ($J\,m^{-2}\,mb^{-1} \times 10^3$)



Observed rate of change of A_Z(——) compared with net rate of change due to above processes (- - - - - - -)

Figure 7.4a. The zonal available potential energy budget for the region 10-2 mb, 14-82 ^{o}N between 11.1.73 - 10.2.73

two component was predominantly downwards at 50 mb during the warming with maximum downward fluxes occurring on 24 and 30.1.73. Whilst during the energy pulse of 11.1.73 and during the warming there was convergence of vertical energy flux throughout the stratosphere, in the intervening period the total flux was downward at 50 mb and upward at 10 mb. There must, therefore, have been an energy source within the 50 - 10 mb region from which energy was propagated to higher levels. The limited accuracy of the components of the kinetic energy budget prevents us from deducing which energy conversion process contributed most to the amplification of the vertical energy flux in this layer; there was weak baroclinic conversion from A_E to K_E at this time but a marked reversal of the $K_E - K_Z$ conversion occurred on 20.1.73 suggesting that barotropic conversion may have been important also.

Latitudinal distributions of vertical energy flux have shown that large fluxes are confined mainly to high latitudes (see, for example, Miyakoda et al, 1970; Labitzke et al, 1975; Quiroz and Nagatani, 1975; Klinker, 1976). The latitude-time section of the flux through 50 mb in Figure 8.4 confirms that this was the case in the 1973 warming. In view of this observation it is more useful to study the behaviour of the vertical energy flux in high latitudes rather than on a hemispheric basis. The time change of the vertical energy flux at $76^{o}N$ for various levels is shown in Figure 8.5. What is significant is that there was energy flux divergence between 50 and 10 mb not only during the period 14 - 22.1.73 indicated by the area-averaged results but throughout the period 11 - 26.1.73. At the same time there was convergence between 10 and 0.4 mb on all days except 15 - 17.1.73. At this latitude increased vertical energy flux preceding the warming

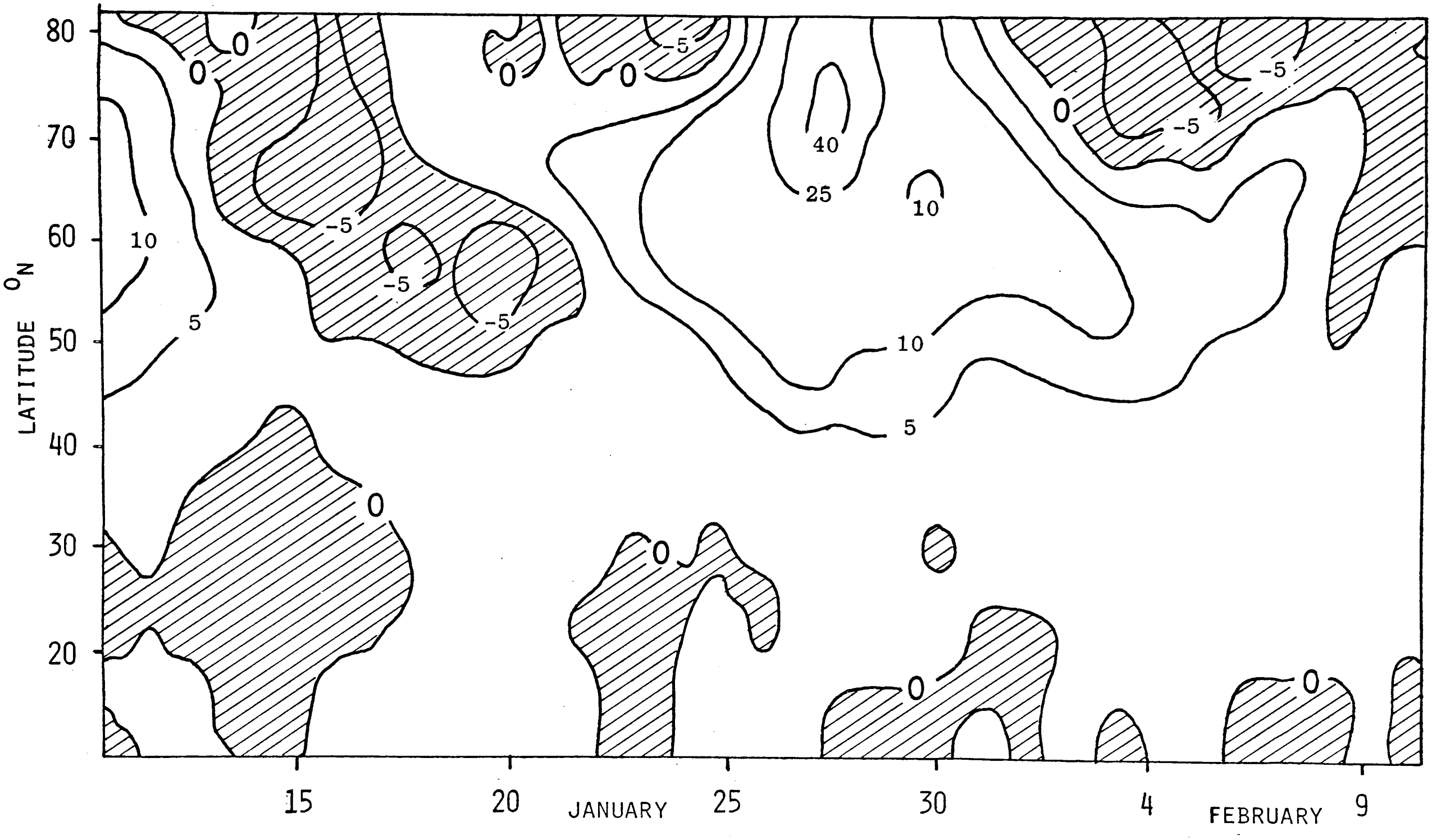
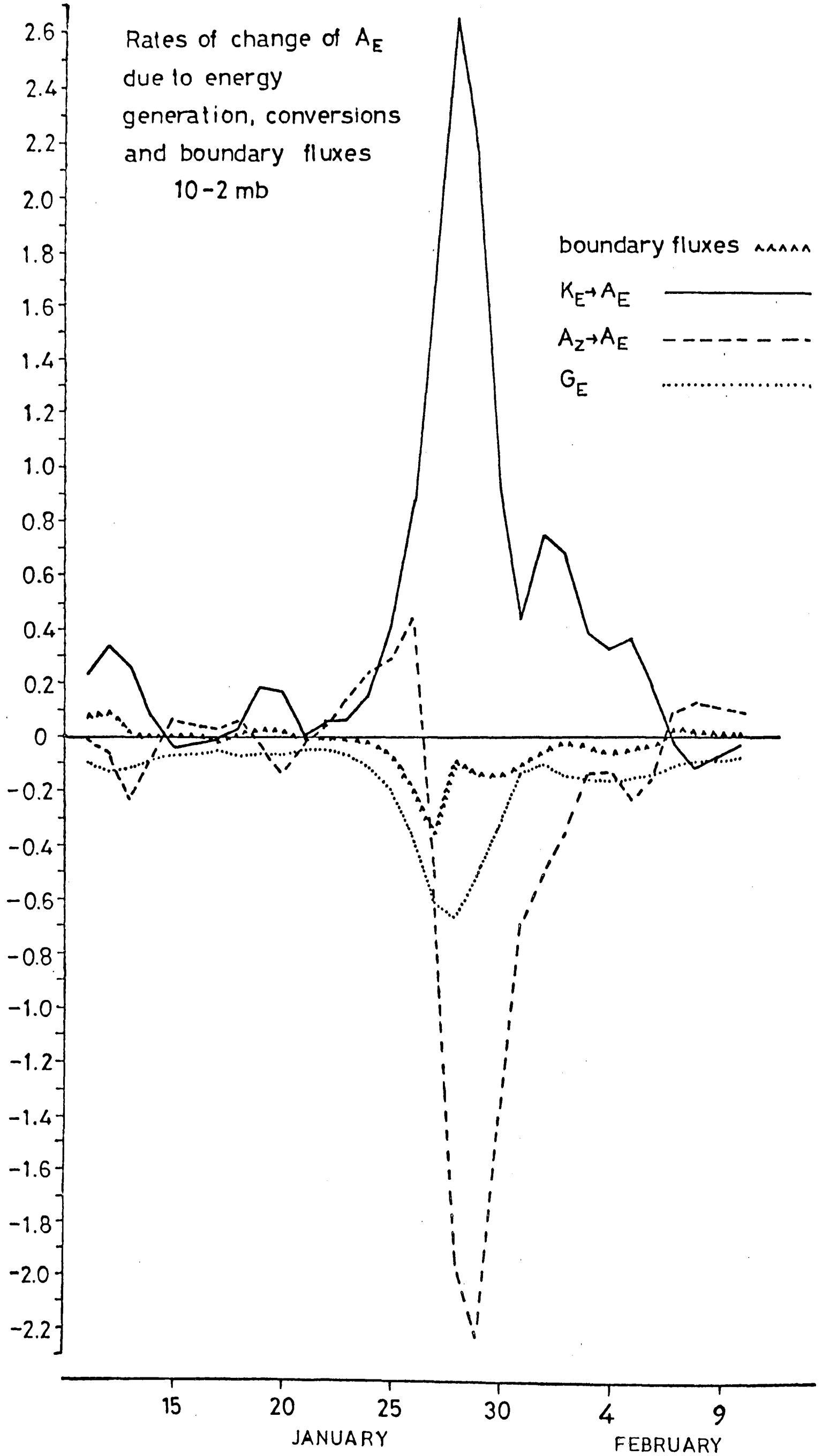


Figure 8.4. Latitude - time section of the vertical eddy energy flux [$-\omega'\Phi'$] at 50 mb for the period 11.1.73 to 10.2.73 ($J\,m^{-2}\,s^{-1} \times 10^{-1}$)


Rates of change of A_E
due to energy
generation, conversions
and boundary fluxes
10-2 mb
boundary fluxes
$K_E \rightarrow A_E$
$A_Z \rightarrow A_E$
G_E
Rate of change of energy ($Jm^{-2}s^{-1}mb^{-1}\times10^{-2}$)
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
-1.4
-1.6
-1.8
-2.0
-2.2
15
20
25
30
4
9
JANUARY
FEBRUARY

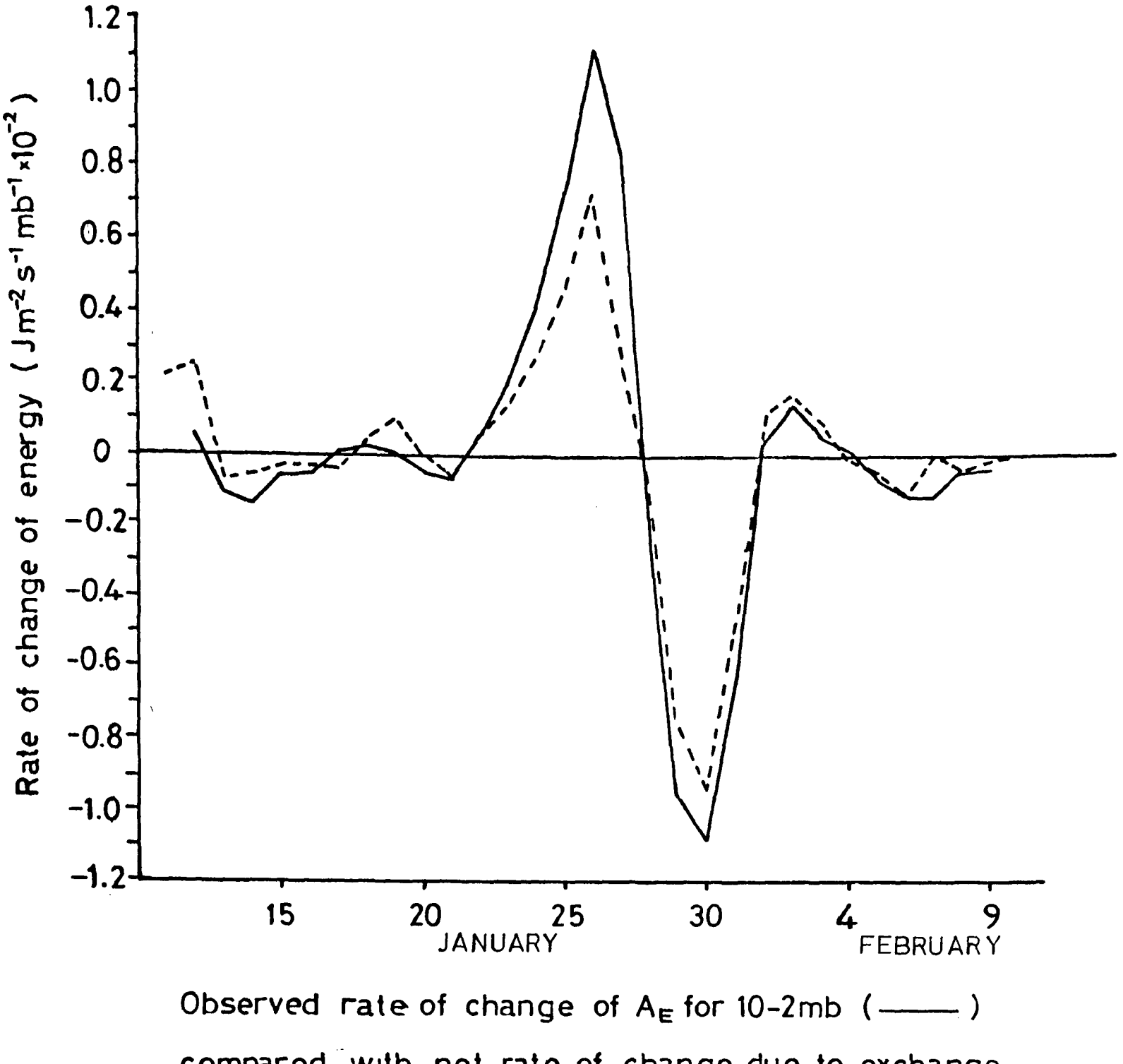


Figure 7.4b. The eddy available potential energy budget for the region 10 - 2 mb, 14 - 82°N between 11.1.73 - 10.2.73

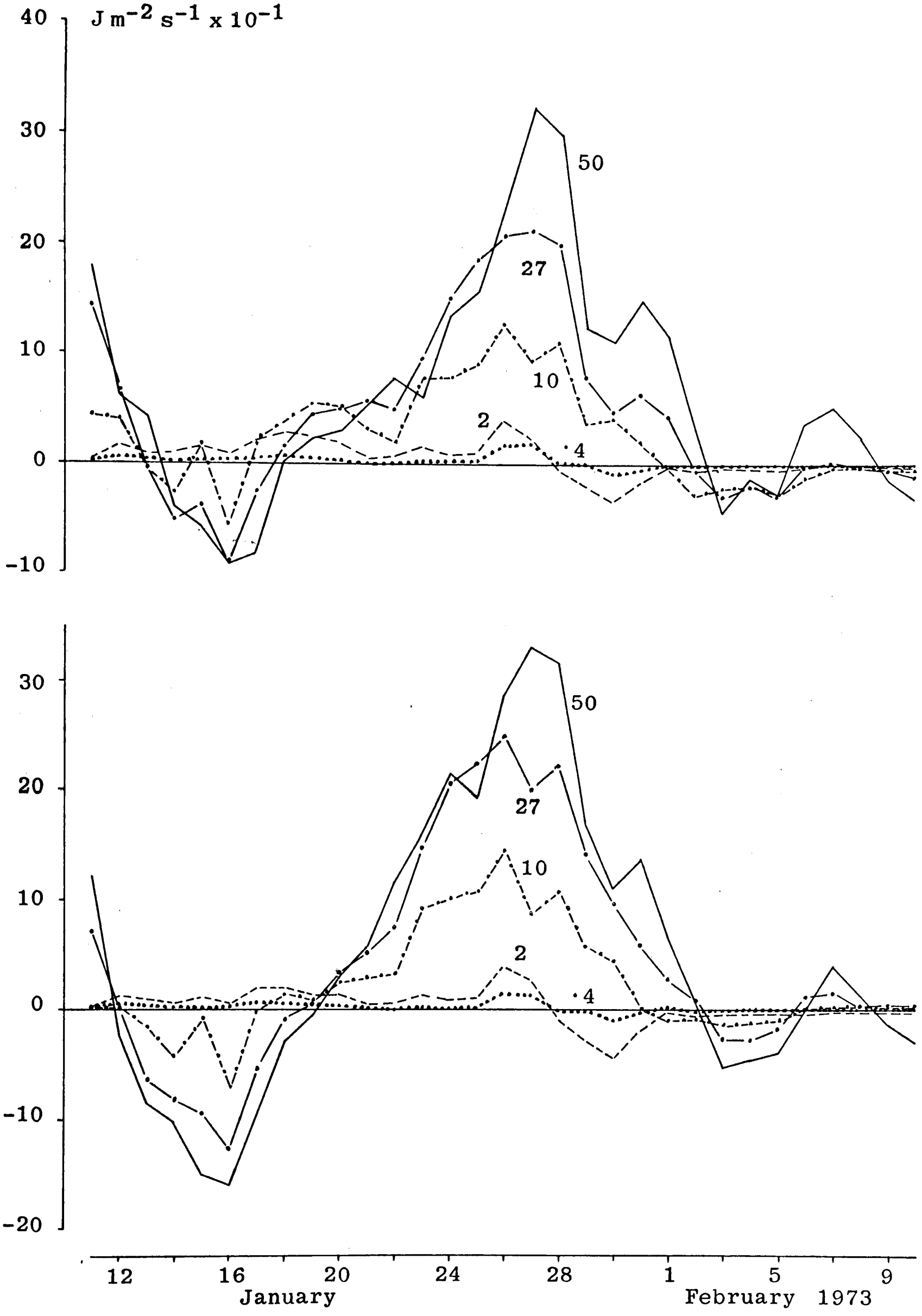
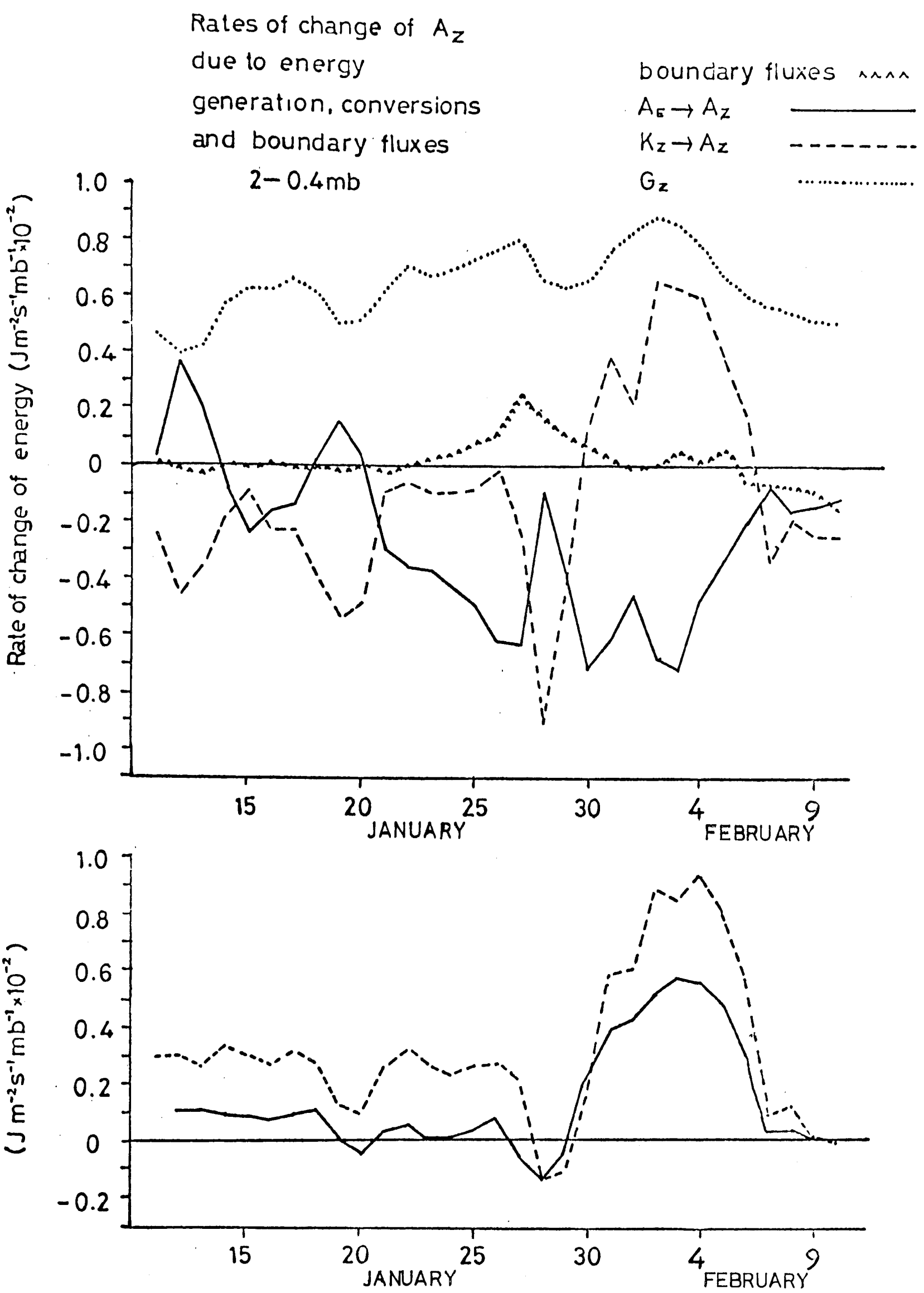


Figure 8.6. As Figure 8.5 but for 68°N



Observed rate of change of A_z (——) compared with net rate of change due to above processes (- - - - - - -)

Figure 7.5a. The zonal available potential energy budget for the region 2 - 0.4 mb, 14 - 82°N between 11.1.73 - 10.2.73

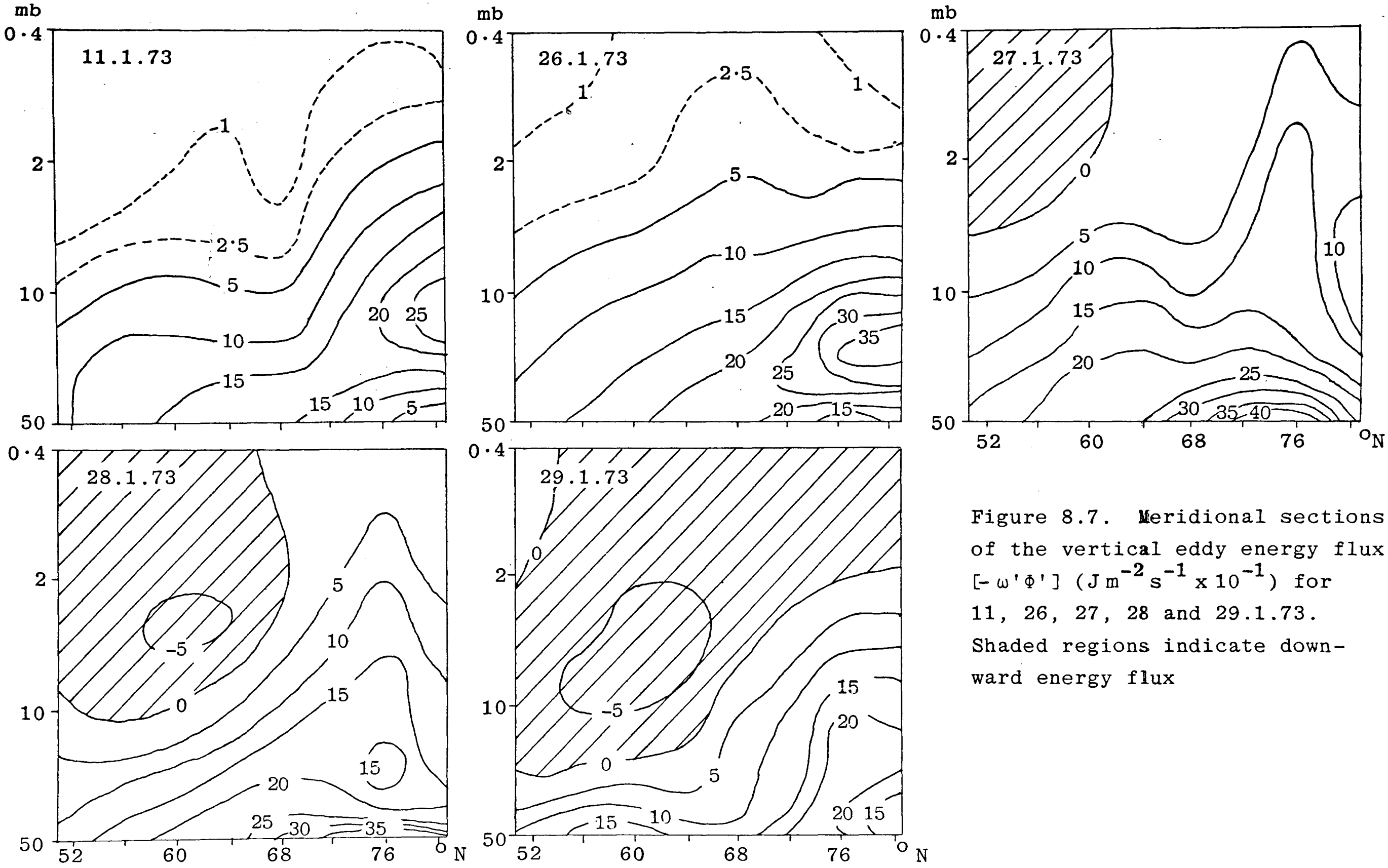
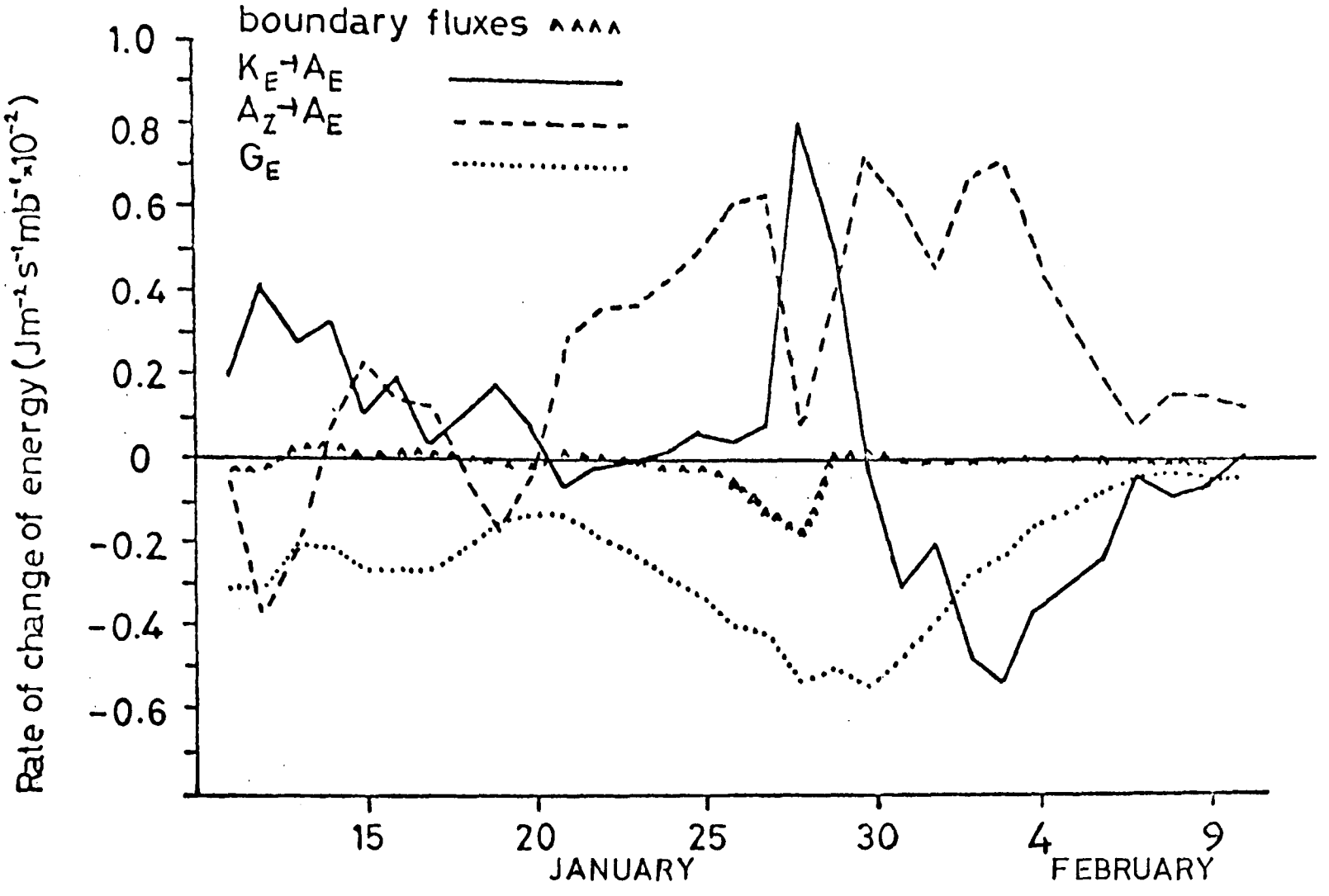
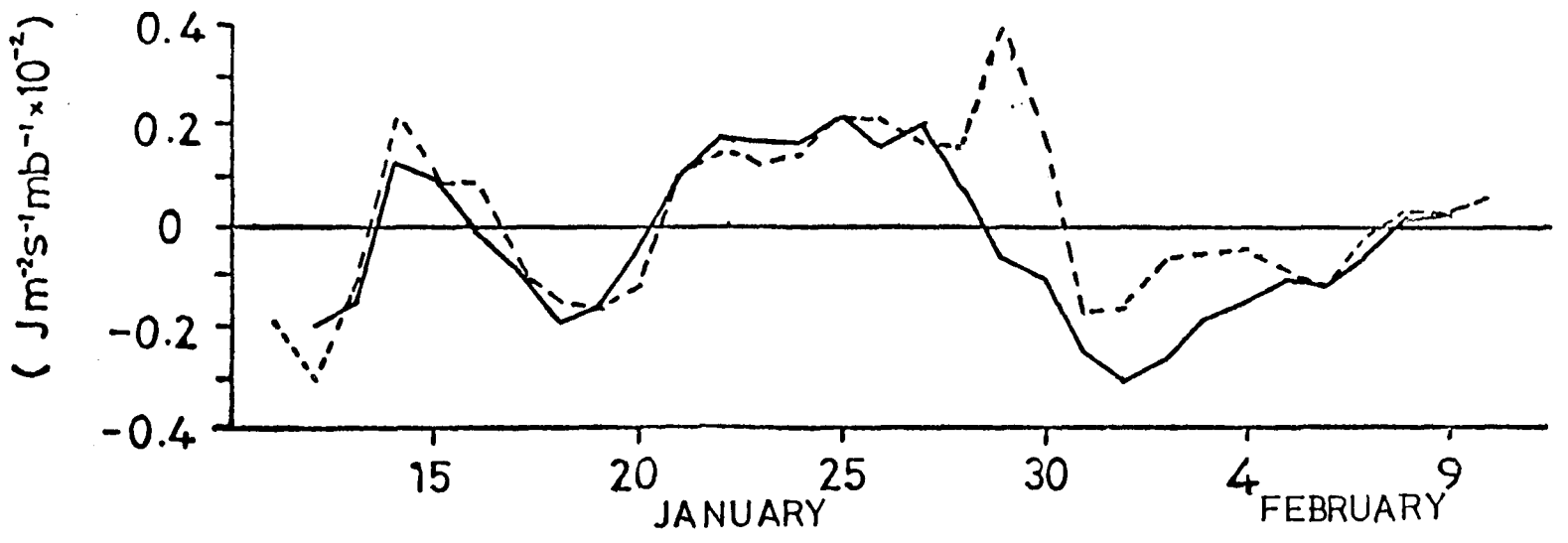


Figure 8.7. Meridional sections of the vertical eddy energy flux $[-\omega'\Phi']$ ($J\,m^{-2}\,s^{-1} \times 10^{-1}$) for 11, 26, 27, 28 and 29.1.73. Shaded regions indicate downward energy flux

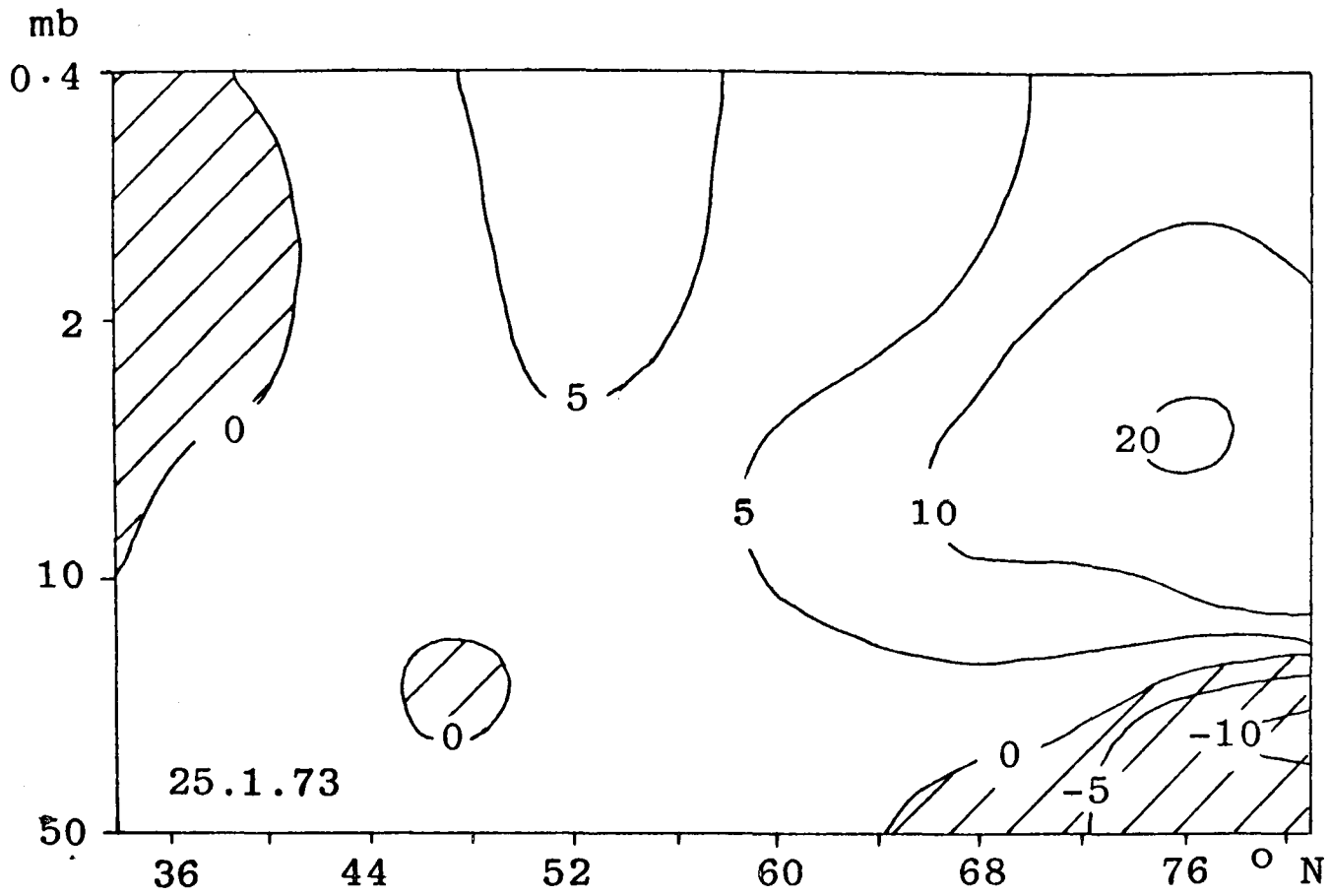


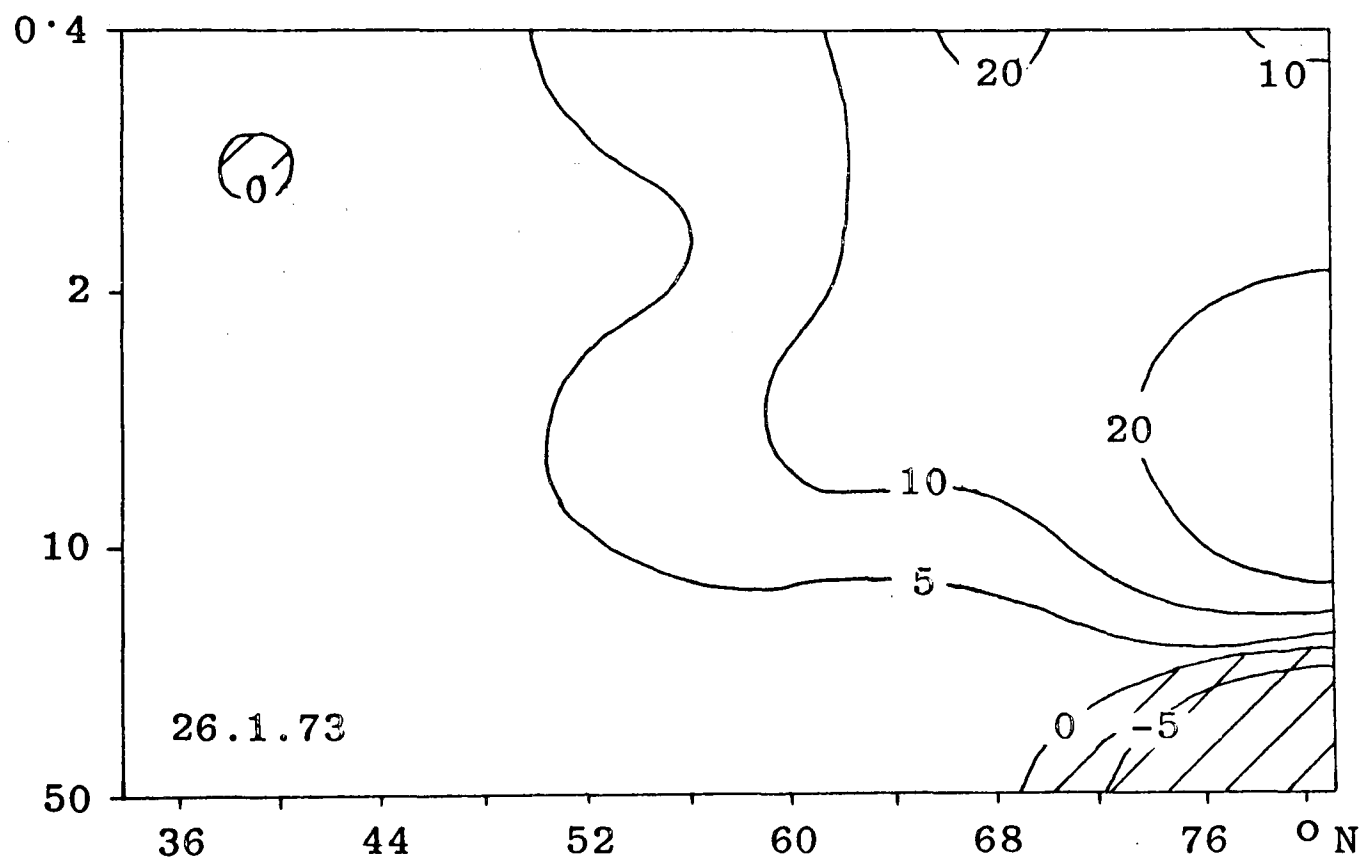
Rates of change of A_E due to energy generation, conversions and boundary fluxes 2 - 0.4 mb

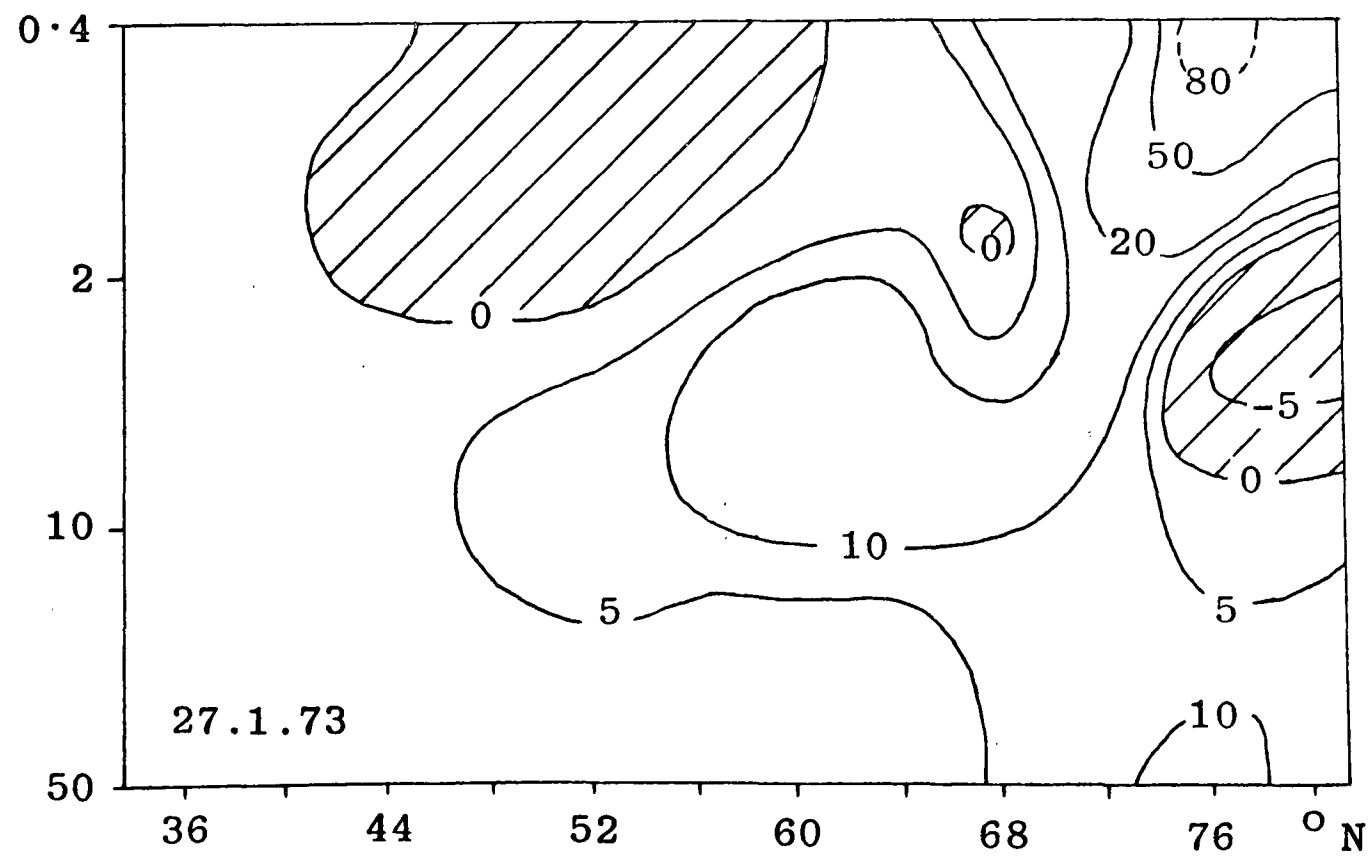


Observed rate of change of A_E (——) compared with net rate of change due to above processes (- - - - - -)

Figure 7.5b. The eddy available potential energy budget for the region 2 - 0.4 mb, 14 - 82°N between 11.1.73 - 10.2.73


mb
0·4
2
10
50
0
5
20
10
0
-10
-5
25.1.73
36
44
52
60
68
76
o N


0·4
2
10
50
0
20
10
5
0
-5
26.1.73
36
44
52
60
68
76
o N


0·4
2
10
50
80
50
20
0
0
-5
0
10
5
5
10
27.1.73
36
44
52
60
68
76
o N

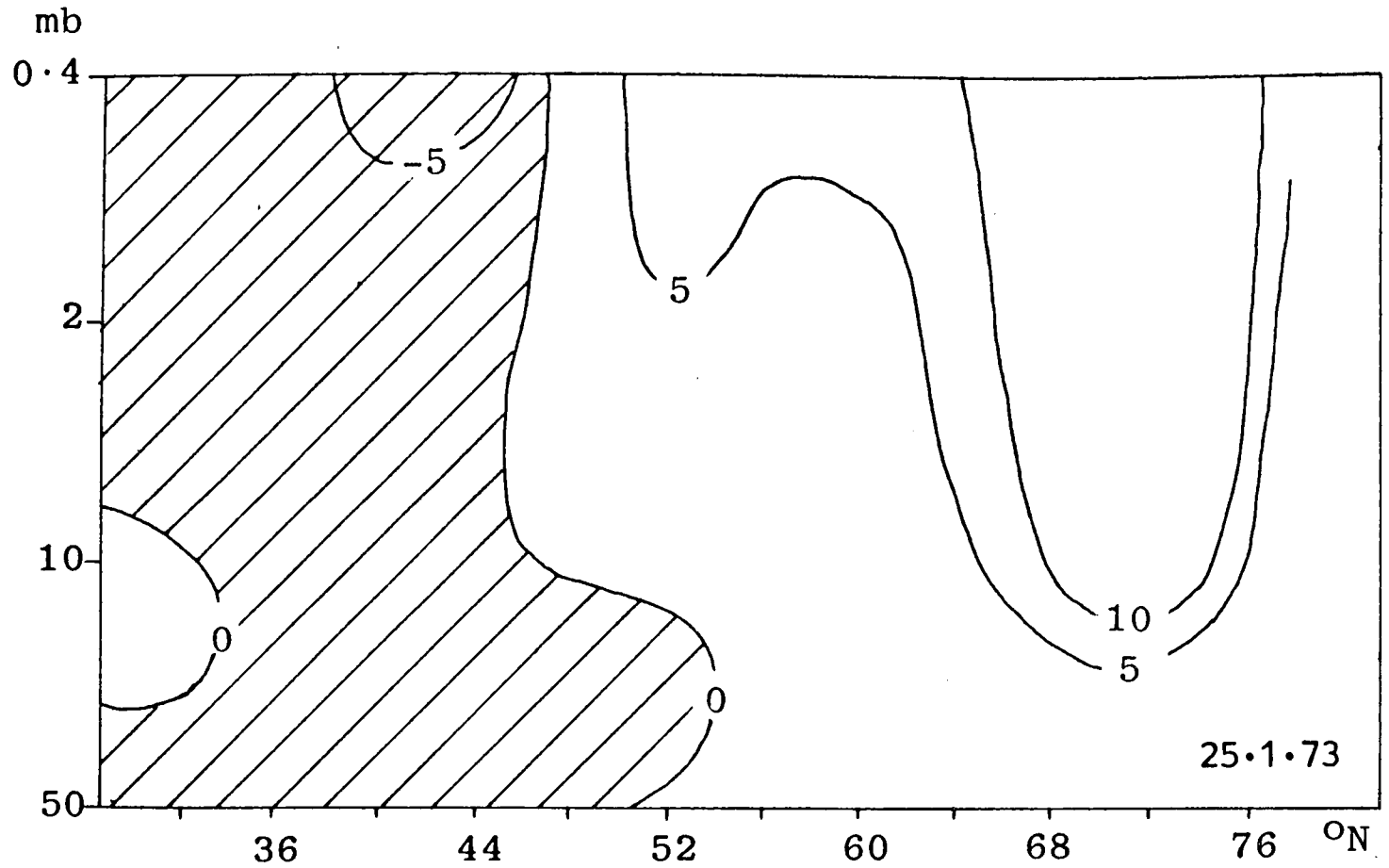
individual processes involved. Surprisingly, good agreement was found during the peak of the warming (28 - 30.1.73) when one would have expected absolute errors in vertical motion calculations to be at their greatest. It is possible that some of this error arises from the approximation used for the generation term (equation 4.9') which is a more important process in the top layer than in the two beneath. As noted in Chapter 4 Newell et al (1970) found large differences between the exact and approximate expressions for this process in the troposphere.

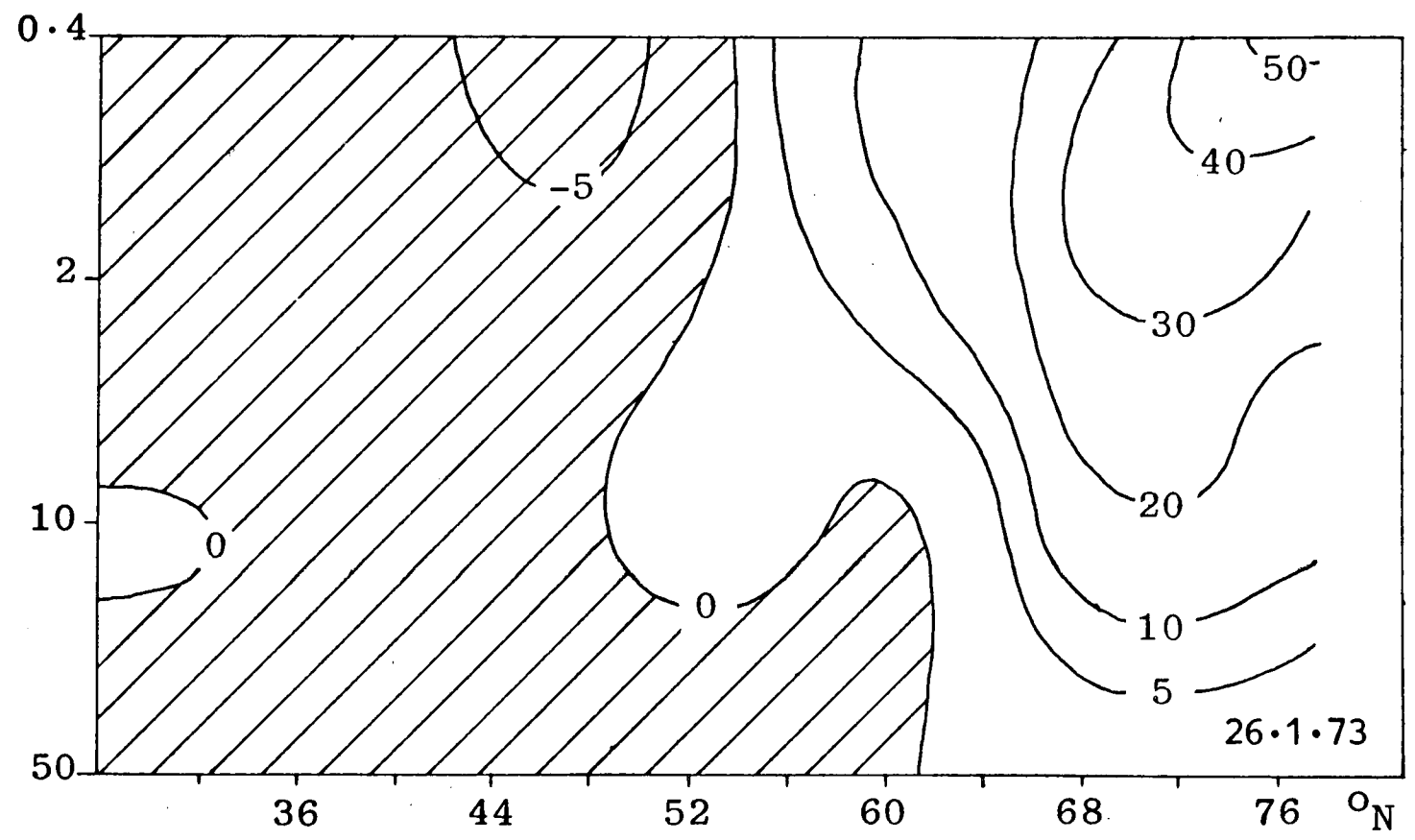
Similar diagrams of the A_E budget are given in Figures 7.3b, 7.4b and 7.5b. Below 2 mb agreement between calculated and observed rates of change is again satisfactory. Above 2 mb there is much better agreement prior to warming than for A_Z but large error is evident after 28.1.73.

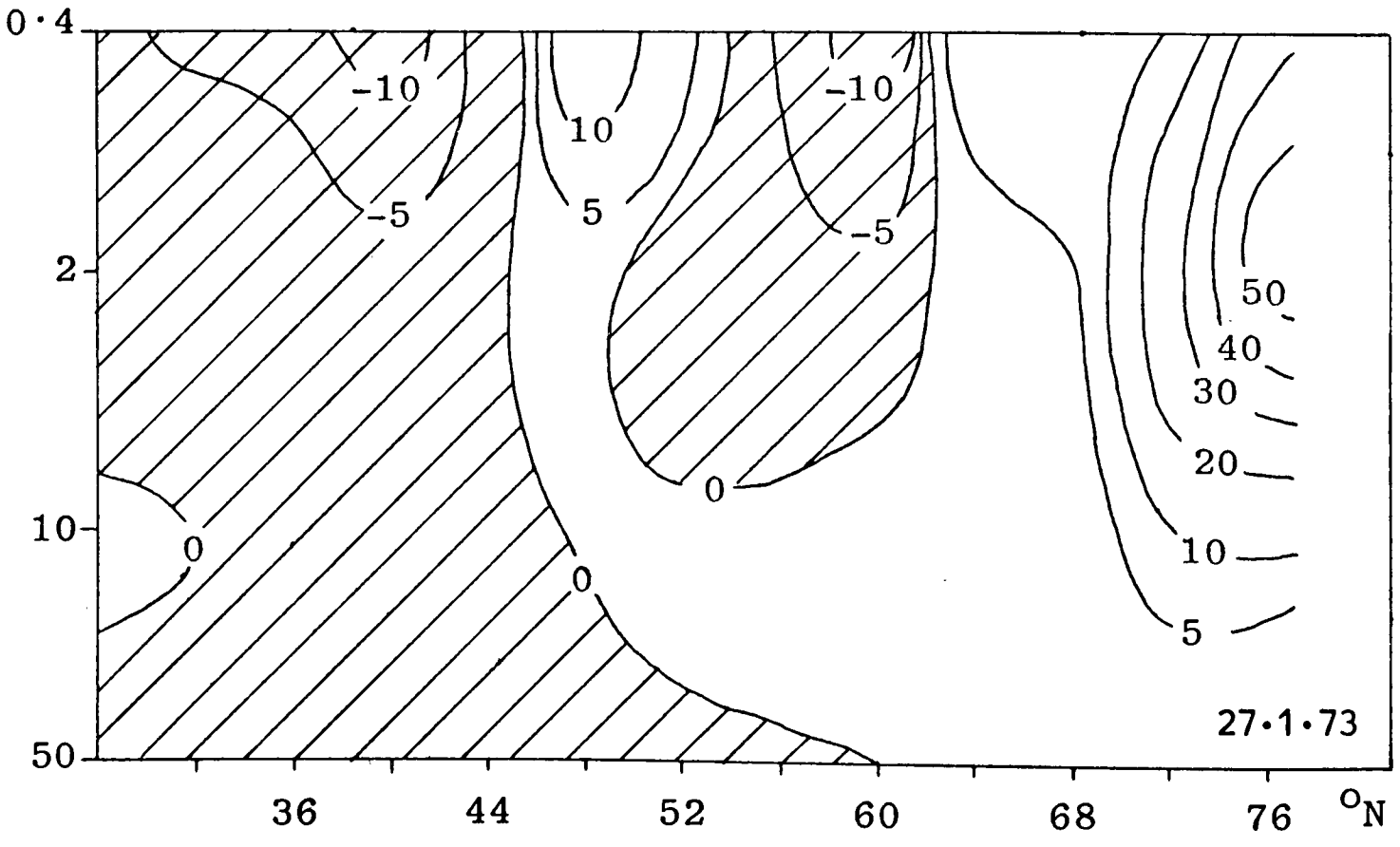
The zonal available potential energy budget

The most obvious features of the A_Z budgets is that for most of the period the net changes of A_Z in each layer were much smaller than the competing conversion and generation processes. Throughout there was a tendency for the $K_Z \rightarrow A_Z$ conversion to be accompanied by $A_Z \rightarrow A_E$. Above 2 mb radiative generation masked this feature to some extent although the two conversions tended to undergo changes of opposite sign. This is a reflection of the process whereby poleward heat transport reduces the equator to pole temperature gradient [1] ($A_Z \rightarrow A_E$) and a mean meridional circulation is induced to offset this process ($K_Z \rightarrow A_Z$). The establishment of the pole to equator temperature gradient below 2 mb between 26 and 30.1.73 is evident in the increase in A_Z which occurred. Above 2 mb A_Z increased rapidly after 30.1.73

[1] Unless stated otherwise a temperature gradient from A to B implies that A is warmer than B.


mb
0·4
2
10
50
-5
5
0
0
10
5
25·1·73
36
44
52
60
68
76
°N


0·4
2
10
50
-5
0
0
50
40
30
20
10
5
26·1·73
36
44
52
60
68
76
°N

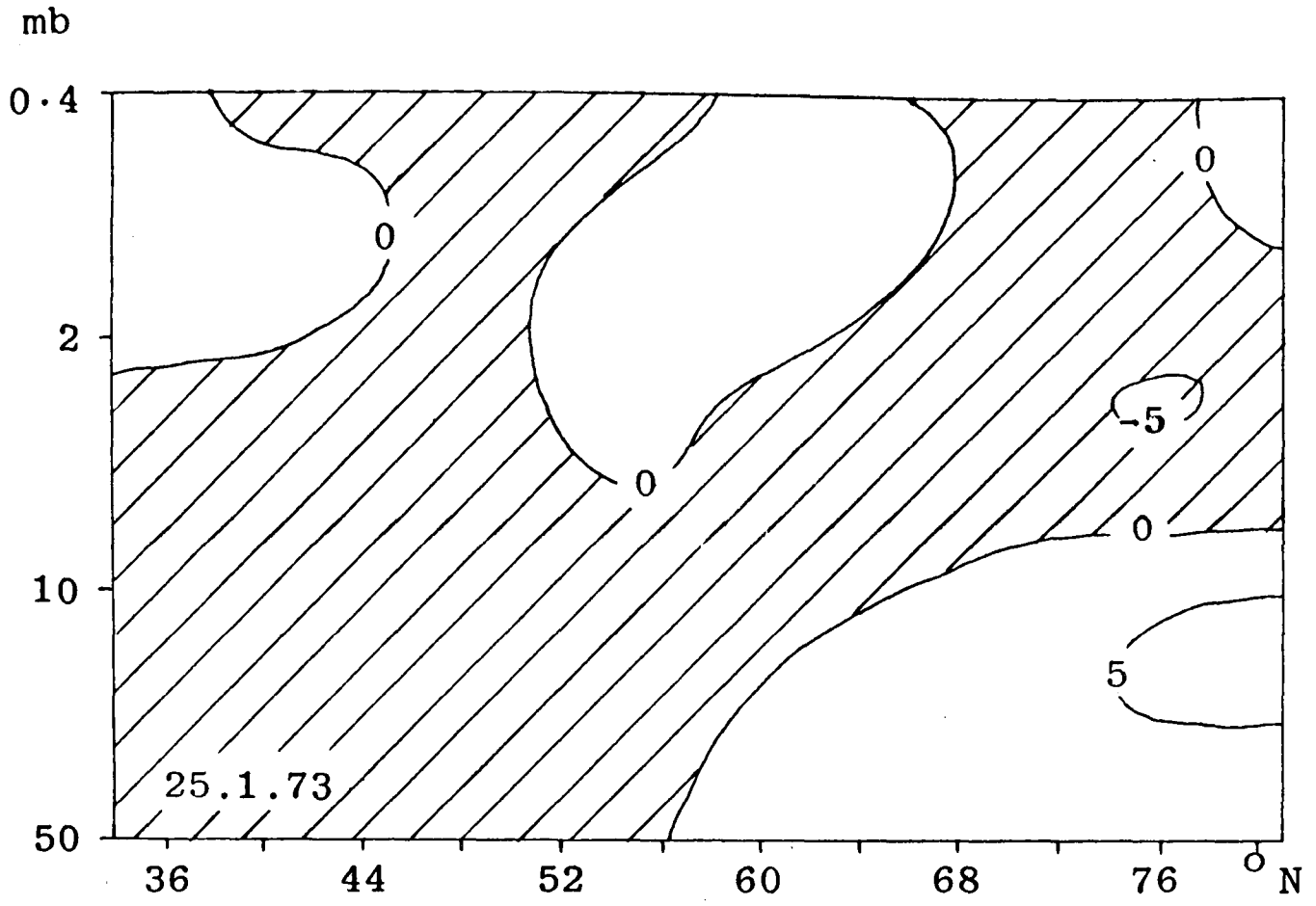

0·4
2
10
50
-10
-5
10
5
-10
-5
0
0
0
50
40
30
20
10
5
27·1·73
36
44
52
60
68
76
°N

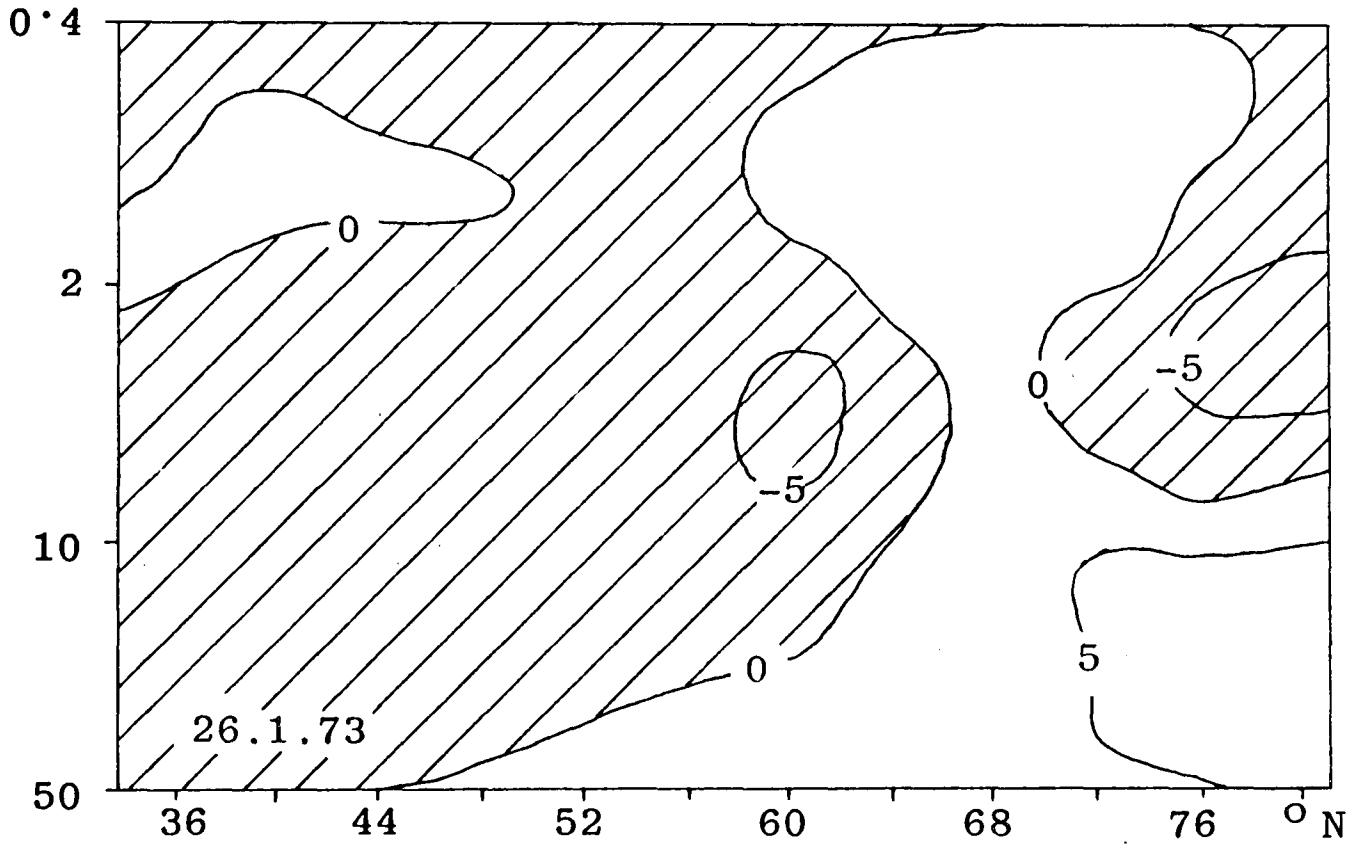
as the pole cooled and a strong equator to pole temperature gradient developed (see Figure 6.25). This latter rate of increase in A_Z was the largest observed despite the fact that the conversion processes between 10 and 2 mb were two to three times larger than those in the top layer.

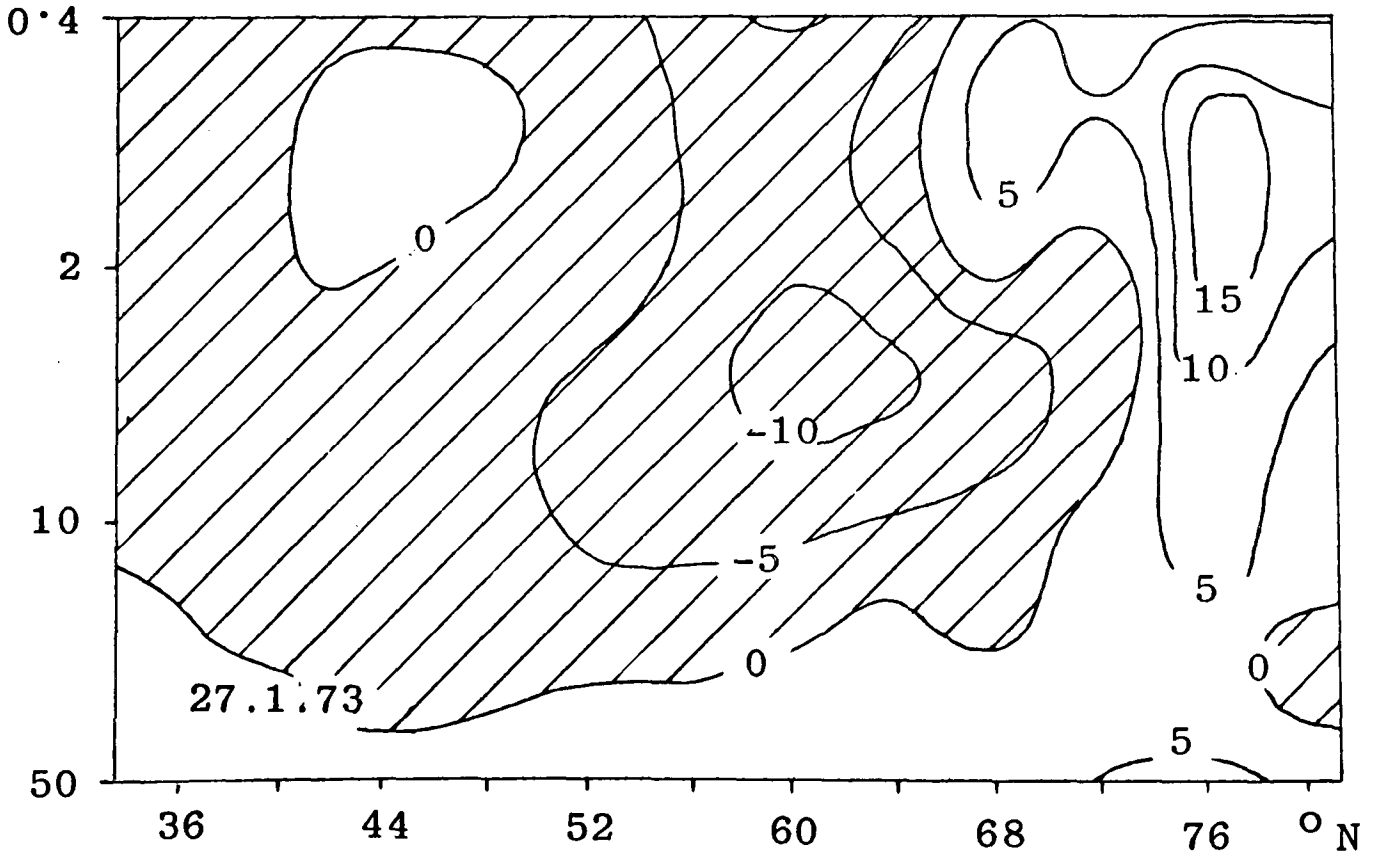
The reversal of the conversion processes below 2 mb at the peak of the warming occurred as a result of the poleward heat flux and the mean meridional circulation remaining in the same sense while the meridional temperature gradient reversed. Above 2 mb polar warming was short-lived and although there was a tendency for the A_E-A_Z conversion to reverse between 27 and 28.1.73 polar cooling quickly became established so that A_E remained a sink for A_Z and K_Z became a strong source via the mean meridional circulation. Radiative generation of A_Z was the major process throughout the warming in the top layer.

The eddy available potential energy budget

A similar tendency for the internal conversions to balance exists in the A_E budget where $A_E \rightarrow K_E$ conversion is usually observed to compensate the changes in A_E caused by conversion from A_Z to A_E.


mb
0·4
2
10
50
0
0
0
-5
0
5
25.1.73
36
44
52
60
68
76
o N


0·4
2
10
50
0
0
-5
0
-5
5
26.1.73
36
44
52
60
68
76
o N

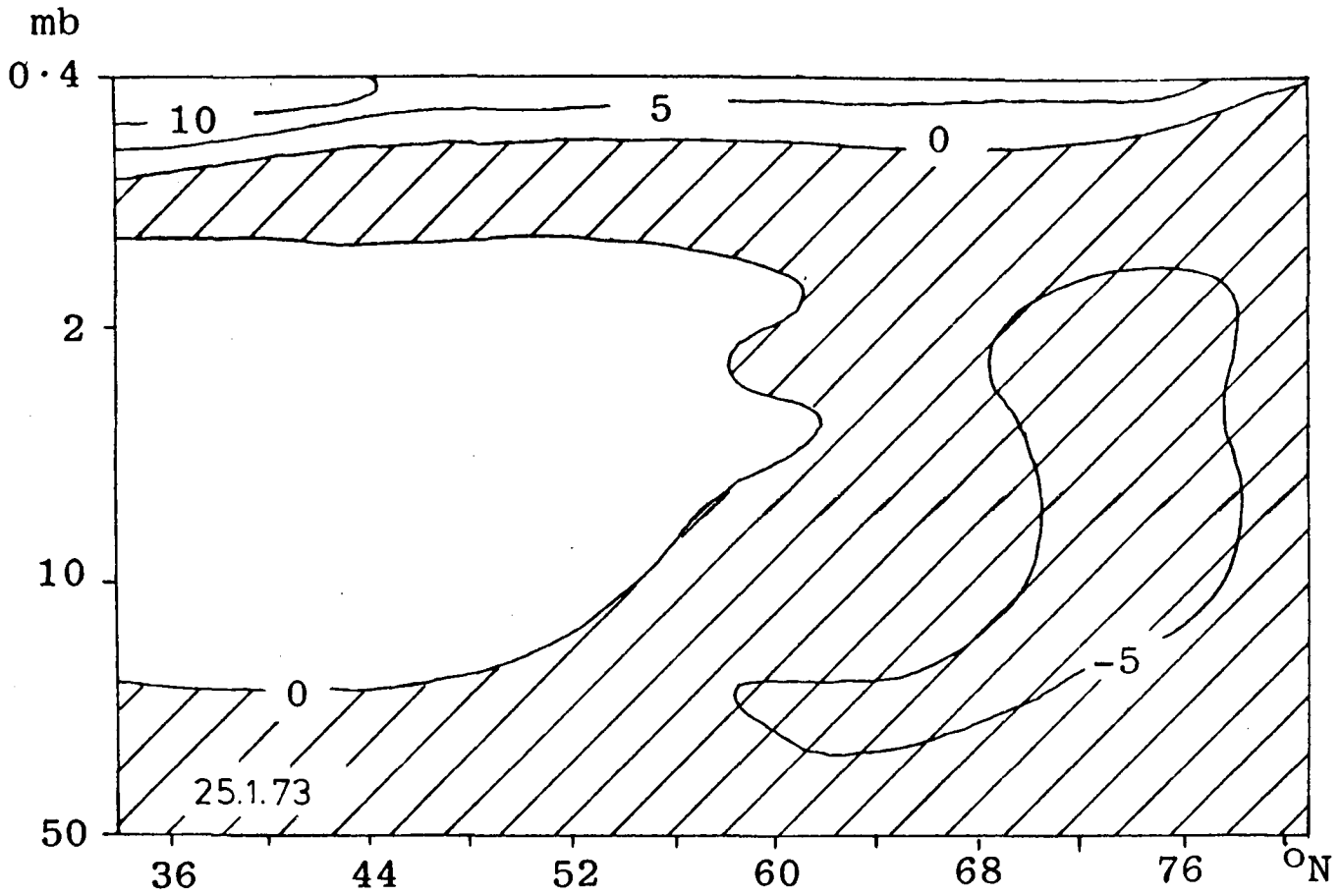

0·4
2
10
50
0
5
15
10
-10
-5
5
0
0
5
27.1.73
36
44
52
60
68
76
o N

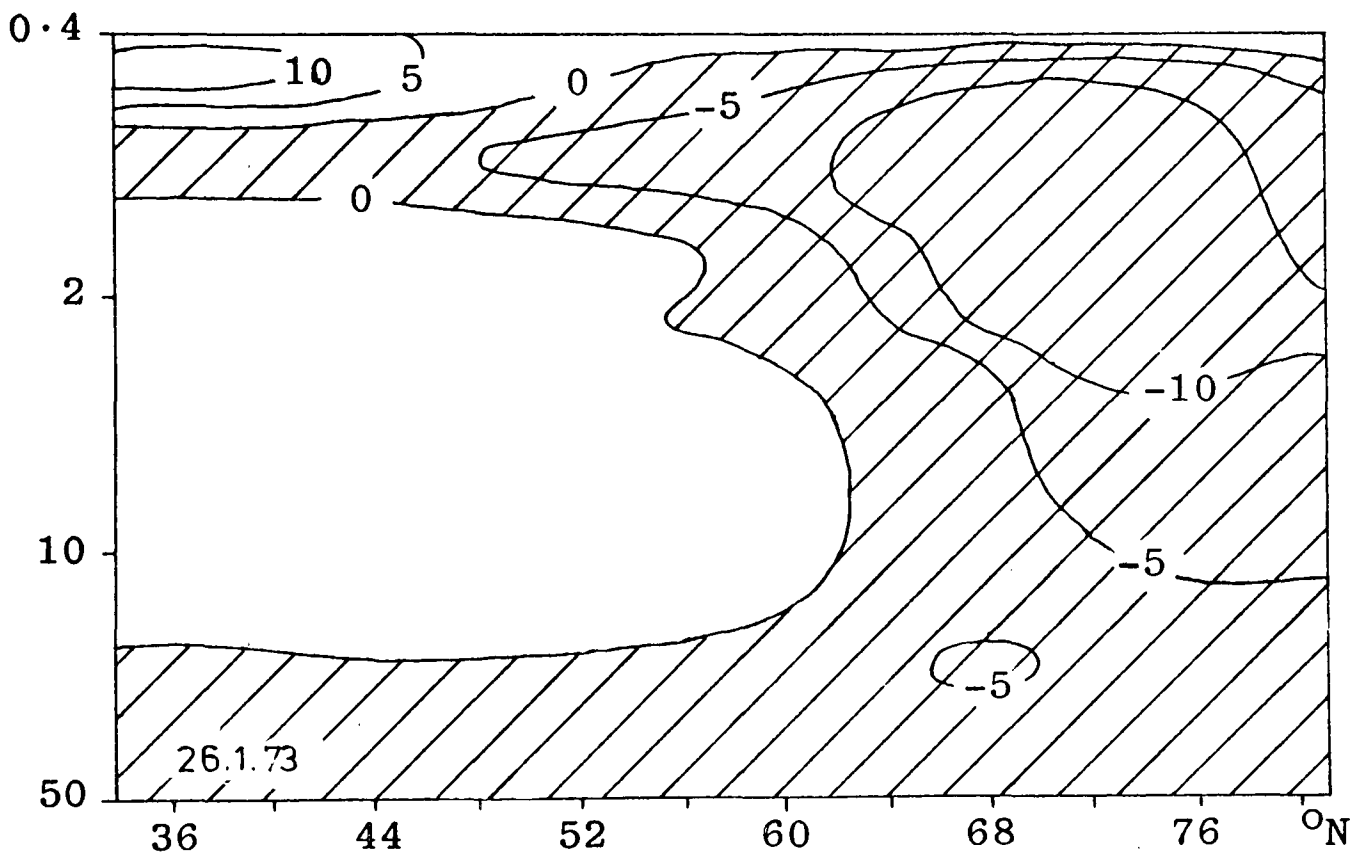
Hartmann (1976) described this process by deriving an expression for the eddy conversion in steady, stationary, long waves analagous to that for vertical energy flux given by Eliassen and Palm (1961):

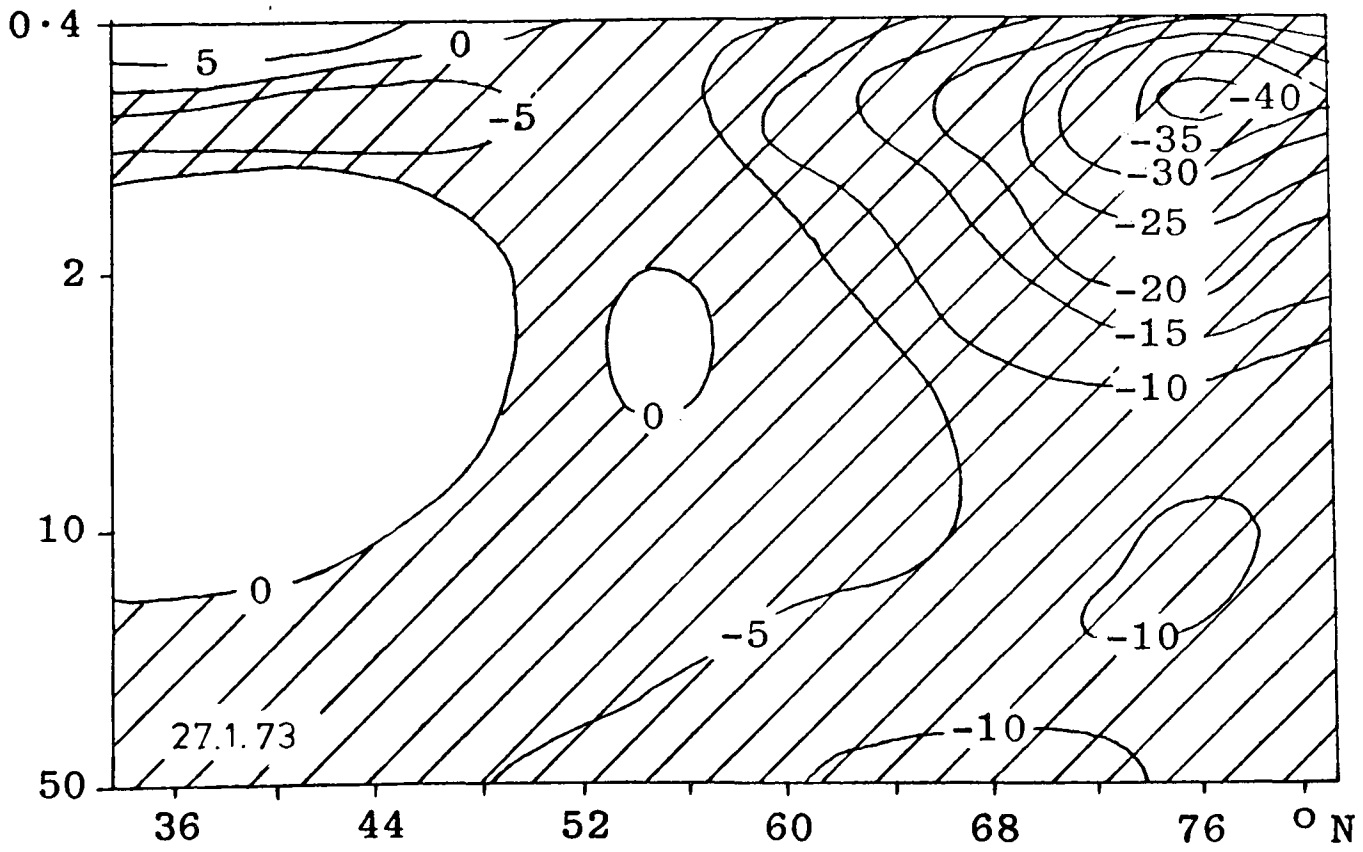
$$[\omega'\alpha'] = f \frac{\partial}{\partial p}[u]\,[v'T'] \,/\, N^2[T] \qquad 7.1$$

where N^2 is the Brunt-Vaisala frequency. The vertical shear of the zonal mean zonal wind, [u], is related to the zonal mean meridional temperature gradient by the thermal wind equation (equation 1.5). If the wave is assumed to be a closed system, positive (northward) eddy heat transport accompanied by increasing westerly flow with height (produced by a meridional temperature gradient from equator to pole) gives an $A_Z \rightarrow A_E$ conversion and negative $[\omega'\alpha']$ - i.e. an $A_E \rightarrow K_E$ conversion.

The A_E budgets in Figures 7.3b, 7.4b and 7.5b confirm that in the 50 - 10 mb region the eddy conversion was consistently in the sense to offset changes in A_E due to conversion to or from A_Z. The energy cycle was thus $K_Z \rightarrow A_Z \rightarrow A_E \rightarrow K_E$ prior to the peak of the warming and $K_E \rightarrow A_E \rightarrow A_Z \rightarrow K_Z$ in the declining phase. Above 2 mb also the $K_E \rightarrow A_E$ and $A_Z \rightarrow A_E$ conversions tended to be in opposition although in the pre-warming period both processes were sources of A_E. Radiative loss of A_E was a significant feature during the warming. In the declining phase the $K_E \rightarrow A_E$ conversion reversed giving a direct baroclinic cycle above 2 mb. Between 10 and 2 mb not only were the conversion processes two to three times greater than those above and below but their net imbalance was large. With both A_Z and K_E supplying A_E and with little radiative destruction of A_E in the period 21 - 26.1.73, a rapid increase in A_E occurred, a result to be expected from the large


mb
0·4
10
5
0
2
10
0
-5
25.1.73
50
36
44
52
60
68
76
°N


0·4
10
5
0
-5
0
2
-10
-5
-5
10
26.1.73
50
36
44
52
60
68
76
°N


0·4
5
0
-5
-40
-35
-30
-25
-20
-15
-10
2
0
10
0
-5
-10
-10
27.1.73
50
36
44
52
60
68
76
° N

rates of change of temperature shown in Figure 6.4b. It is clear that the anomaly of simultaneous increase of the $A_Z \rightarrow A_E$ and $K_E \rightarrow A_E$ conversions was a significant factor in the initiation of the warming. Since the A_Z budget was in near balance throughout the period it would appear that the eddy conversion was directed abnormally at this time. In Figure 7.6 time averaged height - latitude sections of the vertical shear of the zonal wind are compared with the distribution of $-[\omega'\alpha']$ ($A_E \rightarrow K_E$ conversion) in accordance with equation 7.1. Since $[v'T']$ was positive for the most part we should expect positive values of the wind shear to be associated with conversion from A_E to K_E. Although the correlation is far from exact there is broad agreement with this prediction. Note however that between 10 and 2 mb in the period 20 - 26.1.73 there is a significantly larger region contributing to the $K_E \rightarrow A_E$ conversion than would be expected from the zonal wind shear distribution. It therefore seems likely that this process was forced by external means. In the post-warming period very large $K_E \rightarrow A_E$ conversion throughout the 50 - 2 mb region was associated with the negative shear between westerlies at low levels and the strong easterly circulation which had been established in the upper stratosphere. As noted earlier eddy conversion in this sense was observed for a short period above 10 mb during the 1970-1 and 1974-5 warmings by Klinker (1976) with maximum values of -30×10^{-4} $JKg^{-1}s^{-1}$, but its significance was not clear since a detailed budget for this region was not calculated.

(ii) Kinetic Energy

It was shown in Chapter 5 that by using thermodynamically derived zonal mean vertical and meridional velocities a balanced zonal kinetic energy budget would not be obtained. It will be

than required for a balanced K_Z budget. As the region of barotropic conversion spread to the pole the effect of the mean meridional circulation reversed to give increase in K_Z. The middle latitude values of f[u][v] are more difficult to assess. It can be seen in Figure 6.10 that a consistently strong poleward mean meridional velocity was calculated just below 0.4 mb in low and middle latitudes for which poor specification of radiative heating may have been responsible (see Appendix 3). It is clearly this feature which has caused the rapid vertical variation of f[u][v] south of about 45^{o}N. It is likely that the effects of the eddy momentum flux divergence would be opposed by the mean meridional circulation as in high latitudes in which case positive values of f[u][v] would have occurred.

The important results of this analysis are therefore that (1) in high latitudes K_E increased by vertical energy flux convergence and K_Z declined due to the effects of the mean meridional circulation, although the $K_E \rightarrow K_Z$ conversion was opposed to these processes, while (2) in middle latitudes K_E increased and K_Z decreased by barotropic conversion. Barotropic loss of K_Z in middle latitudes occurring before the net loss of K_Z in high latitudes explains why zonal flow deceleration was first observed in middle latitudes.

Between 10 and 2 mb K_E maxima were observed at the same times and latitudes as in the 2 - 0.4 mb layer but were less marked (Figure 8.12). Throughout the period 25 - 29.1.73 in middle latitudes K_E increased via both the barotropic conversion and vertical energy flux convergence and was lost to the eddy potential energy field. In high latitudes vertical energy flux convergence supplied K_E throughout except on 27.1.73. It is significant that the baroclinic conversion reversed to give an

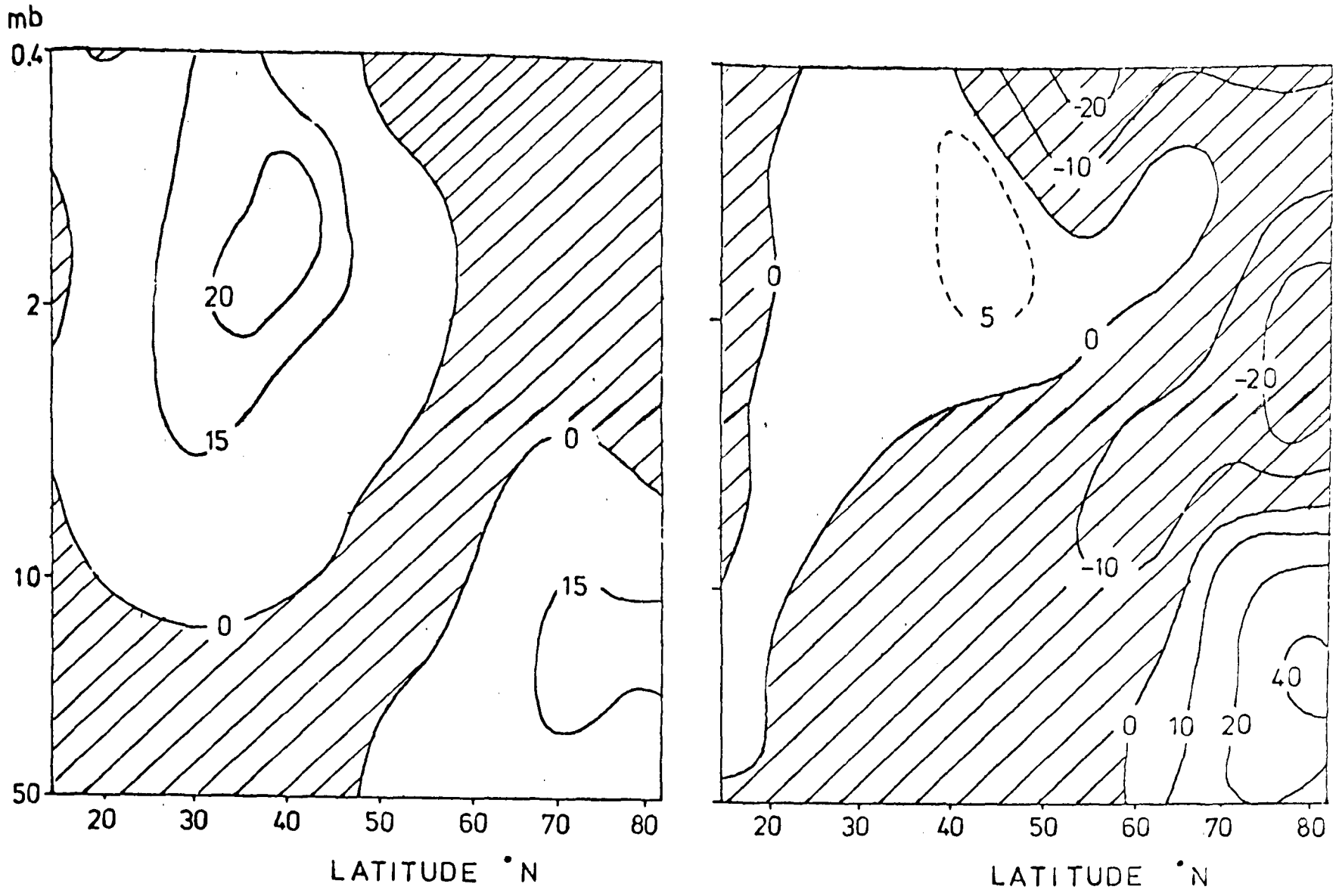


Figure 7.6. Meridional sections of the vertical shear of the zonal mean zonal wind, $\partial[u]/\partial(-\ln p)$, in ms^{-1}(scale height)$^{-1}$ (left), averaged over the periods 20-26.1.73 (upper) and 28.1 - 3.2.73 (lower), and the corresponding sections of $-[\omega'\alpha']$ ($A_E \rightarrow K_E$ conversion) in units of 10^{-4} J Kg^{-1} s^{-1} (right)

increase of K_E on this day throughout the high latitude stratosphere. This again suggests that baroclinic conversion enhanced the vertical energy propagation. The increase in K_E was sustained for a shorter period than at higher levels due to strengthening loss to potential energy on 29.1.73. Changes in K_Z were brought about by the same mechanisms as in the 2 - 0.4 mb layer.

Interaction between K_E and K_Z was small between 50 and 10 mb owing to weak eddy momentum fluxes. Changes in K_E were thus controlled by a balance between the baroclinic conversion and vertical energy flux divergence. The former was maintained throughout the period 25 - 29.1.73 in high latitudes, while in middle latitudes an opposite conversion was established as the negative vertical shear of the zonal flow became progressively stronger in association with easterlies aloft. Although easterlies descended more rapidly in middle latitudes the circulation reversal did not reach the 50 mb level presumably on account of the absence of barotropic conversion. From 30.1.73 onwards the mean meridional circulation decelerated the high latitude zonal flow giving easterly flow north of 72^{o}N from 2 - 5.2.73, significant polar warming not having occurred until after 28.1.73.

8.4. THE SUDDEN WARMING MECHANISM

In the light of the observations of this chapter it is possible to suggest a sequence of events by which the sudden warming and circulation breakdown were achieved:

(1) Prior to warming the structure of a wavenumber one planetary wave was, for reasons unknown, conducive to baroclinic generation of K_E in high latitudes below 10 mb which enhanced the vertical wave energy flux and led to convergence of energy flux between 10 and 2 mb. This energy was manifested in an increase in

seen that there were similar imbalances in the eddy kinetic energy budget, almost certainly due to error in both the $K_E - K_Z$ conversion and in the eddy pressure-work term. In discussing the processes involved a qualitative approach must therefore be adopted. The evolutions of the processes contributing to the K_Z and K_E budgets of the 50 - 10 mb, 10 - 2 mb and 2 - 0.4 mb layers are given in Figures 7.7 - 7.9. On account of the imbalances only results for the period 21.1.73 - 4.2.73 are presented.

Zonal kinetic energy

Three distinct periods of change in the zonal flow are reflected in the observed changes of K_Z in the two upper layers: the increase of westerly winds prior to warming, the decline of these westerlies and the strengthening of the reversed, easterly circulation. Above 2 mb the replacement of polar easterlies with mid-latitude westerlies as the major jet is evident as a fourth phase. Below 10 mb the steady fall of K_Z throughout the period corresponded to steadily declining westerly flow. The changes in K_Z were generally larger than changes in A_Z.

It appears that during the amplifying stage energy was lost via the zonal mean pressure-work term, $B_{\phi z}$, and gained by the $K_E \rightarrow K_Z$ conversion at all levels. Above 2 mb these processes changed sign around 28.1.73. Interaction between K_Z and A_Z did not significantly affect the K_Z budget in the top layer, although between 10 and 2 mb $A_Z \rightarrow K_Z$ conversion was a major source of K_Z in the declining stage on account of the association between the mean meridional circulation and the very large pole to mid-latitude temperature gradient. The action of the mean meridional circulation largely controlled the K_Z budget below 10 mb from 31.1.73 onwards, A_Z being a source of K_Z and $B_{\phi z}$ a sink.

(4) The circulation reversal prevented further propagation of energy to high levels, eddy momentum and heat fluxes declined and the action of the mean meridional circulation alone intensified the easterly circulation. Radiative cooling and the continued rising motion at the pole soon led to a re-establishment of the equator to pole temperature gradient and deceleration of the easterly flow in high levels. Convergence of the vertical energy flux below the zero wind line enhanced a $K_E \rightarrow A_E \rightarrow A_Z$ energy cycle thereby maintaining the high polar temperatures longer in the lower stratosphere. The reversed circulation, as a result, persisted longest between 10 and 2 mb (Figure 6.23), at the top of the layer of maximum polar warming.

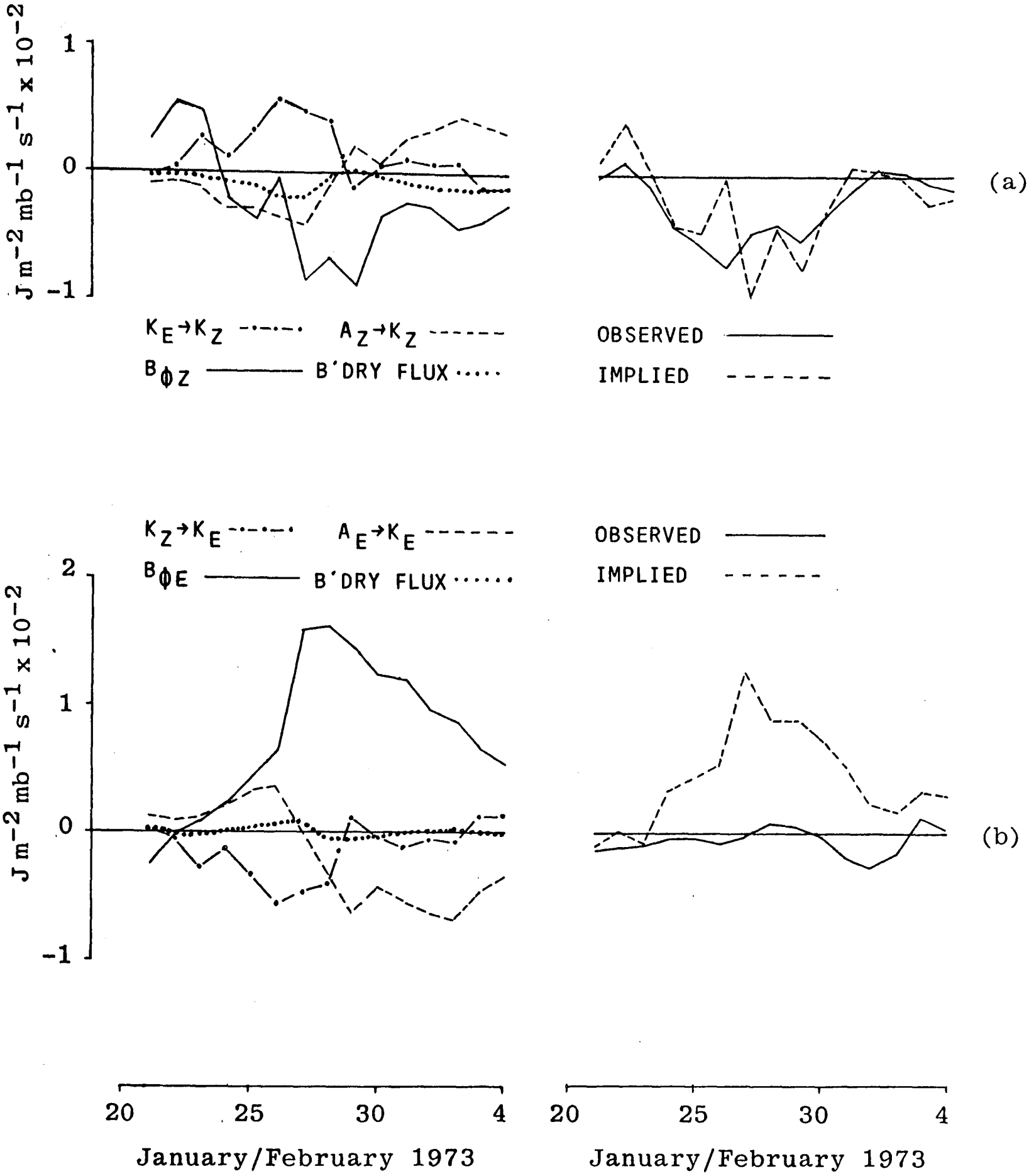


Figure 7.7. The zonal kinetic energy budget (a) and the eddy kinetic energy budget (b) for the region 50 - 10 mb, 14 - 82^{o}N for the period 21.1.73 - 4.2.73. The contributions to the budgets are shown on the left and the implied and observed net changes of the energy type are compared on the right

are discussed in this section.

The magnitude of the vertical energy flux

In Table 9.1 are compared the strengths of the vertical energy fluxes on 11 and 27.1.73 with those observed during other warmings. The values quoted are for a variety of time periods and latitude ranges but considerable similarity between the events is evident. On the basis of magnitude alone, the energy pulse on 11.1.73 was sufficient to initiate a sudden warming. Klinker (1976) concluded that no definite value of the vertical energy flux existed above which a circulation breakdown could confidently be expected and that the structure of the stratosphere and mesosphere determines the effect of the energy pulse. Arpe (1976) noted that vertical energy fluxes during the non-warming winter of 1966-7 were sometimes larger than those which have preceded warmings in other years. It has been shown in the previous chapter that the vertical structure of the wavenumber one disturbance was of crucial importance in initiating the sudden warming and one is led to suppose that the existing state of the stratosphere determines in part the structure of the waves superimposed on the mean flow. A factor which Geisler (1974) showed to be of great importance in determining whether or not a warming occurred in a numerical simulation was the duration of the forcing applied at the lower boundary. Unfortunately the data analysis for this study did not extend back beyond 11.1.73 and it is thus not possible to determine the period over which the vertical energy flux was enhanced at this time.

In view of these observations it seemed appropriate to examine the differences between the zonal flow distributions on 11.1.73 and at the onset of the warming. By way of introduction a brief outline of the theoretical ideas which have been put

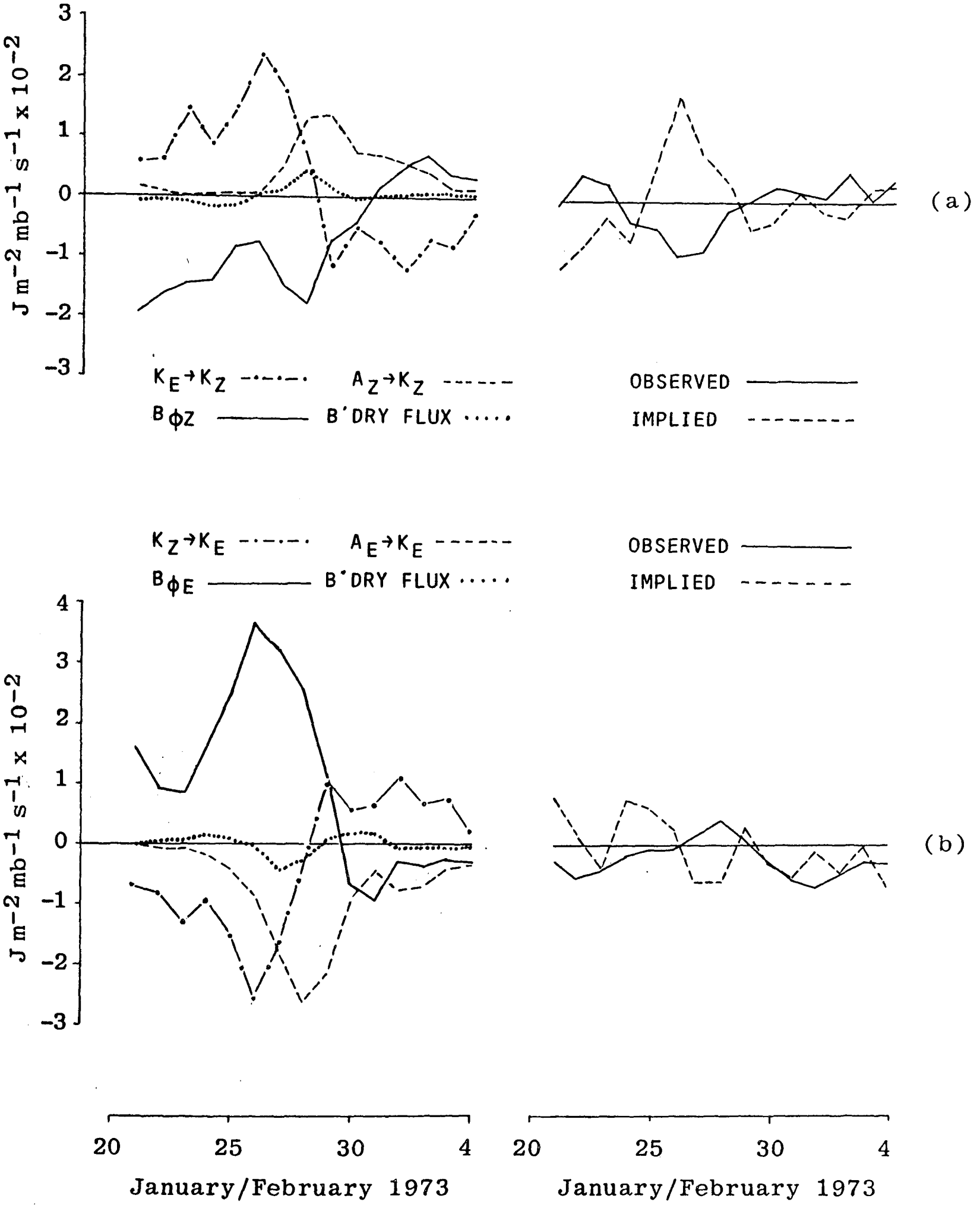


Figure 7.8. As Figure 7.7 but for the 10 - 2 mb layer

forward relating vertical wave propagation to the zonal flow is given.

The effect of the zonal flow distribution on wave propagation

Eliassen and Palm (1961) considered, analytically, energy propagation by steady,[1] quasi-stationary, adiabatic waves in a height and latitude-dependent zonal flow and showed that in westerly flow a poleward (equatorward) eddy heat flux is accompanied by an upward (downward) energy flux, the reverse energy flow directions occurring in easterly flow.

Using a mid-latitude β-plane model[2] with uniform zonal flow, Charney and Drazin (1961) concluded that planetary waves could only propagate vertically if their phase speed relative to the zonal current was westwards and smaller than a critical, wavenumber-dependent velocity, a theory which qualitatively explains the presence only of low wavenumber disturbances in the winter stratosphere and the absence of large scale planetary waves in the easterly flow of the summer stratosphere. Rather different results were obtained when Dickinson (1968) performed similar calculations to those of Charney and Drazin using spherical geometry instead of the β-plane approximation. Some wave modes were shown to be capable of propagating through any observed westerly flow. He further showed (Dickinson, 1969) that radiative damping attenuated wave propagation especially in weak zonal flows, thereby explaining the absence of significant vertical wave propagation at the equinoxes which had been predicted by Charney and Drazin.

---

[1] A steady wave is one in which amplitude is independent of time. Waves with time-dependent amplitude are termed transient.

[2] A β-plane model is one in which the latitudinal variation of the Coriolis parameter, $\beta = df/dy$, is assumed constant. A mid-latitude β-plane model employs the value of β at 45^{o}.

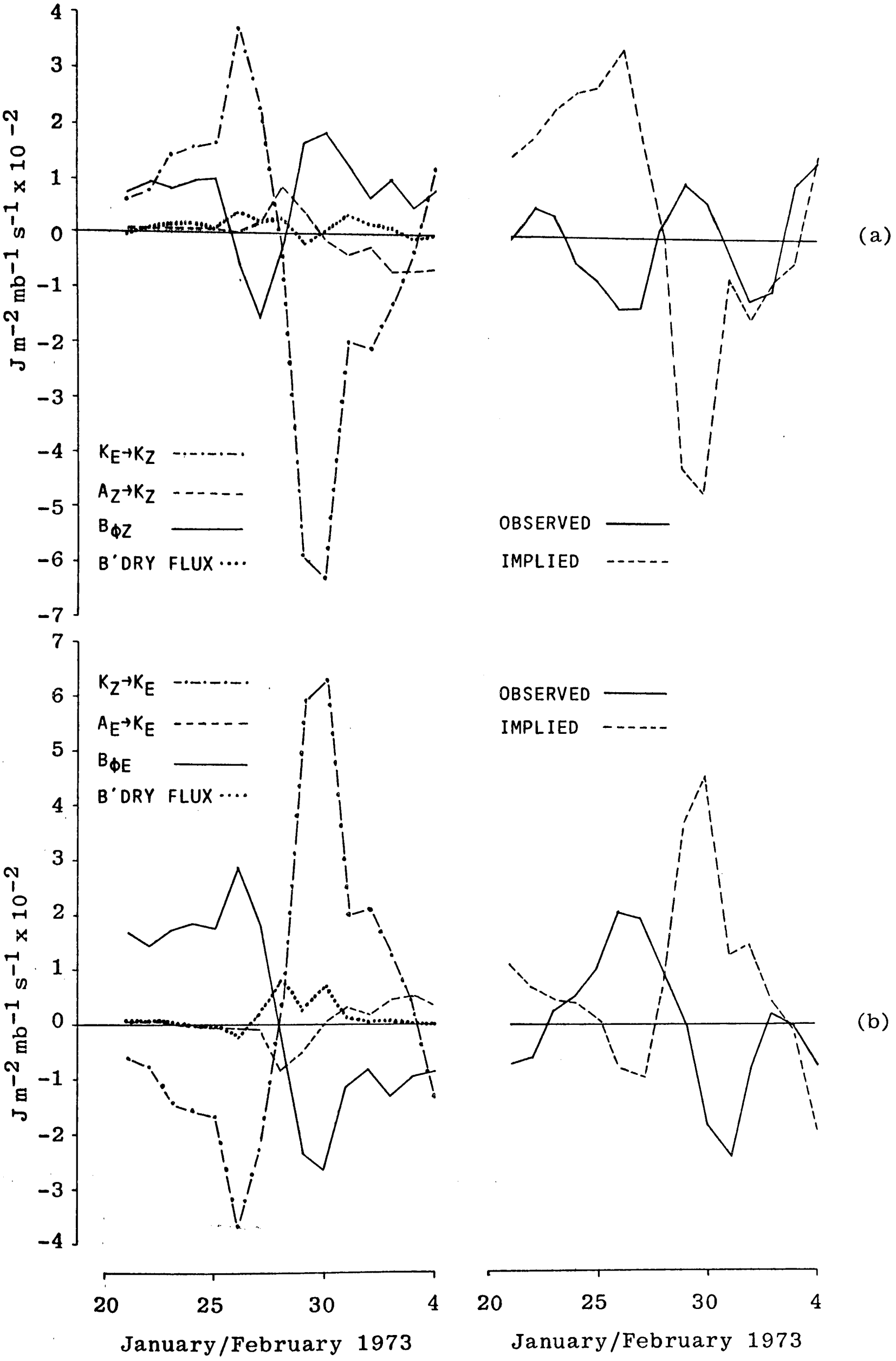


Figure 7.9. As Figure 7.7 but for the 2 - 0.4 mb layer

vorticity, m is zonal wavenumber and other symbols have their usual meaning. When [u] is independent of height and latitude, $\partial[q]/\partial\phi$ is positive giving imaginary μ for negative or sufficiently large positive [u]. This is the Charney and Drazin result. When [u] is a function of height and latitude the sign of $\partial[q]/\partial\phi$, and hence μ, depends on the latitudinal gradient and curvature and the vertical shear of [u]. In Matsuno's model a mid-latitude region of small $\partial[q]/\partial\phi$ gave a low refractive index and wave energy was reflected away from mid-latitudes into the zonal jet. Dickinson, however, used an empirical approximation to $\partial[q]/\partial\phi$ which effectively omitted the zonal wind gradient and curvature terms. He therefore obtained lowest values of μ where [u] was high and waves were channelled away from regions of strong westerly wind. It is worth noting that the effect of including radiative damping is to render μ complex, thus allowing for attenuation of propagating waves.

Although there is good observational evidence in support of the models of Matsuno and Simmons (e.g. Hirota and Sato, 1969) some doubt on the applicability of these theoretical results to the upper atmosphere, however, has been cast recently by Hirota and Barnett (1977). Preliminary analysis of the winter mesosphere using data from the Pressure Modular Radiometer has shown a tendency for maximum wave amplitude to be centred some 15^{o} north of the zonal flow maximum in the middle mesosphere.

Observed relationship between vertical energy propagation and the zonal flow on 11.1.73 and during the warming

The distributions of the mean zonal flow in the meridional plane on 11.1.73 and on the four days 26 - 29.1.73 are presented in Figure 9.1. On 11.1.73 a saddle point in the zonal flow existed near 74^{o}N at about 5 mb. Matsuno (1970) showed that

Eddy kinetic energy

An unusual feature of the 1973 warming was the fact that very little change took place in the K_E content of the 50 - 10 mb layer. This observation contrasts not only with the results for the lower stratosphere during other warmings but also with those of Quiroz et al (1975) who studied the energetics of the 100 - 30 mb layer during this event. They observed a steady increase in K_E between 16 and 29.1.73. It appears that a rise in K_E occurred in the lower and the upper stratosphere but not at intermediate levels. The error in the conversion and generation processes is too large for definitive conclusions regarding their signs to be drawn from the K_E budget alone. However since a much closer balance was obtained for the K_Z budget of this layer it is probable that the $K_E \rightarrow K_Z$ conversion is a reasonable estimate and that the energy gain due to the eddy pressure-work term was overestimated.

In the middle layer energy gained via $B_{\phi E}$ was transferred to both A_E and K_Z, again with little net change in K_E. Above 2 mb, in contrast with lower levels, gain in K_E was very large. The kinetic energy field was largely decoupled from the available potential energy field with energy being transferred between K_E and K_Z and $B_{\phi E}$ acting as a source prior to 28.1.73 and a sink thereafter. Comparing Figures 7.9a and 7.9b it is clear that in the upper stratosphere the K_E-K_Z conversion was overestimated and the pressure-work terms underestimated throughout the warming period.

It should be noted that the horizontal components of the eddy pressure-work term $B_{\phi E}$ (equation 4.25) are identically zero if meridional velocities are calculated by the geostrophic approximation. The effect of these components on the kinetic

saddle points in the distribution of [u] are associated with small values of latitudinal potential vorticity gradient and low refractive index. If this theory applies we should expect to find that little wave energy flow occurred in this region on 11.1.73. The distribution of the vertical energy flux on 11.1.73 shown in Figure 8.7 has a marked minimum not at the latitude of the saddle point but about 6^{o} further south at the latitude of a zonal flow maximum. In Figure 9.2 are plotted similar distributions of the eddy momentum transport [u'v'] for 11, 12, 26, 27 and 29.1.73. According to the Eliassen and Palm approximation the horizontal energy flux [v'Φ'] is directed in the opposite direction to the momentum flux. Thus on 11.1.73 the line of zero momentum flux marks, approximately, the boundary between northward energy flux at higher latitudes and southward energy flux at lower latitudes.[1] This boundary corresponds well with the minimum in the vertical energy flux distribution of Figure 8.7. The energy flow was thus bifurcated and apparently avoided the jet at 68^{o}N in agreement with Dickinson's (1968b) results. However there was clearly a very strong energy flow into the westerly jet at 80^{o}N as indicated by the high values of both the vertical component of energy flux and the southward momentum transport in the region of this jet. It thus seems that the low refractive index region may have effected the splitting of the energy flow. The tendency for energy flow to occur preferentially in regions of strong westerly flow is more evident on 27, 28 and 29.1.73, maximum upward flux occurring at the centre of the high latitude

---

[1] It should be noted that the energy fluxes considered here (due to pressure-work effects alone) are as defined by Eliassen and Palm (1961) and are not the total energy fluxes defined by Holton (1974). This is the interpretation of energy flux adopted by Matsuno (1970).

energy budget of the near-hemispheric region is likely to be small, but the treatment of this process is discussed more fully in Chapter 8 in connection with the latitudinal distribution of kinetic energy.

Since the warming was predominantly a wavenumber one disturbance budgets of individual wavenumber components have not been calculated. In view of the errors obtained firm conclusions would have been difficult to draw and in addition the non-linear interactions between wavenumbers would have been required. However one interesting feature regarding the relationship between wavenumbers one and two eddy kinetic energy (K_{E1} and K_{E2}) is evident in Figure 6.6b. During the period 25 - 30.1.73 it is seen that maxima in K_{E1} were associated with minima in K_{E2}. One interpretation of this observation is that there was direct conversion between K_{E1} and K_{E2}. Although calculation of this conversion has not been made there is some evidence to suggest that this was not the case. Figure 7.10 shows the area-averaged K_{E1} and K_{E2} contents of the 2 - 0.4 mb layer for the period 23.1.73 - 1.2.73, and the evolutions of the $K_Z \rightarrow K_E$, $A_E \rightarrow K_E$ and $B_{\phi E}$ processes for each wavenumber. It appears that the K_{E2} maximum of 30.1.73 may have been caused by marked barotropic conversion $K_Z \rightarrow K_{E2}$ which was stronger than the equivalent conversion in wavenumber one. The wavenumber two components of $B_{\phi E}$ and $A_E \rightarrow K_E$ were relatively weak throughout. A loss of K_{E2} to other K_E components probably occurred after 30.1.73. The different behaviour of different harmonic components of eddy kinetic energy was noted both by Perry (1967), in an analysis of the 1963 warming, and by Arpe (1976), who compared the wavenumber two type warmings of 1967-8 and 1970-1 in the 75 - 7.5 mb layer. During both these latter warmings K_{E2} rose sharply while K_{E1}

jet. The propagation of energy from a high latitude source upwards and southwards into the zonal flow jet at high levels, a feature absent on 11.1.73 due to the lower level energy flow bifurcation, was undoubtedly a necessary precursor of the circulation breakdown. Comparison of the momentum flux distributions on 26 and 27.1.73 shows the effect of the energy propagation to high levels. The changes in momentum flux distribution following the energy pulse of 11.1.73 were not significant at high levels (Figure 9.2).

The common result of the theoretical investigations concerning waves incident on a zero wind line is shown to be well founded. It is seen from Figures 8.7 and 9.1 that there was a close correspondence between the easterly wind line and the zero energy flux line showing that energy associated with the quasi-stationary disturbance could not propagate across the easterly wind boundary.

It is possible that the coincidence of maxima in the zonal flow and vertical energy flux may be a feature peculiar to the sudden warming event. When the latitudinal distributions of wavenumber one vertical energy flux and the mean zonal flow averaged over the period 11 - 27.1.73 are compared it is seen that the maximum vertical energy flux was situated further poleward than the westerly jet at 2, 10 and 50 mb (Figure 9.3a). In Figure 9.3b the latitudes of the maxima in these quantities are plotted against time for the period 15 - 28.1.73. At 50 mb the energy flux was predominantly downwards at 50 mb before 20.1.73 and was therefore not plotted. Double jets occurred at 2 mb on 15.1.73 and at 10 mb on 23.1.73. It is seen that the maximum vertical energy flux was consistently north of the zonal flow maximum but moved to the south at the onset of the warming.

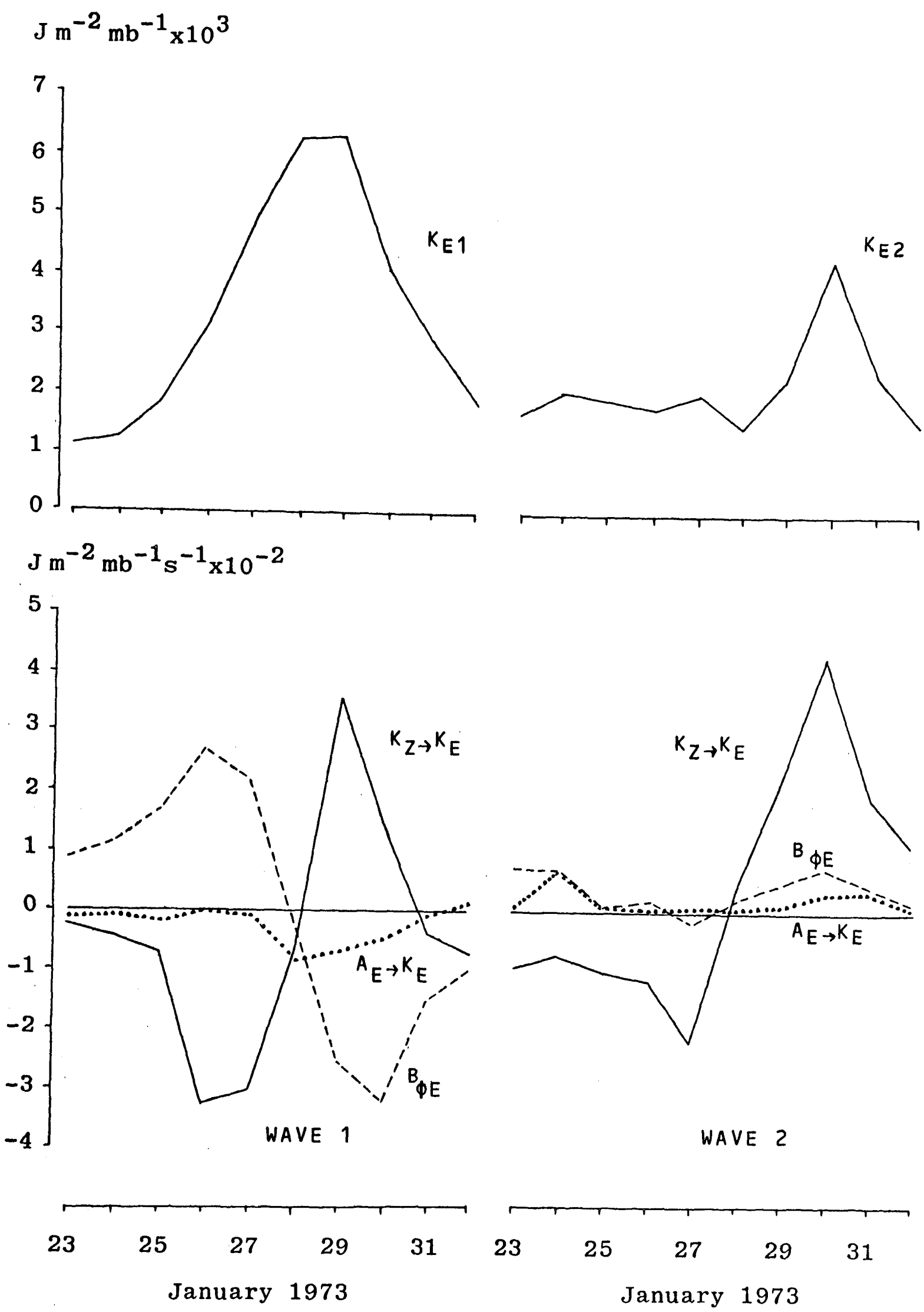


Figure 7.10. Mean content of wavenumber 1 (upper left) and wavenumber 2 (upper right) eddy kinetic energy in the 2 - 0.4 mb, 14 - 82°N region and the wavenumber 1 and 2 processes, respectively, contributing to the change in these energy components for the period 23.1.73 to 1.2.73. Energy exchanges between different wavenumbers are not shown

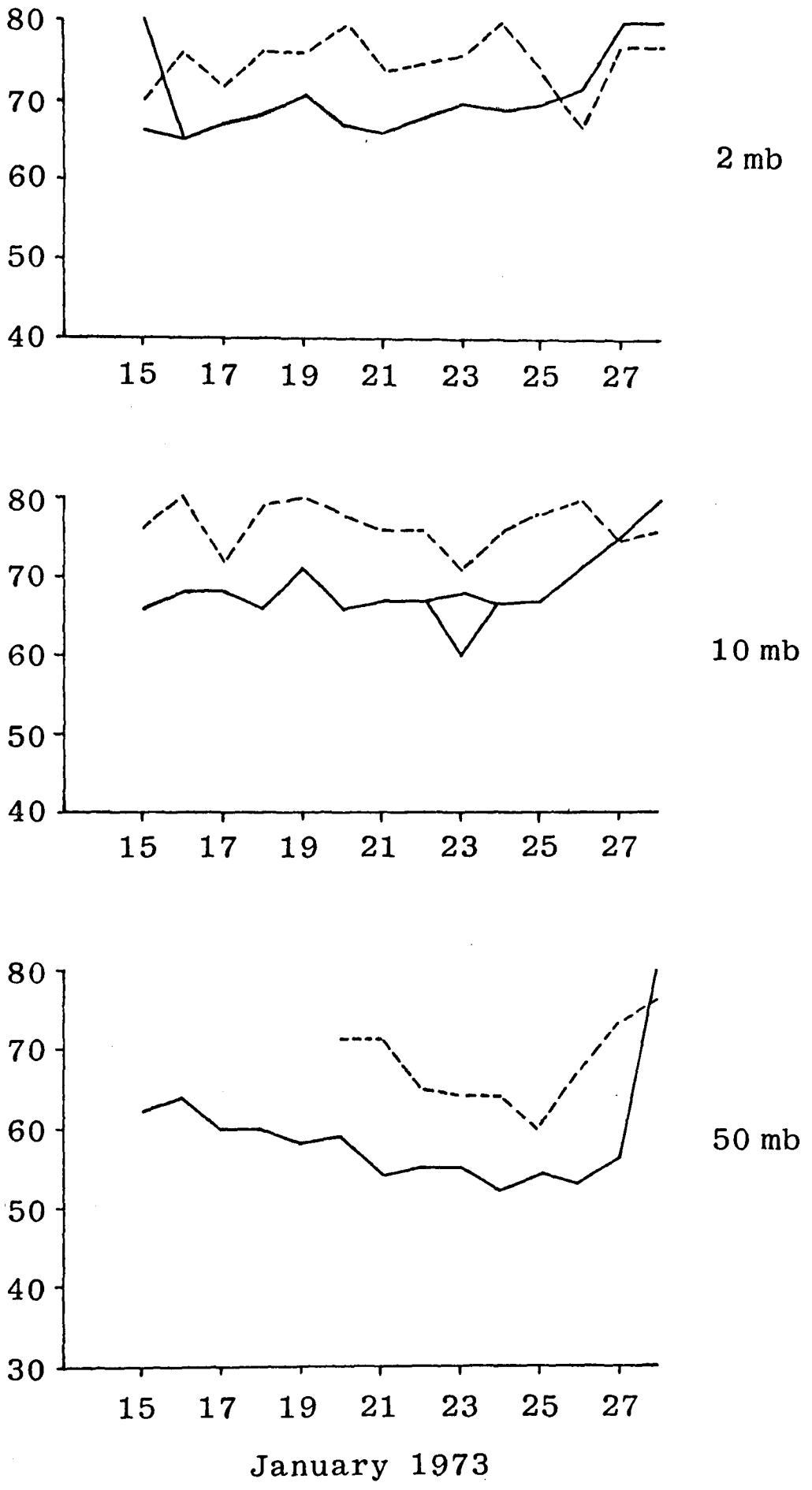


Figure 9.3b. Latitude of maximum zonal mean zonal wind (——) and latitude of maximum vertical (upwards) eddy energy flux (- - - -) against time for the period prior to the warming

declined, although the interaction between these components was in the sense $K_{E2} \rightarrow K_{E1}$. In the 1967-8 event the major source of K_{E2} was A_{E2} while in 1970-1 $B_{\phi E2}$ supplied K_{E2}. In both warmings $B_{\phi E1}$ was a sink for K_{E1}.

(iii) Balance problems in energy budget calculations

Many studies of stratospheric energetics have been based on data with relatively poor temporal resolution, as indicated in Table 7.1. In such cases results have often been averaged over time in order that a coherent picture of the energetics could be presented. Such time-averaging has tended to mask the imbalances which have been found in the calculations for individual days. Both Muench (1965) and Mahlman (1969), for example, commented on the erratic variation of the imbalance in the kinetic energy budgets during January and November-December, 1958, respectively. Miller et al (1972) showed the importance of adequate temporal resolution by finding large differences in the kinetic energy budget of the lower stratosphere using data for four days and that calculated using 15 days' data over the same period. Moreover, if calculations are made on a daily basis a qualitative indication of the random error in the calculated energy conversion and generation processes is given by the degree of continuity from day to day.

In this study accuracy of the initial data rather than data frequency was a problem and it has been suggested that retrieval error has contributed to the imbalances found. Reliability of the data is of course a problem not restricted to users of satellite observations. Arpe (1976) calculated the geopotential flux $[\omega'\Phi']$ at 100 mb during the winter of 1967-8 using data produced by the Institute of Meteorology at the Free University of Berlin and again using data compiled by the German Weather

flux or if there is a vertical shear of the divergence of the eddy heat flux. The first of these mechanisms causes a direct, local acceleration of the mean flow, while the vertical shear of the eddy heat flux divergence brings about a vertical shear in the induced vertical motion which must be accompanied by meridional motion to satisfy mass continuity. It is the action of the Coriolis acceleration of this induced meridional flow which accelerates the zonal flow.

The sudden warming mechanism put forward by Matsuno (1971) is based on the theory first proposed by Charney and Drazin (1961) that in the absence of critical levels the acceleration of the zonal flow by a quasi-geostrophic, adiabatic, steady wave is zero. The accelerations produced by momentum flux divergence and the Coriolis torque therefore cancel each other. During a sudden warming the zonal flow acceleration is not zero. Therefore in developing his theory Matsuno considers two mechanisms by which zonal flow acceleration may occur. He discusses first the consequences of a wave encountering a critical level and secondly the effect of wave transience. The model consists of wave propagation in a zonal flow on a β-plane bounded by vertical walls at two latitudes. For simplicity the wave phase is assumed independent of latitude so that eddy momentum transport is zero and zonal flow acceleration can be effected only by vertical shear in the eddy heat flux. If a wave is incident on a critical level in the zonal flow the poleward eddy heat flux associated with the upward propagating wave jumps to zero. The upward and downward motions forced by the divergence of the heat flux at high and low latitudes respectively therefore diminish above the critical level. To balance mass continuity equatorward flow occurs in the vicinity of the critical level on which the Coriolis acceleration acts to

Service. During a sudden warming period the former values were at times larger by more than 50%. Imbalance in the kinetic energy budget is often attributed to frictional effects. Invariably however such imbalances are much larger than can be explained by this mechanism. Muench (1965), recognising that frictional loss could not explain the error, suggested that use of the adiabatic thermodynamic equation could have been the cause. Although this approximation may have contributed to the error, studies employing the full thermodynamic equation have suffered similar imbalances - for example that of Dopplick (1971) who, like Mahlman (1969), attributed a large residual in the K_Z budget to poor specification of $B_{\phi z}$.

Even if perfect data were available computational errors would remain, of which the use of finite difference techniques is probably the most significant and unpredictable, especially during periods of rapid temporal and spatial change. In any study of energetics it is therefore important to compare observed and implied rates of change of energy and to treat the calculated conversions and generations with an appropriate degree of caution.

(iv) The Lorenz energy cycles

To facilitate comparisons between the stratospheric layers and between pre- and post- warming energetics the conventional box diagrams have been compiled (Figures 7.11a, b). In the boxes are the observed (i.e. directly calculated) rate of change of energy type (top left), the rate of change implied by the conversions, generations and boundary fluxes (top right) and the amount of the energy type present. In order that the time scales of stratospheric energetics may be appreciated the units of conversion, generation, boundary flux and rate of change terms have been converted to $Jm^{-2}mb^{-1}day^{-1}$. The values shown are those

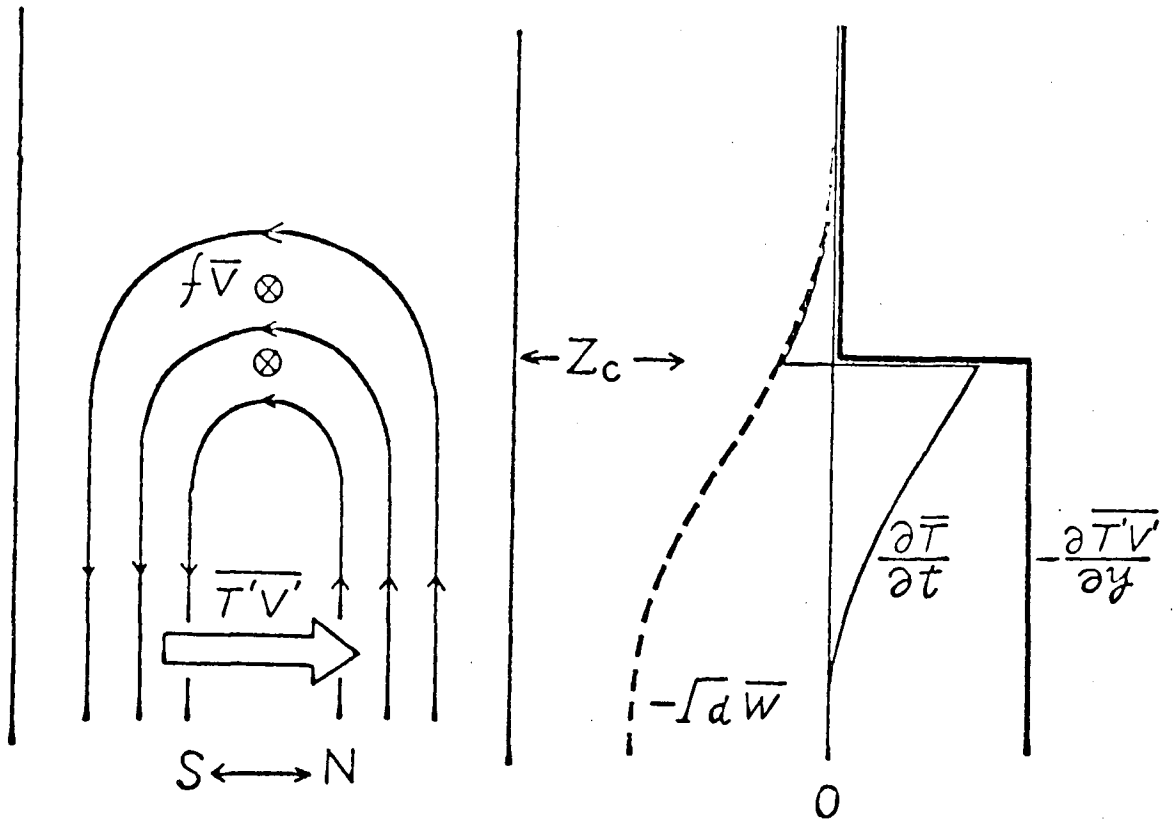


Figure 9.4. Schematic illustration of the mean meridional circulation induced by the waves (left) and the vertical distribution of temperature change components (right) showing eddy heating (thick solid line), mean vertical motion (dashed line) and resultant temperature change (thin solid line). (Taken from Matsuno, 1971). Z_c marks the critical level

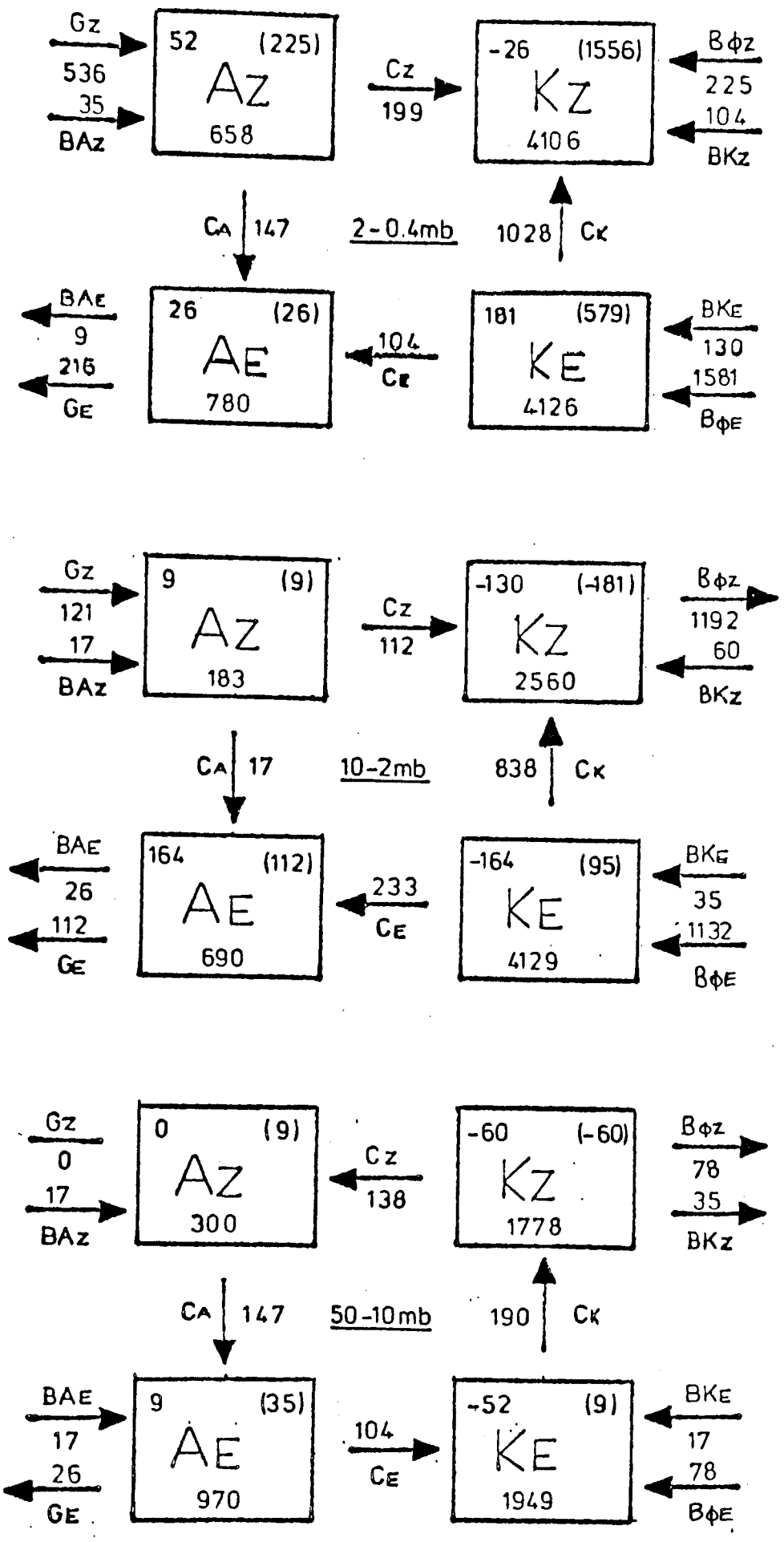


Figure 7.11a. Stratopheric energetics during the period 11 - 27.1.73. The mean energy contents and conversions for the period are shown for the three layers 2-0.4 (top), 10-2 (centre) and 50-10 (bottom) In the boxes are given the directly calculated (top left) and implied (top right) rates of energy change and the amount of energy present. Units are $Jm^{-2}mb^{-1}day^{-1}$ for conversions, generations and boundary fluxes and $Jm^{-2}mb^{-1}$ for energy contents

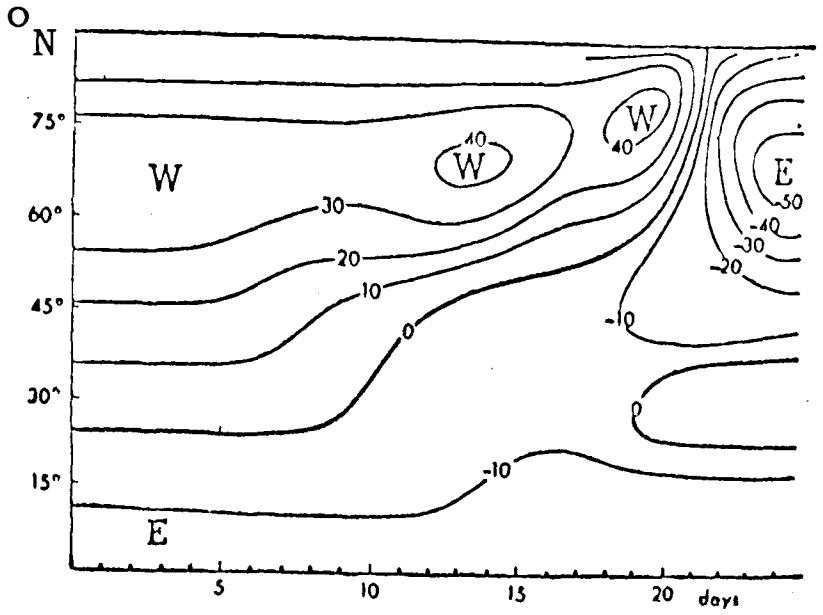


Figure 9.5a. Latitude - time section of the zonal mean zonal wind at 30 km (ms^{-1}) produced in the case C2 warming simulation of Matsuno (taken from Matsuno 1971)

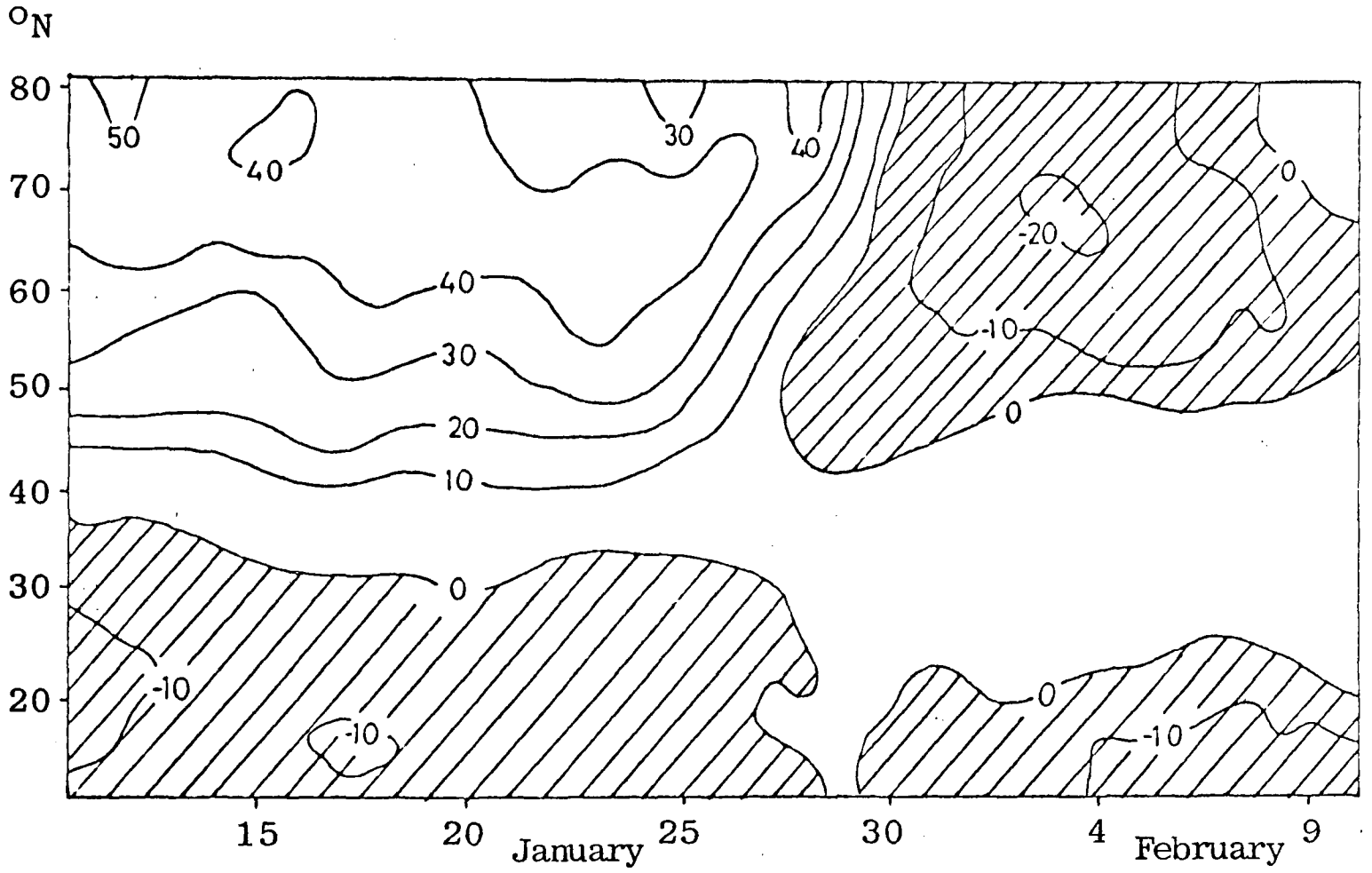


Figure 9.5b. Latitude - time section of the zonal mean zonal wind at 10 mb observed during the period 11.1.73 - 10.2.73 ($m\,s^{-1}$). Shaded regions denote easterly winds

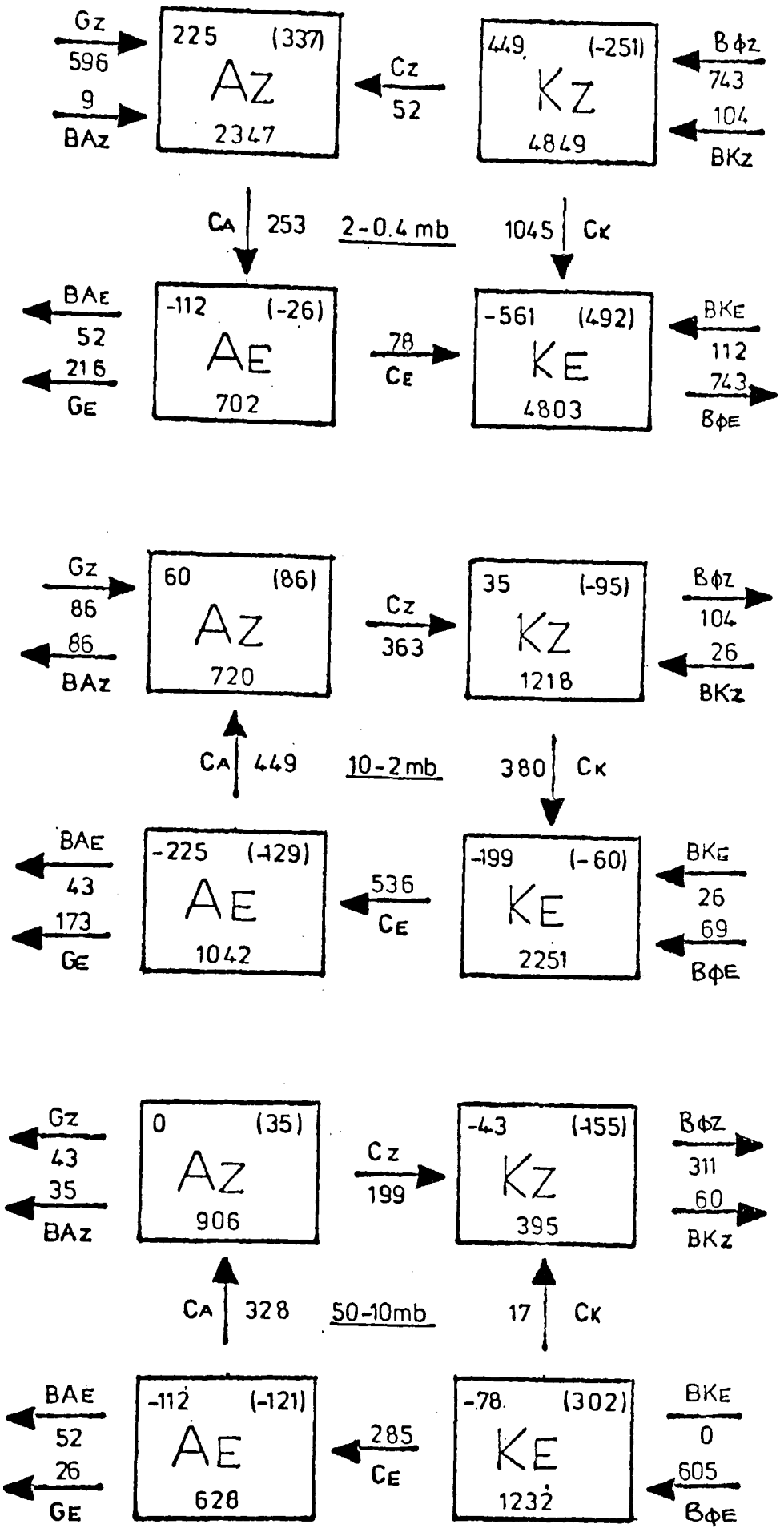


Figure 7.11b. Stratospheric energetics during the period 28.1.73 to 10.2.73. See caption to Figure 7.11a for explanatory notes

oscillations in the poleward heat flux by Madden (1975).

Whilst noting that the rapidity of the descent of the easterly flow observed in high latitudes during the 1973 warming probably lends support to Geisler's theory, it is beyond the scope of the present study to investigate in any detail the mechanism by which this occurred.

Numerical simulation of a warming by Holton (1976)

To assess the possible differences between warmings of wavenumber one and two, Holton (1976) ran two simulations using a primitive equation model with zonal flow distributions and forcing very similar to those used by Matsuno (1971). His results for the wavenumber two event were very like those of Matsuno but a more realistic wavenumber one warming was simulated suggesting that the quasi-geostrophic theory used by Matsuno does not model well the structure of wavenumber one planetary waves. The major difference found between the wavenumber one and two warmings was that in the former the maximum rate of temperature increase occurred at the pole throughout the event, while in the wavenumber two case all latitudes north of 60^{o}N initially warmed at the same rate with a marked temperature maximum appearing first at 50^{o}N and migrating to the pole. Figure 9.6a shows the latitudinal distribution of zonal mean temperature for various days during the development of both warmings plotted as the deviation from the hemispheric mean. The implications of these temperature changes for the zonal flow are illustrated in Figure 9.7a which consists of latitude height sections of the mean zonal wind during the pre-warming and warming stages. In the wavenumber one warming the zonal flow decelerated at all high latitudes, resulting in fairly uniform easterlies, while in the wavenumber two case easterlies appeared first in middle latitudes near 10 mb and moved

previously calculated to the nearest $Jm^{-2}mb^{-1}s^{-1}$ multiplied by 86400 and are thus approximate. Bearing in mind what has been said regarding the accuracy of these calculations, the important features of these figures are the relative rather than absolute values of the energetics.

Assuming these time averaged results to be typical of an active winter it is seen that throughout the stratosphere conversion rates are such that most of the processes above could replace the total amount of an energy type in about two to five days. Since the conversions and generations tend to oppose one another the observed rates of energy change are generally much smaller, an average value of 15 days being observed. The most rapidly changing component, relative to its mean content, during the period of observation was the eddy available potential energy in the 10 - 2 mb layer with an average replacement time of just over four days.

The pre- and post- warming energy conversion cycles for the 50 - 10 mb region were very similar to those found by previous investigations as illustrated in Figures 7.1 and 7.2. Energy rates of change were small with the four internal conversions having similar magnitudes. As has been noted the lack of any significant increase in K_E during the warming is in contrast with the results of earlier warming studies. It should be mentioned that the energy cycle for the 100 - 30 mb layer during this event presented by Quiroz et al (1975) showed conversions in the opposite sense during the amplifying stage. The reason is that they chose to take 5.2.73 as the date dividing the two stages, which, as shown both by their own analyses and by those presented here, was about one week after the peak of the event. The 10 - 2 mb and 2 - 0.4 mb layers had like internal energy

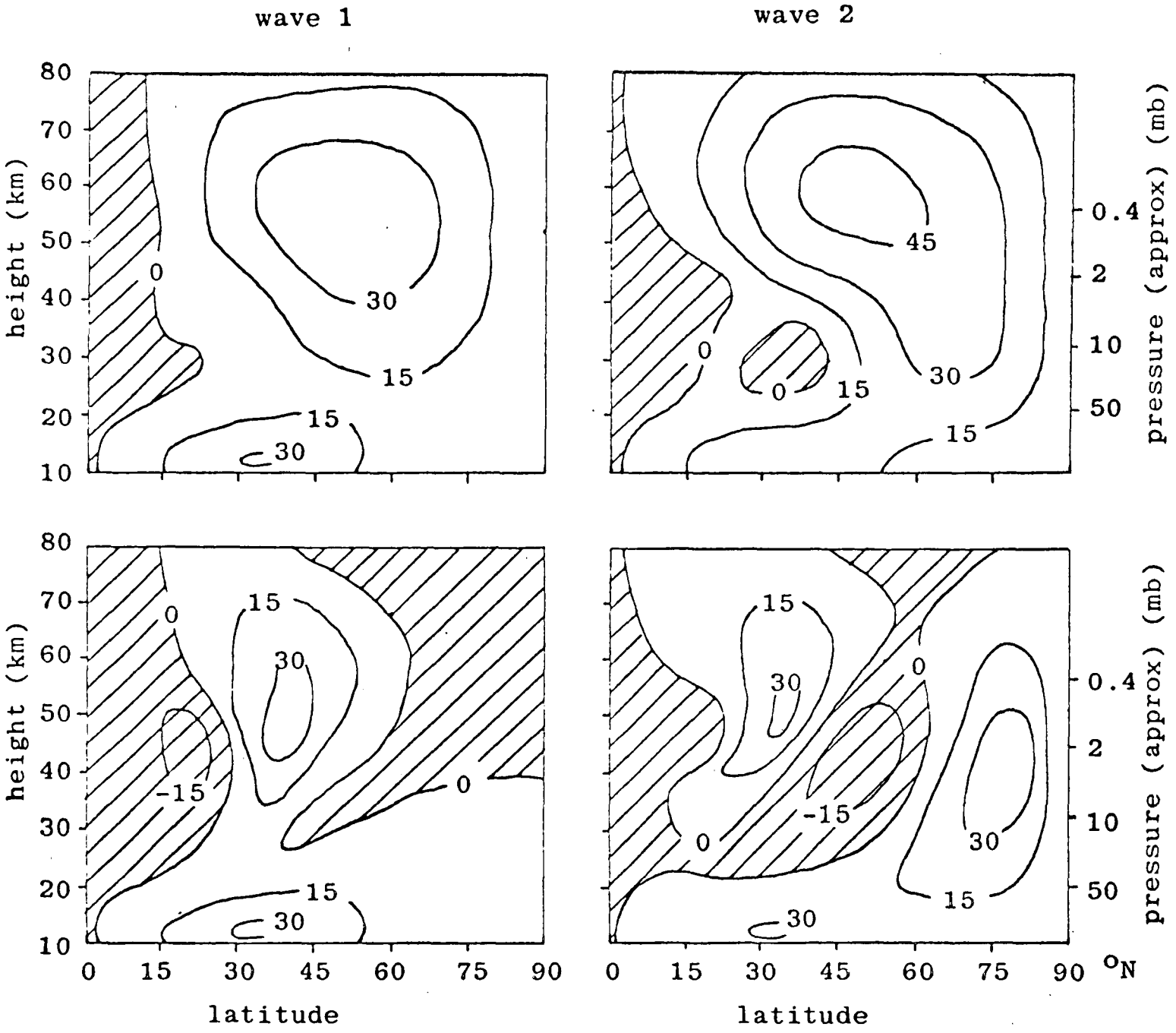


Figure 9.7 a. Latitude - height sections of zonal mean zonal wind ($m\,s^{-1}$) for the pre - warming stage (upper) and warming stage (lower) of wavenumber 1 (left) and wavenumber 2 (right) warmings simulated by Holton, 1976. Easterlies are shaded. (After Holton, 1976)

cycles prior to the warming peak although differences in generation terms and magnitudes of the conversions led to a large increase in A_E between 10 and 2 mb and in K_E between 2 and 0.4 mb. An important feature to note is that the pressure work terms and the barotropic conversion were about four times larger than other processes, excepting the radiative generation of A_Z above 2 mb which was a large energy source throughout. In the declining phase of the event strong reversed cycles below 2 mb maintained the warm pole much longer than in the upper stratosphere (Figure 6.25). A large increase in A_E in the middle stratosphere was observed during the 1969-70 warming by Miller et al (1972). However in that event the energy gain via the eddy pressure-work effect was much smaller than the internal conversions and the gain in A_E occurred via $A_Z \rightarrow A_E$ conversion rather than forced conversion from K_E. The dominance of the $B_{\phi E} \rightarrow K_E \rightarrow K_Z$ cycle above 10 mb was observed by Hartmann (1976) during the southern hemisphere winter of 1973. It is likely that this is a normal wintertime feature of the middle and upper stratosphere.

The overall effect of the warming was to transfer energy from the low and middle stratosphere to the upper stratosphere. This rearrangement of energy is seen clearly both in Figure 7.12, which shows the evolutions of the energy components in each layer, and in Table 7.2 where the sums of available and kinetic energy in each layer on 11.1.73 are compared with values for 10.2.73. The calculations show a 21% fall in energy of the 50 - 0.4 mb region over the period. In view of the changes in individual layers it is probable that this energy was lost to the mesosphere.

A comparison between the observed mean rate of change of A+K during the two phases of the period and the rates implied by conversion, generation and boundary terms is made in Table 7.3.

northwards and upwards as the warming progressed. Corresponding observations for the 1973 warming are presented in Figures 9.6b and 9.7b. The similarity between Holton's wavenumber two warming and the 1973 wavenumber one event is remarkable, and it would appear that the two types of simulated warming are not necessarily associated with disturbances of a particular wavenumber, although Holton suggested that the different boundary conditions appropriate to each wavenumber ultimately determined the different evolutions. The northward movement of easterlies during the 1973 warming is shown clearly in Figure 9.1. Examination of the momentum balance indicates that in the wavenumber one simulation the eddy momentum flux divergence and Coriolis torque had similar latitudinal distributions giving a deceleration of the zonal flow at all latitudes. The Coriolis torque profile was further south in the wavenumber two case and as a result convergence of eddy momentum flux dominated in high latitudes giving an acceleration of the westerlies there. Similar high latitude acceleration occurred initially in the 1973 event but the movement of easterlies from mid-latitudes to the pole was more rapid than in the model. This result lends further support to the view put forward earlier that barotropic conversions of different sign at different latitudes play a major rôle in determining the evolution of a sudden warming.

The general circulation models of Trenberth (1973) and Newson (1974)

Since the atmosphere below 50 mb was not investigated in the current study it has not been possible to ascertain the cause of the marked increase in upward energy flux through the 50 mb level during the warming. As noted in Chapter 1 differences between northern and southern hemisphere winters strongly suggest that both orography and land-sea temperature differences play an

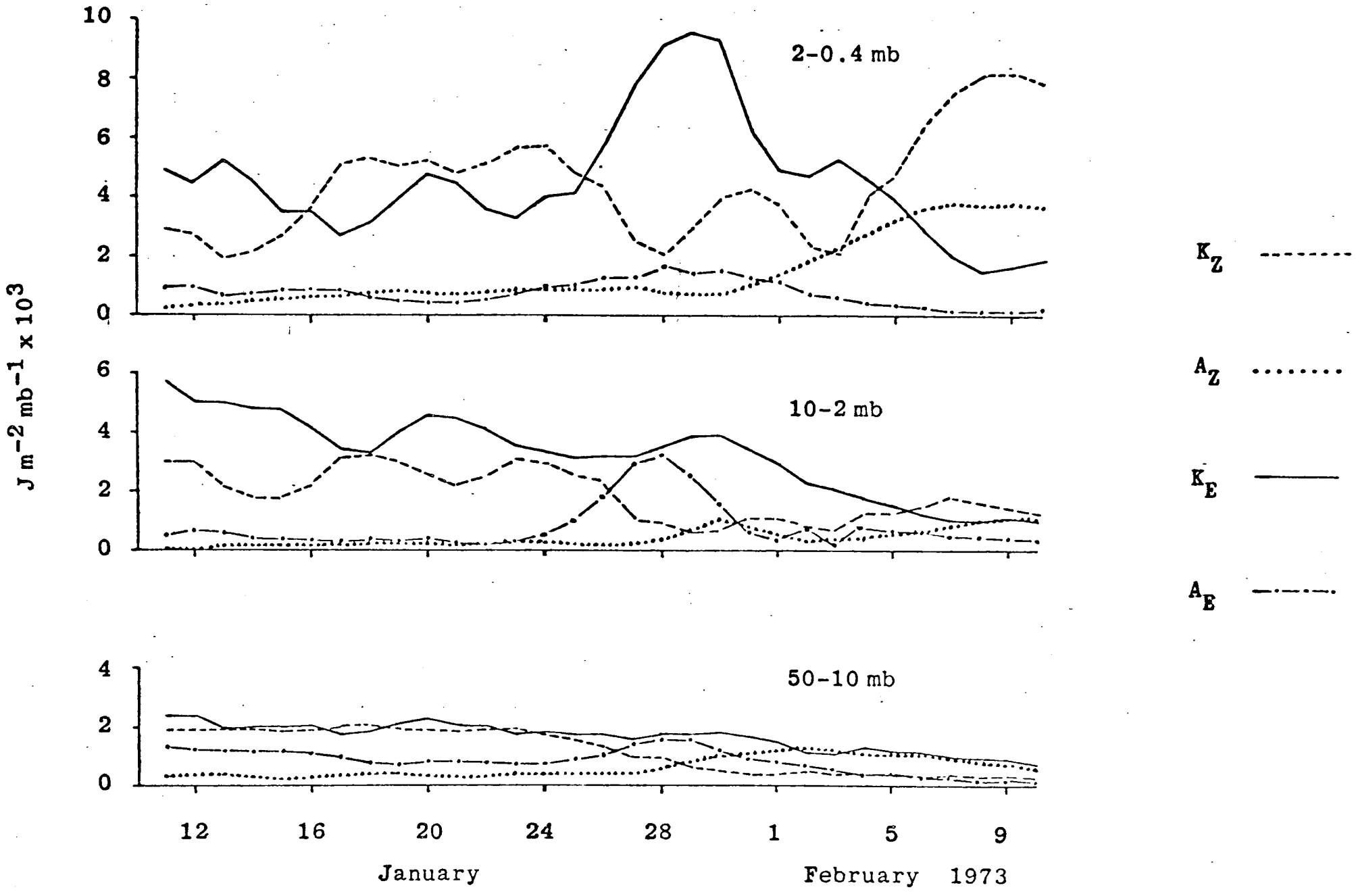


Figure 7.12. Mean energy contents of 2-0.4 mb (top), 10-2 mb (centre) and 50-10 mb (bottom) regions over latitudes 14-82°N plotted against time for the period 11.1.73 to 10.2.73 ($J\,m^{-2}\,mb^{-1} \times 10^3$)

height of the upper stratosphere at the onsets of the model and observed warmings (see Quiroz et al, 1975), the preliminary results of the model energetics for the 3 - 0 mb layer (Newson, 1974) show some significant differences from those of the 1973 event. The pre-warming phase was accompanied by a strong direct baroclinic cycle in the 3 - 0 mb layer which reversed during the warming. It will be recalled that prior to the 1973 warming significant $A_E \rightarrow K_E$ conversion occurred only below 10 mb but was established in the upper stratosphere after the peak of the event. It is also of interest to note that while the height perturbation at 2 mb was dominated by wavenumber one a significant maximum in the tropospheric wavenumber two amplitude was modelled prior to the warming which seemed to propagate upwards in mid-latitudes. Inspection of the 2 mb wavenumber two amplitude at the peak of the model warming, however, shows that it was only about half the amplitude of the wavenumber one perturbation observed in high latitudes during the 1973 event. It remains to be seen from the more detailed analysis of this model warming currently being undertaken whether a more significant, high latitude wavenumber one disturbance was a feature of the simulation.

TABLE 7.2

A+K CONTENT (MEAN OVER 14 - 82°N) ($Jm^{-2}mb^{-1} \times 10^3$)

mb	11.1.73	10.2.73	content on 10.2.73 as % of 11.1.73
2 - 0.4	8.9	13.5	152
10 - 2	9.4	3.9	41
50 - 10	5.8	1.7	29
50 - 0.4	24.1	19.1	79

TABLE 7.3

MEAN RATE OF CHANGE OF A+K ($Jm^{-2}mb^{-1}day^{-1} \times 10^3$)

	11 - 27.1.73		28.1 - 10.2.73	
	observed	implied	observed	implied
mb				
2 - 0.4	0.23	2.39	0.00	0.55
10 - 2	-0.12	0.04	-0.33	-0.20
50 - 10	-0.10	0.01	-0.23	0.06

It is seen that in each layer the implied energy gain was larger than observed in both periods. This suggests energy gains by the generation and/or pressure-work terms were consistently over-estimated, advective boundary fluxes being small.

The redistribution of energy between the different components within each layer and how these distributions compare with other observations is of interest. In Table 7.4 are presented the energy distributions in the stratosphere during a southern hemisphere winter (Hartmann, 1976) and a relatively quiet northern hemisphere winter (Dopplick, 1971), and during the pre- and post-

more qualitative nature than had been envisaged at the outset. Nevertheless some important results were obtained relating more especially to the upper stratosphere for which sudden warming energetics had not previously been analysed.

The sudden warming analysis

The energetics of the lower middle stratosphere during the sudden warming were in broad agreement with observations of previous warmings except in that there was little increase in the eddy energy components. Polar warming occurred here several days later than at higher levels. The vertical energy flux became convergent at these levels only after easterlies established aloft prevented energy propagation to the upper stratosphere.

An interesting result was that the pre-warming period was characterised by a strong baroclinic energy cycle in high latitudes below 10 mb which enhanced the vertical energy propagation in this region and led to a large increase in eddy available potential energy between 10 and 2 mb. Little energy propagated to higher levels at this stage. The structure of the wavenumber one disturbance evolved in such a way that on 27.1.73 baroclinic conversion to eddy kinetic energy occurred in high latitudes at all levels enhancing energy propagation into the upper stratosphere, where a rapid increase in poleward eddy momentum transport was established. Significant barotropic loss of zonal kinetic energy occurred in mid-latitudes of the upper stratosphere as a result and was instrumental in decelerating the zonal flow there. The appearance of easterlies first in mid-latitudes has been observed in numerically simulated sudden warmings but mid-latitude barotropic conversion has not before been suggested as a major factor in bringing about the circulation

warming periods of this study. The southern hemisphere results are averages for July and August, 1973 in the latitude range 15 - 90°S, while Dopplick's results for January and February, 1964 between latitudes 20 and 90°N are used to represent a quiet northern winter.

TABLE 7.4

Quantity ($Jm^{-2}mb^{-1} \times 10^3$)	SH July-Aug '73	NH Jan-Feb '64	NH 11-27.1.73	NH 28.1-10.2.73
	1 - 0.2 mb		2 - 0.4 mb	2 - 0.4 mb
K_Z	20.9		4.1	4.8
K_E	2.9		4.1	4.8
A_Z	0.5		0.7	2.3
A_E	0.4		0.8	0.7
	10 - 1 mb		10 - 2 mb	10 - 2 mb
K_Z	15.6		2.6	1.2
K_E	2.2		4.1	2.2
A_Z	1.4		0.2	0.7
A_E	0.5		0.7	1.0
	100 - 10 mb	100 - 10 mb	50 - 10 mb	50 - 10 mb
K_Z	4.7	2.1	1.8	0.4
K_E	0.7	1.1	1.9	1.2
A_Z	2.0	0.7	0.3	0.9
A_E	0.3	0.7	1.0	0.6
	100 - 0.2 mb		50 - 0.4 mb	50 - 0.4 mb
K_Z	13.7		2.8	2.1
K_E	1.9		3.4	2.7
A_Z	1.3		0.4	1.3
A_E	0.4		0.8	0.8

Considering the warming results first it is seen that in the 50 - 0.4 mb region the event brought about a reduction in the ratio K:A of from 5:1 in the period leading up to the warming to 2:1

of the wave disturbance were found to be of major importance in determining the development of the 1973 warming. The relationship between these variables and the zonal flow distribution has been examined briefly and may be important in connection with the onset of sudden warmings. Future studies of these events, and of quieter winter periods also, should concentrate on these aspects of stratospheric dynamics.

(iii) It has been shown that significant enhancement of vertical energy flux may occur through internal energy conversion processes. However the very large increase in upward propagation of energy at the peak of the warming appeared to be associated with increased energy flux through the 50 mb level. It has only been possible in this study to describe the stratospheric response to this energy pulse and not its source. Continued analysis of the behaviour of the troposphere and lower stratosphere is thus essential if complete explanations of these phenomena are to be found.

afterwards, the region below 2 mb contributing more to this reduction than that above. These values compare with a ratio of 9:1 for the southern winter, the difference being due not so much to smaller amounts of available energy but to the very large K_Z component there. From Figure 7.12 it is seen that the ratio of eddy to zonal energy at the peak of the warming was a little less than 4:1 between 2 and 0.4 mb and nearly 6:1 between 10 and 2 mb. These ratios fell substantially after the event, but only from about 4.2.73 onwards and above 2 mb did the zonal components become substantially larger than the eddy components, with what was probably a short-lived maximum Z:E ratio of 7:1 occurring on 8.2.73. Over the 50 - 0.4 mb region the time averaged value of the E:Z ratio was 1.3:1 in the pre-warming period and 1:1 in the post-warming phase. In the southern hemisphere winter Z:E ratios of between 6:1 and 7:1 were observed for all three layers. Comparing the Z:E ratios for the lower stratosphere only, values of 6.7:1, 1.6:1 and 0.7:1 were obtained for the southern winter, the quiet northern winter and the whole of the sudden warming period respectively. If these results are typical it would appear that only occasionally does the stability of the northern hemisphere winter zonal circulation approach that of its southern counterpart. Moreover the relative stability of the circulation in a quiet northern winter compared with a disturbed northern winter is much less than that of a quiet southern winter compared with a quiet northern winter.

by using the continuity equation 5.9,

$$= -\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}([v][\Phi]\cos\phi) - \frac{\partial}{\partial p}([\omega][\Phi]) - [\omega][\alpha]$$

where $\alpha = \rho^{-1}$, by using the hydrostatic equation 1.3. The right hand side of this equation is the integrand of expression 4.26 and the left hand side the integrand of expression 4.27.

Chapter 8

VERTICAL ENERGY FLUX, THE KINETIC ENERGY BUDGET AND THE SUDDEN WARMING MECHANISM

Introduction

The calculations of hemispheric energetics have shown that in the upper stratosphere the sudden warming was dominated by changes in kinetic energy. The study of the kinetic energy budget is therefore extended in this chapter. The rôle of the eddy pressure-work effect at various latitudes and its relation to the forcing at 50 mb is discussed. The latitudinal distribution of the processes affecting the kinetic energy budgets are presented and it is shown that increases in eddy kinetic eddy in middle and high latitudes were probably brought about by different mechanisms, and that baroclinic conversion appeared to be an important factor in the early stages of the event. An attempt is made to link these observations to form a coherent picture of the way in which the circulation breakdown was produced.

8.1. INTERPRETATION OF WAVE ENERGY FLUX

Up to now the pressure-work terms have been considered to represent a mechanism whereby kinetic energy appears or disappears within a region due to pressure forces on its boundaries. An alternative interpretation, and one which will be useful here, is to regard $[\omega'\Phi']$ and $[v'\Phi']$ as eddy fluxes of geopotential or eddy energy fluxes whose convergence or divergence leads to local changes in the eddy kinetic energy. The analyses of wave - mean flow interaction in the literature have, however, given rise to some confusion as to exactly what constitutes the vertical and horizontal wave energy fluxes. Eliassen and Palm (1961) derived from the perturbation equations of a stationary mean zonal flow

$$= \frac{gn}{2af\cos\phi}\left(-a^{\phi}_{n,h}\, b^{\phi}_{n,T} + a^{\phi}_{n,T}\, b^{\phi}_{n,h}\right)$$

The 'total' eddy heat flux is

$$[v'T']^{\phi} = \sum_{n=1}^{4} [v_n T_n]^{\phi}.$$

Finite differences were used to evaluate latitudinal gradients of geopotential height. Thus

$$u_n^{\lambda,\phi} = -\frac{g}{af}\frac{1}{2\delta}\left((a^{\phi+\delta}_{n,h} - a^{\phi-\delta}_{n,h})\cos n\lambda + (b^{\phi+\delta}_{n,h} - b^{\phi-\delta}_{n,h})\sin n\lambda\right)$$

where δ is the latitude grid interval. The zonal mean wavenumber n eddy momentum flux at latitude ϕ is therefore

$$[v_n u_n]^{\phi} = -\left(\frac{g}{af}\right)^2 \frac{1}{2\delta\cos\phi}\frac{1}{2\pi}\int_0^{2\pi}\Big(n\,(-a^{\phi}_n \sin n\lambda + b^{\phi}_n \cos n\lambda)$$
$$\left((a^{\phi+\delta}_n - a^{\phi-\delta}_n)\cos n\lambda + (b^{\phi+\delta}_n - b^{\phi-\delta}_n)\sin n\lambda\right)\Big)\, d\lambda$$

$$= \left(\frac{g}{af}\right)^2 \frac{n}{4\delta\cos\phi}\left(a^{\phi}_n(b^{\phi+\delta}_n - b^{\phi-\delta}_n) - b^{\phi}_n(a^{\phi+\delta}_n - a^{\phi-\delta}_n)\right)$$

where the geopotential height subscript 'h' has been omitted for the sake of clarity. The 'total' eddy momentum flux is

$$[u'v']^{\phi} = \sum_{n=1}^{4} [v_n u_n]^{\phi}.$$

in geostrophic balance the wave energy equation

$$\frac{\partial}{\partial y}[v'\Phi'] + \frac{\partial}{\partial p}[\omega'\Phi'] = -[u'v']\frac{\partial}{\partial y}[u] - [u'\omega']\frac{\partial}{\partial p}[u] + f\sigma^{-1}\frac{\partial}{\partial p}[u]\ [v'\frac{\partial}{\partial p}\Phi'] \qquad 8.1$$

where Φ' is the perturbation geopotential, σ is a static stability parameter and other terms have their usual meanings. They regarded the terms on the right hand side as conversions from zonal kinetic and zonal potential energy to the corresponding eddy forms, and the left hand side as wave energy flux divergence. Using the thermal wind equation (equation 1.5) to replace $\partial[u]/\partial p$ in the last term it is seen that these conversion terms are those given in Chapter 4 as being appropriate only for a closed system, i.e. for a region with no flow through the boundaries. In order to present in a physically more meaningful way the conversions between the wave and the zonal flow Holton (1974) recast the equation in the form

$$\frac{\partial}{\partial y}([v'\Phi'] + [u][u'v']) + \frac{\partial}{\partial p}([\omega'\Phi'] - f\sigma^{-1}[u][v'\frac{\partial}{\partial p}\Phi'] + [u][u'\omega']) = [u]F \qquad 8.2$$

where F, the 'eddy forcing', is given by

$$\frac{\partial}{\partial y}[u'v'] + \frac{\partial}{\partial p}[u'\omega'] - \frac{\partial}{\partial p}(f\sigma^{-1}[v'\frac{\partial}{\partial p}\Phi'])$$

The third term in F is a rewritten form of the Coriolis torque term $f[u][v]$ which represents the total loss of zonal kinetic energy through the action of the mean meridional circulation. Holton thus regarded eddy forcing as all processes which alter the zonal kinetic energy while Eliassen and Palm considered wave - mean flow interaction as involving interchange between both eddy and both zonal mean energy components. As a result of the different conversion terms the energy flux divergences are different: the left hand side of equation 8.1 is the divergence of eddy kinetic energy flux due to pressure-work terms

using a zonally constant heating rate. The theoretical basis of the scheme is given below.

Consider first the 'oxygen-only' photochemical scheme (Chapman, 1930). The basic reactions are:

$$O_2 + h\nu \rightarrow O + O \qquad (J_2) \qquad \text{R1}$$

$$O + O_2 + M \rightarrow O_3 + M \qquad (k_2) \qquad \text{R2}$$

$$O_3 + h\nu \rightarrow O + O_2 \qquad (J_3) \qquad \text{R3}$$

$$O_3 + O \rightarrow 2O_2 \qquad (k_4) \qquad \text{R4}$$

$$O + O + M \rightarrow O_2 + M \qquad (k_5) \qquad \text{R5}$$

The symbols in brackets denote dissociation or reaction rates. Reaction R5 is not important in the stratosphere. From reactions R1 - R4 may be deduced the rate of change of 'odd' oxygen, i.e. O and O_3:

$$\frac{\partial}{\partial t}([O] + [O_3]) = 2J_2[O_2] - 2k_4[O][O_3] \qquad \text{A3.1}$$

where [ ] denotes concentration. Examination of the magnitudes of the various terms shows that $k_2[O_2][M] \gg k_4[O_3]$. Thus at equilibrium ($\partial/\partial t([O_3]) = 0$) we have

$$J_3[O_3]_e = k_2[O_2][O][M] \qquad \text{A3.2}$$

From A3.1 and A3.2 we obtain the ozone equilibrium concentration

$$[O_3]_e = [O_2] \left(\frac{J_2 k_2 [M]}{J_3 k_4}\right)^{\frac{1}{2}} \qquad \text{A3.3}$$

The main temperature dependence of the right hand side of A3.3 lies in k_2 and k_4. Recent estimates of these temperature dependences are *exp* (510/T) for k_2 and *exp* (-2300/T) for k_4 giving

$$[O_3]_e \propto exp\ (1405/T) \qquad \text{A3.4}$$

if other temperature dependences are ignored.

The next stage is to introduce catalytic chains considered

only, while that of equation 8.2 is the divergence of the flux of total eddy energy (see Andrews and McIntyre's (1976) equation 4.2). Now for quasi - geostrophic, adiabatic motions Holton shows that both sides of equation 8.2 are zero for non-zero [u], i.e. the momentum flux divergences balance the Coriolis torque in F and the working by pressure forces is balanced by fluxes of eddy kinetic and potential energy. Eliassen and Palm obtained the same result but, of course, their interpretation was different, namely one of non-zero vertical and horizontal energy fluxes, the familiar 'Eliassen and Palm approximations':

$$[v'\Phi'] \simeq -[u][u'v'] \text{ and } [\omega'\Phi'] \simeq f[u]\sigma^{-1}[v'\frac{\partial}{\partial p}\Phi']$$

During a sudden warming the zonal flow and the wave disturbances are not steady (i.e. are not independent of time). However it has been shown in Chapter 6 that even then there may be a near balance between the eddy and the mean meridional circulation terms in both the heat and momentum balances, so that F may be much smaller than its components. Thus to a fair approximation we may assume that the total wave energy flux as defined in equation 8.2 is zero even in a warming situation. The close agreement between the calculated $[\omega'\Phi']$ and the steady state Eliassen and Palm approximation is indicated in Figure 8.1 where the convergence of $[\omega'\Phi']$ in the three stratospheric layers is plotted. Since both methods of calculation are based on heat balance the differences, presumably, are due simply to the inclusion of radiative heating and the local rate of change of temperature in the calculation of the ω's used in the direct calculation, and computational errors. The implications of this discussion for the eddy kinetic energy balance to be investigated later in this chapter are as follows: (a) since the vertical

Depending on the relative importance of the hydrogen and nitrogen oxides cycles it may be expected that $[O_3]_e$ will be proportional to a value in the range *exp* (1400/T) to *exp* (300/T).

Measurements from the SCR and the back scatter ultra-violet spectrometer on Nimbus 4 were used to investigate the relationship between ozone concentration and temperature near the stratopause. A temperature dependence of $[O_3]$ in the region *exp* (1100/T) to *exp* (900/T) was found suggesting that the hydrogen oxides scheme is important in the upper stratosphere (Barnett et al, 1975b).

On the basis of the theory and observational evidence presented, a value of *exp* (1000/T) was taken in this study. The height - latitude distribution of temperature, T_e^*, corresponding to the equilibrium ozone mixing ratios, $[O_3]_e^*$, given in Table A3.1 was obtained from the model run which produced the mixing ratios. The equilibrium ozone mixing ratio, $[O_3]_{e,\lambda}$, for a given level at longitude λ was then determined from the relation

$$[O_3]_{e,\lambda} = \frac{[O_3]_e^*}{exp\ (1000/T_e^*)}\ exp\,(1000/T_{e,\lambda})$$

where $T_{e,\lambda}$ is the temperature at longitude λ, for levels above 5.75 scale heights. Below this level dynamics play a large part in controlling the ozone distribution and it was not possible to specify longitudinal variation in solar heating. However heating rates are much smaller there so the variation is less important.

To illustrate the effect of this scheme height - longitude sections of the temperature and solar heating at 60^{o}N calculated for 28.1.73 have been plotted in Figure A3.1.

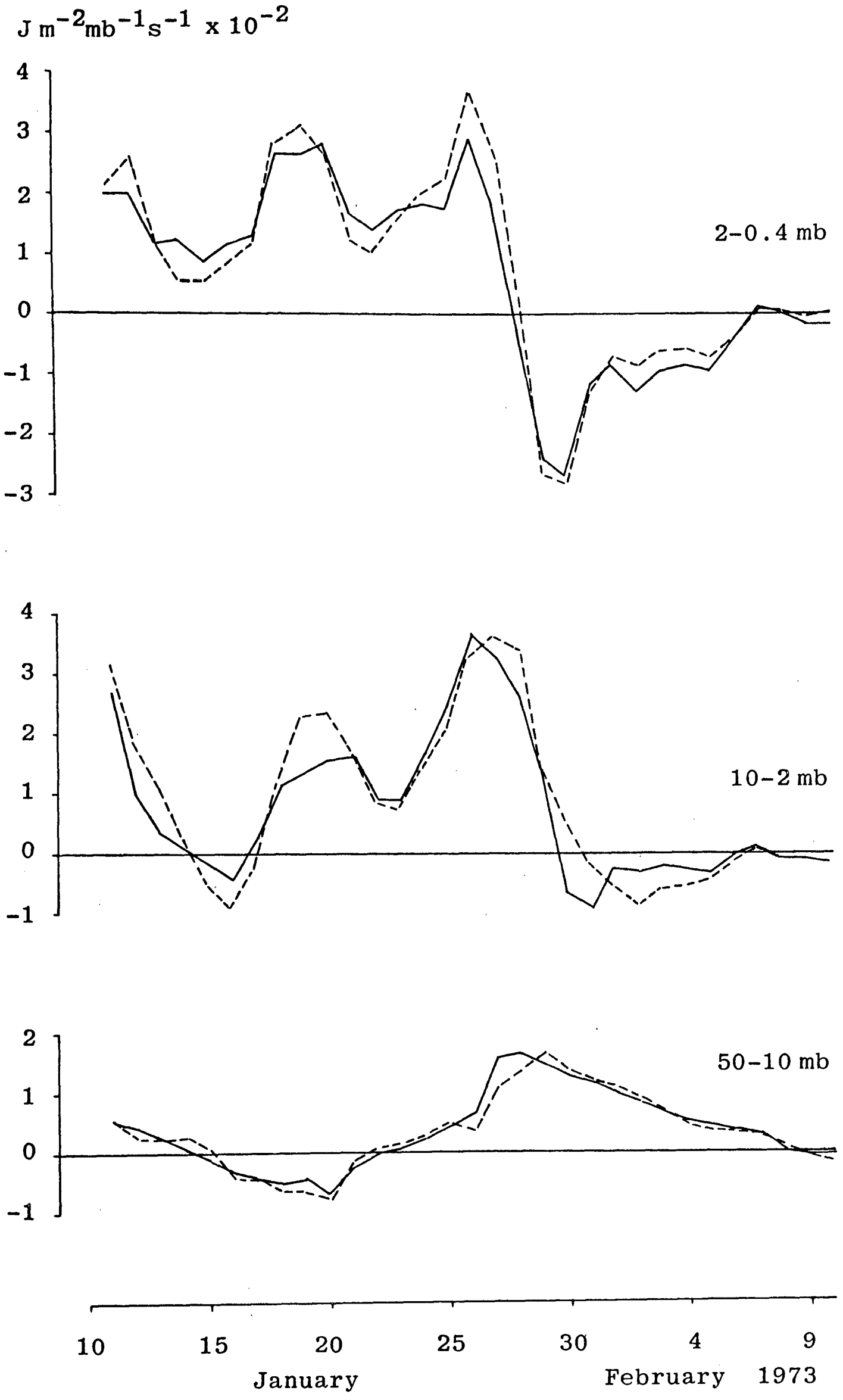


Figure 8.1. Comparisons of the convergence of the eddy vertical energy flux $B_{\Phi E}$ calculated directly (——) and calculated using the Eliassen and Palm (1961) approximation (----) for the 2 - 0.4, 10 - 2, and 50 - 10 mb layers during the period 11.1.73 to 10.2.73 (J m^{-2} mb^{-1} s^{-1} x 10^{-2})

Long wave cooling

Infra-red cooling in the 15μm CO_2 band and 9.6μm O_3 band was calculated using the relationship

$$\frac{dT}{dt} = k_1(z,X)\,B(T) + k_2(z,X)$$

where k_1 and k_2 are functions of height and emitting constituent (Harwood and Pyle, 1975). Table A3.2 below gives the values of k_1 and k_2 for these emission bands at levels 8.25 to 3.75 scale heights. The units of k_1 are $K\,day^{-1}(W\,m^{-2}\,(cm^{-1})^{-1})^{-1}$ and the units of k_2 are $K\,day^{-1}$.

Table A3.2 (see text above)

Height	CO_2 (15μm)		O_3 (9.6μm)	
	k_1	k_2	k_1	k_2
8.25	-58.64	-1.840		
7.75	-23.98	3.158	-7.02	-0.011
7.25	-24.05	1.238	-12.46	0.129
6.75	-26.58	1.946	-13.52	0.098
6.25	-21.64	0.863	-16.31	-0.076
5.75	-18.63	0.409	-20.22	0.046
5.25	-18.75	0.576	-22.30	0.224
4.75	-14.42	0.116	-13.10	0.316
4.25	-12.89	0.117	-6.93	0.504
3.75	-11.96	0.084	-1.13	0.462

Figure A3.2 shows height - longitude sections at 60°N of the infra - red cooling and the net heating due to both solar heating (Figure A3.1) and infra - red cooling on 28.1.73. Note the very large cooling rates in the high temperature region between 0 and 90°E and the net warming in the relatively cold upper stratosphere at western longitudes.

component of the eddy pressure work term $[\omega'\Phi']$ effects changes in eddy kinetic energy but is balanced in the steady state by an eddy potential energy flux its convergence must be included in the eddy kinetic energy budget; and (b) the horizontal component of the eddy pressure-work term $[v'\Phi']$ is balanced by an eddy kinetic energy flux term and thus their net effect on the eddy kinetic energy budget is zero. Therefore in the following discussions $[\omega'\Phi']$ is the only eddy energy flux considered.

8.2. ASSOCIATION BETWEEN VERTICAL ENERGY FLUX AND EDDY KINETIC ENERGY

The response of the upper stratosphere to convergence of the vertical eddy energy flux is shown clearly in the latitude-time section of the eddy kinetic energy between 2 and 0.4 mb in Figure 8.2. In middle and high latitudes three periods of high eddy kinetic energy corresponded well with the three 'energy pulses' around 12, 19 and 26.1.73 shown in Figure 8.1. Divergence of the vertical energy flux during the declining phase of the warming accompanied substantial loss of eddy kinetic energy. At lower levels correlations were less marked owing to the greater relative importance of other processes affecting the eddy kinetic energy.

An interesting feature of Figure 8.1 is that during the period of energy convergence above 10 mb around 19.1.73 energy was being lost by vertical flux divergence between 50 and 10 mb. The way in which this occurred is illustrated in Figure 8.3 in which are plotted daily values of both the total eddy and wavenumber one area-averaged vertical energy fluxes through the 50, 27, 10 and 2 mb surfaces. It is clear that wavenumber one was responsible for the upward energy propagation. The wavenumber

Radiative heating and vertical motion

It was pointed out in Chapters 6 and 8 that anomalously strong poleward mean meridional velocities were calculated for low and middle latitudes near 3 mb and 0.6 mb and that poor specification of the radiative heating was probably the cause. Figure A3.3 shows height - latitude sections of (a) zonal mean solar heating, (b) zonal mean infra-red heating and (c) the zonal mean net heating on 28.1.73. It is seen that very strong gradients of the solar heating, and hence of the net radiative heating, occurred in low latitudes at about 3 mb (just below 6 scale heights). This effected a large gradient of the vertical motion in this region which consequently produced a large gradient of the mean meridional velocity, as calculated by equation 5.9, with strong poleward motion in low latitudes.

A close inspection of Figures 6.10b and A3.3 shows the dependence of the sign of the meridional motion on the sign of the vertical gradient of the net radiative heating in low latitudes at all levels. The heat budget of low latitudes is characterised by a near balance of the radiative heating and cooling due to vertical motions. The mean meridional circulations at low latitudes calculated in this study were thus largely determined by the radiation scheme.

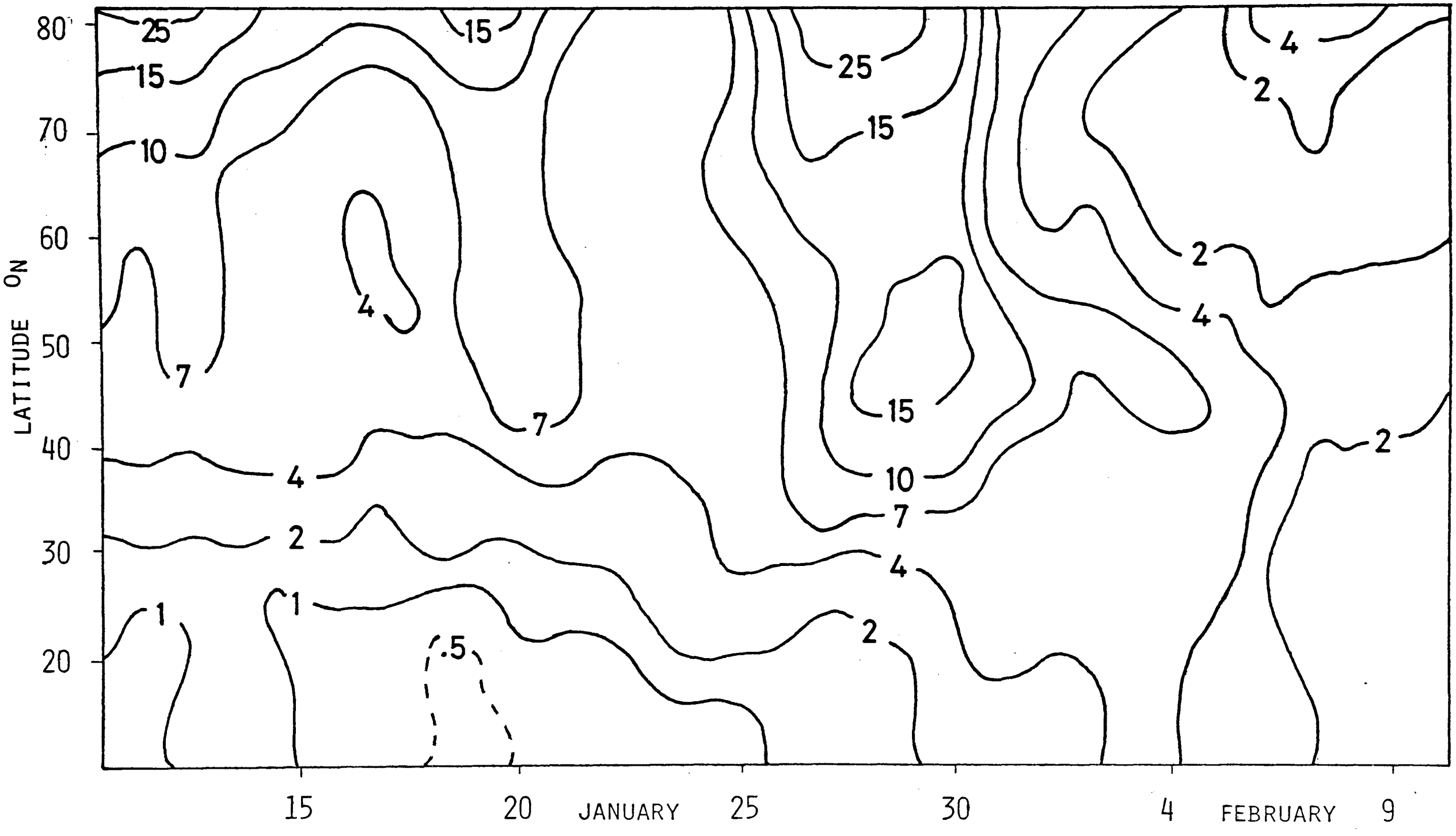


Figure 8.2. Latitude - time section of mean eddy kinetic energy content of the 2 - 0.4 mb layer for the period 11.1.73 to 10.2.73 ($J\,m^{-2}\,mb^{-1} \times 10^3$)

REFERENCES

Abel, P.G., Ellis, P.J., Houghton, J.T., Peckham, G., Rodgers, C.D., Smith, S.D. and Williamson, E.J. (1970). "Remote sounding of atmospheric temperature from satellites: II. The selective chopper radiometer for Nimbus D". Proceedings of the Royal Society, A320, 35-55.

Andrews, D.G. and McIntyre, M.E. (1976). "Planetary waves in horizontal and vertical shear: the generalized Eliassen-Palm relation and the mean zonal acceleration". Journal of the Atmospheric Sciences, 33, 2031-2053.

Arpe, K. (1976). "Energetics of the stratosphere, a comparison between one winter without and two winters with major warmings". Beiträge zur Physik der Atmosphäre, 49, 189-211.

Austen, M.D., Barnett, J.J., Curtis, P.D., Houghton, J.T., Morgan, C.G., Rodgers, C.D. and Williamson, E.J. (1976). "Satellite temperature measurements in the 40-90 km region by the pressure modulator radiometer". Presented at the 19th COSPAR Meeting, Philadelphia, U.S.A.

Barnett, J.J. (1973). "Remote sounding of the atmosphere". D. Phil. Thesis: Oxford University.

Barnett, J.J. (1975). "Large sudden warming in the southern hemisphere". Nature, 255, 387-389.

Barnett, J.J. (1975b). "Hemispheric coupling - evidence of a cross-equatorial planetary wave-guide in the stratosphere". Quarterly Journal of the Royal Meteorological Society, 101, 835-845.

Barnett, J.J. (1976). "Smoothing of gridded data". Unpublished departmental document, Atmospheric Physics Department, Oxford University.

Barnett, J.J., Harwood, R.S., Houghton, J.T., Morgan, C.G., Rodgers, C.D. and Williamson, E.J. (1975). "Comparison between radiosonde, rocketsonde and satellite observations of atmospheric temperatures". Quarterly Journal of the Royal Meteorological Society, 101, 423-436.

Barnett, J.J., Houghton, J.T., Morgan, C.G., Pick, D.R., Rodgers, C.D., Williamson, E.J., Cross, M.J., Flower, D., Peckham, G. and Smith, S.D. (1973). "Stratospheric observations from Nimbus 5". Nature, 245, 141-143.

Barnett, J.J., Houghton, J.T. and Pyle, J.A. (1975b). "The temperature dependence of the ozone concentration near the stratopause". Quarterly Journal of the Royal Meteorological Society, 101, 245-257.

Chapman, S. (1930). "A theory of upper atmospheric ozone". Quarterly Journal of the Royal Meteorological Society, 3, 103-125.

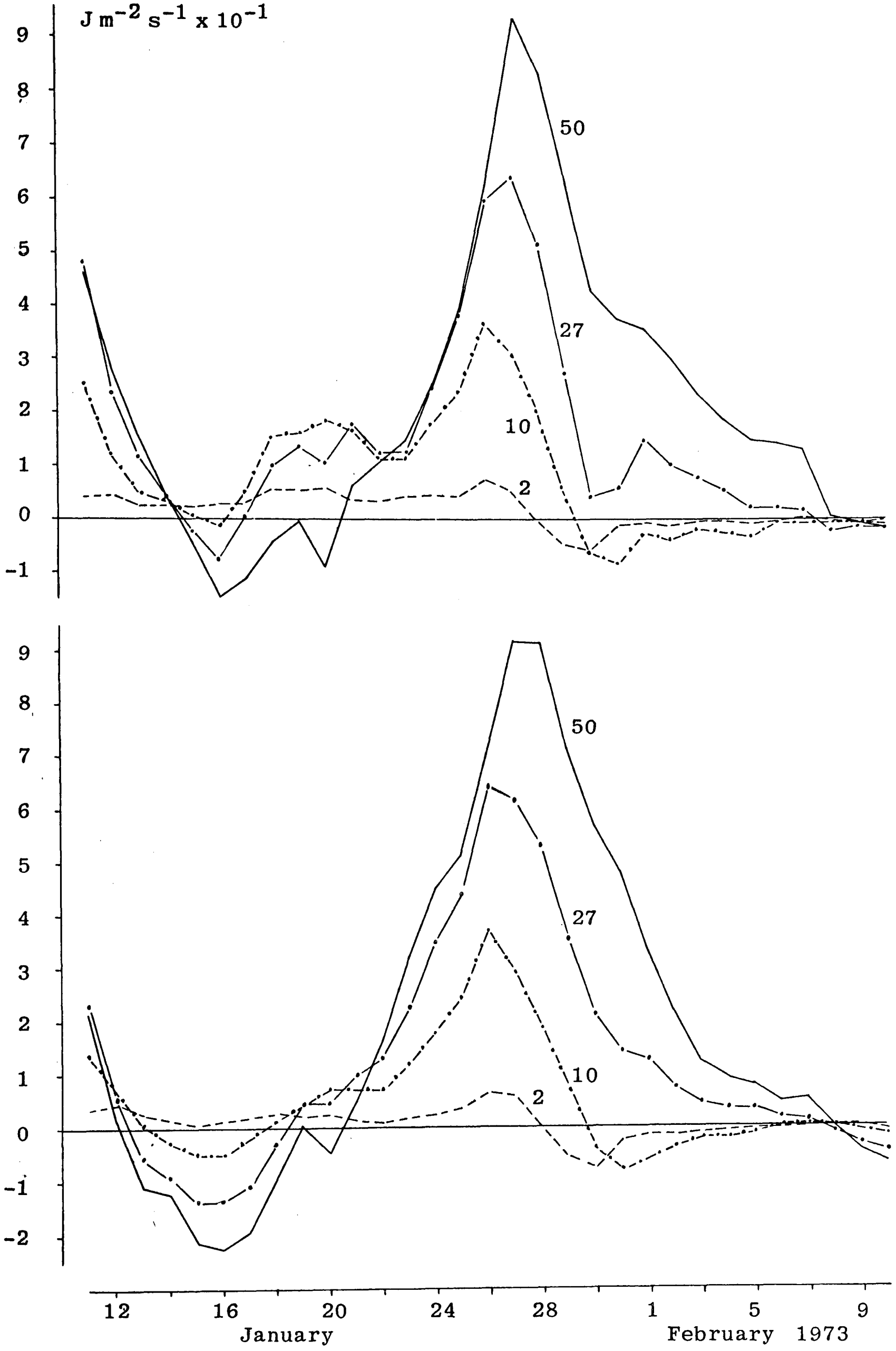


Figure 8.3. Total eddy (upper) and wavenumber 1 (lower) area-averaged vertical energy flux ($\overline{-\omega'\Phi'}$) ($Jm^{-2}s^{-1} \times 10^{-1}$) at 50, 27, 10 and 2 mb plotted against time for the period 11.1.73 to 10.2.73

Geisler, J.E. and Dickinson, R.E. (1974). "Numerical study of an interacting Rossby wave and barotropic zonal flow". Journal of the Atmospheric Sciences, 31, 946-955.

Groves, G.V. (1972). "Atmospheric structure and its variations in the region from 25 to 120 km". COSPAR International Reference Atmosphere 1972, Part 2, Akademie-Verlag, Berlin.

Hartmann, D.L. (1976). "The dynamical climatology of the stratosphere in the southern hemisphere during late winter 1973". Journal of the Atmospheric Sciences, 33, 1789-1802.

Harwood, R.S. (1975). "The temperature structure of the southern hemisphere stratosphere August - October 1971". Quarterly Journal of the Royal Meteorological Society, 101, 75-91.

Harwood, R.S. and Pyle, J.A. (1975). "A two-dimensional mean circulation model for the atmosphere below 80 km". Quarterly Journal of the Royal Meteorological Society, 101, 723-747.

Hirota, I. (1967). "The vertical structure of the stratospheric sudden warming". Journal of the Meteorological Society of Japan, 45, 422-435.

Hirota, I. (1968). "Planetary waves in the upper stratosphere in early 1966". Journal of the Meteorological Society of Japan, 46, 418-430.

Hirota, I. (1971). "Excitation of planetary Rossby waves in the winter stratosphere by periodic forcing". Journal of the Meteorological Society of Japan, 49, 439-448.

Hirota, I. (1975). "Spectral analysis of planetary waves in the summer stratosphere and mesosphere". Journal of the Meteorological Society of Japan, 53, 33-44.

Hirota, I. and Barnett, J.J. (1977). "Planetary waves in the winter mesosphere - preliminary analysis of Nimbus 6 PMR results". Quarterly Journal of the Royal Meteorological Society, 103, 487-498.

Hirota, I. and Sato, Y. (1969). "Periodic variation of the winter stratospheric circulation and intermittent vertical propagation of planetary waves". Journal of the Meteorological Society of Japan, 47, 390-402.

Holton, J.R. (1974). "Forcing of mean flows by stationary waves". Journal of the Atmospheric Sciences, 31, 942-945.

Holton, J.R. (1975). "The dynamic meteorology of the stratosphere and mesosphere". Meteorological Monographs, 15, No. 37, American Meteorological Society, Boston, Mass., U.S.A.

Holton, J.R. (1976). "A semi-spectral numerical model for wave - mean flow interactions in the stratosphere: application to sudden stratospheric warmings". Journal of the Atmospheric Sciences, 33, 1639-1649.

two component was predominantly downwards at 50 mb during the warming with maximum downward fluxes occurring on 24 and 30.1.73. Whilst during the energy pulse of 11.1.73 and during the warming there was convergence of vertical energy flux throughout the stratosphere, in the intervening period the total flux was downward at 50 mb and upward at 10 mb. There must, therefore, have been an energy source within the 50 - 10 mb region from which energy was propagated to higher levels. The limited accuracy of the components of the kinetic energy budget prevents us from deducing which energy conversion process contributed most to the amplification of the vertical energy flux in this layer; there was weak baroclinic conversion from A_E to K_E at this time but a marked reversal of the $K_E - K_Z$ conversion occurred on 20.1.73 suggesting that barotropic conversion may have been important also.

Latitudinal distributions of vertical energy flux have shown that large fluxes are confined mainly to high latitudes (see, for example, Miyakoda et al, 1970; Labitzke et al, 1975; Quiroz and Nagatani, 1975; Klinker, 1976). The latitude-time section of the flux through 50 mb in Figure 8.4 confirms that this was the case in the 1973 warming. In view of this observation it is more useful to study the behaviour of the vertical energy flux in high latitudes rather than on a hemispheric basis. The time change of the vertical energy flux at $76^{o}N$ for various levels is shown in Figure 8.5. What is significant is that there was energy flux divergence between 50 and 10 mb not only during the period 14 - 22.1.73 indicated by the area-averaged results but throughout the period 11 - 26.1.73. At the same time there was convergence between 10 and 0.4 mb on all days except 15 - 17.1.73. At this latitude increased vertical energy flux preceding the warming

Madden, R.A. (1975). "Oscillations in the winter stratosphere: Part 2. The role of horizontal eddy heat transport and the interaction of transient and stationary planetary-scale waves". Monthly Weather Review, 103, 717-729.

Mahlman, J.D. (1969). "Energetics of a "minor breakdown" of the stratospheric polar night vortex". Journal of the Atmospheric Sciences, 26, 1306-1317.

Matsuno, T. (1970). "Vertical propagation of stationary planetary waves in the winter northern hemisphere". Journal of the Atmospheric Sciences, 27, 871-883.

Matsuno, T. (1971). "A dynamical model of the stratospheric sudden warming". Journal of the Atmospheric Sciences, 28, 1479-1494.

Miller, A.J. (1970). "The transfer of kinetic energy from the troposphere to the stratosphere". Journal of the Atmospheric Sciences, 27, 388-393.

Miller, A.J. (1974). "Periodic variation of atmospheric circulation at 14-16 days". Journal of the Atmospheric Sciences, 31, 720-726.

Miller, A.J., Brown, J.A. and Campana, K.A. (1972). "A study of the energetics of an upper stratospheric warming (1969-1970)". Quarterly Journal of the Royal Meteorological Society, 98, 730-744.

Miller, A.J. and Johnson, K.W. (1970). "On the interaction between the stratosphere and troposphere during the warming of December 1967 - January 1968". Quarterly Journal of the Royal Meteorological Society, 96, 24-31.

Miyakoda, K., Strickler, R.F. and Hembree, G.D. (1970). "Numerical simulation of the breakdown of a polar-night vortex in the stratosphere". Journal of the Atmospheric Sciences, 27, 139-154.

Muench, H.S. (1965). "On the dynamics of the wintertime stratospheric circulation". Journal of the Atmospheric Sciences, 22, 349-360.

Muench, H.S. (1968). "Large-scale disturbances in the summertime stratosphere". Journal of the Atmospheric Sciences, 25, 1108-1115.

Murgatroyd, R.J. (1970). "The structure and dynamics of the atmosphere". In 'The global circulation of the atmosphere', ed. G. A. Corby, pp. 159-195. Royal Meteorological Society, London.

Newell, R.E. (1964). "The circulation of the upper atmosphere". Scientific American, 210, 62-74.

Newell, R.E. Vincent, D.G., Dopplick, T.G., Ferruzza, D. and Kidson, J.W. (1970). "The energy balance of the global atmosphere". In 'The global circulation of the atmosphere', ed. G.A. Corby. Royal Meteorological Society, London.

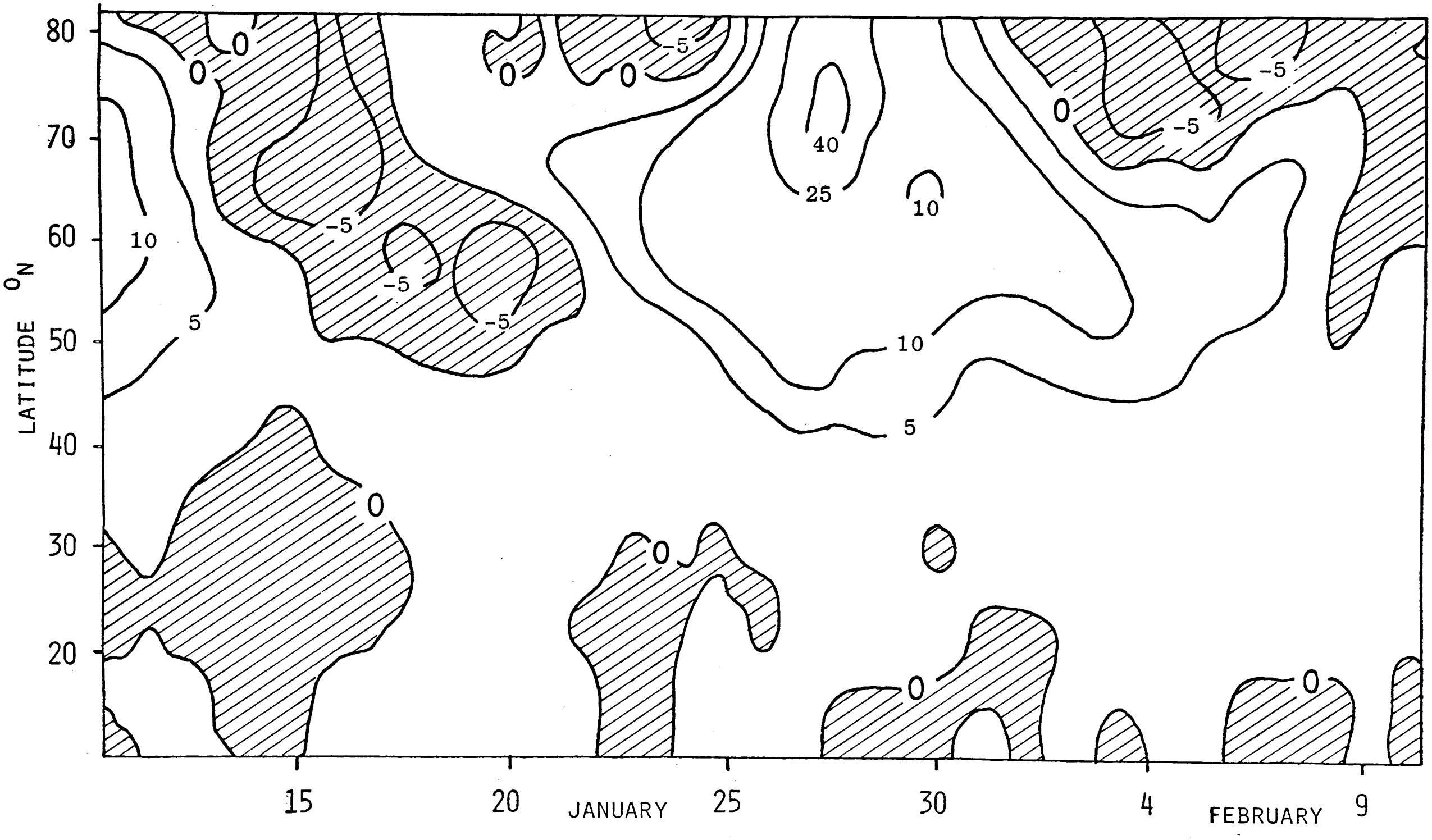


Figure 8.4. Latitude - time section of the vertical eddy energy flux [$-\omega'\Phi'$] at 50 mb for the period 11.1.73 to 10.2.73 ($J\,m^{-2}\,s^{-1} \times 10^{-1}$)

Scherhag, R. (1952). "Die explosionsartigen Stratosphärenerwarmungen des Spätwinters 1951-52". Bericht des Deutschen Wetterdienstes, 6, 51-63.

Simmons, A.J. (1974). "Planetary scale disturbances in the polar winter stratosphere". Quarterly Journal of the Royal Meteorological Society, 100, 76-108.

Smith, P.J. (1969). "On the contribution of a limited region to the global energy budget". Tellus, 21, 202-207.

Strand, O.N. and Westwater, E.R. (1968). "Statistical estimation of the numerical solution of a Fredholm integral equation of the first kind". Journal of the Association for Computing Machinery, 15, 100-114.

Trenberth, K.E. (1973). "Dynamic coupling of the stratosphere with the troposphere and sudden stratospheric warmings". Monthly Weather Review, 101, 306-322.

van Loon, H., Jenne, R.L. and Labitzke, K. (1973). "Zonal harmonic standing waves". Journal of Geophysical Research, 78, 4463-4471.

van Loon, H., Labitzke, K. and Jenne, R.L. (1972). "Half-yearly wave in the stratosphere". Journal of Geophysical Research, 77, 3846-3855.

van Mieghem, J. (1973). "Atmospheric Energetics". Oxford Monographs on Meteorology, Clarendon Press, Oxford.

Vincent, D.G. (1968). "Mean meridional circulations in the northern hemisphere lower stratosphere during 1964 and 1965". Quarterly Journal of the Royal Meteorological Society, 94, 333-349.

Fritz, S., Wark, D.Q., Fleming, H.E., Smith, W.L., Jacobowitz, H., Hilleary, D.T. and Alishouse, J.C. (1972). "Temperature sounding from satellites". NOAA Technical Report NESS 59, Washington D.C.