Tellus (1987), *39A*, 376–384

Baroclinic instability as a mechanism for the serial development of polar lows: a case study

By RICHARD J. REED[1], *European Centre for Medium Range Weather Forecasts, Shinfield Park, Berkshire RG2 9AX, England* and CHARLES N. DUNCAN, *Department of Meteorology, The University, James Clerk Maxwell Buildings, King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, Scotland*

(Manuscript received 18 June; in final form 18 November 1986)

ABSTRACT

Satellite infrared images of a series or train of four polar lows are presented, and the synoptic environment of the lows is documented with the help of operational analyses prepared by the European Centre for Medium Range Weather Forecasts (ECMWF). The analyses reveal that the disturbances formed in a shallow baroclinic zone of small static stability close to the region of maximum 1000–500 mb thermal wind and that the flow was of the reversed shear type.

The observed characteristics of the disturbances were in partial agreement with those obtained from a linear, quasi-geostrophic, dry baroclinic model. The wavelength of maximum instability in the model of approximately 500 km was in satisfactory agreement with the observed wavelength. The computed phase speed was somewhat larger than the observed speed, a result attributed to the effect of organized deep convection in retarding the actual motion. The large growth rate of the 500 km wave was qualitatively consistent with the rapid growth of the systems as seen on the satellite images. It is argued, however, that baroclinic instabilty alone was not capable of accounting for the rapidity of the developments and that some other mechanism, presumably latent heat release in organized deep convection, also played an important part in the growth.

1. Introduction

The small synoptic or subsynoptic-scale cyclones known as polar lows are a phenomenon of common occurrence during the winter half year in the broad oceanic region bounded by Norway, Spitsbergen, Greenland, Iceland and the British Isles (Rasmussen, 1983). Satellite observations have been particularly helpful in identifying these systems and in revealing the variety of forms that they take (Forbes and Lottes, 1985). Numerous cases exist in which the systems appear multiply or in series (e.g., Rasmussen, 1985). An outstanding example of serial behavior took place during the period 29-31 January 1983 when four vortical cloud patterns or disturbances formed sequentially over or near the Norwegian Sea and moved southwestward towards Iceland while maintaining a remarkably uniform separation or wave length (Fig. 1). It is a matter of considerable interest to document the large-scale environment within which such a train of polar lows developed and to identify the mechanism responsible for the regular behavior.

One possible explanation for the regular behavior is that the disturbances were initiated by baroclinic instability and that the observed scale represents the wave length of maximum instability for the given basic state. Several authors have suggested baroclinic instability as a major mechanism for polar low formation (Harley, 1960; Harrold and Browning, 1969; Mansfield, 1974; Duncan, 1977; Reed, 1979). The purpose of the present paper is twofold: first, to document

[1] On leave from Department of Atmospheric Sciences, University of Washington, Seattle, WA, 98195, USA.

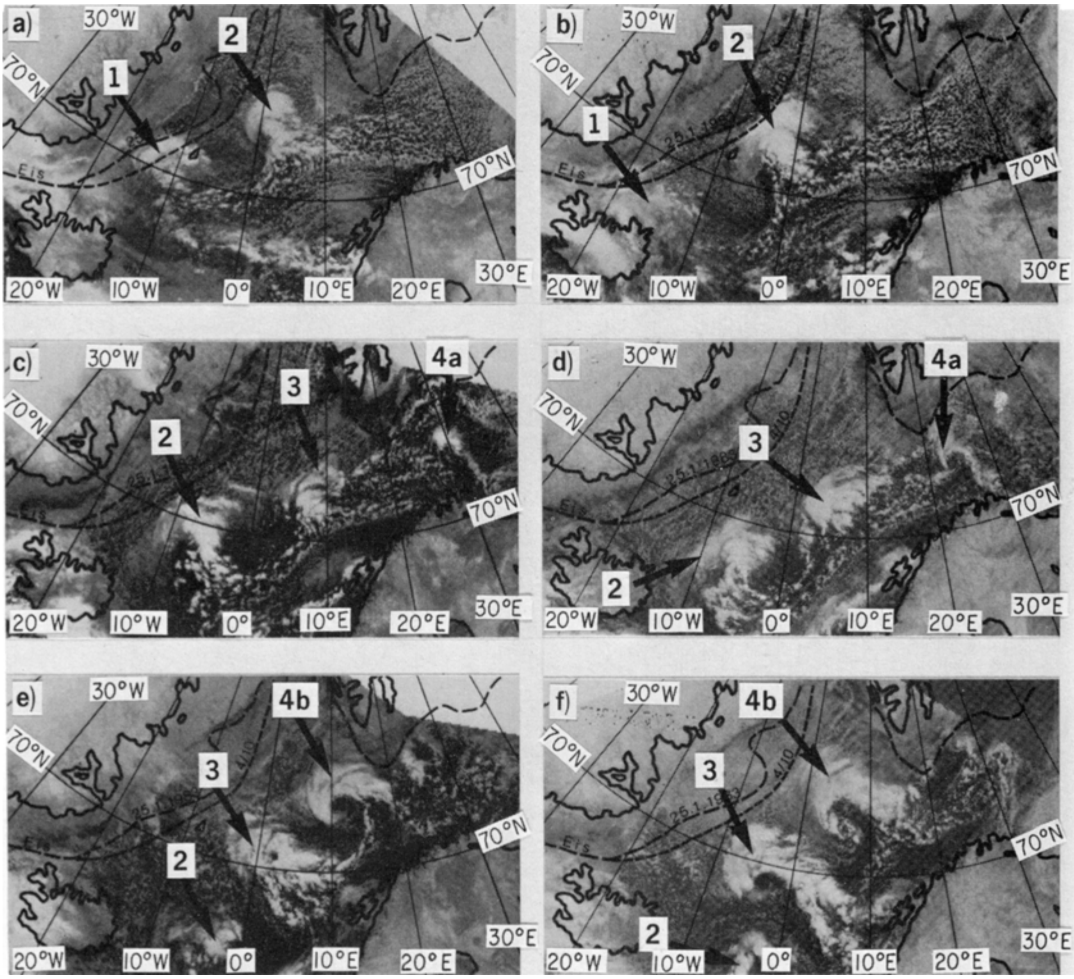

Fig. 1. NOAA-7 satellite infrared images. (a) 01.38–05.10 GMT, (b) 11.30–15.04 GMT, 29 January 1986; (c) 01.26–04.57 GMT, (d) 11.18–15.43 GMT, 30 January 1983 and (e) 01.13–04.45 GMT, (f) 11.07–14.39 GMT, 31 January 1983. Arrows indicate positions of polar lows 1–4. Dashed lines indicate positions of ice edge on 25 January 1983 and on 10 April. (Courtesy of M. Eckardt, Institute for Meteorology, Free University of Berlin.)

the large-scale wind and thermal structures attending the formation of the wave train and, second, to determine how well the wave characteristics fit those obtained from solutions of a linear, quasi-geostrophic baroclinic model that has been applied previously by Duncan (1977, 1978) to the problem of polar low development.

2. Observed wave characteristics

Four waves or vortices, labelled 1 to 4 in Fig. 1, occurred during the 3-day period. Wave 4 displayed an anomalously rapid movement between midday of 30 January (frame d) and early morning of 31 January (frame e). We attribute this apparent rapid movement to the development of a new disturbance (labelled 4b in frames e and f) ahead of the original disturbance (labelled 4a in frames c and d). The cusp-shaped field of enhanced convection which lies between waves 3 and 4a in frame d is believed to be the seat of the new formation.

From the pictures, five determinations of wavelength were made and six determinations of wave speed. The following waves and frames

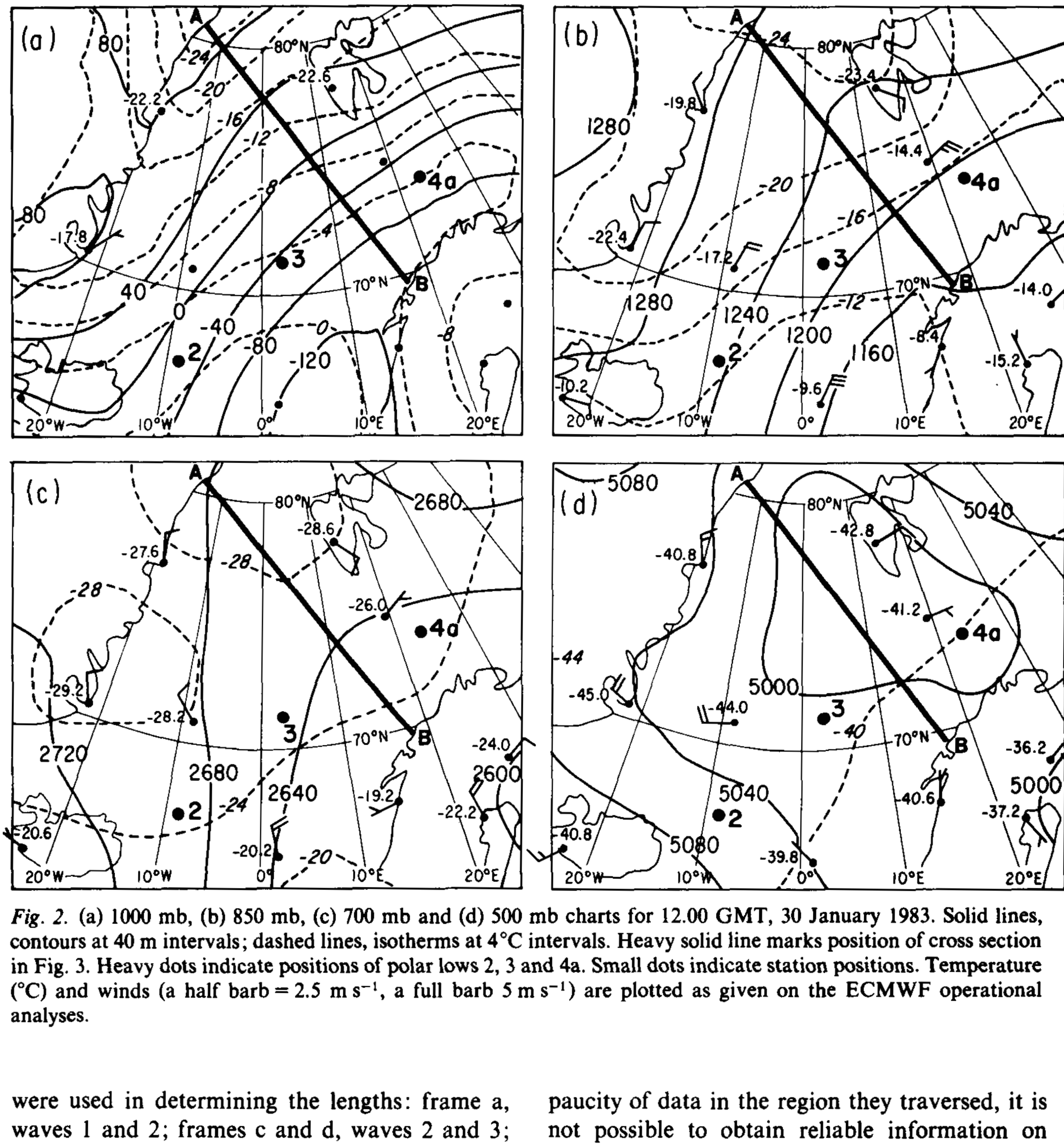

Fig. 2. (a) 1000 mb, (b) 850 mb, (c) 700 mb and (d) 500 mb charts for 12.00 GMT, 30 January 1983. Solid lines, contours at 40 m intervals; dashed lines, isotherms at 4°C intervals. Heavy solid line marks position of cross section in Fig. 3. Heavy dots indicate positions of polar lows 2, 3 and 4a. Small dots indicate station positions. Temperature (°C) and winds (a half barb = 2.5 m s^{-1}, a full barb 5 m s^{-1}) are plotted as given on the ECMWF operational analyses.

were used in determining the lengths: frame a, waves 1 and 2; frames c and d, waves 2 and 3; frames e and f, waves 3 and 4b. The following waves and pairs of frames were employed in determining the speeds: wave 2, frames a and c and frames b and d; wave 3, frames c and e and frames d and f; wave 4a, frames c and d; wave 4b, frames e and f. The mean and standard deviation of wavelength determined from the foregoing information were 570 km and 58 km, respectively. The mean and standard deviation of wave speed were 4.9 m s^{-1} and 1.0 m s^{-1}, respectively.

Because of the small size of the systems and the paucity of data in the region they traversed, it is not possible to obtain reliable information on their structures and growth rates, as desirable as such information would be for determining whether they fully fit the description of baroclinic waves. It should be noted, however, that the growth rates as deduced from cloud patterns, were large. Disturbance 3 grew from a barely detectable perturbation in the cloud field (frame b) to a fully mature system characterized by a vortical appearing cloud mass with eye-like center (frame e) in about 36 h; disturbance 4b, if our interpretation of its history is correct, experienced a similar evolution in only 24 h.

A crude quantitative estimate of the growth rates can be made from the 850 mb wind perturbation observed at Jan Mayen (71°N, 9°W) at 00.00 GMT 30 January when wave 2 was situated close to the station (Fig. 1, frame c). The observed 850 mb wind was 080° 25 m s^{-1}; the wind given by the corresponding large-scale or smoothed 850 mb contour and isotach analyses (not shown) of the European Centre for Medium Range Weather Forecasts (ECMWF) was 020° 12 m s^{-1}. (Further mention of ECMWF analyses appears in the next section.) The magnitude of the vector difference, 22 m s^{-1}, provides a rough measure of the amplitude of the wind perturbation at the given time and level and suggests that 20 m s^{-1} is a conservative estimate of the amplitude of the wind perturbation of this and the companion disturbances when fully developed. Amplitudes of this size, or larger, have been observed in other cases (Rasmussen, 1985; Shapiro et al., 1987). In the case of disturbance 4b, which appeared to grow from a slight perturbation to a fully developed system in 24 h, an assumed 20 m s^{-1} amplitude yields growth rates (e-folding) in the range 1–3 d^{-1} for initial wind perturbations in the range 1–7 m s^{-1}. For disturbances that required 36 h to develop an upper limit of 2 d^{-1} seems more appropriate for the growth rate.

3. The large-scale wind and thermal fields

Figs. 2a–d depict fields of geopotential and temperature at 1000, 850, 700 and 500 mb, respectively, for 12.00 GMT 30 January. The fields were slow varying throughout the period so that Fig. 2 is representative of the period as a whole. The analyses shown on the figures are the routine operational analyses prepared at ECMWF by the N48 model that was in use at the time. Descriptions of the analysis system can be found in Arpe et al. (1986) and Lönnberg and Hollingsworth (1986). Neither at the hour in question nor at other hours were the analyses able to resolve the polar lows, but for present purposes this inability is not a disadvantage, since our aim here is to establish the large-scale or smoothed fields, not to examine individual disturbances. The inability to represent the lows stemmed both from deficiencies in the observational network and from a lack of sufficient resolving power in the analysis system. The positions of the disturbances (2, 3 and 4a) based on satellite identification are marked on the charts.

From Fig. 2a (and a cross section to be shown later) it is apparent that the lows formed and travelled within the region of strongest surface winds (contour gradient) and within or near the warm edge of the region of strongest surface thermal gradient. Two of the three systems (3 and 4a) and the forming system 4b are located in or near a region of weak geopotential height gradient and wind at 500 mb. The third system (2) lies beneath a region of stronger 500 mb winds that are directed nearly perpendicular to the 1000 mb winds. It is possible that the pronounced turning of wind with height was a factor in the subsequent dissipation.

Comparison of Figs. 2a–d reveals that the strongest thermal gradient occurred near the surface (1000 mb) and that the gradient diminished monotonically with height becoming nearly zero at 500 mb. Because the surface flow and thermal wind were oppositely directed, the basic current also diminished with height. The maximum vertical shear (geostrophic) of the basic current of necessity took place near the surface where the thermal gradient was largest.

In view of the flatness of the contour gradient at 500 mb, the 1000–500 mb thickness field in the region surrounding 72°N and 20°E, where the main wave amplification occurred, is well represented by the 1000 mb contours. Making use of this fact, we note that the disturbances formed within the belt of strongest thickness gradient or thermal wind. Moreover their west-southwestward movement, in conformity with the low-level wind flow, was in the direction opposite to the thermal wind. Thus the case in question provides an example of what Duncan (1978) has termed a reversed shear flow, i.e., a flow in which the direction of wave propagation is opposite to the thermal wind direction. The study of Forbes and Lottes (1985) reveals that the reversed shear case is not unusual. Indeed nearly 80% of the polar vortices classified by these authors as "developing" vortices, systems shown by their data to represent the strongest type of polar vortex, were characterized by reversed shear flow.

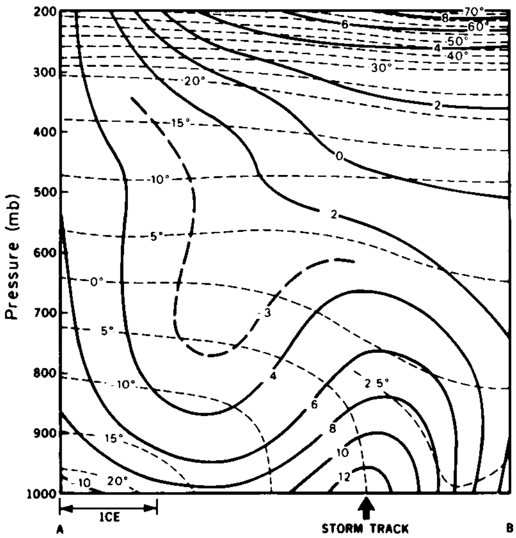

Fig. 3. Cross section along line AB in Fig. 2. Solid lines are isotachs (m s^{-1}) of the geostrophic wind component perpendicular to the section. Dashed lines are potential isotherms (°C). The extent of the sea ice and positions of the storm track are indicated at bottom.

The propagation of the disturbances in the direction opposite to the shear is explained by the fact that short baroclinic waves move approximately with the mean wind in the layer in which they are embedded. In the case of the typical mid-latitude baroclinic disturbances the shear and mean wind are in the same direction. In the case of polar lows of the type under investigation, the shear and mean wind are in opposite directions.

A feature of the low-level flow pattern that is not so evident in Fig. 2a as in the corresponding charts for earlier hours is the confluence of the flow (contours) in the direction normal to the isotherms and to the direction of travel of the waves. Thus the wave growth took place in a region where the thermal gradient was being enhanced by the geostrophic deformation. The enhanced thermal gradient that is evident in the area to the northeast of Iceland (70° N, 10° W), is at least in part a consequence of this frontogenetical effect. Another important contributor to the enhanced thermal gradient in the region of wave growth is the differential diabatic heating in the cross-flow direction. Air parcels to the north of the storm track followed trajectories that were entirely or mainly over the ice. Air parcels to the south had long histories over the open water of the Barents Sea and were warmed by the upward flux of sensible heat.

A final picture of the thermal and wind structure appears in a vertical cross section (Fig. 3) taken along the line AB in Figs. 2a and b. Solid lines depict the geostrophic wind component normal to the section, dashed lines potential temperature isotherms. The analysis is based on temperatures and geopotential heights extracted from the 1000, 850, 700 and 500 mb maps shown in Fig. 2 and from 300 mb and 200 mb maps appearing in the European Meteorological Bulletin published by the Deutschen Wetterdienstes. Temperatures and heights were read at intervals of 150 km; geostrophic winds were computed at the same intervals with use of centered (300 km) differences.

A shallow layer of large static stability, horizontal temperature gradient and vertical wind shear is seen on the poleward side of the ice edge. This shallow feature is no doubt much stronger in reality than shown on the smoothed ECMWF analyses. The large temperature gradient and wind shear extend seaward in an increasingly deep layer of much lesser static stability. The changing structure reflects the effect of the greater modification that the air on the right-hand side of the section has undergone than the air on the left-hand side. Note, however, that the airflow was primarily perpendicular to the plane of the cross section so that it is not strictly correct to view the air mass on the right-hand side as being simply a modified version of that on the left-hand side. A fairly abrupt edge to the low-level baroclinic zone exists beyond (to the southeast of) the position of the storm track. A rapid drop in wind speed occurs in the same region, especially near the surface, placing the storm track close to the core of a low level jet (and the point of maximum tropospheric thermal wind). Above 500 mb a moderate to large static stability prevails everywhere. A second region of enhanced baroclinicity is found in the lower stratosphere where it is unlikely to affect surface development.

4. Model solutions

A quasi-geostrophic baroclinic model (Duncan, 1977) was used to investigate the structure and

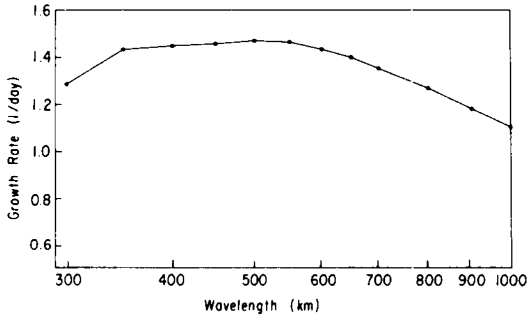

Fig. 4. Growth rate as a function of wavelength for the experiment with no horizontal shear and vertical wind profile observed at the storm track (see Fig. 3).

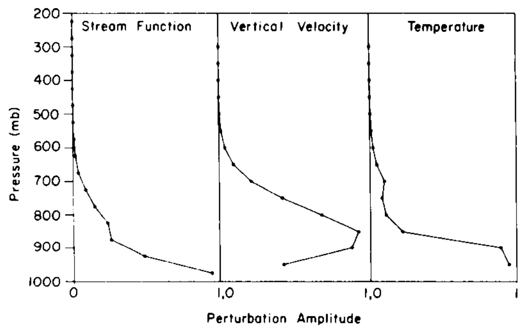

Fig. 5. Perturbation amplitudes for the streamfunction, vertical velocity and temperature, for the 500 km wave of Fig. 4. The magnitudes are in normalized arbitrary units.

development of baroclinic waves in the flow shown in Fig. 3. As there are considerable variations in static stability across this flow, the air being quite stable over the ice and near the ice edge and becoming less stable with distance from the ice edge, it is not possible within the quasi-geostrophic assumption to comprehensively model the development. In order to study the most basic properties of the development the wind field was chosen to have the same vertical profile as that observed along the storm track (see Fig. 3) but no horizontal shear was included. The static stability used was also derived from analyzed temperatures in the vicinity of the storm track. Near the surface, where the lapse rate is close to dry adiabatic, the likely uncertainty of the static stability is of the order of 50 percent. Details of the model parameters are given in the appendix.

Solutions were obtained for the fastest growing mode for several wavelengths in the range 1000 km to 300 km. The growth rates of these modes are shown in Fig. 4 which illustrates the wavelengths of about 500 km are the most unstable. At 500 km wavelength the waves move with a phase speed of 8.7 m s^{-1}. This speed is rather greater than the observed 4.9 m s^{-1} phase speed. While it is not possible to make a direct comparison of the growth rate with observations, the maximum growth rate of 1.47 day^{-1} is consistent with the observed rapid development of these polar lows. The effect of doubling or halving the low-level static stabilities would be an uncertainty of ± 100 km in wave length and $\pm 25\%$ in growth rate. A more important factor in determining these parameters is the depth of the disturbances. The depth is limited by the upper level static stabilities which are known with a high degree of confidence.

The slower movement in nature than in the model may be an indication that the motion of the lows was retarded by the associated organized deep convection. From the infrared imagery (Fig. 1) it is apparent that the convection penetrated the slower moving upper layers during much of the lifetimes of the systems. The model disturbances, on the other hand, were confined almost entirely to the faster moving lower layers, as shown by the amplitude profiles in Fig. 5. It is also possible that the disturbances developed more rapidly than can be accounted for by baroclinic instability alone. The model solutions neglect friction and lateral shear and use the maximum baroclinicity shown on Fig. 3. All of these factors cause the baroclinic growth rate to be overestimated. It seems likely that latent heat release in cumulus clouds is required in order to explain fully the very rapid growth seen in the satellite pictures. An important role for convective heating or conditional instability of the second kind (CISK) in polar low development has been suggested by Økland (1977) and Rasmussen (1979) and verified in model experiments by Sardie and Warner (1983, 1985).

Examination of the energetics of the disturbances reveals that the conversion of eddy available potential energy to eddy kinetic energy occurs entirely within the 1000 to 800 mb layer. A small eddy kinetic energy flux divergence results in some evidence of the perturbation above the

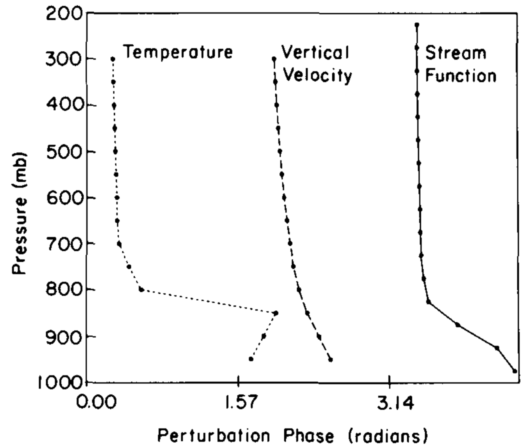

Fig. 6. Perturbation phases corresponding to the amplitudes in Fig. 5. Solid line, streamfunction; short dashed line, temperature; long dashed line, vertical velocity.

800 mb level but its amplitude is almost insignificant above 700 mb, as is seen in Fig. 5. The phase relationship between the streamfunction and vertical velocity perturbations is particularly interesting in the reversed shear case since the upward motion is located behind the trough. However, the warm part of the thermal perturbation is also behind the trough as a result of the thermal wind direction being directed opposite to the direction of motion, and the usual baroclinic signature with warm air rising and cold air descending is evident (Fig. 6).

Experiments carried out to determine what effect horizontal shear had on the development show the inadequacy of linear, quasi-geostrophic models in cases such as this where horizontal gradients in static stability are also present. The effect of horizontal shear on baroclinic waves is to reduce their energy conversion efficiency, resulting in a reduced growth rate in regions of strong horizontal shear. When the model was used with the realistic velocity field shown in Fig. 3 the disturbance developed most strongly in the area of weak horizontal wind shear just off the ice edge. The observed polar lows were unable to develop in this area because the static stabilty was much greater than in the model.

An experiment to include horizontal shear was performed by restricting the width of the channel to 525 km centred on the observed storm track. This was the region in which the static stability profile given in the appendix was valid. The results of this experiment were not very different from that in which no horizontal shear was included except that growth rates were somewhat reduced (0.92 day^{-1} for the 500 km wave compared with 1.47 day^{-1} in the no-shear case). The structure of the wave in this experiment reveals the reduced development in regions of large horizontal shear. Fig. 7 shows the stream function perturbation amplitude for the two cases with and without horizontal shear. Another feature of the case with horizontal shear is that the vertical extent of the perturbation is somewhat less. This highlights the need for high resolution in the vertical in models which simulate polar low development.

5. Summary and conclusions

We have presented a series of satellite images depicting the development of a train of four, more or less evenly spaced, polar lows over the Norwegian Sea and documented, with the help of ECMWF analyses, the large-scale flow pattern associated with their development. The disturbances formed in a shallow baroclinic zone of small static stability that extended outward from the boundary of the arctic ice pack through a layer of increasing depth. The track of the disturbances was situated close to the zone of strongest baroclinicity in the layer between the surface and 500 mb. It appeared that the temperature contrast was produced in large part by differential diabatic heating in the cross stream direction and that the contrast was enhanced by a confluence of the wind in the direction normal to the isotherms.

The flow was of the reversed shear type described earlier by Duncan (1978). The strongest (west-southwestward) flow was near the surface; the 500 mb flow was nearly stagnant. Accordingly, the disturbances were steered west-southwestward by the low-level winds in the direction opposite to the 1000–500 mb thermal wind.

The observed characteristics of the disturbances were compared with those obtained from a quasi-geostrophic, dry baroclinic model (Duncan, 1977) with use of basic state parameters taken from ECMWF operational analyses. The wavelength of maximum instability in the model

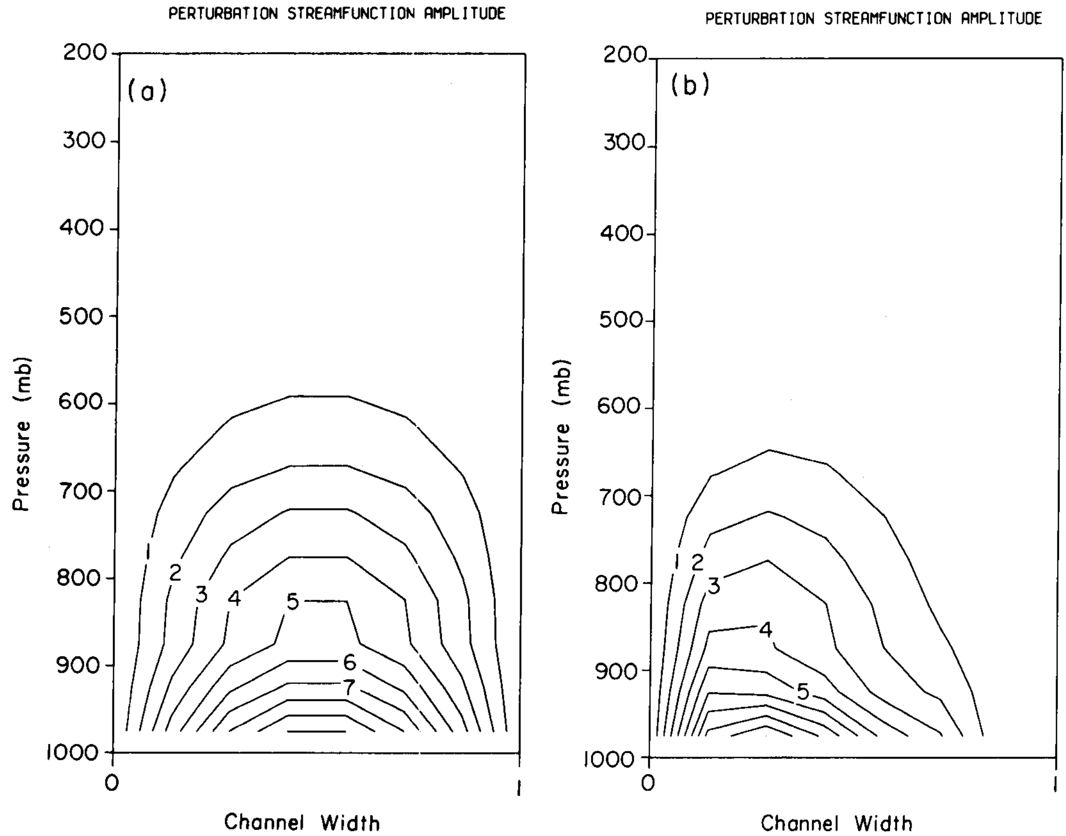

Fig. 7. Perturbation streamfunction magnitude (arbitrary normalized units) as a function of height and channel width for (a) no horizontal shear, (b) observed wind field with width of 525 km centred on the storm track. Both diagrams refer to the 500 km wave.

of about 500 km compared favourably with the observed wavelength of 570 km. The calculated phase speed for the 500 km wave of 8.7 m s^{-1} was considerably faster than the observed speed of 4.9 m s^{-1}. It is suggested that the slower movement in nature resulted from the presence of organized deep convection. An effect of the convection was to extend the domain of the disturbances into the nearly stagnant upper layers.

The rapid growth rate (1.47 d^{-1}) of the 500 km wave is qualitatively consistent with the observed rapid growth. Satellite imagery revealed that only 24–36 hours were required for the disturbances to reach full maturity. However, in view of the neglect of friction and lateral shear in the model and the use of the largest observed vertical shear, it seems unlikely that the observed rapid growth, estimated to be of the order 1–3 d^{-1}, can be fully explained by baroclinic instability. Convective heating may also have contributed importantly to the amplification.

In summary, from the available evidence it appears that baroclinic instability can explain the initial growth of the disturbances and the observed wavelength of 500–600 km but that some other mechanism, presumably deep convection, is required in order to account for the observed speed of travel and for the extremely fast growth rate.

6. Acknowledgements

The authors wish to thank Dr. M. Eckhardt of the Institute of Meteorology, Free University of Berlin, for calling their attention to, and providing, the satellite images and Dr. A. Hollingsworth of ECMWF for his encouragement and interest. One of the authors (R.J.R.) received partial support from the U.S. National Science Foundation and National Oceanic and Atmospheric Administration (NOAA) under grant number ATM-8421396. The authors also acknowledge Dr. J. N. Pitts, Jr. and his research group for their assistance during the experiment at the Statewide Air Pollution Research Center, Riverside, California.

7. Appendix

Model parameters

Unless otherwise stated in the text the parameters used by the model for the experiments described in Section 4 were as follows:

number of pressure levels = 33
increment between pressure levels = 25 mb
number of horizontal points = 15
increment between horizontal points = 75 km
Coriolis parameter = 0.000137 s^{-1}
β parameter = 0.0 m^{-1} s^{-1}
convergence criterion = 0.02 m s^{-1}

Pressure (mb)	Static stability (m^2 s^{-2} mb^{-2})	Pressure (mb)	Static stability (m^2 s^{-2} mb^{-2})
1000	0.0023	550	0.018
950	0.003	500	0.024
900	0.004	450	0.037
850	0.005	400	0.05
800	0.006	350	0.085
750	0.007	300	0.13
700	0.01	250	0.24
650	0.011	200	0.48
600	0.014		

Although the vertical grid increment is 25 mb, the staggered form of the grid makes the effective increment 50 mb for any variable.

REFERENCES

Arpe, K., Brankovic, C., Orial, E. and Speth, P. 1986. Variability in time and space of energetics from a long series of atmospheric data produced by ECMWF. *Beitr. Phys. Atmos. 59*, 321–355.

Duncan, C. N. 1977. A numerical investigation of polar lows. *Q. J. R. Meteorol. Soc. 103*, 255–267.

Duncan, C. N. 1978. Baroclinic instability in a reversed shear flow. *Met. Mag. 107*, 17–23.

Forbes, G. S. and Lottes, W. D. 1985. Classification of mesoscale vortices in polar airstreams and the influence of the large-scale environment on their evolutions. *Tellus 37A*, 132–155.

Harley, D. G. 1960. Frontal contour analysis of a "polar" low. *Met. Mag. 89*, 141–147.

Harrold, T. W. and Browning, K. A. 1969. The polar low as a baroclinic disturbance. *Q. J. R. Meteorol. Soc. 95*, 710–723.

Lönnberg, P. and Hollingsworth, A. 1986. The structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus 38A*, 137–161.

Mansfield, D. A. 1974. Polar lows: The development of baroclinic disturbances in cold air outbreaks. *Q. J. R. Meteorol. Soc. 100*, 541–554.

Økland, H. 1977. On the intensification of small-scale cyclones formed in very cold air masses heated by the ocean. *Institute Rept. Ser. No. 26*, Institutt for Geofysikk, Universitet I Oslo [Boks 1022, Blindern, Oslo 3, Norway], 31 pp.

Rasmussen, E. 1979. The polar low as an extratropical CISK-disturbance. *Q. J. R. Meteorol. Soc. 106*, 313–326.

Rasmussen, E. 1983. A review of meso-scale disturbances in cold air masses. In: *Mesoscale meteorology—theories, observations and models* (eds. D. K. Lilly and T. Gal-chen). Boston: D. Reidel Publishing Company, 246–283.

Rasmussen, E. 1985. A case study of a polar low development over the Barents Sea. *Tellus 37A*, 407–418.

Reed, R. J. 1979. Cyclogenesis in polar airstreams. *Mon. Wea. Rev. 107*, 38–52.

Sardie, J. M. and Warner, T. T. 1983. On the mechanism for the development of polar lows. *J. Atmos. Sci. 40*, 869–881.

Sardie, J. M. and Warner, T. T. 1985. A numerical study of the development mechanisms of polar lows. *Tellus 37A*, 460–477.

Shapiro, M. A., Fedor, L. S. and Hampel, T. 1987. Research aircraft measurements of a polar low over the Norwegian Sea. *Tellus 39A*, 272–306.