Doctoral Thesis

Structure and dynamics of distinctive flow anomalies in the lowermost stratosphere

Author(s):
Kew, Sarah Frances

Publication Date:
2007

Permanent Link:
https://doi.org/10.3929/ethz-a-005483580

Rights / License:
In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.
Structure and Dynamics of Distinctive Flow Anomalies in the Lowermost Stratosphere

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by

SARAH FRANCES KEW

MSc, University of Reading
MSci Physics, Imperial College London

born 12.10.1979

citizen of the United Kingdom

accepted on the recommendation of

Prof. Dr. H. C. Davies, examiner
Dr. C. A. Davis, co-examiner
Dr. M. A. Sprenger, co-examiner

2007
For my thoughts are not your thoughts, neither are your ways my ways, declares the LORD.
As the heavens are higher than the earth, so are my ways higher than your ways
and my thoughts than your thoughts.

As the rain and the snow come down from heaven, and do not return to it without watering the earth and making it bud and flourish,
so that it yields seed for the sower and bread for the eater,
so is MY WORD that goes out from my mouth:
It will not return to me empty, but will accomplish what I desire and achieve the purpose for which I sent it.

Isaiah 55:8-11, NIV
# Contents

Abstract iv

Zusammenfassung vi

1 Introduction 1
   1.1 Motivation .......................................................... 1
   1.2 Aims and outline ...................................................... 3

2 Theory, Data Sets and Tools 5
   2.1 The potential vorticity perspective ............................... 5
   2.2 Data sets .............................................................. 6
      2.2.1 The ERA-40 data set ........................................... 6
      2.2.2 ECMWF forecasts and analyses ................................. 7
   2.3 Tools ................................................................. 7
      2.3.1 Lagrangian trajectories ........................................ 7
      2.3.2 Contour dynamics ............................................... 7

3 Case Study 9
   3.1 Isentropic evolution ................................................... 9
   3.2 Cross sections .......................................................... 10
   3.3 Further comments .................................................... 13
   3.4 Summary and conclusions ............................................ 14

4 Isentropic Tracking 19
   4.1 Introduction ............................................................ 19
   4.2 Methodology ............................................................. 20
      4.2.1 Feature identification ......................................... 21
      4.2.2 Feature tracking ............................................... 23
      4.2.3 Identification of physical events ............................. 24
   4.3 Demonstration of tracking on case study ............................ 27
      4.3.1 Demonstration of identification and advection ................ 27
      4.3.2 Evolution of tracked characteristics for case study .......... 27
      4.3.3 Single events ................................................. 29
      4.3.4 Time evolution of tracked characteristics ................... 30
      4.3.5 Summary of evaluation ....................................... 31
7.3.4 Vertical sections .............................................. 91
7.3.5 Estimating the strength of upper level PV penetration .......... 97
7.4 A potential vorticity perspective on forecasting .................... 102
7.4.1 Forecasting challenges ......................................... 102
7.4.2 Traditional forecasting methods .................................. 103
7.4.3 Prospects for use of isentropic PV in forecasting ............... 103
7.4.4 Influence of upper level PV on polar low development in ECMWF polar low forecasts ............................................ 104
7.4.5 Case study ...................................................... 104
7.4.6 Discussion ..................................................... 105
7.5 Summary and outlook .............................................. 105

8 Summary ...................................................................... 109

A Linear Instability of a Polar β-Plane Annulus ......................... 111
A.1 Basic state .......................................................... 111
A.2 Normal mode analysis of ring instability .......................... 111

B Contour Dynamics ................................................................ 113
B.1 From the vorticity equation to the contour integral ............... 113
B.2 Including the planetary vorticity f ................................. 114
B.3 Numerical scheme and initialisation .............................. 115
B.4 Validation and sensitivity ............................................. 115
B.5 Contour surgery ...................................................... 115
B.6 Visualisation .......................................................... 116

References ......................................................................... 117

Curriculum Vitae .................................................................. 123

Acknowledgements ................................................................ 125
Abstract

Localised mesoscale cyclonic flow anomalies are a prevalent feature of the lowermost stratosphere. They are related to a variety of dynamical phenomena, such as stratospheric streamers, cut-offs, tropopause folds and jet-streaks. Viewed from the PV-θ perspective, the anomalies often take the form of horizontally and vertically coherent PV structures with amplitudes of a few PVU above the ambient isentropic values. These anomalies are accompanied above and below by respectively warm and cold potential temperature anomalies which are characteristic of the patterns seen in idealised models of isolated anomalies in a uniform background flow. The structures have a vertically penetrating cyclonic flow field which can induce circulations and/or interactions with other flow anomalies on lower-lying levels.

In this study, the PV perspective is adopted to explore several aspects of lower-stratospheric flow anomalies and their ambient environments using the ERA-40 data set.

In the region of the tropopause, such flow anomalies can have a profound dynamical impact on the surface through their acting as instigators and intensifiers of extra-tropical cyclones. A case study of one such cyclogenesis precursor is presented to emphasise their longevity (6 days) and origin well within the stratosphere. A feature-tracking tool is developed and applied to 10 consecutive winters to examine the general characteristics of lower stratospheric PV anomalies (geographical distribution, structure, life-cycle) on isentropic surfaces and to concomitantly shed light on an earlier phase of the cyclogenesis. The tool is able to identify small scale anomalies within a background PV gradient and to detect genesis, lysis, merger and splitting events. The anomalies are found to have a exponentially decaying lifetime frequency distribution. Over 25% survive at least 4 re-analysis periods (one day), which suggests they have a basis in real observations. The geographical distributions reveal a tendency for events to be most numerous near high topography and in an annular band centred around 70°N.

In a second part, consideration is given to aspects of the dynamical distinctiveness of the stratospheric so called 'middleworld' in which the anomalies reside. A climatological investigation reveals that the maximum stratospheric PV values are distributed in the form of a polar maximum on upper middleworld isentropes and as an annular circumpolar band on the lower isentropes. Diagnostics developed to explore temporal variability of the band indicate a peak in amplitude occurs in late spring, followed by a rapid decay in summer.

A distinctive structure is also detected in spring in the time mean PV distribution of the lower stratosphere that reflects the annular anomaly distributions and is consistent with the region being dynamically unstable. A barotropic linear instability analysis suggests that a circumpolar ring of enhanced vorticity of the climatological dimensions would be subject to a wavenumber 1 instability. Contour dynamics simulations are used to illustrate the non-linear evolution of an idealised ring and observe its break down into vortices and filaments. The results suggest that the lowermost stratosphere is not simply a passive receptor or transmitter of planetary waves but can itself be an instigator of dynamical activity.

The time averaged tropopause height falls to a minimum over the polar regions. It is asserted that prominent PV anomalies of the lowermost stratosphere will be of significance for polar meteorology, in particular for the development of polar mesocyclones. In a third part, a study is undertaken to examine the evolution and roles both of individual anomalies and their synoptic environments in the pre-genesis and development phase of polar lows in the Norwegian and Barents Seas. In this region, conditions conducive for polar low development are the presence of a mature synoptic cyclone over Scandinavia whose circulation directs cold, Arctic air over relatively warm ocean, destabilising
the air column. From a PV perspective, this synoptic situation is frequently accompanied by stratospheric intrusions behind the cold front, transporting PV anomalies into a suitable position to trigger development.

A geographically fixed composite of the isentropic PV field for 21 dates on which mature polar lows were present in the Norwegian and Barents Seas reflects this picture. Inspection of the individual cases shows that about half fit well the expected scenario. The prominent synoptic features of the remaining cases illustrate other important processes and features. These include 1) a tongue of high PV flanking the east side of a high pressure ridge over Greenland, 2) baroclinic waves along a secondary PV gradient, 3) low level vorticity induced by orographic flow and 4) isolated anomalies which originate from stratospheric streamers. Consideration is given to the performance of ECMWF forecasts and their sensitivity to the presence and structure of polar low precursor features in the (re-)analysis.
Zusammenfassung


In dieser Arbeit werden diese Strukturen und ihre Umgebung, wie sie im ERA-40 Datensatz detektiert werden, genauer untersucht.

Wenn dieses Strömungsanomalien in der Nähe der Tropopause auftreten, können sie Tiefdruckgebiete am Boden auslösen und / oder verstärken. Dies wird anhand einer Fallstudie aufgezeigt, in der auch deutlich wird, dass die Strömungsanomalien langlebig sind (6 Tage) und ihren Ursprung in der polaren Stratosphäre haben. In dieser Arbeit ist eine automatisierte Routine entwickelt worden, welche es erlaubt, die Bewegung dieser Strömungsanomalien auf isentropen Flächen zu verfolgen. Diese Routine wird während zehn aufeinander folgenden Wintern eingesetzt, um die geographische Verteilung, die Struktur und den Lebenszyklus der Strömungsanomalien zu untersuchen und so eine sehr frühe Phase der Zyklogenese besser zu verstehen. Die Routine kann kleinskalige Strömungsanomalien in Anwesenheit eines PV-Gradienten identifizieren, sowie deren Entstehung, Auflösung, Teilung und / oder Verschmelzung bestimmen. Die Verteilung der Lebensdauer dieser Strukturen ist exponentiell abfallend. Mehr als 25% der Strukturen existieren jedoch mindestens während 24 Stunden, was dafür spricht, dass diese Strukturen auf realen Beobachtungen basieren und nicht künstlich durch die Modellphysik generiert werden. Die geographische Verteilung zeigt, dass die Anomalien vermehrt über Gebirgen und in einem Ring um den Pol entlang 70°N auftreten.

In einem zweiten Teil wird die so genannte 'stratosphärische Mittelwelt' genauer untersucht. Es zeigt sich, dass sich die auf einen klimatologischen Zeitraum maximale PV in den tiefen isentropen Lagen dieser Mittelwelt in einer ringförmigen Struktur um den Pol anordneten, während sie in den höheren Lagen direkt über dem Pol liegt. Der zirkumpolare Ring ist Ende Frühling am stärksten ausgeprägt und zerfällt relativ rasch Anfang Sommer.

Das Vorhandensein des Ringes in der mittleren PV-Verteilung, besonders im Frühling, bedeutet implizit, dass sich in dieser Region Instabilitäten ereigen können. Eine lineare Analyse der barotropen Instabilität der Ringstruktur zeigt, dass diese anfällig ist für Wellenzahlen 1 Instabilitäten. Idealisierte Konturdynamic-Simulationen zeigen, dass nichtlineare Instabilitäten den Ring in kleine Vortices und kleinere Strukturen zerkleinern lassen. Diese Resultate zeigen auf, dass die untere Stratosphäre nicht nur als passiver Wellenleiter fungiert, sondern selbst die Quelle von neuen Instabilitäten sein kann.

In einem dritten Teil wird der Einfluss der kleinskaligen Strömungsanomalien auf die Bodenwetterentwicklung in der Nordpolarregion untersucht. Da die Tropopause in diesen Regionen relativ tief liegt, wird der Einfluss dieser Anomalien dort besonders deutlich spürbar. Insbesondere werden meso-skalige Zyklonen untersucht. Die Rolle der Strömungsanomalien und der synoptischen Umgebungen in der Entwicklungphase von diesen so genannten 'Polar Lows' wird für eine Anzahl (21) von Fällen untersucht. Im Nordmeer und in der Barentssee können solche Polar Lows entstehen, wenn synoptische Tiefdruckgebiete über Skandinavien kalte und trockene Luft aus der Arktis in Richtung Meer leiten. Dabei wird die vertikale Stabilität der Lufssäule reduziert. Von einem PV-
Standpunkt aus gesehen findet man häufig stratosphärische Intrusionen in der Tropopausenregion hinter der Kaltfront am Boden, die PV-Anomalien in eine geeignete Lage für Polarlow-Bildung führen können.

Chapter 1

Introduction

1.1 Motivation

Atmospheric divisions

The atmosphere is conventionally divided into layers according to the vertical temperature gradient. The two lowermost layers which comprise most of the mass are the most important for influencing weather and atmospheric dynamics. These are the troposphere (0-10 km as a global average), for which temperature decreases with height, and the stratosphere (10-50 km), for which the temperature is constant or increasing with height. The troposphere and stratosphere are also chemically and dynamically distinct from one another. The troposphere is characterised by high humidities, low ozone concentrations and low static stability, whereas the stratosphere is characterised by low humidities, high ozone and stable stratification.

Due to the high humidities and low stability in the troposphere, latent heat release is of high importance for surface weather and can, in regions of surface baroclinicity, give rise to significant vertical motion. The very low stratospheric humidities, on the other hand, imply that the dynamical effect of latent heat release can be neglected at upper levels. Radiative transfer however has a much more important role in the stratosphere due to short-wave heating from ozone and long-wave emission from carbon dioxide (Andrews et al., 1987). Radiative transfer acts to reduce the amplitude of temperature fluctuations. Estimations of thermal damping rates based on vertical length scales of several kilometres are 20-40 days in the lower stratosphere and decrease to 5 days in the upper stratosphere (Haynes, 2005, and references therein). On short timescales, flow in the lowermost stratosphere can be assumed largely adiabatic and quasi confined to 2D surfaces of constant entropy, or 'isentropes'.

With a focus on the dynamics of the lowermost stratosphere, it makes sense to follow the dynamic developments on isentropic surfaces using a materially conserved quantity such as potential vorticity (PV). In the atmosphere, isentropes correspond to surfaces of fixed potential temperature. Climatological distributions of potential temperature show that isentropic surfaces present in the lower subtropical troposphere slope upwards to the tropopause in midlatitudes and then level off towards the pole. The lowest isentropes cap the poles, and the highest isentropes have no contact with the surface.

A physically meaningful division of the atmosphere based on the climatological distribution of potential temperature was first proposed by Sir Napier Shaw (1930). Hoskins (1991), building on the ideas of Shaw, suggested a 3-fold division, as illustrated by figure 1.1, into 1) the overworld, which consists of all isentropic surfaces which are everywhere above the tropopause\(^1\); 2) the middleworld, where isentropic surfaces intersect the tropopause and 3) the underworld, which consists of the lowermost surfaces which are connected with the Earth's surface. In this perspective, the middleworld is dynamically isolated from the diabatic and frictional effects of the underworld and is distant from the radiative processes of the overworld.

\(^1\)The dynamical tropopause is defined here by the 2 PVU isosurface in the extra-tropics and the 380 K isentrope in the tropics.
The middleworld

The middleworld contains the lowermost stratosphere and is the focus of this study. Figure 1.2 displays the instantaneous PV distribution on a typical winter middleworld isentropic surface of 320 K. Several interesting features can be seen, some of which are not evident in the climatological picture, and they serve to highlight topics of ongoing research:

- The tropopause region is identifiable as a region of strong PV gradients close to the 2 PVU isoline, which exhibits large meridional fluctuations in the form of planetary waves. This region acts as a wave guide (e.g. Schwierz et al., 2004) and is important in baroclinic development (Eady, 1949, Hoskins et al., 1985).

- Elongated protrusions of stratospheric air into the troposphere, known as streamers, are evident (prominent examples are labelled S₁ and S₂ in figure 1.2). As the 320 K isentropic surface in the vicinity of the tropopause tilts in physical space towards the ground (figure 1.1), streamers also penetrate into the troposphere vertically, as regions of anomalously high PV. Localised, highly positive anomalies are often located at the extremities of streamers and may act as precursors of cyclogenesis as they induce spin-up of tropospheric air below.

- A tropospheric streamer in the process of being sequestered to form a tropospheric cut-off (labelled Ct) is present. Streamer S₁ can be identified with a stratospheric cut-off on a

Fig. 1.1: The atmosphere according to Hoskins (1991). Shading distinguishes the three suggested atmospheric regions (underworld, middleworld and overworld) and the partition of the middleworld into a stratospheric and tropospheric part. The divisions in this figure are determined from the ERA-15 zonally averaged potential temperature (contour interval: 20 K) for January and its relationship to the tropopause (dashed 2 PVU isoline for the extra-tropics and 380 K in the tropics). The vertical pressure scale is logarithmic. See text for further details.
1.2. AIMS AND OUTLINE

lower isentrope. Both streamers and cut-offs have been linked with the transfer of mass and chemical constituents between the troposphere and stratosphere, known as exchange events (Stohl et al., 2003, Sprenger et al., 2007). Some of the small PV structures detected in the ERA-40 and contained within streamers can be linked with measurable features such as dry intrusions in water vapour imagery (see figure 1.3).

- Within the stratospheric portion of the middleworld, many mesoscale PV structures are present (cyan shaded in figure 1.2). Until now, studies have concentrated on understanding the roles of specific anomalies in cyclogenesis (e.g. Fehlmann and Davies, 1999) or their general properties on the tropopause (e.g. Hakim, 2000). The anomalies in the interior of the stratospheric body, however, have received less attention. It is these features which will be the focus of the present study.

The tropopause and the stratospheric portion of the middleworld are climatologically lowest over the poles and are not far from the surface (less than 10 km). The circulation associated with isolated anomalies can penetrate to around this height. Thus it may be that the study of anomalies well inside the stratospheric body will be of significance for polar meteorology.

1.2 Aims and outline

The overall objective of this thesis is to study the structure, dynamics and impact of the flow anomalies in the middleworld using the PV framework.

The ERA-40 data set is the main data source for this project. It is described together with the basic notions of PV in chapter 2. In chapter 3, the instantaneous structure and characteristics
are explored for a prominent tropopause anomaly which instigates a rapid surface cyclogenesis and which can be subjectively traced back deep into the stratospheric portion of the middleworld. This sets the scene for an investigation into the population life cycle characteristics and geographical distribution of middleworld anomalies using a feature tracking tool in chapter 4. In chapter 5, isentropic climatologies of the middleworld PV structures in which these anomalies reside are presented. The interannual variability and dynamical stability of the climatological structures are considered in chapter 6. Chapter 7 concerns the impact of lower middleworld PV anomalies on the development of polar mesoscale cyclones and discusses the benefits of the upper level PV perspective in the forecasting of these storms. Summarising remarks are given in chapter 8.
Chapter 2

Theory, Data Sets and Tools

2.1 The potential vorticity perspective

In recent years, particularly following the seminal review of Hoskins et al. (1985), the PV perspective has become widely known for offering an attractive, dynamically meaningful means to analyse atmospheric flow.

The definition of PV according to Ertel (1912) is a combination of kinematic (absolute vorticity) and thermodynamic (static stability) terms:

\[ PV = \frac{1}{\rho} \eta \cdot \nabla \theta, \]

where \( \rho \) is the density, \( \eta \) is the absolute vorticity and \( \theta \) is the potential temperature.

Using the hydrostatic approximation and pressure coordinates on a spherical earth, PV can be approximated for synoptic and planetary scales as

\[ PV = (\zeta_\theta + f) \left( -g \frac{\partial \theta}{\partial p} \right), \]

where \( \zeta_\theta \) is the vertical component of the relative vorticity, \( f \) is the Coriolis parameter, \( g \) is the gravitational acceleration and \( p \) is the pressure.

PV is expressed in the convenient PV units (PVU) defined by 1 PVU = \( 10^{-6} \) m\(^2\) s\(^{-1}\) kg\(^{-1}\).

Climatological PV distributions exhibit low values in the troposphere, typically near 1 PVU or below, and high values greater than 4 PVU are typical in the stratosphere, due to the increased static stability there. The latitudinal dependence of the Coriolis parameter \( f \) contributes to a general increase in the absolute vorticity towards the polar regions. In broad terms, the combined effect is a gradual increase of PV from the sub-tropics towards the poles, superposed on a near discontinuity at the tropopause. We choose to define the dynamical tropopause by the 2-PVU surface, which is co-located within the zone of enhanced PV gradient on isentropic surfaces.

Three important principles underlie the advantages of the PV perspective:

- **Conservation**
  
  In adiabatic and frictionless conditions, the PV of an air parcel is conserved:

  \[ \frac{D}{Dt} PV = 0. \]

  Under these conditions, time evolution of the PV field is confined to isentropic surfaces.
These conditions generally hold at tropopause levels and in the lowermost stratosphere, where humidity is low and frictional surface effects are too distant to be of influence. At upper levels, PV can therefore be used as a quasi passive tracer of air parcels.

On lower levels and in the upper stratosphere (where radiative processes become important), PV conservation should not be assumed. The creation or destruction of PV by diabatic and frictional processes is expressed by

$$\frac{D}{Dt} PV = -\nabla (gh(\zeta + f k) + gF \times \nabla \theta),$$

(2.1)

where $\dot{\theta}$ is the diabatic heating and $F$ is the frictional force. A negative PV tendency is observed where the flow leaves a region of impulsive diabatic heating and a positive PV tendency is present in the entrance zone of the heating region.

- The invertibility principle

A remarkable property of PV is that the 3D distribution can be inverted within a chosen domain to diagnostically deduce the flow and temperature fields within that domain, provided that appropriate boundary conditions, a suitable reference state and balance conditions are defined (Kleinschmidt, 1950).

- The partition principle

The partition principle states that a PV distribution can be decomposed into individual PV anomalies and that the isolated contribution of each anomaly to the flow field may be inferred by performing an inversion for each separately.

The partition and invertibility principles have been applied, for example, to the study of extratropical cyclogenesis (e.g. Davis and Emanuel, 1991, Davis, 1992) and tropical cyclones (e.g. Wu and Emanuel, 1995) to reveal the relative influence and roles of the salient PV elements of these systems.

2.2 Data sets

2.2.1 The ERA-40 data set

All data sets used in this thesis originate from the ECMWF (European Centre for Medium-range Weather Forecasts). The so-called 'ERA-40' comprises the ECMWF Re-Analyses of 45 years of observations from September 1957 to August 2002 from all conventional platforms. A detailed report on the data set is given by Uppala et al. (2005). The ERA-40 employs a 3D variational data assimilation scheme at each time step which combines observations with a model-derived background field to produce an optimised representation of the atmosphere for the information available. The same assimilation scheme is applied to the entire data set, permitting and facilitating comparisons of data from different years. However, although the data assimilation scheme is fixed, the number, distribution and type of observations changed considerably over the time period. The most noteworthy change is the assimilation of satellite data after 1979, which resulted in significant improvements in the quality of the data, particularly in the Southern Hemisphere and in the upper atmosphere. Climatological analyses performed in this study use data after these changes were made, in the last decade of the data available (1991-2001).

The ERA-40 data set provides global re-analysis data on a 6-hour temporal resolution (00, 06, 12 and 18 UTC) and has a spectral resolution of triangular truncation T159, which corresponds to about 125 km in the horizontal and 60 hybrid vertical levels from the surface to 1 hPa. The vertical resolution near the tropopause is approximately 32 hPa and about 14 model levels span the lowermost stratosphere (290 K - 380 K). The data used in this thesis have been interpolated on to a $1^\circ \times 1^\circ$ longitude-latitude grid. The re-analysis provides the following primary (prognostic) variables: surface pressure (p), temperature (T), specific humidity (q), wind components (U, V),
vertical velocity ($\omega$). Additional variables relevant for this study such as PV, potential temperature, static stability, and vorticity, are computed from the primary variables and interpolated onto isobaric and isentropic surfaces. The 3D PV field is calculated using the hydrostatic definition of Ertel PV (equation 2.2).

### 2.2.2 ECMWF forecasts and analyses

ECMWF forecasts and analyses are retrieved from the ECMWF for case studies in chapter 7. The ECMWF changed the resolution of their forecast model grid several times over the time span of the ERA-40\(^1\). For the dates relevant to this study, the forecast and analysis fields have a spectral resolution of T511 and 60 vertical levels. For use in conjunction with ERA-40 data, the forecast and analysis fields are also interpolated onto the same 1° x 1° grid.

### 2.3 Tools

#### 2.3.1 Lagrangian trajectories

Lagrangian backward and forward trajectories are applied in this study to trace the history and evolution of the flow of air parcels coincident with PV structures. The trajectories are analysed using the trajectory tool LAGRANTO developed by Wernli and Davies (1997), see also Wernli (1997). The 3D fields required for the calculations are taken from the ERA-40.

#### 2.3.2 Contour dynamics

The contour dynamics technique is a numerical method for examining the evolution of 2D inviscid flows. The technique is ideally suited to studying the evolution of simple potential vorticity configurations where the flow field can be divided up into separate regions of constant vorticity. A contour dynamics tool is used in this thesis to explore PV configurations of the lowermost stratosphere. The theory of contour dynamics and the specifics of the tool used are outlined in appendix B.

\(^1\)See http://www.ecmwf.int/products/forecasts/guide/user-guide.pdf for a detailed guide
Chapter 3

Case Study - A Lower-Stratospheric PV Anomaly

In this chapter, a case study is introduced which will be referred back to in chapter 4. It concerns a PV anomaly which initiated a rapid cyclogenesis event in October 2000. The anomaly is depicted with a series of isentropic maps and vertical sections. Lagrangian trajectories are employed to trace the flow of air parcels involved in the formation of the precursor and its interaction with the induced cyclone.

The case study will be used to

- illustrate that isentropic PV anomalies seen in the ERA-40 data, of the form presented in figure 1.2, can potentially serve as precursors of synoptic scale cyclogenesis
- show that precursors can exist several days ahead of a cyclogenesis event as coherent anomalies
- describe the formation of the coherent anomaly and its fate after cyclogenesis
- make observations on the nature of the isentropic precursor evolution which will aid interpretation of their isentropic tracking.

The isentropic evolution will first be presented, followed by inspection of vertical cross sections and finally a discussion and summary of findings relevant for subsequent chapters.

Aspects of the same case were examined by Dirren (2002) to illustrate an upper level vortex-jet interaction. He presented the upper and lower level evolution from 14th - 17th October 2000 using the analyses produced for the ECMWF forecast. Here, the period of study is extended further back to 12th October 2000 to capture an earlier phase of the cyclogenesis using the ERA-40 data set.

3.1 Isentropic evolution

The evolution of upper level PV is depicted on the 320 K isentrope (figure 3.1), where the precursor displays its maximum amplitude for the large part of its lifetime. Sea level pressure below 1000 hPa and two instructive isolines of potential temperature at 850 hPa are shown to illustrate the surface development. An arrow in each panel points to the location of the precursor.

At the earliest date shown, 12th October 12 UTC (a), the anomaly is not distinct but is contained within a larger elongated PV structure near the centre of the stratospheric portion of this middle-world isentrope, over the Arctic. The approach of an intruding lobe of tropospheric air from the south (approximately 160 W) increases the local (North-South) PV gradient and is accompanied by a local wind maximum (jet streak) of over 40 ms⁻¹ on the southern side of the PV structure (not shown). The elongated PV structure is positioned in a region of strong (isentropic) shear and
as the tropospheric anomaly moves north, it is deformed and splits into two parts. The part later identified as the cyclogenesis precursor (positioned at approximately 86N, 110W) travels south over Canada, adjacent to the negative anomaly, whilst the other part remains on the opposite side of the date line (b). In figure 3.2, forward trajectories starting within the elongated PV structure seen in (a) demonstrate the splitting and show that trajectories moving towards Canada are co-located with the PV precursor.

Between 13th October 12 UTC (c) and 14th October 00 UTC (d), the precursor moves south, forming the west side of a synoptic scale Rossby wave. Its aspect ratio becomes smaller as it curls into a coherent structure. The PV amplitude is seen to increase from 8 PVU to over 10 PVU between these two times. The reason for this increase is discussed in section 3.3.

Twelve hours later on 14th October 12 UTC (e), there are early indications of surface development. The northerly 850 hPa isentrope (the one closest to the anomaly) has been deflected to the north on the east side of the anomaly and to the south on the south-west side. This suggests that the cyclonic wind field associated with the anomaly has penetrated to lower levels and is forcing warm air advection downstream.

The amplitude of the low-level temperature perturbation continues to increase and the anomaly moves into a position immediately north of the tropopause (f). By 15th October 18 UTC (g), the PV precursor can be seen at the onset of cyclogenesis as a distinct PV structure near the tropopause, approaching the North Canadian Atlantic Coast. It has induced a Rossby wave on the tropopause and surface pressure falls below 995 hPa downstream of the precursor, at the location of the low level warm anomaly. Its horizontal extent is approximately 500 km and the maximum PV of 10 PVU is about 4 PVU above the surrounding values.

A day later, on 16th October (h), a fully developed synoptic cyclone is present. The southerly 850 hPa isentrope provides a good indication of the cyclone's frontal system. The tilt between the surface pressure minimum and the PV anomaly is reduced, the PV anomaly being almost directly over the pressure minimum.

At this stage, diabatic processes are active and condensational heating is detected in the ERA-40 up to 500 hPa (not shown). The region of negative PV tendency, above the region of heating, approximately coincides with the tropospheric intrusion seen to circulate the low cyclonically on the 320 K surface in (h). As the tropospheric intrusion impinges on the positive precursor anomaly, a small section of the upper PV maximum is sheared away from the precursor and becomes part of the tail of the dry slot behind the cyclone. The larger remaining part moves directly over the surface low (i). This redistribution of the anomaly is illustrated by forward trajectories in figure 3.3. A day later on 18th October (j), surface pressure is increasing and the upper PV precursor has been disbanded into smaller fragments. The maximum PV within the precursor on 320 K decreases from 9 PVU to 7 PVU. The PV precursor is no longer a dominant feature of the flow on this isentropic surface.

### 3.2 Cross sections

To observe also the vertical coherency of the anomaly and its relation to the underlying atmosphere, the evolution of the vertical structure of the precursor is presented with the help of North-South PV cross sections in figure 3.4. The sections are taken through the centre of the anomaly at 320 K. Isentropes are superimposed at regular intervals and, in addition, the two isentropes selected for figure 3.1 are emphasised to ease comparison between the two figures.

The first cross section (a) reveals that the PV structure over the Arctic, from which the precursor was formed (see 3.1 (a) and (b)), has a local maximum in the vertical between the 320 K and 330 K surfaces (a). This feature moves towards lower latitudes and retains its amplitude of 9 PVU until 13th October 12 UTC (c).

Between (c) and (d), the anomaly's amplitude increases (as was also detected in figure 3.1 (c) to (d)), however its centre remains on the same isentropic level. If the height of the 2 PVU surface beneath the anomaly is taken as a measure of its vertical intrusion into the troposphere, the intrusion appears to remain steady at around 450 hPa until 14th October 00 UTC (d). By
3.2. CROSS SECTIONS

Fig. 3.1: PV on 320 K in colour with isobars of sea level pressure below 1000 hPa at intervals of 5 hPa (blue) and two isentropes at 850 hPa of potential temperature for 277 K (north) and 287 K (south) (red). An arrow in each panel points to the location of the anomaly of interest.
14th October 12 UTC (c), the anomaly has encountered a region of strongly sloping isentropes in the underlying atmosphere - the main baroclinic zone. In moving adiabatically further south, the precursor is forced to descend along the isentropes to lower altitudes. It is at this stage that a temperature perturbation on 850 hPa was seen in figure 3.1. The intrusion extends at least down to 550 hPa (g). The deepest intrusion was actually closer to 600 hPa and occurred to the south-east of the 320 K PV maximum, i.e. not in the plane of the cross section.

Just north of the stratospheric intrusion, a tropospheric intrusion into the stratosphere can be seen to reach about 300 hPa (h). Directly below this is a low-level PV maximum, which is a product of the diabatic processes associated with the surface cyclone (Dirren, 2002). At this time, the cross section reveals that the anomaly centre has moved from 320 K to a higher isentropic level near 330 K, which results in the decrease in amplitude seen on the 320 K map (figure 3.1 (h)), and the density of isentropes in the stratospheric intrusion below the anomaly decreases. This is consistent with the anomaly being positioned in a region of moist ascent - above the surface cyclone. Diabatic heating in the lower troposphere will cause the isentropes above the region of heating to dip downwards towards the temperature anomaly.

To further illustrate the adiabatic nature of the flow during the first part of the precursor's lifetime, a different representation of its cross section is given in figure 3.5. Cross sections of the 9 PVU
3.3. Further comments

Let us return briefly to the issue of the increase in PV seen between panels (c) and (d) of figures 3.1 and 3.4. The fixed position of the precursor, with respect to the isentropes seen in the cross sections, indicates that the amplitude increase is not due to diabatic processes. Either there is a fluctuation in the consistency/accuracy of the assimilated data or higher PV has been advected into the centre of the structure from elsewhere on the 320 K surface. A parallel neighbouring anomaly was seen to approach on the east in figure 3.1 but it was of a lower magnitude. As

Fig. 3.2: Forward trajectories starting on 320 K within the local PV maximum (PV > 8.5 PVU) identified with the precursor on 12th October, 12 UTC, and going forward to 14th October, 06 UTC. The background PV map is valid on 14th October, 06 UTC, at 316 K. Black rings with white centres mark the beginning of the trajectories and white rings with black centres mark the ends.
Fig. 3.3: 2-day forward trajectories starting on 320 K with the local PV maximum (PV > 8 PVU) identified with the precursor on 15th October 12 UTC and going forward to 17th October 12 UTC. The background PV map is valid on 17th October 12 UTC at 317 K. Black rings with white centres mark the beginning of the trajectories and white rings with black centres mark the ends.

there were no other anomalies of similar or greater magnitude in the vicinity prior to this time, the amplitude increase is most probably an artifact of the data assimilation scheme (the magnitude was 'corrected' as the anomaly was advected into a denser upper-air observing network over Canada) and the resolution of the grid (prior to this time, the anomaly is very narrow and its maximum is not well resolved by the ERA-40 grid) rather than the result of a physical process.

Indeed, inspection of backward trajectories (figure 3.6), started on 13th October, 12 UTC, when the increase in PV was detected in the ERA-40, reveals no evidence of merger having taken place with the PV structure lying to the east. The trajectories in fact remain together in a coherent flow for about 4 days. They fan out over Greenland, indicating that the region to the north of Greenland was where the precursor air parcel was initially formed.

3.4 Summary and conclusions

The evolution in the ERA-40 of a mesoscale PV anomaly that acted as a precursor to surface cyclogenesis has been presented. The anomaly was generated by its splitting from a larger PV structure over the Arctic, after which it rolled up into a coherent structure and was advected adiabatically towards the jet. This particular precursor had a maximum amplitude near 320 K during the adiabatic stage of its life time.

The precursor remained coherent until shortly after cyclogenesis. The growing wave on the tropopause initiated by the presence of the precursor and amplified by baroclinic instability resulted in a tropospheric intrusion which rotated about the cyclone cyclonically. The flow associated with this negative anomaly split the precursor into a smaller yet coherent anomaly, which travelled in phase with the surface disturbance, and an elongated structure, which could be associated with a dry intrusion into the cyclone.

After the surface cyclone had reached maturity, the amplitude of the (smaller) coherent anomaly diminished on 320 K and appeared to dissipate. Cross sections revealed the anomaly's maximum
Fig. 3.4: North-South PV cross sections (colour) and contours of potential temperature at intervals of 10 K (black) and additionally at 277 K and 287 K (red). Latitude is given on the horizontal axis and pressure (hPa) on the vertical.
had shifted to higher isentropic levels.

The precursor was also long-lived. From the moment of splitting from a larger PV structure over the Arctic (13th October, 00 UTC), to the moment of dissipation in the presence of diabatic processes (18th October, 18 UTC), the PV anomaly survived for over 5.5 days. It was present, as a coherent structure, 3 days before cyclogenesis. Backward trajectories revealed that the air parcel which formed the precursor could be traced back to a region of convergent flow near northeast Greenland (9th October, 00 UTC) at almost one week before cyclogenesis.

The long period of adiabatic motion for this PV precursor leading up to cyclogenesis motivates the use of isentropic surfaces for tracking of potential precursors.

It can be expected that coherent precursors form in regions of shear by breaking off from larger high PV structures. Diabatic processes can be responsible for the decay of anomalies on one surface and their creation on another. This should be reflected on isentropic surfaces by non-conservation of PV of an order larger than the fluctuations observed in a generally adiabatic flow in the ERA-40. The determination of the time of genesis and lysis, and hence the life time, of isentropic anomalies will however be sensitive to method used to identify them and to the ability of the model to resolve fine-scaled PV structures.

Care should be taken in interpreting the detection of sudden changes in PV of an anomaly. Due to the nature of the data set, not all changes can be attributed to physical processes.

This particular feature's amplitude, scale and coherent structure suggest that the accompanying...
SUMMARY AND CONCLUSIONS

The dynamics of such PV anomalies can play a major role in cyclogenesis. Upper-level induced cyclogenesis of the form portrayed here carries with it implications for accurate prediction and prompts further dynamical considerations. From a predictive standpoint, it is important that the upper-level precursor's amplitude and structure be accurately captured in the analysis. From a dynamical standpoint, it is of interest to determine the origin and dynamical evolution of the precursors.

Fig. 3.6: 5-day backward trajectories starting on 320 K within the local PV maximum (PV > 8 PVU) identified with the precursor on 13th October, 12 UTC, and going back to 8th October, 12 UTC. The background PV map is valid on 8th October, 12 UTC, at 326 K. Black rings with white centres mark the beginning of the trajectories and white rings with black centres mark the ends.
Isentropic Tracking of Stratospheric PV Anomalies

A tracking tool has been developed to enable the general characteristics of isentropic PV anomalies to be established. As a prelude, a brief review of tracking techniques and relevant climatologies from the literature is presented in section 4.1 and is followed by a description of the tracking tool in section 4.2. The capabilities of the tool are demonstrated in section 4.3 by its application to the case studied in chapter 3. In section 4.4, the tool is applied to 10 years of ERA-40 data to produce a winter climatology of isentropic anomalies for the period December 1991 - February 2001 on four isentropic surfaces spanning 310 K - 340 K. Findings are summarised in section 4.5.

4.1 Introduction

It was noted in chapter 1 that inspection of isentropic PV maps reveal many small scale positive PV structures and it was seen in chapter 3 that anomalies of this kind can play a role in instigating surface cyclogenesis. It is these structures of dynamical impact which we wish to capture objectively. From this standpoint it is clear that anomalies and their amplitudes should be defined with reference to their immediate surroundings rather than with respect to a climatological mean or a constant threshold.

The importance of upper-level cyclonic vorticity maxima for surface development in mid-latitudes is well recognised and serves as a motivation for studying their dynamics and climatological distribution. In addition to increasing understanding of an early phase of surface development, such climatologies have another important application in providing a means to assess the performance of numerical models in their handling of upper level flow anomalies. The numerical models should reproduce similar distributions, motions and intensities to those given by climatologies.

Earlier studies of upper level flow anomalies were typically concerned with the distribution of troughs in geopotential height at 500 hPa (e.g. Bell and Bosart, 1989), though it was recognised that mobile upper troughs were accompanied by significant downward extrusions of the tropopause and attained a maximum amplitude near the tropopause (Sanders, 1988). Recently, more studies have been carried out at tropopause level (encouraged by the availability of improved upper level data) using pressure, vorticity, or potential temperature to define anomalies (e.g. Lackmann et al., 1997, Hakim and Canavan, 2005). Upper-level cyclonic maxima have also been linked to the presence of jet streaks (e.g. Reiter, 1969) and mid-tropospheric frontogenesis (Lackmann et al., 1997).

The signature of an upper level flow anomaly is thus evident in a variety of dynamical fields. However, this does not imply that every field is appropriate for tracking. Pressure and geopotential height, for example, are not conserved following the upper-level flow and are often subject to strong gradients in the background field, which can introduce a bias favouring deeper and slower moving local maxima (Sinclair, 1994).
In contrast, PV has a number of qualities which make it particularly suitable for tracking upper level flow features: It is a variable whose distribution can be inverted to deduce the flow field associated with an anomaly. It has a vorticity component which, being a derivative of the flow, will enhance local maxima even against a strong background gradient. Therefore potential precursors can be identified at an earlier stage in their life cycle. Large gradients are typically found near the tropopause but, poleward of the tropopause, the relatively weak large scale PV gradient eases identification of anomalies. Most importantly, the near conservation of PV can be expected following the flow on middleworld isentropes. The PV framework is being increasingly applied to describe the upper level contributions to cyclogenesis (e.g. Davis, 1992). It is therefore advantageous if the analysis of the pre-genesis phase take place in the same framework, using middleworld surfaces to follow the dynamics.

An objective analysis technique such as tracking is necessary to explore properties of a large populations of anomalies. A variety of tracking techniques are documented in the literature. A simple and common technique often applied to the tracking of surface cyclones is the nearest neighbour search (Blender et al., 1997), in which a cyclone is sought at the new time step within a search radius centred on the position at the previous time step. However, for the tracking of upper-level potential vorticity anomalies, such a simple method would be inadequate, due to the anomalies being far more numerous than cyclones and of a smaller scale. More sophisticated techniques have been developed, some of which involve a matching of the vortex properties such as intensity (Sinclair, 1994) between time steps, searching backward as well as forward in time (e.g. Dean and Bosart, 1996), and predicting the new position based on the most recent observations of the systems' velocity (Sinclair, 1994). Here a relatively simple method will be sought which performs well at upper levels.

As well as determining genesis and lysis locations, a tracking procedure can be employed to find instances of splitting and merger. Few studies however have dealt with merger and splitting of upper level vorticity maxima. One example is Dean and Bosart (1996) who objectively studied 500 hPa trough merger and fracture with the use of a tracking tool. They motivated their study by observations that both merger and splitting of 500 hPa troughs can be instrumental in cyclogenesis. In a study of a single case (see Hakim et al., 1996), where merger was responsible for cyclogenesis, poleward and equatorward anomalies were reported to have interacted with one another whilst rotating cyclonically in a background confluent flow. Splitting, on the other hand, was associated with formation of cut-off lows positioned equatorward from the main westerlies. It would be of interest to see if isentropic tracks initiated by splitting exhibit different distributions to those initiated by merger. An analysis of merger and splitting events will therefore be part of this study.

A few studies have used a composite approach to determine the life cycle characteristics of upper level vorticity maxima. Lackmann et al. (1997) for example, were able to associate the beginning of the life cycle of precursors of rapid cyclogenesis with the development of an elongated, lowered region of the tropopause, i.e. an elongated local PV maximum. This subsequently compacted into a more circular configuration just prior to initiating low level development. Hakim and Canavan (2005) also used a composite approach to study the subset of arctic tropopause vortices poleward of the jet, presenting the composite temporal evolution of potential temperature, amplitude, radius and latitude.

The aim is that the method developed will successfully identify geographical regions where isentropic PV structures form, intensify and decay, and through a composite analysis, provide insight into temporal evolution of their general characteristics. Attention will be drawn to changes in the vertical distribution by intercomparison of results for different isentropic levels.

4.2 Methodology

In this section, the objective technique for tracking of stratospheric PV structures on isentropic surfaces is outlined. The description of the algorithm is split into two stages, the first being the identification of the structures and the second being the tracking, i.e. the procedure used to transfer the identity of an observed structure at one time step to an observed structure at the subsequent time step. The capabilities of the tool, including the detection of merger and splitting, are also
4.2. METHODOLOGY

described.

4.2.1 Feature identification

The identification procedure seeks to select localised PV structures on an isentropic surface whose PV is elevated above the ambient or 'background' PV field.

The selected maxima are assigned a finite area, which permits an assessment of the anomaly's shape and is integral to the tracking technique. The assigned area is defined by a PV isoline, referred to below as the 'edge PV'. In order to assess the amplitude of the anomaly with respect to its local environment, a local 'background PV' for each structure is determined. The identification procedure also groups the structures into 'classes', according to their vicinity to the tropopause. Figure 4.1 schematically illustrates some parameters and terms used in the identification procedure.

Defining the outer boundary or 'edge PV'

The outer boundary for a potential centre is determined by seeking the PV value ('edge PV') such that the summed area of grid cells surrounding a local maximum and exceeding the 'edge PV' is approximately equal to a predefined area threshold.

The area threshold which determines the 'edge PV' is chosen to be $170 \times 10^3 \text{ km}^2$, which is equivalent to the area of about 33 longitude-latitude grid cells at 65°N and approximately corresponds to the physical dimensions of the observed anomalies. The chosen area is large enough to capture the form of the features yet small enough for isolated maxima to remain separated.

The 'edge PV' is found by applying a binary search algorithm to the PV field until the PV of the bounding contour has been determined to the nearest 0.01 PVU. This limiting step, $\Delta PV$, is small enough such that any further decrease would make no significant change to the area of the majority of (isolated) identified features.

The potential centres are assigned their defining boundaries, with the strongest maxima being considered first. If the bounding contour assigned to a particular maximum incorporates a stronger maximum, the 'edge PV' of the lesser maximum is raised until the two maxima become distinct, resulting in an area reduction of the lesser maximum. Lesser maxima within the bounding contour of a strong maximum are disregarded. In this way, preference is given to the strongest maxima but

Fig. 4.1: A schematic to illustrate some parameters and terms involved in the identification of an isentropic PV anomaly.
weaker maxima may still be included. The method of Hakim and Canavan (2005) would overlook these weaker maxima by the requirement that a maximum exceeds the values of all other points within a radius of 650 km. The inclusion here of weaker maxima which neighbour stronger maxima increases the chance of capturing an earlier stage of the lifetime evolution.

Defining the background PV

In the same way as for the 'edge PV', the local background PV is determined by the application of an area criterion. The threshold area is set to a suitably large (large synoptic) size such that, due to the relatively shallow large-scale PV gradients in the lower stratosphere, even a moderate increase in area will make little difference to the PV value of the bounding contour. The minimum value for the background PV is limited to 2 PVU. The background PV area criterion was set to 10'000 x 10^3 km^2, which is equivalent to the area of about 19 grid boxes of 10° longitude x 10° latitude centred at 65°N.

Determining the amplitude

The amplitude of a feature is defined as the difference in PV between the local maximum and the local background. An amplitude threshold can be introduced at this stage as for most procedures in the literature (e.g. Bell and Bosart, 1989, Dean and Bosart, 1996) so that only strong anomalies are tracked. However, in this study, we choose to set the amplitude threshold to zero for the sake of retaining properties of structures in low-amplitude stages of their lifetimes. A weaker restraint on amplitude thresholds is re-introduced at the analysis stage, i.e. amplitude is required to exceed a threshold during some point in the lifetime and not at every instant.

Minimum and maximum area constraints

The method described above could occasionally lead to the identification of rather large structures (substantially larger than the area threshold). If, for instance, the local maximum is very shallow, i.e. the discretisation ΔPV is large with respect to the local variation in PV, the defining area of the identified structure will be highly sensitive to a single step change in the PV value at the boundary. This situation could arise in a plateau of large PV values. If such a maximum is included in the set of selected maxima, a loss of information will result on all the surrounding weaker maxima. From a different perspective, if the area assigned to the maximum is approaching the background area, it can no longer be considered different or anomalous to the background. Extremely small structures may also be identified (if either the maximum PV is close to the minimum accepted value of 2 PVU or if they closely border a stronger PV maximum) which will not have a significant dynamical influence and may be due to non-meaningful fluctuations, i.e. noise inherent in the data set. Including such structures will decrease the tracking efficiency for larger structures as the chance of overlap between the advected structure and the identified structure is reduced.

A final check is therefore made on the area of the identified structures. Absolute minimum and maximum area thresholds of 40 x 10^3 km^2 (8 grid cells of 1° x 1° at 65°N) and 500 x 10^3 km^2 (a 10° longitude x 10° latitude box, centred at 65°N) respectively are imposed to prevent identification of unacceptably small or large areas.

Class

The identified structures are classified into the following three classes according to their vicinity to the dynamical 2 PVU tropopause:

I polar class - well inside the stratosphere,

II near-tropopause class - inside the stratosphere but close to the tropopause,

III cut-off class - outside the main stratospheric body.

A clustering of the isentropic PV field into areas with PV > 2 PVU enables cut-offs to be distinguished from the main stratospheric body. The latter is defined as the largest cluster. The
remaining clusters are considered to be stratospheric cut-offs. If any point of an identified strato-
spheric structure is within a distance of 3° latitude (3x11000 km) from the tropopause, it is said
to be near the tropopause and interaction between the structure and the tropopause is likely.
The classification of the stratospheric population of vortices into those near the tropopause (jet)
and those further north is motivated by the assumption of different behaviour in the different
domains. Horizontal and vertical shears associated with the jet stream are likely to have an effect
on the properties of vortices close by. In contrast, the cut-off anomalies far removed from the
tropopause will be less influenced by shear and, due to their isolation, might be expected to have
longer lifetimes.

Tuning for identification parameters
The parameters used in the identification procedure are listed in table 4.1 together with the values
chosen.
Suitable thresholds were chosen by systematically varying the parameter values and comparing the
objectively identified structures with those identified subjectively in the PV field. The effect of
changing the thresholds of selected parameters on the identification process is illustrated in figure
4.2. In the central panel (a) of the figure, the structures identified using the chosen parameter
values are displayed and referred to as the 'control'. In the surrounding panels, the effect of
changing single parameters from their control values are illustrated.
If the target structure area is reduced (b), the structures' forms are not as well resolved and con-
sequently many show an increase in isotropy (become more point-like). If the target structure
area is increased (c), elongated structures may be better resolved but large neighbouring struc-
tures become very close in places. Intermittent contact between neighbouring structures will lead
to reduced track lengths and false splitting and merger statistics. Reducing the minimum area
threshold (d) results in an increase in the number of smaller maxima which are unlikely to be
dynamically significant. For example, the extremities of thin filaments are resolved as tiny isolated
maxima, whereas in the control, just the main body of the filament is registered. Increasing the
minimum area threshold (e) has resulted in forced merger of a few structures as well as the loss
of features of interest including some near the pole (a potentially important source region). In
the control, the small polar features are captured for the date investigated. If the maximum area
threshold is reduced (f), suitably sized structures may broken down to reveal additional secondary
maxima which tend to be very small (see the elongated cut-off over the Atlantic). If the maximum
area threshold is raised (g), structures which were clearly isolated in the control become merged
and, due to their large size, will be favoured over smaller structures for continued tracking. A
broader distribution of sizes is noticeable. An extremely small limiting step (h) is ineffective in
narrowing the size distribution toward the target size as this level of accuracy has no significance
for the chosen size and resolution of the identified structures. With a large limiting step (i), the size
distribution broadens, and some structures which were reasonably sized in the control are forced
to become smaller.
Overall, it can be seen that the values selected in the control are reasonable. Although changes
in the size, shape and total number of identified structures can be detected, the distribution is
relatively insensitive to small changes in the chosen parameter values.
An example of the identified PV structures for the case study will be presented later in section 4.3
(figure 4.4).

4.2.2 Feature tracking
The flow in the lower stratosphere is quasi adiabatic. Our method exploits this property of adiabatic
flow, that is the quasi-material advection of PV on isentropes. Isentropic winds are used to advect
all points belonging to an identified structure at time \( t_n \), thereby predicting the new location and
form of the structure at the next time period, \( t_{n+1} \), 6 hours later. An outer boundary is assigned
to the advected points by defining the convex hull of the cluster. Isolated points which lie at a
radial distance of more than 2.5 standard deviations from the mean cluster centre, are removed.
Chapter 4. ISENTROPIC TRACKING

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Purpose of threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target structure area</td>
<td>$170 \times 10^3$ km</td>
<td>Determines the 'edge PV'</td>
</tr>
<tr>
<td>Minimum structure area</td>
<td>$40 \times 10^3$ km</td>
<td>Defines the smallest accepted structure area</td>
</tr>
<tr>
<td>Maximum structure area</td>
<td>$500 \times 10^3$ km</td>
<td>Defines the largest accepted structure area</td>
</tr>
<tr>
<td>Local background area</td>
<td>$10000 \times 10^3$ km</td>
<td>Determines the background PV</td>
</tr>
<tr>
<td>$\Delta$PV</td>
<td>0.01 PVU</td>
<td>Limiting step in binary search for edge and background PV</td>
</tr>
</tbody>
</table>

Table 4.1: Selected values for thresholds and constraints used by the identification procedure. See also figure 4.1.

If a significant overlap (an area overlap of 10 percent was found to be appropriate) exists between the advected structure and a structure identified at $t_{n+1}$, then the identified structure is given the same identity as its counterpart at $t_n$, and is said to have been tracked to $t_{n+1}$ (figure 4.3 (a)).

To improve the accuracy of the advection, it is carried out in two (3 hour) steps. The wind at the intermediate time step $t_{n+1/2}$ is calculated by linearly averaging the $t_n$ and $t_{n+1}$ isentropic wind fields.

Generally, this method exhibits excellent matches between predicted and observed structure contours (an example will be presented in figure 4.4 of section 4.3).

If the overlap does not satisfy the area criterion, the structure is given a new identity. However, it is also tagged as having had a small overlap, to distinguish it from the case of completely pure genesis, where no overlap with advected structures whatsoever is detected.

**Vortex properties**

The tracking procedure records the coordinates, PV and velocity of all points of each identified structure as well as properties representative for the whole structure: mean coordinates, mean PV, maximum PV, edge PV, amplitude, velocity at the PV maximum, isotropy (detailed below), and class (position with respect to the tropopause). The advected structure’s mean coordinates are also noted.

**Isotropy**

To assess the two-dimensional shape of the identified anomalies, a simple measure of isotropy is sought. The measure chosen is the squared ratio of the longest dimension of the structure to the diameter of a circle possessing the same area as the structure. If the structure has an elliptical form, the chosen measure is equivalent to the ratio of the minor to major axis length, i.e. the aspect ratio. If the true form departs from a perfect ellipse, the chosen method is biased towards lower isotropy because the major axis length always takes the maximum length scale, whereas the calculated minor axis length will be an average, rather than a maximum, of the shape’s extension in the perpendicular direction. Still, if there is a trend in aspect ratio over the lifetime evolution, this method should be capable of detecting a signal.

**4.2.3 Identification of physical events**

In addition to the tracking of the aforementioned properties at each time step for each vortex, the grouping of the individual vortex observations into tracks permits a detection of genesis and lysis events and also splitting and merger.

**Genesis and lysis**

The beginning track is defined as a genesis event and the end as lysis. Cases of splitting and merger constitute genesis. Details are given below. The means of lysis is not determined.

Genesis and lysis events are a small subset of the total number of observations. Therefore their
Fig. 4.2: An illustration of the effect of changing the thresholds of selected parameters on the bounding contours of the identified structures on the 320 K isentrope for 16th January 1992, 06 UTC. The central image (a) shows the structures identified when the selected values in table 4.1 are used for the thresholds. Note that areas are expressed in units of $1 \times 10^3 \text{ km}^2$. The top row shows the effect of changing the target structure area to 80 (b) and to 300 (c). The left hand panel shows the effect of setting the minimum structure area to 10 (d) and 120 (e). The right hand panel shows the effect of setting the maximum structure area to 400 (f) and 5000 (g). The lowermost row illustrates the effect of changing the limiting step size for the binary search to 0.0001 PVU (h) and 0.5 PVU(i).

climatological distributions tend to have a noisier appearance. For this reason, genesis and lysis distributions are additionally smoothed with a Gaussian filter, tailing off to zero at a radius of 2°.

Splitting and merger

The tracking methodology, i.e. matching advected and identified structures, lends itself to a simple detection of splitting and merger events (see figure 4.3 (h) and (c)). Two or more features at $t_n$
are said to have merged into a single identified structure at \( t_{n+1} \) if the latter has significant overlap with the advected areas of the former. The reverse is the case for splitting, where more than one identified structure at \( t_{n+1} \) experiences significant overlap with the same advected structure. For overlap to be considered significant in merger, the total overlap of all advected areas with the identified structure must exceed 10% of the identified structure's area. For splitting, the 10% overlap ratio threshold should be satisfied for each identified fragment individually.

An uneven merger (splitting) is said to occur if the area ratio of a small component (fragment) to the largest component (fragment) is smaller than a threshold of 0.65. In the case of uneven merger (splitting) events, the identity of the largest component (fragment) is transferred to (retained by) the merged (split) feature. Otherwise, both merger and splitting constitute genesis, and the participating identified structures receive new identities.

Occasionally there are 'hybrid' genesis events resulting from splitting from one structure and merger with another. In these cases, the genesis is labelled either as a splitting or a merger, depending on which process is most dominant. In the case that two identified structures are assigned the same (tracked) identity, they are submitted to further testing. The identified structure possessing the greater overlap with the common advected area is tracked further. The other is given a new identity.

**Track properties**

Once a track is complete, a summary of the properties relevant over the lifetime of a feature may be made, e.g. lifetime, changes of class, the proportion of time in each class, conservation of maximum PV, and velocity with respect to the background flow.

Tracks which start before the onset of the observation period but which terminate within it are excluded in their entirety and those starting within the observation period are included in their entirety. This ensures there is an equal number of genesis and lysis events.

Tracks may be selected and visualised according to a range of criteria e.g. lifetime, amplitude, genesis method or near-tropopause position.

**Tunable parameters**

There are two parameters involved in the tracking process. The first, is the threshold for overlap ratio of the advected structure(s) to the identified structure area. The same parameter features in the identification procedure and its value is kept the same (10%) for simplicity. The second parameter is the threshold defining an uneven merger or splitting. This is the ratio of the second largest fragment (component) to the largest involved in cases of splitting (merger). The value 65% was found to be reasonable.
4.3 Demonstration of tracking on case study

In this section, the identification and tracking procedures are applied to the dates of the case study of the cyclogenesis precursor introduced in chapter 3 (12th October 2000, 12 UTC, to 19th October 2000, 00 UTC).

Firstly, two consecutive time steps are shown to provide an illustration of the identified and advected structures. This will be followed by a presentation of the tracked characteristics of the precursor.

4.3.1 Demonstration of identification and advection

Figure 4.4 presents the PV field in panels (a) and (b), and identified structures in panels (c) and (d), on the 320 K surface just before the genesis of the precursor (12 UTC, 12th October 2000) and at subsequent time steps (18 UTC). The contours of the structures advected from 12 UTC to 18 UTC are superimposed in red at 18 UTC (d). The total number of identified structures is 45 at 12 UTC and 44 at 18 UTC. A total of 11 structures undergo genesis, of which 5 are due to splitting and 7 are not identified with a previously existing structure. A total of 10 structures undergo lysis. Merger is detected twice but identities from previously existing structures are transferred to the merged structures. Identity is transferred also for 2 structures resulting from splitting.

The figure provides good examples of all of the main processes described above (merger, splitting, pure genesis, lysis), except that for these dates there is no case of an identified feature with too small an overlap with an advected area. Examples of these processes are briefly pointed out below, along with one tracked feature which demonstrates a change of class.

- The precursor of the case study provides an illustration of a splitting event. At 12 UTC, the structure from which the precursor is formed (labelled ‘P’ in panel (a)) is situated at about 80°N and crosses the date line. At 18 UTC, a splitting (labelled ‘S’ in panel (d)) of this structure is detected by the routine, illustrated by the same advected contour encompassing the two identified fragments. The fragment labelled ‘P’ in panel (b) to the east of the date line develops into the precursor.

- An example (labelled ‘M’) of merger (two advected contours coinciding with a single identified structure) can be observed near 60°N, 80°W in panel (d). This structure remains merged beyond the shown time steps.

- A new structure (labelled ‘N’) has been identified at the tip of a streamer near 50°N, 160°N in panel (d) (no advected contour exists). The PV charts support the detection of a new structure, showing a ridge of PV at 12 UTC (a) but no local maximum. The latter is seen to appear at 18 UTC (b) and is of a relatively low amplitude.

- An advected contour overlapping no identified structure (labelled ‘L’ in panel (d)) signals that a lysis event has occurred near 45°N, 130°E, for a low-amplitude structure.

- The re-inclusion of a cut-off (labelled ‘C1’ in panel (c)) into the main stratospheric body (where it is labelled ‘C2’ in panel (d)), i.e. a change from class III to class II, can be seen near 70°N, 30°E.

The identified structures are small but do closely match the visual features. Additional anomalies of too small an amplitude to be detected subjectively are also identified. Note in (d) the close agreement in location and form between many of the advected and identified structures.

4.3.2 Evolution of tracked characteristics for case study

We now focus attention on to the single anomaly studied in chapter 3, starting with a description of single events encountered and then examining the evolution of tracked properties. The track of the anomaly is displayed in figure 4.5 and is shaded by the anomaly's class. The anomaly was
Fig. 4.4: Illustration of identification and advection for 12th October 2000 12 UTC (left panels) and 18 UTC (right panels). The top panels display PV on 320 K with identified structures outlined in black. The lower panels display the same identified structures with the PV background removed. The 2 PVU isoline is contoured in black in each panel. In (c) the identified structures are shaded according to their amplitude. In (d), the identified structures are shaded by the class (green - class I, yellow - class II, blue - class III). The outlines of the structures advected from 12 UTC to 18 UTC are shown in (d) in red. Event examples are labelled as M (merger), S (splitting), N (new i.e. ‘paw’), L (lysis). A structure observed to re-enter the stratosphere is labelled in (c) as C1 and in (d) as C2. The pre-precursor is labelled in (a) as pP and the precursor is labelled as P in (b).

identified and tracked for 6 days, i.e. 24 data times. Reference will be made to PV maps for this period from figure 3.1.
4.3. DEMONSTRATION OF TRACKING ON CASE STUDY

Fig. 4.5: The entire trajectory of the precursor produced by the tracking routine, from 12th October, 18 UTC, to 18th October, 18 UTC. The position of the precursor is marked with a point at 6 hour intervals. Grey points indicate class I classification (inner stratosphere), white points indicate class II classification (near tropopause).

4.3.3 Single events

Genesis

The tracking procedure detects genesis by splitting on 12th October 18 UTC (figure 4.4 (a), (b) and also figure 3.1 (a), (b)) in accordance with subjective observations and results from backward trajectory calculations in chapter 3. After the splitting, the two identified fragments increase their separation. The precursor moves towards Canada and the other fragment moves towards the pole.

Lysis

Lysis is registered on 18th October 18 UTC (see figure 3.1 (j)). At this time, diabatic processes associated with the occluded cyclone directly beneath it are active. The disappearance of the anomaly is due to its cross isentropic motion, discussed in chapter 3. No splitting or merger is observed to take place. The most noticeable changes in the tracked characteristics at the time of lysis is a sudden decrease in amplitude, the most rapid decrease over the entire lifetime.

Changes of class

The structure is initially classified as an inner stratospheric structure, i.e. class I. On 14th October, 06 UTC, it changes status to class II, indicating it is within 3° latitude of the jet. On 17th October, 06 UTC, the status is restored to class I, at which stage the anomaly is positioned over the surface cyclone and moving back towards the North.

Splitting

After the initial splitting which results in the genesis of the precursor, one further splitting is observed within the lifetime. A fragment of a very small size relative to the precursor is detected to split from the precursor at 12 UTC on 17th October, 12 hours after it is reclassified as a polar class
anomaly. It splits on the east side of the precursor and rotates about it cyclonically for 12 hours before disappearing, probably due to diffusion. The precursor retains its initial identity which seems reasonable as it is clearly the more dominant structure.

Other interactions
A weaker quasi stationary structure near the Labrador Sea (visible in figure 3.1 (g) and (h) with a maximum PV above 9 PVU) is attracted into co-rotational motion with the precursor as the latter approaches, from 12 UTC on 15th October to 00 UTC on 16th October. They remain close for a further 18 hours. The initially stationary structure is then incorporated into the dry intrusion behind the synoptic cyclone (figure 3.1 (i)). There is no noticeable deviation to the track of the precursor during this time.

There is a kink in the track on 17th October from 12 UTC to 18 UTC, which is when the structure centre is positioned directly over the surface pressure minimum of the induced surface cyclone (figure 3.1 (i)).

4.3.4 Time evolution of tracked characteristics

The tracked values of maximum PV, amplitude, isotropy and speed are presented in figure 4.6. A fit through the data points using a smoothing spline is overlaid.

Maximum PV
A rapid increase of PV occurs over the first two days, which was also noted in chapter a. Following the peak PV, an average decrease of around 1.5 PVU is recorded over about 1.5 days, after which the value fluctuates around 10 PVU.

Anomaly amplitude
In recognising that mobile upper troughs are regions of strongly sloping tropopause, Lackmann et al. (1997) related trough genesis to processes that lower and steepen the dynamic tropopause. In our study, if an anomaly is positioned near a steepening tropopause or is moving towards a steep tropopause, an increase in its amplitude may be registered due to a decrease in the background PV. In the case study, the initial increase in amplitude experienced by the anomaly is partly due to a decrease in the background PV, caused by a tightening of the PV gradient in its vicinity. A hint of change toward smaller amplitudes between the final two times is seen.

Isotropy
Initially isotropy is low, i.e. the structure is elongated (e.g. figure 3.1 (c)). Within about one day and on its approach toward the jet, the anomaly forms into a coherent structure and isotropy increases (figure 3.1 (d)). When the anomaly is within close proximity to the jet it enters a region of strong zonal flow and experiences large zonal shear to the south (14th October 18 UTC). Consequently, the trend in isotropy is reversed whilst the anomaly moves into the base of a trough, translating to the east (figure 3.1 (e) – (f)). As the anomaly turns to the north and enters the front of the trough, it starts to roll up again, taking the form of a small compact head with a sheared out extended tail (figure 3.1 (i)). The tail region, being a thin filament with no significant PV maxima, is not detected by the identification procedure and is excluded from the identified structure. The head becomes positioned over the surface cyclone and remains so until dissipation. The isotropy thus increases once again. These fluctuations in isotropy occur on a time scale of about 2.5 days and are superimposed on an overall positive trend over the lifetime.

Wind speed
The measure of speed presented is the average value over the identified area of the precursor of the isentropic wind field. The precursor experiences a faster flow as it approaches the jet. The speed is subsequently reduced as the anomaly moves into the downstream side of the large scale trough.
and turns to the north. A large decrease occurs on 17th October. The minimum recorded speed is close to zero at 12 UTC. This point in time coincides with the timing of the minimum tilt between the precursor and its surface cyclone (figure 3.1 (i)).

**Background PV**

The background PV is not shown explicitly in figure 4.6, but is included implicitly as the difference between the maximum PV and the amplitude. The general trend is an increasingly gradual decrease with time as the structure moves from being contained well inside the stratospheric body to a location closer to the tropopause.

**4.3.5 Summary of evaluation**

The identification procedure was shown to be capable of selecting the structures identified by eye in a background field of gentle gradient. Correspondence between predicted positions by forward advection and the identified structures was very high, supporting the appropriateness of the advection technique for isentropic tracking and its application for the detection of merger and splitting.

The developed methods force the identified structures to be of similar area. Fixing the size of the features is of benefit to the tracking procedure. If the size were allowed more freedom to fluctuate, the efficiency of tracking isolated small features compared to isolated large features would be reduced, due to the reduced overlap between identified and advected areas. If larger structures were permitted, their advected areas would be more likely to encounter and overlap neighbouring structures, rendering a correct labelling less likely and leading to reduced lifetimes due to the increased number of interactions. The possibility of monitoring a realistic size evolution of tracked features is therefore eliminated. However, changes in size might be inferred by changes in amplitude. Observations suggest, there is a slight increase in the size of near-tropopause coherent structures with amplitude (Hakim, 2000).

With reference to the case study of the cyclogenesis precursor, the objective tracking in this chapter and the subjective tracking of chapter 3 both suggest genesis on 12th October 18 UTC. The lysis times suggested by the subjective and objective tracking differ by just 6 hours. The objective tracking identifies a final structure associated with the precursor when subjectively it appears that the precursor has been disbanded.

Objective tracking supports the trajectory analysis that genesis was initiated by splitting. The later splitting of the coherent anomaly into a compact head and elongated tail detected by forward trajectories is not registered by the objective tracking, due to the narrow width of the tail. The identity is thus continued from the coherent anomaly to the compact head, which is not unreasonable. It can be assumed however, that some splitting and mergers involving narrow filaments will be neglected by the procedures used here.

The objective tracking records the trends observed subjectively in the evolution of maximum PV as can be expected. The objective tracking shows the initial increase in PV, as discussed in chapter 3, to be the largest change in the lifetime. In the second part of the lifetime, the PV is more conserved, except for a hint of decrease over the last 6 hours. The tracked amplitude and velocity seem reasonable. Changes in isotropy support the impression gained from subjective observations, despite the simple approach.

**4.4 Climatology of isentropic PV anomalies**

In this section, the general characteristics of isentropic PV anomalies are explored by applying the tracking routine to ten consecutive winters (1991/1992 to 2000/2001) to produce event climatologies (all events, genesis, lysis, merger and splitting) and composites in time of a selection of tracked characteristics.

The identification and tracking algorithms are run over an entire winter season covering the period
Chapter 4. ISENTROPIC TRACKING

Fig. 4.6: Time evolution over the precursor lifetime of a selection of tracked variables. The hour and date (dd/mm) are given on the horizontal axes. A smoothing spline is fitted through the data points.

1st December to 28th February. The distributions for the individual years are combined to form the 10-year climatology. Climatologies on four isentropic surfaces (310 K, 320 K, 330 K and 340 K) representative of the lowermost stratosphere in winter are produced. Each structure is represented just once per time step at its central longitude and latitude. Distributions therefore depict probabilities of an event occurring at a specific position and do not account for the structure size or the fraction of time that a particular location is under a PV structure.

4.4.1 Presentation of climatological distributions

Three types of distribution are calculated to represent geographical climatologies. A description of the measures follows.

Event density

Event density is calculated as the number of events (identified structures) falling within a circle centred on each grid point with a radius of 5° latitude.

Track density

To produce the track density, the number of tracks observed over the whole climatology passing within a circle with a radius of 5° latitude are counted. Each track is counted just once per grid point. To allow for centres that travel more than this radius within a time step, the track is interpolated onto the grid so that all points closest to the line of the great circle connecting track nodes become part of the track. This representation of results will decrease counts for quasi stationary
4.4. CLIMATOLOGY OF ISENTROPIC PV ANOMALIES

features and will increase counts for highly mobile features. The number of times a track is counted is proportional to its length, whereas in an event frequency distribution, the number of times a track is counted is proportional to its lifetime. Comparing the two types of representation, the ratio of the number of times a track is counted in the track density to the number of times it is counted in the frequency distribution will be proportional to its speed (provided the track centre does not double back on itself at any point). Maxima present in the frequency distribution which are not present in the track density are therefore mainly due to quasi stationary features. Other maxima in track density are likely to be shifted further towards the jet with respect to maxima in the frequency distribution.

Displacement velocity

The displacement velocity distribution is calculated by linearly interpolating the displacement velocity (i.e. the ratio of displacement between nodes to the time step) along the tracks and taking the average of all observations within a circle with a radius of 5 latitude degrees centred on a grid point as the displacement velocity at that grid point. Due to the highest velocities being close to the southern limit of the observations, there are sharp gradients at these locations. The figures should therefore be interpreted as providing the best estimate, given the observations available, of the displacement velocity for a PV structure at that location rather than the actual distribution.

4.4.2 Presentation of time composites

A lifetime threshold of 4 time steps (one day) is imposed for the time composites to eliminate transient structures and provide at least 4 points per track to view trends. Trends in tracked variables were inspected as a function of absolute age (relative to the instant of genesis) and as function of age normalised by lifetime.

4.4.3 Results and discussion

The presentation of the results are mainly restricted to the 320 K surface. Any striking differences between the results on different isentropic levels will be mentioned.

Lifetime frequency distribution

Figure 4.7 displays the frequency distribution of anomaly lifetimes ranging from 6 hours to 12 days for the four isentropic surfaces separately (a). The gradient of each curve has an almost constant exponent after an age of about 2 days until about 9 days, showing that, over this period, the frequency distribution is well modelled by exponential decay and that the probability of survival is constant per unit time. Lifetime frequency distributions presented by Lefevre and Nielsen-Gammon (1995) and Hakim and Canavan (2005) similarly show exponential decay. Isolated observations of longer lived features extend the tail of the distribution. The four curves in figure 4.7 show similar distributions but, by comparing the gradients of the curves, it can be seen that lower level features have greater survival probabilities once they have survived for over a day. The tendency of the 310 K frequency distribution to deviate from the exponential, for lifetimes less than one day, is possibly caused by the minimum PV threshold acting to remove low-PV features and thereby splitting otherwise continuous tracks into smaller segments of reduced lifetime. The longest lived feature was on 330 K and had a lifetime of 18.75 days. This was closely followed by the second longest lived feature, on 310 K, with a lifetime of 17 days.

A linear regression is applied to the 320 K curve as shown in figure 4.7 (b). Hakim and Canavan (2005) and Lefevre and Nielsen-Gammon (1995) found 6-hour survival probabilities of 90% (for tropopause potential temperature extrema) and 90% (for 500 hPa geostrophic curvature vorticity maxima) respectively. Our corresponding figure at 80% is somewhat lower, which maybe related to a slightly reduced tracking efficiency, due to tracking on the order of two times as many features in a smaller area (the structures are limited to having stratospheric PV), or it may be due to the smaller size and therefore inherently shorter lifetimes of our features or the higher altitude, or a combination of all the preceding factors. About 27% of structures on 320 K survive longer than
Geographical distributions

In figure 4.8, the geographical distributions of event density, track density and mean displacement velocity are presented for all tracks with a minimum lifetime of one day and which attain a minimum amplitude of 1 PVU.

Event density

The largest event densities are observed over north Greenland, close to the pole and the north Canadian Arctic. The distribution of secondary maxima over Alaska, Siberia and the Barents Sea form an annular band at about 70°N. Further maxima are located over the northeast of China at the entrance to the Pacific storm track, and over the southern part of the Sea of Okhotsk (northeast of Japan). With the exception of the pole and seas, these locations are all near high terrain. The largest densities are located over Greenland, which is also the region with the highest terrain north of 50°N and the third largest mountain in the Northern Hemisphere. A strong minimum occurs over the Arctic Ocean. The density there is reduced by around 50% in comparison to the mean density at 70°N.

Other climatologies also show correlations with high terrain, although different behaviours result according to the altitude of the observations. The 500 hPa climatologies show that maxima tend to be located downstream of major orographic barriers (e.g. Bell and Bosart, 1989, Hakim, 2000, Lefevre and Nielsen-Gammon, 1995, Dean and Bosart, 1996), particularly when there is a northerly component of mid-tropospheric flow (Sanders, 1988). In tropopause climatologies, they tend to be co-located with or downstream of high topography (Hakim and Canavan, 2005).

Climatologies by Dean and Bosart (1996) and (Bell and Bosart, 1989) show similarly positioned maxima over central Siberia and the north Canadian Arctic and all aforementioned climatologies produce maxima near the entrance to the Pacific storm track, but the 500 hPa troughs also show maxima over the far east Pacific.

The tropopause climatology of Hakim and Canavan (2005) corresponds the closest to our results. In particular, the maximum over north Greenland, absent in the 500 hPa climatologies, is produced as well as a minimum over the Arctic Ocean. Different features include a maximum positioned near the Aleutian Islands and the presence of a SW-NE oriented band across the North Sea. Their maxima near Alaska and across Siberia are shifted further to the south than ours which may be due to bias inherent in their grouping data into longitude-latitude grid boxes (Taylor, 1986), or to the tropopause being closer to the surface (by about 100 hPa at the location of Alaska) where frequencies are reduced over topography.

Another prominent feature of our event frequency distributions is the increased frequencies very close to the pole. The maximum has an increasing dominance the higher the isentropic surface, is less evident in the track densities, and has a mean displacement velocity of zero or close to zero. These observations suggest that feature is a quasi stationary intrusion of the polar vortex1. The proximity to the Greenland maximum suggests that the two features may interact or be related. Close inspection of figure 2 of Hakim and Canavan (2005) suggests they may have observed the same feature, but the latitudes very close to the pole have been obscured. The NCEP track densities (on 330 K) of Hoskins and Hodges (2002) show a ridge of high densities over the pole, resembling our track density at 330 K, while their ERA-15 results show an even more prominent polar maximum.

Track density

Track density differs from event density in that every tracked feature is only counted once per area bin (here a 'polar cap' of 5° radius). Large counts from slow moving systems remaining in the same region are reduced, and the track is interpolated along its entire length to the grid (not just

1Note that counts directly on the pole are not included in the climatology and so the value there is an average taken from the surrounding grid points.
4.4. CLIMATOLOGY OF ISENTROPIC PV ANOMALIES

Fig. 4.7: Frequency distribution of anomaly lifetimes for the entire climatology. (a) The four curves show the frequency distributions on the different isentropic levels 310 K (dot-dashed), 320 K (dotted), 330 K (dashed) and 340 K (solid). (b) The curve for 320 K is shown with a linear regression applied between ages of 2.25 and 8.5 days. \( R^2 \) takes the value 0.949. The 6-hour survival probability over this age range, is 80.0%.

at its location at the data time), thereby increasing counts for faster moving systems. The effect of this change is that maxima in track density occur further to the south, closer to the jet, than maxima in the event density.

The broad pattern of the track densities (figure 4.8 (b)) reflects the shape of the planetary scale troughs of the climatological winter isentropic circulation (see for example, Martinus et al. (2005) or chapter 5) with the exception of the strong minimum over the Arctic Ocean.

More specifically, the track density reveals a high concentration of tracks to the north of the Atlantic and Pacific jet entrance regions. It is the mobile disturbances which are more likely to have an effect on surface development (Sinclair, 1994) as they have greater chance of encountering lower-level baroclinicity or interacting with the jet. It is therefore not entirely surprising to see that there are high track frequencies directly upstream of the surface storm tracks.

Elsewhere, large densities are found in a band oriented northwest-southeast over Northern America and in a second band connected with the maximum over Greenland and extending from north Iceland to Scandinavia and into central Russia. The northwest-southeast orientation of the high track densities over Northern America has a link with documented phenomena: Lackmann et al. (1997) point out that mid-tropospheric frontogenesis is favoured in northwesterly flow over central North America. In most of the cases they observed, tropospheric disturbances acting as precursors of rapid cyclogenesis were formed into coherent structures and were advected to the coast in such flow.

Track densities of PV anomalies on 330 K are presented by Hoskins and Hodges (2002) for a winter season climatology of the years 1979–1999 for the ERA-15, supplemented with ECMWF analyses for the later years, and NCEP data. As they do not restrict PV to stratospheric values, their maxima can occur further south and, due to the removal of a planetary wave background (wavenumbers less than or equal to 5), are detectable over jets and the storm tracks (the focus of their study). Our results on higher isentropes (340 K), where the stratosphere extends further south, show a better resemblance to their 330 K field. The maximum over the Pacific jet entrance is observed in the NCEP track density and also in their track densities of negative potential temperature anomalies on the 2 PVU tropopause, but is absent in the ERA-15. They suggest the reason for this absence is that tracks were rejected by their failure to satisfy their imposed track smoothness criterion.

The (NCEP) fields of Hoskins and Hodges (2002) show a weak ridge of higher frequencies turning toward Greenland from the west Atlantic storm track and a local maximum over north Greenland. The clearer local maximum over Greenland in their track density is possibly due to the removal of a background field. Their data also confirm the presence of an area of low frequencies over the Arctic
and Aleutian Islands and a local minimum in frequencies between Greenland and the Labrador Sea. This latter minimum, seen in both our track and event density, is absent in the event density climatologies at 500 hPa, but has a weak amplitude in the tropopause event frequency climatology of Hakim and Canavan (2005). The main differences between the results of Hoskins and Hodges (2002) and ours are thus due to the detection by their method of anomalies near and to the south of the jet, whereas we intentionally restrict investigations to stratospheric PV anomalies.

Displacement velocity

The maximum amplitudes of displacement velocity (figure 4.8 (c)) are co-located with jet entrance regions over the Pacific and Atlantic. The Pacific maximum is noticeably stronger. Ridges of large displacement speeds on the downstream side of these maxima extend well out into the storm tracks. Vortices accelerate (tracks converge) into the Atlantic maximum from the Pacific belt of maxima or from the Arctic Ocean. Tracks converge (vortices accelerate) into the Pacific maximum from the Atlantic belt and a lower latitude branch over the Middle East. Displacement velocity amplitudes are very low between Greenland and the Canadian Arctic and over the pole, east Siberia and Alaska, indicating that a substantial proportion of the maxima in vortex density observed in these regions are due to quasi-stationary high PV structures.

Event genesis and lysis densities

Event genesis and lysis densities are presented in figure 4.9. Genesis densities are high over the Greenland landmass (particularly to the north), the Canadian Arctic, and Alaska. Secondary maxima are distributed about the annular band of high vortex density with isolated maxima near the entrance to the Pacific storm track and over the North Sea and Scandinavia. Lower genesis frequencies within the annular band are seen in the lee of the Rockies over Canada into the Arctic and from the Bering Sea into the Arctic. Lysis regions are similar to genesis regions but are slightly downstream and more evenly distributed around the longitudes. In addition to the aforementioned genesis locations, high lysis frequencies are found over Labrador, south Greenland and in the lee of the Rockies over northern Canada. When switching back and forth between genesis and lysis distributions one obtains a subtle impression that there is a shift from a more zonal distribution (genesis) to a more perturbed wavenumber 2 distribution with more southerly lysis regions over the storm track entrances.

The 'genesis minus lysis' frequency distributions show that genesis broadly dominates over lysis on the eastern part of the Pacific and Atlantic Oceans and western sides of the American and European continents, i.e. at the end of the storm tracks. Genesis strongly dominates over the west coast of North America, northwest of Greenland and north of the Tibetan Plateau. Lysis strongly dominates in the Sea of Okhotsk and northeast of Greenland, northwest Canada in the lee of the Rockies and in the Arctic Ocean close to eastern Siberia.

Winter climatologies of troughs and closed circulations at 500 hPa also show high genesis frequencies over the Canadian Arctic and at the entrance to the Pacific storm track e.g. (Lefevre and Nielsen-Gammon, 1995, Bell and Bosart, 1989) but the maxima associated with the high terrain over Greenland and Alaska are absent. At 500 hPa, frequency maxima occur preferentially downstream of more southerly orographic barriers (east of the Rockies and Tibetan Plateau), whilst high lysis frequencies occur upstream of orographic barriers, over the two ocean storm tracks and central Siberia. Lefevre and Nielsen-Gammon (1995) and Sanders (1988) report that there is a tendency for genesis to exceed lysis over land (upstream of oceanic storm tracks) and lysis to exceed genesis over oceans (within and downstream of storm tracks). However, there is also a region of genesis recognisable in the east Atlantic 'genesis minus lysis' distributions of Lefevre and Nielsen-Gammon (1995) and Dean and Bosart (1996). Lysis is also observed to dominate over the pole (Dean and Bosart, 1996). The apparent inconsistency that genesis dominates at the end of the storm tracks and upstream of raised topography for our climatology, where in other climatologies lysis dominates, can be at least partially explained by our inclusion of splitting events into the genesis climatology: The end of the storm tracks and upstream of orographic barriers are regions of diffluent flow, frequented by streamers (Martius et al., 2005). Wave breaking and the redistribution of positive PV by diabatic processes will result in the elongation and splitting of PV structures.
4.4. CLIMATOLOGY OF ISENTROPIC PV ANOMALIES

(a) Event density  (b) Track density

(c) Displacement velocity

Fig. 4.8: Winter climatology of PV anomalies identified on 320 K: Event density (a), track density (b), mean displacement velocity in m s\(^{-1}\) (c). Densities show 10-year winter-season totals over circular areas of 5° radius centred on each grid point. Lifetimes are at least 1 day. Amplitudes of the included structures attain at least 1 PVU within a lifetime.

Where splitting dominates merger, genesis will dominate lysis. Splitting is also responsible for producing smaller troughs and vortices which may miss detection in other climatologies.

Splitting and merger

Event distributions for genesis by splitting and merger are shown in figure 4.10 (a) and (b). Both distributions exhibit high frequencies over Greenland, although splitting is more common in the northern part, and Alaska. Secondary maxima are found in the zonal band of high vortex frequencies. Differences are that merger tends to reflect the undulation of the climatological troughs, whilst the splitting distribution is more zonally oriented.

In section 4.2 it was explained that if an uneven merger occurs, the identity of the more dominant constituent part is transferred, i.e. one track is terminated but no new track is created. The merger event is still registered but at an intermediate stage of the lifetime of the continued track. Similarly, there can also be splitting events recorded at intermediate stages in the lifetime of a track.
Fig. 4.9: Frequency distributions on 320 K for genesis events (a), lysis events (b) and 'genesis minus lysis' (c). Densities show 10-year winter-season totals over circular areas of 5° radius centred on each grid point.

Results for all merger and splitting events are displayed in figure 4.11 (a) and (b). The climatological distributions of these events mirror their distributions at genesis but are suggestive of a small displacement to the east of merger frequency maxima compared to lysis maxima. In contrast to splitting events, the frequency of merger is noticeably reduced over the eastern Atlantic, the eastern Pacific and Hudson Bay. Thus the ratio of splitting to merger events increases in these locations. A tendency of increased frequencies of splitting in comparison to merger over the east ocean basins is also evident in the 500-hPa trough climatology of Dean and Bosart (1996).

There is a detectable difference, particularly over North America, between the orientation and origin of tracks generated by splitting and those generated by merger (figure 4.11 (c) - (d)). Splitting events exhibit a more zonal region of high track densities over the east coast. The orientation of the ridge of high track densities extending back toward the west coast is suggestive that tracks entering the zonal region have an origin over the Pacific basin. For merger events, the tracks entering the east coast density maximum appear to come more from the west Canadian Arctic and the trough in track density is more prominent, which is linked to the relative reduction of merger events over the east Pacific Ocean.

In addition to the maximum over the east American coast (Atlantic storm track entrance), both
4.4. CLIMATOLOGY OF ISENTROPIC PV ANOMALIES

(a) Splitting

(b) Merger

(c) Pure genesis

(d) Small overlap

Fig. 4.10: Genesis event distributions on 320 K for the four different categories of splitting (a), merger (b), pure (c), new with less than the threshold overlap (d). Densities show 10-year winter-season totals over circular areas of 5° radius centred on each grid point.

splitting and merger distributions show a maximum associated with the Pacific storm track entrance; a result also true of Dean and Bosart (1996). The merger distribution has a third prominent maximum in the Eurasian trough over northwest Russia.

Dean and Bosart (1996) link a spring maximum in blocking activity to spring merger events being more common than autumn merger. It would be of interest to see if the isentropic merger and splitting distributions presented here change with season and support their suggestion.

Pure genesis

Another category of genesis registered by the tracking procedure is 'pure genesis'. These events constitute identified structures which have zero overlap with advected structures from the previous time step. It is possible that a number of events are incorrectly registered as 'pure genesis' where the routine fails to connect two local maxima between time steps, either due to inaccuracy of the advection method or non-conservation of PV. If this is the case, we might expect to see a correlation between the climatologies of pure genesis and the climatologies low overlap ratio. Both climatologies are presented in figure 4.10 (c) and (d). The two climatologies share frequency max-
Fig. 4.11: Genesis and merger events on 320 K. Top panels show all splitting events (a) and all merger events (b). Lower panels show track densities for tracks initiated by splitting (c) and merger (d). Lifetime (1 day) and amplitude thresholds (exceeds 1 PVU within lifetime) apply. Densities show 10-year winter-season totals over circular areas of 5° radius centred on each grid point.

ima over the central Canadian Arctic and on the north coast of Alaska but otherwise the maxima are located in different regions, lending credibility to the assumption that the features are new. That they share a maximum over north Canada is probably due to there being a greater frequency of vortices there. The distribution of pure genesis is actually very similar, more similar than our total event density, to the annual cyclone density of tropopause vortices obtained by Hakim and Canavan (2005). Another explanation for the detection of 'new' structures however needs to be considered, namely that the appearance of these new vortices could be due to structures being transported by the mean flow into areas of increased resolving ability of the assimilation network, where they are suddenly detected. This is possibly the cause of the ridge of high values along the western coast of North America and maybe the strip along the northern boundary of west Siberia. The remaining regions are an elongated band extending from east of Hudson Bay past the Greenland tip and Ireland and on to Spitzbergen (a similar belt is observed by Hakim and Canavan (2005)), a region to the north of the west Himalaya and east of Mongolia. Non adiabatic effects (latent heat release and friction) could be responsible for PV generation in the Himalayan and Mongolian locations. Notably, frequencies are low over central Greenland, showing that the
large genesis densities there are mainly due to splitting and merger. The pure genesis frequencies are relatively low overall, in comparison with the other climatologies and account for just a small proportion of total genesis. The period of study should be expanded to test the suggested explanations in this paragraph.

Long-lived structures
A second climatology was composed of structures lasting a minimum of four days and obtaining an amplitude of at least 1 PVU in their lifetimes. The results are shown in figure 4.12. The longest-lived structures have a preference to be located over the Eurasian continent. The event and track density distributions are suggestive that a frequent path way is zonal motion across southern Europe with a transition to higher latitudes over central Russia into Siberia and then into the Pacific storm track. Frequencies are substantially reduced over the major ocean basins and the American continent, although there is still a maximum in the Canadian Arctic and north Greenland. A strong maximum exists at the pole, revealing that the quasi stationary centres there tend to have long lifetimes. The southerly latitudes of the distribution maxima over southern Europe give an indication that the features there are cut-offs. They re-enter the stratosphere over Russia.

Large amplitude anomalies
A climatology requiring an increase in amplitude of least 3 PVU within the lifetime was made for structures surviving at least one day. Results (figure 4.13) show high densities are primarily distributed over North America (excluding the Canadian Arctic), the North Pole, Greenland and the Mediterranean. The highest event frequencies are over Turkey and Greenland. There is a very dominant region of high track densities coincident with the Atlantic storm track, fed by tracks from the Pacific region.

Results from Hoskins and Hodges (2002), who show mean intensities of PV anomalies on 330 K, support the tendency for larger intensities to be observed over the Mediterranean and North America and reduced intensities over Asia. Their maxima are located to the south of ours, which is probably due having no restriction on the lower limit of PV values, and they show an additional region of intense structures to be located over the east sector of the Pacific storm track. This region could be absent in our data due to the small total numbers of tracks there, presumably related to a decreased likelihood of observing large amplitude features in that location.

Classes
A climatology of strong amplitude (> 3 PVU) structures (not shown) which remain as polar class vortices (class I) shows preferred locations over north Greenland and secondary maxima centred north of Hudson Bay and over eastern Siberia, north of the Sea of Okhotsk. Thus the vortices have a tendency to be located in either of the two major troughs of the climatological winter circulation. Frequencies fall to zero in the centre of the Arctic Ocean.

Different isentropic levels
A comparison follows of the distributions on other isentropic levels with those on 320 K. Figures are not presented. Inspecting the event densities from 320 K to 340 K in sequence, the Arctic maximum to the north of Greenland seen in figure 4.8 (a) becomes the most dominant feature. This would seem to suggest the increasing domination of a persistent polar vortex, although there is still a strong minimum over the central Arctic Ocean. The polar maximum becomes stronger again at 310 K, due to the restricted equatorward extent of the stratosphere. Preferred regions poleward of 60°N on other levels are similar to those on 320 K, but as the stratospheric area increases with higher isentropes, new maxima appear at lower latitudes over central Europe and close to the tropopause over Asia. At 340 K, the maximum over the Canadian Arctic is no longer a prominent feature.

Large track densities occur in tighter bands, moving to lower latitudes with the tropopause over the Atlantic and Pacific jet entrances. On changing surface from 330 K to 340 K, the preferred
regions over central Asia shift from a zonal band near 60°N to a zonal band 20 degrees further south.

Translation velocities increase with isentropic level by, for example, 1 ms$^{-1}$ per 10 K on average for the Atlantic maximum over the range of surfaces studied.

The 'genesis minus lysis' frequency distributions show interesting changes ascending the levels. At 310 K, the major patterns are similar to 320 K except for over Greenland. The most dominant genesis location is over the Greenland tip on the 310 K surface whereas at 320 K, lysis dominates in this location. At 330 K, the region where genesis dominates over western America becomes stronger and extends further south. On 340 K, it becomes evident that there is a zonal wave-train like pattern extending from North Africa through to east China, close to 40°N, coincident with the region of increased track density over Asia on this level immediately to the north of the climatological tropopause. The pattern exists on lower levels, phase shifted to the east, but is not as striking.

Merger events have high frequencies near the entrance to the Pacific storm track, Alaska and Scandinavia for all levels. However the Greenland maximum weakens and frequencies increase over the US (330 K) and by 340 K a strong band extends from central US northwestward to Alaska, coincident with the Rockies. Over Europe, frequencies are raised over the Alps. The

---

(a) Event density  
(b) Track density  
(c) Displacement velocity

*Fig. 4.12: As for 4.8 but for features lasting at least 4 days.*
4.4. CLIMATOLOGY OF ISENTROPIC PV ANOMALIES

Fig. 4.13: As for 4.8 but for features attaining an amplitude of at least 3 PVU in their lifetimes.

most noteworthy changes to the splitting distribution are the reduction of the Greenland and Alaskan maxima, the emergence of a zonally oriented maximum over the central US, and increased frequencies over Europe but located to the north of the merger maximum, over north Germany. The most noticeable change in the distribution of pure genesis vortices is a shift in the maximum over North America from Alaska on 320 K to the east coast on 330 K. The maximum at the start of the Pacific storm track remains.

4.4.4 Time evolution characteristics

Composites of several tracked characteristics for the 320 K surface (PV, amplitude, isotropy and speed) are displayed as a function of a structure’s fractional age in figure 4.14 and as a function of absolute age in figure 4.15.

Amplitude

A maximum amplitude is experienced during the first half of the lifetime, 2-3 days after genesis on the lower levels, and later for longer lived features on the highest levels. The starting amplitude is noticeably larger than the final amplitude. This is likely to be due to splitting and merger
Chapter 4. ISENTROPIC TRACKING

events, which tend to produce new structures of relatively strong amplitude, being responsible for a substantial proportion of the geneeses.

The relative changes in maximum and minimum amplitude over the lifetime are similar for all levels but larger absolute amplitudes occur on lower levels. Short lived structures decay most rapidly on lower levels.

It is evident that the amplitude does not remain constant over the lifetime and that lysis takes place following a reduction in amplitude.

Maximum PV

Maximum PV remains large during the first part of the lifetime (for about 3 days for lower levels, and 4 days for 340 K). A decrease is then observed. Longer-lived vortices experience greater decreases in maximum PV.

Isotropy

The change in composite isotropy over the lifetime is small but exhibits a small positive trend, which is more pronounced on lower levels. A vortex will undergo changes in isotropy when it encounters a region of shear. One explanation for the reduced trend at higher levels could be the increase in the area of the stratosphere relative to the length of its boundary (the tropopause), i.e. there is a relative increase of the number of structures existing away from the near tropopause shear zones.

That the trend is not larger, could be explained by the occurrence of several large fluctuations between high and low values over a typical lifetime, as observed for the case study. If a vortex is generated by splitting or merger, the isotropy might be expected to be initially small, to increase as the vortex wraps up, then to decrease again if the vortex encounters a new region of shear. Splitting or diffusion of thin PV filaments at the extremities of an elongated structure will act to decrease the length of the major axis and thus return the isotropy to larger values.

Speed

A local minimum in background velocity (not displacement velocity) is experienced at about 2 days after genesis, which coincides with the time at which the maximum amplitude is observed. If the cyclonic rotation associated with an anomaly strengthens compared to the background flow, the average speed at the vortex centre will be reduced. In other words, the decrease in speed at the vortex centre observed in the composite could be a signal that the anomaly has its greatest impact on the surroundings when the amplitude is largest. Speed is more closely (inversely) correlated to changes in amplitude than to changes in PV, implying that a decrease in local background PV (e.g. as a vortex approaches the tropopause or enters a tropopause fold) is instrumental in augmenting the impact of the anomaly on the background flow.

4.5 Summary and outlook

In this chapter, a tool developed for the tracking of isentropic PV anomalies has been presented, applied to a case study and used to produce a 10-year winter climatology for lower middleworld isentropes.

A decaying exponential lifetime distribution was found on each level, consistent with other results from upper level feature tracking in the literature. Most structures are short lived but over 25% of vortices on 320 K survived longer than four time periods (one day), and 14 tracks survived longer than 10 days, suggesting that the anomalies observed on ERA-40 isentropic maps cannot be assumed to be transient structures due to interpolation or assimilation errors.

Geographical distributions reveal a tendency for high event frequencies to occur near high topography and in an annular band centred around 70°N. The more mobile features have high frequencies upstream of storm tracks. Prominent genesis locations are Greenland, the Canadian Arctic and
4.5. SUMMARY AND OUTLOOK

Fig. 4.14: Time evolution of a selection of tracked variables for structures on 320 K. A minimum lifetime of 4 days and a lifetime minimum amplitude of 1 PV are enforced. Tracks are grouped into sets with the same lifetime. Evolution is displayed as a function of fractional age, i.e. the absolute age normalised by the structure's lifetime. Averages are made for all available results between 0 days and an upper limit of 4, 5, 6, 7, 8, 9 and 10 days separately, accounting for the 7 curves in each panel. The set of curves show the effect of changing the upper limit used in the averaging. Curves representing structures of shorter lifetimes tend to be smoother.

Alaska. Genesis dominates lysis near storm track exits, northwest Greenland and north of the Tibetan Plateau. The distribution of tracks initiated by a merger of two or more structures compared to the distribution of tracks initiated by a splitting, suggest that merger is favoured in deep planetary troughs and splitting in weaker troughs with stronger zonal flow.

A composite of the time evolution of vortex properties was constructed for vortices surviving more than 4 days. A shallow and constant increase in isotropy resulted. The amplitude and PV peaked during the first half of the lifetime and a sharp decrease in amplitude before lysis appeared to be common. The background velocity composite revealed a mid-life minimum, which was hypothesised to be a signal of the time of strongest influence of the anomaly on its background flow and in the case of mid-latitude anomalies, would coincide with the vertical alignment of an anomaly with a surface cyclone.

The current capabilities of the presented tracking tool open up many further analysis avenues for the investigation of stratospheric PV anomalies which have not yet been explored or commented on in this text. Some suggestions for further investigation are to

- Extend the analysis to all seasons, higher isentropic levels and the Southern Hemisphere
- Explore the distributions and vortex property evolutions for tracks originating, terminating or passing through localised regions of interest e.g. the Canadian Arctic
- Analyse inter-annual variability and links to common indices of teleconnective activity e.g.
Chapter 4. ISENTROPIC TRACKING

Fig. 4.15: The same as for figure 4.14 except that lifetime is expressed as absolute age (no normalisation is performed).

- Investigate differences in behaviour and properties for the different classes of structures (classes are defined in section 4.2)
- Explore the connection between track position and temporal evolution of e.g. isotropy or PV conservation
- Use the climatology to single out cases with particular characteristics for further study.

The tool could be further developed to

- Assess isentropic interactions between anomalies, as was observed in the case study
- Connect tracks of the same 3D anomaly appearing on different isentropic levels, with view to exploring the isentropic lysis process
- Track isentropic local PV minima.

In the next chapter, a complimentary climatology is presented to reveal preferred geographic locations of the uppermost stratospheric PV values. The resulting frequency distributions will be proportional to the fraction of time that a specific point can be expected to be under the influence of a region of strong PV.
Chapter 5

Middleworld Climatologies

In this chapter, two ERA-40 monthly climatologies of isentropic PV are presented for the 10-year period (December 1991 – November 2001) for a large range of isentropic levels. The first climatology is a frequency distribution, showing the temporal frequency with which stratospheric PV exceeding a certain threshold is observed at each position. The purpose of this climatology is to help identify locations within the stratosphere where anomalously high PV is most likely to be found. The second climatology is simply the mean PV distribution and is constructed for comparison. A brief review of PV climatologies in the literature is given in section 5.1. The methodology and presentation of the climatologies follow in sections 5.2 and 5.3. The results are discussed in section 5.4 and a summary is given in section 5.5.

5.1 Introduction

The broad scale distribution and seasonal variability of the mean isentropic PV for both hemispheres has been documented recently for the troposphere and lowermost stratosphere. Climatological mean patterns strongly resemble the stationary wave patterns seen in climatologies of geopotential height.

The following brief description is a summary of remarks of Liniger and Davies (2004) and Martius (2005). In the winter, strong troughs and PV gradients occur over the North-Eastern American continent and North-Eastern Asia upstream of the storm tracks, whilst a weaker trough occurs over Eastern Europe. Together they produce a wavenumber 3 pattern. Ridges occur downstream of the stronger troughs over the eastern Pacific to North-Western North America and over the Eastern Atlantic to Western Europe. In the summer, a less pronounced pattern of wavenumber 5-6 is observed and strong areas of anticyclonic wave breaking are evident over the subtropical central Atlantic and Pacific. Strong PV gradients occur in the region of the tropopause break in conjunction with the strong troughs and lack of ridging across the Asian continent and the strong Asian monsoon anticyclone. PV gradients are on the whole weaker on the lower isentropes, in the exit regions of the storm tracks and in the summer.

Recently, subsequent to the initiation of this study, the ECMWF issued an ERA-40 Atlas (Källberg et al., 2005).1 The Atlas describes the climate of the period 1979 – 2001. Among many other fields, seasonal and annual PV means are mapped for both hemispheres on 300 K, 315 K, 330 K, 350 K, 530 K and 850 K. The interannual variability is also included. The larger variabilities are found over the poles on every isentrope, as well as over the storm tracks for 315 K to 350 K, in mountainous regions for 300 K, and over the equator and to a lesser extent near the pole on 530 K and above.

In previous studies, typically climatologies only for winter or summer months are computed and only one or two surfaces in the range 310 K to 350 K are selected for presentation (e.g. Brunet et al., 1995), depending on the application. The 310 K surface, for example, has been used to represent extra-tropical conditions e.g. Martius et al. (2005) and the 350 K surface has been used to depict the subtropical flow setting e.g. Postel and Hitchman (1999).

1The Atlas is available at http://www.ecmwf.int/research/era/ERA_40_Atlas/index.html
The variability of the climatological PV distribution is usually presented as the standard deviation of the data (both positive and negative anomalies) from the climatological mean. Fewer climatologies have determined the distribution of positive anomalies separately.

Several studies have addressed this issue indirectly by investigating the distribution of features related to positive PV structures. Positive PV structures on isentropic surfaces are mainly due to vertical undulations of PV isosurfaces which intersect the isentrope (for example, follow in figure 3.4 (a) the undulations in PV along the 320 K isentrope), alternatives being the presence of a dominant maximum from the stratospheric polar vortex near the pole, or diabatically produced PV anomalies. Dynamical structures associated with such undulations are streamers (stratospheric intrusions), tropopause folds and cut-offs. Isentropic climatologies of these structures are available (see Martius et al., 2005, Croci-Maspoli, 2002, Wernli and Sprenger, 2006) and are likely to be related to the climatology presented in this chapter near tropopause latitudes.

A particularly illuminating technique for assessing local high-frequency isentropic PV variability which does treat positive and negative anomalies separately was presented by Liniger and Davies (2004). Rather than fixing a specific isobaric or isentropic height for the climatology, interseasonal comparison is eased by the individual selection of a representative isentrope according to its mean latitudinal tropopause position. The PV anomalies are however defined with respect to a climatological mean and areas of high PV occurring in a region of where the climatological value is also high would be excluded from the distribution.

In this study we would like to be able to detect areas of high PV even if they occur in a region of where the climatological value is also high. The method is outlined in the next section.

5.2 Methodology

5.2.1 Data set and range of isentropes

Global isentropic climatologies are constructed for each month from the 6-hourly fields of the ERA-40 data for the period December 1991 to November 2001. The hydrostatic form of Ertel PV is used in these climatologies and has been interpolated from pressure levels onto isentropic surfaces.

The chosen isentropic surfaces are concentrated over a range of levels which lie within the middleworld or just above it. The lower surfaces in particular undergo a seasonal cycle in altitude and stratospheric area, and some surfaces are only part of the middleworld for part of the year. For example, the PV climatology study of Liniger and Davies (2004), shows for the period studied that the lowest middleworld isentrope in the northern hemisphere winter is near 293 K, whilst in the summer, the lowest isentrope is about 324 K. For simplicity, a potential temperature range which embodies the middleworld for all months is fixed for the entire year. The levels 290 K to 420 K, at 10 K intervals were selected.

In addition to these levels, a few further surfaces were selected (separated by wider intervals) to represent the middle atmosphere (overworld). The dynamics of the polar vortex, which can influence the lower stratosphere particularly during late winter and spring (during sudden stratospheric warmings (SSWs)) is often depicted on the 10 hPa or 850 K surface (e.g. O'Neill and Pope, 1985, Baldwin and Holton, 1988). The 850 K surface is chosen as the limiting upper level in the climatologies, with additional surfaces at 460 K, 500 K, 600 K and 700 K inserted to represent the transition from the middleworld into the polar vortex of the overworld.

5.2.2 Stratospheric high PV frequency climatology

The purpose of this climatology is to identify locations where anomalously high stratospheric PV is most likely to be found. High PV will be defined here as PV which exceeds a fixed threshold for the month.

The PV threshold is determined for each selected surface for each month of every year following the steps below:
5.2. METHODOLOGY

(1) At each data time, the 2D PV field is clustered into regions with PV > 2 PVU.

(2) It is determined which clusters are considered to belong to the main stratospheric body. These clusters are selected for the climatology and the remaining cut-off structures are excluded. The details of this step are outlined in section 5.2.3.

(3) A threshold is selected for each month individually such that, in the monthly mean, 20% of the total stratospheric area (as defined below in section 5.2.3) possesses a PV exceeding that threshold.

Figure 5.1 provides an illustration of the areas on the instantaneous 330 K PV field of 1st May 1995, 00 UTC, which satisfied the threshold for that month and were included into the climatology.

Once the threshold has been determined, it is applied to the PV field at all times over the month on the relevant surface. The total number of times at which the threshold is exceeded at a point is counted and normalised by the number of observations in the month at that point, giving a frequency in the range of 0 to 1. Points exceeding the threshold which lie outside of the stratospheric body are excluded from the climatology.

Monthly high PV frequency climatologies are initially constructed for each year separately. The results for all years are then combined with equal weighting into a 10-year monthly climatology. The resulting distribution indicates where the largest PV values are most frequently found.

5.2.3 Identification of the stratospheric body

The following explanation specifies how the stratospheric body on an isentropic surface is defined for the frequency climatology described above and presented in section 5.3 of this chapter.

The PV field on each isentropic surface is clustered into areas of PV > 2 PVU. At upper levels, the stratospheric body is easily identified as the cluster whose area is at least twice as large as all other cluster areas. The remaining 'cut-off' clusters are excluded from the climatology. At lower levels, however, the stratosphere often breaks up into large, almost equally sized regions, as well as some smaller cut-off like areas. Rather than exclude these larger regions, two different criteria are used to determine which clusters will be considered as part of the stratosphere at these levels.

1) We make use of results from a 3D clustering of the 3D PV field which distinguished those regions of PV > 2 PVU connected to the main stratospheric body from those belonging to cut-offs entirely encased by tropospheric PV and also from those in contact with the surface. This procedure was documented by Sprenger et al. (2003) and further details can be found in Croci-Maspoli (2002). All isentropic clusters which are identified as surface or mid-tropospheric structures are rejected, and only those connected to the stratospheric reservoir are retained.

2) The stratospheric area of the overlying (10 K above) isentropic surface is used as a mask. Areas completely or partially covered by the mask are included, whilst any areas lying outside the mask are treated as cut-offs and excluded.

5.2.4 PV mean climatology

A climatology of mean PV is created in this study to 1) enable the high PV frequency climatology to be directly compared to the mean PV for the same period, data set and levels and 2) to enable viewing and presentation of the results with an increased number of contours in comparison to the quantity typically used in the literature over the polar regions.

As for the high PV frequency climatology, the mean monthly PV climatologies are initially constructed for each year separately. The results for all years are then combined with equal weighting into a 10 year monthly climatology.

An additional problem in the construction of the PV mean climatology is that lower isentropic surfaces occasionally intercept the surface, especially in the summer when their height reaches an
Chapter 5. MIDDLEWORLD CLIMATOLOGIES

Fig. 5.1: The instantaneous distribution of stratospheric PV (shaded) for 1st May 1995, 00 UTC, on 330 K. Areas satisfying the threshold PV (required for inclusion into the climatology) are shaded in light grey.

annual climatological minimum. In this case, rather than setting undefined values to zero, averages over the month are calculated excluding the counts at these times and locations, i.e. the total number of records in these locations are reduced. This approach will result in a slightly higher PV (positive bias) than otherwise in mountainous regions of the climatologies such as the Himalayas and Greenland but makes no difference to the resulting distribution patterns or conclusions. An alternative and more complex approach, as employed by the ECMWF, would be to interpolate PV onto isentropes extrapolated below the land surface.

5.3 Climatologies

Climatologies have been constructed for both hemispheres. Results for the Northern Hemisphere alone will be presented and described here.

Note that a temporary website² has been created where it is possible to see the full climatologies (including results for individual years and for both hemispheres) and to select any two figures for comparison.

5.3.1 Isentropic climatologies

The results for the high PV frequency climatology and PV mean climatology for a representative selection of middleworld levels spanning 300 K – 400 K are presented in figures 5.2 and 5.3 respectively for the mid-season months January, April, July and October. For the same months, the results for the upper levels 460 K – 850 K are shown in figure 5.4 for the high PV frequency and in figure 5.5 for the mean PV.

Both the middleworld high PV frequency climatology presented here and the tracking event fre-

²http://iacweb.ethz.ch/staff/skw/mwclim/
5.3. CLIMATOLOGIES

Frequency climatologies from chapter 4 indicate that, on isentropic surfaces of the lower middleworld, high PV values and coherent PV structures tend to be distributed in a ring-like band and a local frequency minimum is positioned near the pole.

Framed panels in figure 5.2 emphasise the most prominent ring for January, April, July and October for the levels shown. The middleworld ring appears most prominent in late spring to summer, when it is also detectable in the mean climatology as a local increase of PV at the latitude of the ring (figure 5.3).

Above these levels and into the overworld, a different configuration is observed where, on the whole, PV increases towards the pole (see and compare figures 5.4 and 5.5). There are strong seasonal differences however in the overworld. The polar maximum becomes stronger on ascending the levels into the overworld in January, reflecting the dominance of the polar vortex in this month. In April, the polar vortex is much more disturbed, especially at 850 K where two regions of high frequency are observed (over Canada and over Scandinavia–west Siberia). It is at this time of year, that the upper atmosphere undergoes a reversal of flow which remains until the reformation of a strong stratospheric vortex in early winter. In July and October, a disturbed vortex is detected at intermediate levels (500 K in July and 460 K in October) whilst the distribution on the highest levels considered reveal a maximum is returning over the pole.

The level of the vertical transition from the ring-like configuration to a near pole centred maximum higher up has a seasonal cycle in altitude, mirroring the seasonal cycle observed for the lowermost isentrope of middleworld, i.e. the altitude peaks in the summer and falls to a minimum in winter.

Another intriguing feature, which once again is more noticeable in the high PV frequency distribution but is also present in mean PV climatologies (including the seasonal mean PV of the ERA-40 Atlas), is that the form of Greenland is detectable, at least in part, on some isentropic surfaces in the middle of the range representing the middleworld. Topography is apparently influencing the distributions on middle levels more than lower levels. It may be that upward propagating gravity waves which are linked to topography can act to influence the distribution of high PV values in the stratosphere when the stratosphere extends far enough equatorward to interact and support wave breaking.

5.3.2 Vertical distributions

The isentropic distributions exhibited a largely circularly symmetric distribution centred near to the pole. It is therefore not unreasonable to represent the vertical distribution by stacking the 2D fields into a 3D array and displaying zonally averaged fields in vertical sections (potential temperature by latitude). The frequencies may then be interpreted as the probability that PV at a particular latitude is stratospheric and within the top 20 percent of area-weighted stratospheric PV values observed on an isentropic surface for that month. Note that the PV threshold changes with the vertical coordinate as does the area of the stratospheric air mass.

The vertical distribution for high PV frequency and mean PV of the lower levels for levels (300 K – 420 K) are displayed in figure 5.6. The frequency distributions (top panel) indicate the presence of the isentropic ring by an area of raised frequencies situated in between mid- and polar latitudes over a sub-range of the lower isentropic surfaces. The vertical extent of the feature is greatest in spring, extending from 300 K (where the frequency maximum is on average at 72°N) up to 350 K (where the frequency maximum is on average at 69°N). In summer, an off-polar maximum can be detected on the levels 330 K – 370 K.

The vertical distribution for the mean PV for April reflects the presence of the ring by a region of negative, zonally averaged, poleward PV gradients from about 72°N to the pole, over the vertical range 320 K – 340 K.

It seems that a further sub-division of the middleworld could be meaningful i.e. into an upper region of positive poleward PV gradients and a lower region which encounters a reversal in gradient.

Also noticeable in the cross sections of mean PV is the large region of reduced vertical gradients north of 45°N between 350 K and 390 K in July. This is suggestive of mixing in these layers. This may be due to interior wave breaking or radiative processes.
Sections for the upper levels are shown in figure 5.7. Note that, due to the differences in the forms of frequency and mean distributions, the vertical range used for each climatology differs. The top panel shows the frequencies for 300 K – 850 K. The lower panel shows the mean PV for the levels 400 K – 850 K.

The ring of the middleworld and the disturbed polar vortex of the overworld do not seem to be linked in a coherent manner. The stratospheric polar vortex is seen to be at its weakest in summer and strongest in winter.

5.3.3 An illustrative year

As an illustration of the year-round persistence of the middleworld ring, monthly frequency climatologies for a single year (December 1998 – November 1999) are presented in figure 5.8. The isentropic level on which the feature is most prominent is selected for each month individually.

The annual cycle and interannual trends will be described and discussed further in the following chapter, where an amplitude-related index is developed for this purpose and to facilitate comparisons with other time series.

5.4 Discussion

The middleworld climatologies presented in this chapter confirm the presence of a persistent local PV minimum near the pole, as suggested from the tracking climatology, and that it is due to a reduced temporal frequency of high PV in the region rather than being due to one large dominating positive feature (such as the polar vortex) reducing the event frequency per unit area.

Despite being detectable in the mean PV climatology, the ring-like feature of the middleworld has probably not received attention in the literature before due to a number of reasons as now discussed. The ring structure in the mean PV climatology becomes less discernible the longer the length of the climatology. The ECMWF have recently issued the ERA 40 Atlas (Källberg et al., 2005) containing seasonal climatologies of many standard variables for the best ERA-40 years 1979–2002. The PV climatologies do not clearly reveal a band structure as observed here. This is most likely due to their averaging over entire seasons rather than months and using a longer period of data. A long period of averaging causes the PV minimum near the pole to disappear due to displacements of the centre of the distribution which cause overlap and cancellation between high values in the rings and low values in the holes. The choice of large contour intervals in the Atlas also helps to conceal shallow minima near the pole. As a measure of verification, we constructed a winter season mean PV climatology for the years of the ERA-40 Atlas and were able to reproduce their isentropic distributions. Of the other climatological results presented in the literature, there is a tendency to focus on summer or winter seasons alone and on levels below 340 K – 350 K, where the ring is less pronounced.

A ring of strong vorticity around a low vorticity centre is known to be a conditionally unstable configuration (e.g. Schubert et al., 1999). If the band of high PV evident climatologically were present instantaneously, it could support the mutual amplification of waves along its boundaries in turn prompting larger perturbations and the subsequent break up of the ring into smaller scale fragments. If the instability occurs on a fairly short time scale, relative to the time scale of the processes controlling the formation of the structure, then instantaneous fields would be in a state of continual dynamical readjustment (equilibrium) and it will never be possible to observe the initial state. The climatological mean would then give a smoothed out picture of an intermediate configuration. In the next chapter, results from preliminary experiments using a contour dynamics tool to explore the transition from a ring of high vorticity into a more stable configuration are presented. An idealised representation of the climatological structure is taken as a proxy (first guess) for the initial state.

It is hypothesised that the break up of the ring could be partly responsible for the creation of the small scale structures seen in the instantaneous PV field, including those that spawn cyclogenesis. It is therefore important that numerical weather prediction models correctly represent this dynamical
5.4. DISCUSSION

The band observed in the lower middleworld may present an obstacle for the transmission from the troposphere into the stratosphere of vertically propagating planetary scale waves - for which a positive ratio of poleward zonal PV gradient to zonal mean wind is expected (Andrews et al., 1987). The band itself however, as just described, may have the potential to be an instigator of wave activity, counteracting the view that the lowermost stratosphere is a predominantly passive

---

Fig. 5.2: High PV frequency climatology displayed for mid-season months on isentropic surfaces ranging from 300 K to 400 K. The mean position of the 2 PVU isoline is contoured in thick black. For each displayed month, the panel with the most prominent ring-like distribution is framed.
Fig. 5.3: Mean PV climatology for a selection of middleworld isentropes ranging from 300 K to 400 K. Note that the contour spacing increases for PV > 10 PVU. The mean position of the 2 PVU isoline is depicted in thick black.

receptor and transmitter of vertical propagating waves.

Finally, care should be taken if using metrics based in PV, such as the technique of equivalent latitude. The technique of equivalent latitude is often used to understand the distributions of stratospheric tracers such as ozone and is based on the presumption that the mean PV increases poleward, with higher values taking up an ever smaller area. The equivalent latitude associated with a PV value is the latitude found bounding a polar cap of the same area as the summed area
of all grid cells possessing an equal or greater PV. The application of the technique would not be so appropriate for isentropic levels where the PV ring is pronounced. If used without regard to the climatological structure being an annular rather than a polar maximum, results could be misleading (i.e. tracer values occurring where PV is highest would be incorrectly mapped to the pole).
Fig. 5.5: Mean PV climatology for a selection of isentropes from 460 K to 850 K. The thick black contour shows the location of the mean 2 PVU isoline.

5.5 Summary

A 10-year frequency distribution climatology has been created on isentropic levels from 300 K - 850 K in order to identify locations within the stratosphere where anomalously high PV is most likely to be found.

A high latitude (near pole) maximum is the dominant configuration for the overworld, (with the exception of late spring where the position of the maximum is far more variable).
Fig. 5.6: Cross sections for zonally averaged high PV frequency (top) and mean PV (bottom). The position of the 2 PVU isoline of the zonally averaged PV is shown in thick black. The horizontal axis spans from the equator (far left) to the pole (far right). The vertical axis gives the isentropic level and spans 290 K to 420 K.

Fig. 5.7: Cross sections for zonally averaged high PV frequency (top) and mean PV (bottom) for upper levels. The position of the 2 PVU isoline of the zonally averaged PV is shown in thick black. The horizontal axis spans from the equator (far left) to the pole (far right). The vertical axis gives the isentropic level and spans 300 K to 850 K in the top panel and 400 K to 850 K in the lower panel.

In the lower midlatitude there is a change of configuration. Instead of a polar maximum, a near polar minimum is found. Higher frequencies are distributed in a band with a maximum in the zonally averaged field near 70°N. The structure can be seen all year round in the frequency climatologies but with a seasonal dependence on altitude and amplitude. A weaker signal of this ring-like structure is also observed in the mean PV climatology in spring.
Fig. 5.8: Frequency of high PV on individually selected isentropic levels (see labelling on panels) to illustrate presence of the middleworld ring throughout one year (December 1998 - November 1999). The mean position of the 2 PVU isole is overlaid in thick black.
Chapter 6

The Middleworld and the PV-Ring

“For good or for ill it [the ring] belongs to Middle-earth; it is for us who still dwell here to deal with it.”

_The Lord of the Rings_, J.R.R. Tolkien

6.1 Introduction

In chapter 5 it was shown that the time mean PV distribution on isentropic surfaces in the lowermost stratosphere was characterised by a quasi-annular band of higher PV values. This PV ‘pattern’ is notable from both a phenomenological and theoretical standpoint. From a phenomenological standpoint, it is noteworthy since it is co-located with the zonal band where localised PV anomalies are most prevalent. From a theoretical standpoint, it is of significance that the pattern satisfies a necessary condition for pseudo-barotropic instability.

Here, two aspects of the ring are studied. Firstly, the seasonal and interannual variability of the high-PV pattern is described using diagnostics. Secondly, the possibility of instability is examined further by conducting a linear instability analysis and exploring the non-linear evolution using contour dynamics.

6.2 Seasonal and interannual variability

In this section, a set of diagnostics are calculated for the observed ring in order to assess its temporal variability.

6.2.1 Ring diagnostics

The diagnostics are calculated for several middleworld isentropic surfaces (290 K – 420 K, at 10 K intervals) for the period January 1991 to December 2000, based on the monthly high PV frequency distributions calculated in chapter 5. Some selected stages involved in obtaining the diagnostics are illustrated in figure 6.1.

Where a ring is observed, its centre does not in general lie directly over the pole but is displaced from it by a few degrees. An idealised, circularly symmetric ring will be fitted to the observed ring and used to compute diagnostics based on the zonally averaged field. To obtain the best fit, the observed field is first shifted such that the centre of the distribution lies over the pole.

The centre of the observed distribution is determined objectively as follows:

- A set of equidistantly spaced points provides an initial set of potential centres.
The point at which the average field value within a radius of 5° degrees latitude is minimised, is selected as the best guess for the distribution’s centre.

The pole is placed at the new centre and the ring diagnostics are calculated.

If the ring radius (as defined in point 3. below) does not exceed 10°, or if there is only one maximum along the suggested ring latitude (wavenumber 1), the presence of a ring is no longer assumed. The procedure is repeated but instead of locating a local minimum in the field, the mean longitudinal variance is minimised over a circular area of 30° radius. Thus if the observed distribution contains a polar maximum rather than a ring, the new centre will be placed in the region of the maximum.

All resulting distributions are inspected to check that a reasonable centre has been located.

Several parameters can now be determined (see figure 6.1 (d)):

1. The ring amplitude is defined as the difference between the maximum zonal mean high PV frequency and the minimum zonal mean to the poleward side of the maximum.

2. The inner and outer radius of the ring are determined to be where the zonal mean is equal to the half-height of the ring.

3. The ring radius is the radius within the ring where the mean zonal amplitude is maximised.

4. Completeness of the ring is the fractional area of the observed ring (i.e. where the field exceeds the ring half-height value) which lies within the modelled ring. This gives a measure of the ring’s deviation from a circular symmetry.

5. The vertical extent of the ring pattern is determined as the range of isentropic levels for which a ring structure can be detected.

6.2.2 Results

Seasonal variability

Figure 6.2 illustrates aspects of the seasonal variability taken from the 10-year climatology. The maximum elevation of the ring of the widest radius (red curve) in isentropic coordinates occurs in August, whilst the minimum occurs in early winter. The seasonal variation in this elevation is similar to the seasonal change in tropopause height for a fixed latitude (see e.g. Liniger and Davies, 2004), although the ring reaches its minimum isentropic level about 2 months earlier.

The maximum vertical extent of the annular structure (beige shading in figure 6.2) is greatest in early summer and least in autumn.

The largest radii occur in spring and summer. The vertical section for the mean PV (see July inset in figure 6.2 or refer back to figure 5.6) shows a large area of weakened vertical PV gradients north of 55°N and between about 350 and 300 K at this time.

There is a noticeable reduction in the ring’s radius and vertical extent in August and the radius remains small through to autumn. Correspondingly, a positive poleward PV gradient is restored.

Time series of diagnostics can be made for a representative surface or for a fixed level, for the entire year or for selected months. If the level is fixed, a seasonal cycle linked to the seasonal variation in altitude of the isentropic surface will also be contained in the trends. The isentropic surface with the largest ring radius is selected to be the representative surface for each month on which diagnostics will be compared.

The seasonal trends of ring amplitude and ring radius of the 10-year climatological ring are displayed in figure 6.3 (a) and (b) respectively. Note that the time of occurrence of the largest amplitude does not correspond to that of the largest radius. The peak maximum is found in April/May and a secondary maximum in October/November. The seasonal variation of the ring completeness (not shown) shows a peak in May (the mean exceeds 0.8) and a minimum (the mean
6.3. RING INSTABILITY

Fig. 6.1: An illustration from May 1995 of the various stages in calculating the index. Instantaneous isentropic PV fields (a) are combined into the monthly high PV frequency distribution (b), according to the method in chapter 5. The resulting high PV frequency distribution is centred on the pole (c). The zonal mean field (d) is used to determine the maximum ring amplitude (maximum difference between the zonal mean frequency within the hole and at the peak), the ring radius at maximum amplitude, the inner and outer ring radii and the amplitude at the ring boundary. The three black zonal rings superimposed in (c) correspond to the three diagnostic radii. The contourd isolines in (c) have a zonal mean high PV frequency equal to the value at half-height and, when compared to the inner and outer zonal rings, illustrate the slight deviation of the ring from circular symmetry.

falls to about 0.65) over the winter season.

Interannual variability

A 10-year time series for the ring amplitude and ring radius are shown in figure 6.1 (a) and (b), with a running mean of 1.5 years (17 months) superimposed (time series of other diagnostics are not shown here). Both diagnostics show raised values towards the end of the period (1999-2000). In addition, the ring amplitudes are more elevated in the years 1992-1993.

6.3 Ring instability

A simple analytical model is set up for an idealised axi-symmetric ring on the polar $\beta$-plane (see figure 6.5) using cylindrical coordinates. The climatological ring, represented by suitable values
Fig. 6.2: Seasonal variability. The isentropic level of the most prominent ring (largest radius) at each month is shown in red. The vertical extent of the annular structure is shown in beige, and the height of the shaded blue strip is proportional to the ring radius. Thumbnail images of the zonal mean PV are shown for January, April, July and October (larger versions are shown in figure 5.6).

Fig. 6.3: Seasonal variations for each year (grey) in the period 1991-2000 at the isentropic level of the most prominent ring in (a) amplitude, (b) ring radius. The bold contour depicts the 10-year average.

for these parameters, is tested for instability.

6.3.1 The polar $\beta$-plane

In a $\beta$-plane, the variation of the Coriolis parameter $f$ with latitude is approximated by the first two terms of a Taylor expansion of the latitudinal dependence of $f$ about a reference latitude
6.3. RING INSTABILITY

Fig. 6.4: Time series of ring diagnostics: amplitude (a) and ring radius (b). A 1.5-year running mean is overlaid in bold.

(Holton, 2004). For a polar $\beta$-plane, the reference latitude is the pole. Thus

$$f = 2\Omega \sin \theta$$
$$= 2\Omega \cos \Phi$$
$$\approx 2\Omega \left[1 - \frac{\Phi^2}{2}\right],$$

(6.1) (6.2)

where $\Omega$ is the Earth's angular velocity, $\theta$ is the latitude in radians and $\Phi = \frac{\pi}{2} - \theta$ is the latitudinal deviation from the pole.

Cylindrical coordinates are appropriate for a $\beta$-plane. In cylindrical coordinates, the radial coordinate $r$ is equated to $\Phi$ via the approximate relation $\Phi \approx \frac{r}{r_e}$, where $r_e$ is the radius of the Earth. Substituting this expression into (6.2), we obtain the variation of $f$ with $r$:

$$f \approx 2\Omega - \frac{\Omega}{r_e^2} r^2$$
$$\approx f_0 - \alpha r^2 \quad (f_0, \alpha, \text{constant}).$$

(6.3)

The Coriolis parameter takes the value $f = f_0 = 2\Omega$ at the pole and decreases with the square of the radial coordinate.

The polar $\beta$-plane approximation is appropriate provided $\Phi^2 \ll 2$.

Equating the expressions for $f$ on the sphere (6.1) and $f$ on the plane (6.3), it is seen that features on the sphere are line mapped onto the polar $\beta$-plane by the transformation

$$r \approx \left[2(1 - \sin \theta)\right]^{1/2} r_e.$$

(6.4)

6.3.2 Analysis of annular instability on the polar $\beta$-plane

The observed ring is idealised by an axisymmetric annulus of uniform absolute vorticity with relative amplitude $\varepsilon$, of inner radius $a$ and outer radius $b$, creating three regions of uniform absolute vorticity (see figure 6.6). The barotropic vorticity equation is formulated for these three regions - $r < a$, $a < r < b$, $r > b$ and matched across the boundaries $a$ and $b$. The outline of the derivation of the stability condition is given in appendix A. The resulting stability equation is:
Fig. 6.5: The idealised ring on the polar $\beta$-plane. The inner and outer radii of the ring are denoted by 'a' and 'b' respectively.

\[
\frac{-\sigma^2}{f_0} = \frac{1}{16} \epsilon^2 [m^2 (1-p)^2 (1+\chi)^2 - 4m(1-p)(1+\chi) + 4(1-p^m)]
\]  

(6.5)

with

- $\sigma_i$ = exponential growth rate
- $m$ = azimuthal wave number
- $\bar{\epsilon}$ = $\epsilon/f_0$
- $f_0$ = $2\Omega$
- $p = \left(\frac{a}{b}\right)^2$
- $\chi = \frac{1}{4} \left(\frac{b}{r_e}\right)^2 \bar{\epsilon}$
- $r_e$ = radius of the Earth.

Instability occurs if the right hand side of 6.5 is negative.

Fig. 6.6: The idealised ring. See section 6.3.2 for an explanation of the symbols.
6.4. CONTOUR DYNAMICS IDEALISED EXPERIMENTS

Stability of the climatological ring

The climatological ring of the largest amplitude occurs in April or May. In April, the ring inner and outer radii on the 340 K surface, taken from the 10 year mean PV climatology, are positioned approximately at 74° and 62° latitude respectively, which correspond to radii of about 0.25 re and 0.5 re in the cylindrical coordinate system. The amplitude of the ring in PV units is between 1 and 2 PVU. The amplitude of the absolute vorticity step is of the order of $1 \times 10^{-5}$ s$^{-1}$ near the 340 K surface.

Figure 6.7 (a) shows the fastest growing wave numbers calculated from equation 6.5 for a ring with a vorticity step of $1 \times 10^{-5}$ s$^{-1}$, approximately of the climatological magnitude. For the climatological ring dimensions estimated for April, the analysis suggests the zonal wave number $m = 1$ for the fastest growing mode. The growth rates for $m=1$ are depicted in figure 6.7 (b). For the climatological ring dimensions for April, the analysis suggests a growth rate $\sigma$ of approximately $2 \times 10^{-6}$ s$^{-1}$ (e-folding time, $\sigma^{-1} \sim 6$ days).

For a growth rate of about 1 day$^{-1}$, one would need an absolute vorticity step of about $10 \times 10^{-5}$ s$^{-1}$ (when $m = 2$) for the climatological dimensions (see figure 6.8 (a) and (b)).

Generally from figures 6.7 and 6.8, it is seen that higher wavenumber instabilities occur with narrower rings. Increasing the radius increases the wavenumber of the most unstable mode for larger amplitudes. An effect of the circular geometry can be seen by the dominance of wavenumber 1 instabilities at large outer radii and large ring widths. This occurs when the diameter of the hole is smaller than the ring width and $m = 1$ instabilities develop from interactions of the inner contour on one side of the pole with the same contour on the other side of the pole. Largest growth rates are found for rings with large radii and small widths.

6.4 Contour dynamics idealised experiments

Having derived the linear instability conditions for a ring of enhanced vorticity, the evolution of the ring into the non-linear instability regime will now be illustrated using 2D contour dynamics simulations. The aim is to explore whether the instability and subsequent break up of a ring of enhanced vorticity could be an explanation for the distribution of flow structures we observe instantaneously and climatologically in the lowermost stratosphere. The contour dynamics tool used for this study has been modified to enable representation of dynamics on the polar $\beta$-plane as well as on the $f$-plane. The theory behind contour dynamics and details of the contour dynamics tool and modifications are outlined in appendix B.

![Figure 6.7: Linear instability for a vorticity step of $1 \times 10^{-5}$ s$^{-1}$: (a) wavenumber of the fastest growing mode and (b) growth rates for $m=1$ as a function of the outer radius 'b' and the ring width 'b-a'. Radius units are given as multiples of the Earth's radius, $r_e$.]
6.4.1 Methodology

Parameter model

Simple idealised experiments are carried out to demonstrate the effect of the ring geometry on the non-linear dynamics.

The total set of parameters, expanding on the parameter space used in the linear stability analysis, are

- Ring radius
- Ring width
- Vorticity step at the inner edge of the ring
- Ring edge gradient (represented as a series of smaller steps discontinuities) at the inner and outer boundary.

The additional parameters allow for an improved representation of the observed climatological structure. The vorticity step on the inner edge of the ring is varied between the background and the ring vorticity, such that the ambient vorticity can be larger on the poleward than the equatorward side of the ring. The edge gradient is represented by several small steps which replace an unrealistically large single discontinuity. The inner and outer edge gradients are varied independently.

Each parameter is varied in turn whilst the other parameters take the values for the reference state. Experiments are first carried out on the $f$-plane and then some selected experiments are repeated on a polar $\beta$-plane, where a background velocity field is used to represent the planetary vorticity.

Contour dynamics representation

The contour dynamics tool solves the barotropic vorticity equation assuming non-divergent 2-dimensional flow. This is reasonable for upper-levels where flow is quasi-adiabatic and confined to quasi-2D surfaces. The vorticity distribution is modelled by a series of contours separating regions of constant vorticity, i.e. by a series of finite discontinuities.

For a ring with a single vorticity step, as was the case for the linear stability analysis, the ring is represented by just two contours. The experiments modelling a sloped gradient at the ring edges use 5 contour steps to represent the transition between the maximum amplitude and ambient background vorticity.
6.4. CONTOUR DYNAMICS IDEALISED EXPERIMENTS

Integrations are carried out for $8 \times 10^5$ s (about 9 days) or until a limit on the total contour length is reached.

Reference state

As an instability on the ring evolves, contours separating the high vorticity of the ring and low vorticity of the surroundings will be strongly distorted. The redistribution of vorticity will result in the widening of the ring and a decrease in amplitude in the zonal mean vorticity. If such an instability is at work in the middleworld, the climatological observations will be a smeared out picture of an intermediate stage of the transition from an initially unstable state to a state of dynamic equilibrium, with a wider ring of decreased amplitude compared to that of the initial state.

A reference state is chosen with a larger radius ($r_e$) and a narrower ring width (such that $a/b = 0.6 r_e$) and a larger amplitude ($2/3 f$), to obtain a more unstable configuration which might relax toward the mean climatologically observed state. The following results relate to changes made with respect to the reference state. The reference state gives a wavenumber 3 perturbation within the integration time and the ring breaks up into 3 vortices.

6.4.2 Results

Radius

From the linear stability analysis (see figure 6.8) it can be seen that, for the chosen vorticity step and range of ring dimensions, increasing the ring radius will increase the wavenumber of the most unstable wave (except where the ring is wide enough to favour a wavenumber 1 instability). An illustration of the evolution for two rings of different radii is shown in figure 6.9. It can be seen that the unstable rings deform into geometrical (polygonal) shapes before breaking into individual vortices. The outer extremes of the fragmented ring segments are sheared out into long filaments. The ring with the initially largest radius (a) develops a wavenumber 4 instability and high vorticity has been advected to the pole by the end of the integration. The ring with a wave number 3 instability (b) approaches at a slower rate toward the pole. The ring with the smallest radius remains whole within the integration time although a slight asymmetry on the inner edge indicates there may be a developing instability (not shown).

The evolution of the zonal mean vorticity for the wavenumber 3 perturbation is shown in figure 6.10. The tendency for vorticity to be redistributed toward the pole is clear. The zonal mean ring decreases in amplitude; the maximum is skewed to the poleward side, closer to the pole. Equatorward of the maximum, the vorticity decreases with radial distance from the pole. Filaments of high vorticity will be able to remain distant from the pole in this vorticity profile. The radial profile becomes closer to the observed climatological profiles.

Ring width

Three examples of the ring evolution for various initial widths shown in figure 6.11, illustrate that the narrower rings are subject to higher wave number instabilities.

The ring in row (a) is of similar relative dimensions (although of larger amplitude) compared to some climatological observations on lower isentropic surfaces. It exhibits a wavenumber 2 instability but a slow growth rate. The instability is evident only towards the very end of the integration. However, the difference between the start and end configurations is not unlike the subtle change observed in the tracking climatology (chapter 4) between the large scale density distributions of genesis and lysis events (see figure 4.9).

A ring width one third as wide (b), undergoes a wavenumber 4 instability. Low vorticity is drawn out of the hollow of the ring and high vorticity is drawn inwards. For a ring width reduced again by one third (c), the instability sets in with one third of the time. In this case, a wavenumber 7 instability has developed.

The ring of intermediate radius (b) is the only example shown where high vorticity structure has
Fig. 6.9: Variations of the ring radius. The initial vorticity configurations are shown on the left, an intermediate stage in the centre, and the stage at the end of the integration on the right. The radius is increased by one sixth from the top row (a) to the lower row (b). Time $t$ is expressed in the units $1 \times 10^4$ s.

Fig. 6.10: Evolution of the zonal mean vorticity for the ring depicted in figure 6.9 (a). The initial profile is a step function. Subsequent profiles are drawn every $5 \times 10^4$ s (approximately 0.6 days) once noticeable instability occurs. With time, the maximum mean zonal amplitude decreases and the peak is shifted poleward.

been advected into the centre of the domain. Contraction of the zonal distribution towards the pole occurs once the ring has broken up into individual vortices and some high vorticity has been sheared away.

As the climatological ring changes width during the seasons, there could also be a change in the size and poleward migration of individual vortices. It could be of interest to see if there is any evidence of a seasonal cycle in the size of localised anomalies.
6.4. CONTOUR DYNAMICS IDEALISED EXPERIMENTS

Fig. 6.11: Variations of the ring width. Rings with successive reductions of 1/3 in width are shown from the top row (a) to the bottom row (c). The initial vorticity configuration is displayed in the left-hand column and the state at the end of the integration in the right-hand column. Time \( t \) is expressed in the units \( 1 \times 10^4 \) s.

**Vorticity step at inner ring edge**

The increase of vorticity within the hollow of the ring from the external background vorticity (zero) to the ring vorticity brings the distribution into a more stable configuration. For a small hole vorticity, a quarter the magnitude of the ring vorticity, wavenumber 3 instability results (figure 6.12) as for the reference state. If the hole vorticity is raised to half the ring vorticity, the ring remains stable for the reference width and amplitude, due to the relative increase in differential rotation.

**Vorticity gradient**

Four different configurations of gradients at the ring edges are shown in figure 6.13. The width of the high vorticity band is kept the same as for the reference ring and the decrease to zero vorticity on either side is permitted to occur over an increasing range.

In panel (a), there is a shallow gradient on both edges and the configuration remains stable. If the gradient is increased (b), a wavenumber 3 perturbation develops, as for the reference state.

In (c), there is a single discontinuity on the inner edge and a shallow gradient on the outer edge. In
Chapter 6. THE MIDDLEWORLD AND THE PV-RING

(a) \( t = 80 \)  
(b) \( t = 80 \)

Fig. 6.12: The vorticity distribution at the end of integration of a ring with raised vorticity within the hollow, depicted (a) on the plane and (b) in the zonal mean. The vorticity step on the inner edge is raised by 1/4, with respect to the ring amplitude, above the background. Time \( t \) is expressed in the units \( 1 \times 10^4 \) s.

In this case, the instability is most evident along the inner edge where the poleward vorticity gradient becomes negative. The outer boundary is also perturbed but there is no differentiation between the behaviour of the contours. When the single discontinuity is on the outer edge and a shallow gradient on the inner edge (d), the outermost contour of inner edge becomes perturbed, where the poleward vorticity gradient turns negative.

The ring observed in the mean PV climatologies has a weaker gradient on the poleward side. The most realistic simulated situation reflecting the mean PV distribution is (d), where the inner gradient is shallower than the outer gradient.

Increased amplitude with gradient

Doubling the amplitude of the reference ring simply doubles the growth rate on the \( f \)-plane. Two interesting evolutions are demonstrated for a raised amplitude of a thin ring with shallow (figure 6.14) and steep boundary gradients (figure 6.15). Both configurations initially exhibit a wavenumber 5 instability along the boundaries of the central band with the most elevated vorticity. For the ring with the tighter gradient (figure 6.15), the entire ring is broken up into 5 coherent vortices. For the ring with the weaker gradient (figure 6.14), the edge instability is stabilised, possibly by the relative increase in vorticity poleward on the maximum. Confined by the edge gradients, interactions between the vortices within the ring take place. The five vortices rearrange into two larger circulations. The band at this stage has a wave number 3 structure.

Inclusion of planetary vorticity

Three examples are given in figure 6.16 where \( f \)-plane results are compared to polar \( \beta \)-plane results for identical ring configurations after the same length of integration time. In the first and second rows, comparisons can be made between \( f \)- and polar \( \beta \)-plane results for rings with differing radii with a single contour representing the edge, and in the last row, for a thin ring with edge gradients with a wavenumber 5 instability.

The first two examples indicate that the effect of the planetary vorticity on the instability is to dampen the growth rate and increase the wave number of the fastest growing instability. For the thin ring, 11 vortices develop in place of 7, and for the wider ring, 5 vortices develop in place of 4. The dimensions of the vortices appear more isotropic. Advection of high vorticity to the pole is constrained by the additional zonal velocity from the planetary vorticity. The constraint on meridional motion could account for the isotropic appearance of the vortices.

In the third example, where five contours are used to represent each ring edge, both rings are radially restrained by the shallow edge gradients and develop the same azimuthal wavenumber instability. Little difference is seen in the radial vorticity profile (not shown). Small differences are
6.5 Summary

A seasonal and interannual variability of the time mean distribution of high PV in the lowermost stratosphere has been objectively detected with a set of ring-diagnostics. A seasonal cycle is evident in the ring's vertical isentropic range, isentropic level of the maximum amplitude and the ring radius. The annular structure has a maximum amplitude in late spring and a secondary maximum in late autumn, while its vertical isentropic range is greatest in the summer. In August, the ring contracts and its vertical extent is reduced.

Linear stability analysis suggests that this prevalent structure is subject to a wavenumber one instability even in its climatological form. However, contour dynamics integrations show that a...
Fig. 6.14: Evolution of a narrow ring with shallow edge gradients. Vorticity increases in steps of $4 \times 10^{-5} \text{s}^{-1}$ from a background of $0 \text{s}^{-1}$ to the ring interior of $20 \times 10^{-5} \text{s}^{-1}$. Time $t$ is expressed in the units $1 \times 10^4 \text{s}$.

A more realistic representation of the climatological ring (edge gradients and raised vorticity in the hollow) is less likely to be unstable and break into small scale structures. If the amplitude is raised above the typical climatological values, it is seen that the pseudo barotropic instability of a ring of elevated vorticity could be responsible for both (i) the climatological structure of a region of high time-mean PV values confined to a zonal band, and (ii) the creation of small scale PV structures, seen in the instantaneous fields, as the ring relaxes to a more stable state.
Fig. 6.15: Evolution of a narrow ring with steep edge gradients. Vorticity increases in steps of $4 \times 10^{-5} s^{-1}$ from a background of $0 s^{-1}$ to the ring interior of $20 \times 10^{-5} s^{-1}$. Time $t$ is expressed in the units $1 \times 10^4 s$. 
Fig. 6.16: Comparison of evolutions on the $f$- and polar $\beta$-plane. The upper two rows (a) and (b) show $f$-plane and polar $\beta$-plane developments for the same configurations as the unstable rings in rows (b) and (c) of figure 6.11. The lowest row (c) shows the $f$- and polar $\beta$-plane developments for a ring with the same radius as for row (b) but with edge gradients. Time $t$ is expressed in the units $1 \times 10^4$ s.
Chapter 7

Polar Lows and Stratospheric PV Anomalies

The objective of this chapter is to review and study the PV perspective of polar low development, with a focus on the pre-genesis phase and the accompanying conditions for development.

Firstly, a general description of polar lows is given with reference to climatologies and development mechanisms, and the application of the PV perspective to polar lows in the literature is explored. Then the synoptic PV environment of polar lows in the Norwegian and Barents Sea is presented using a composite analysis. Two cases are studied to highlight the origin and role of stratospheric PV anomalies in polar low cyclogenesis. Finally it is discussed how the PV perspective could be exploited in polar low forecasting.

7.1 Introduction to polar mesoscale vortices

Polar lows are intense, short-lived mesoscale cyclones which develop in outbreaks of cold polar air masses over water. They have horizontal scales approximately between 200 and 1000 km and surface winds near or above gale force (i.e. greater than 15 m s$^{-1}$). Additional severe conditions associated with polar lows are large amplitude ocean waves (13 m waves have been reported), near-zero visibility, heavy snowfall (10 cm snow has been reported), thunder and lightning (Rasmussen and Aakjaer, 1992) and tornadoes (Businger and Reed, 1989). Polar lows tend to decay rapidly upon landfall or otherwise when the source of upper-level forcing is removed.

Favoured locations for formation are regions where sea surface temperatures are relatively high during outbreaks of cold polar air, e.g. the northeast Atlantic, Norwegian and Barents Seas, North Pacific, Sea of Japan, Labrador Sea, Alaska, near the Antarctic coast and New-Zealand (Rasmussen and Turner, 2003). Thus favourable conditions are commonly fulfilled at the rear of synoptic-scale cyclones, where flow is directed from polar latitudes to lower latitudes over relatively warm water. The typical synoptic environment for polar low development changes from area to area depending on the topography of local land masses. The typical synoptic conditions for Norwegian and Barents Seas polar lows will be described further in section 7.2.

In the Northern Hemisphere, polar lows are primarily a winter phenomenon. In the Norwegian and Barents Seas, polar lows are most frequent in the months November-January and March (Rasmussen and Turner, 2003). A frequency minimum occurs in February and is generally associated with the intense strength of the high pressure region over Siberia, extending to Scandinavia and bringing southerly winds to the region (Kolstad, 2006). Antarctic mesoscale cyclones however can occur at any time of year and at any longitude, even over the ice shelves (Carrasco et al., 2003). The Antarctic lows are generally weaker the Arctic lows. This is thought to be due to the strong zonal circumpolar flow, which tends to hinder the formation of the large air-sea temperature differences found in the Arctic (Rasmussen and Turner, 2003).

There has been much controversy (e.g. Rasmussen and Cederskov, 1994, Businger and Reed, 1989)
over the mechanisms responsible for polar low cyclogenesis and growth; the proposed mechanisms including baroclinic instability as growth from infinitesimal (e.g. Reed, 1979, Harrold and Browning, 1969) or finite disturbances (e.g. Montgomery and Farrell, 1992), conditional instability of the second kind (CISK) (e.g. Rasmussen, 1979), sea-air interactions (e.g. Emanuel, 1994), barotropic instability where the local topography favours convergence lines, and orographic forcing (e.g. Klein and Heinemann, 2001). Consequently numerous classifications and definitions of polar mesocyclones have been suggested (e.g. Businger and Reed, 1989, Rasmussen and Cederskov, 1994). Historical reviews can be found in Nordeng (1990) and Rasmussen and Turner (2003).

It is now generally accepted that a whole spectrum of lows exist (e.g. Renfrew, 2003) ranging from purely baroclinic to purely convective, the majority of cases exhibiting both baroclinic and convective behaviour. Some rare examples of types at the extreme ends of the spectrum are given in Rasmussen and Cederskov (1994).

Two main cloud signatures have come to be associated with polar lows and reflect the processes most active in the development.

- The *comma cloud* variety is the largest-scale system in the polar low spectrum (Businger and Reed, 1989) and appears to be a smaller version of the extratropical cyclone (Bresch et al., 1997, Nordeng, 1999). This type occurs when baroclinic processes are dominant and is often situated ahead of an upper level trailing trough. The comma tail may be aligned like a front with a trough in the surface pressure. Baroclinicity is deep. Some examples of comma clouds are found in (Reed, 1979).

- The *spiraliform* variety tends to be a shallower system, occurring further poleward in the cold air mass where convective processes dominate, and often forming close to the ice edge. The cloud structures take a more cellular appearance and sometimes a clear region in the centre can be observed, resembling the eye of tropical storms. Examples can be seen in Nordeng and Rasmussen (1992) and Reed (1979).

In some cases, surface vorticity or mesoscale features are supplied by pre-existing structures e.g.

- A frontal wave (polar front) may supply warm and cold fronts (Reed, 1979) to a mesocyclone creating an instant occlusion

- The centre of old occlusions can be regenerated, often by upper-level forcing (e.g. Rasmussen and Aakjaer, 1992)

- Remnants of synoptic cyclones can provide the initial circulation: A synoptic scale cyclone split into two parts due to interaction with topography in southern Greenland. The circulation remaining over the Labrador sea was transformed into a polar low with help of an upper-level forcing (Moore and Vachon, 2002).

It is now recognised that most polar lows are triggered by upper-level disturbances in the form of a mid tropospheric trough, upper level vortex or a positive PV anomaly approaching a low level baroclinic zone. Only very occasionally have polar lows been observed to form without significant upper-level support (e.g. Craig et al., 1994). The next sections look at polar low development with a focus on the upper levels using the PV perspective.

**PV perspective**

The PV perspective presented by Hoskins et al. (1985) offers a constructive framework with which to view polar low development (Bresch et al., 1997). It has been applied in polar low research by e.g. Nordeng (1990), Nordeng and Rasmussen (1992), Montgomery and Farrell (1992), Rasmussen et al. (1992), Moore et al. (1996), Ræsting and Midttøm (1996), Bresch et al. (1997), Nordeng (1999), Browning and Dicks (2001) and Claud et al. (2004).

Prior to Hoskin's 1985 paper, upper level forcing was mostly explored using quasi geostrophic theory without reference to PV. The 500 hPa level was usually taken to be representative upper
level dynamics and it was noted that baroclinically generated polar lows could often be associated with troughs or cut off lows at this level.

Many recent works acknowledge or assume the presence and crucial role of multi-body PV interactions in the genesis process. Few however have focussed on the time leading up to genesis (e.g. Moore et al. (1996) claim there are no published cross-sections showing the PV distribution prior to the development of a Norwegian Sea polar low), and choose to leave aside a discussion of the origin of the upper-level triggering disturbance (e.g. Montgomery and Farrell, 1992, Nordeng and Rasmussen, 1992). A later exception is Nordeng (1999) who simulated a mesoscale low in the Bay of Biscay, possessing similar properties to polar lows.

**Evolution of PV in polar low development**

A description of the typical sequence of events of polar low development as seen from the PV perspective follows. The sequence has been compiled from many different case studies from the literature of Northern Hemisphere polar lows.

- A near tropopause-level PV disturbance descends to the west of the surface cyclone (e.g. Nordeng and Rasmussen, 1992, Moore et al., 1996, Claud et al., 2004) as part of a dry intrusion into the synoptic scale cyclone (Claud et al., 2004).
- There is ascent upstream of the upper-level anomaly and decent behind (Montgomery and Farrell, 1992).
- Upstream ascent stimulates latent heat release, creating a lower level warm anomaly (Browning and Dicks, 2001) which enhances lower tropospheric baroclinicity (Renfrew, 2003). A positive surface PV anomaly is created beneath the warm anomaly and a mid-level negative PV anomaly above and to the east/downstream of the polar low (Montgomery and Farrell, 1992, Bresch et al., 1997).
- The upper PV and lower warm anomalies approach each other. If penetration is deep enough, they can phase lock (Claud et al., 2004) and are subject to baroclinic instability (Nordeng and Rasmussen, 1992).
- Surface PV greatly increases due to intensive latent heat release (Montgomery and Farrell, 1992) as the polar low is generated.
- The polar low's circulation may act to deflect the lower part of the upper level PV anomaly to the west of the region of decreased static stability associated with the mid-level negative anomaly (Nordeng, 1999, Claud et al., 2004). The lower part of the upper level PV anomaly rotates around the polar low cyclonically, moving to the south of the surface polar low (Claud et al., 2004, Nordeng, 1999, Montgomery and Farrell, 1992), in the same manner as a synoptic scale intrusion (Bresch et al., 1997). It helps to maintain the circular form of the lower level cyclonic development.
- The upper part of the upper PV anomaly becomes vertically aligned with the surface polar low and tracks with it (Businger and Balk, 1991, Montgomery and Farrell, 1992). Due to the low static stability beneath, enhanced by the presence of diabatic heating, the strong vorticity field associated with the upper anomaly can penetrate to even lower levels Nordeng and Rasmussen (1992) and induce cyclonic spin up. The polar low reaches maturity.
- The baroclinic role of the PV anomaly ceases as the lack of tilt to the axis implies that transfer from potential to kinetic energy is no longer favourable (Claud et al., 2004, Businger and Balk, 1991) and diabatic process become dominant in maintaining the polar low (Claud et al., 2004).
- Some have noted that the upper level PV anomaly retains its phase to the polar low (Bresch et al., 1997), others notice additionally there is a decrease in amplitude of the upper PV anomaly or that the forcing diminishes (Claud et al., 2004, Nordeng and Rasmussen, 1992, Montgomery and Farrell, 1992), whilst others note that it loses its influence and moves away from the polar low (Nordeng, 1999).
From this summary, it would appear that the sub-structure of an upper level PV precursor can be important. Therefore, representing an anomaly on a single isentropic level could restrict understanding of the dynamics, unless the level is carefully chosen. In the above mentioned literature, single isentropic surfaces were chosen in the range 278 K–300 K, levels which are usually typical of the upper troposphere and often chosen because of their proximity to the 500 hPa surface (Nordeng and Rasmussen, 1992). One exception was found in Moore et al. (1996) who used the 290–300 K layer (as opposed to a single level). Vertically integrating PV over a relevant isentropic range can improve the representation of the 3D structure of the anomaly at genesis. Indeed, it can come about, that upper level PV anomalies of different origins and isentropic levels superimpose or become aligned in a way which triggers the polar low or enhances its development.

This was the case for the spiraliform polar low studied by Nordeng and Rasmussen (1992) and Grønås and Kvamsted (1995), which Nordeng and Rasmussen termed 'most beautiful'. Grønås and Kvamsted (1995) looked at the height of the dynamical tropopause (2 PVU surface) which, at the location of the polar low, lay even lower in altitude than the 278 K isentropic surface selected by Nordeng and Rasmussen (1992) and so sampled the lower part of the upper anomaly only. Nordeng and Rasmussen (1992) did comment that a small scale anomaly was superimposed (height was not given) on the large scale anomaly where the polar low started to develop and suggested it might have been responsible for the deep circulation in the vertical. Further inspection of this case using ERA-40 data shows there was indeed an anomaly, at 330 K, which came to be directly over the surface pressure minimum. This anomaly had a different history to the anomaly at 278 K. A confluent flow brought the 330 K anomaly from near Greenland and the 278 K anomaly from north of Spitzbergen into the polar low genesis area. Whilst the 278 K anomaly became positioned above the polar low, the 330 K anomaly was, in agreement with Nordeng and Rasmussen (1992), diverted to the west along with the 500 hPa low and appeared to co-rotate with the surface low.

Role of the upper level PV anomaly

The upper level PV anomaly is often simply envisaged or understood as having a role in baroclinic instability. Its role there, as outlined in the sequence above, is to organise vertical motion, ascent ahead and descent behind (e.g. Lambert et al., 2004; Browning, 1993; Businger and Baik, 1991), and for there to be mutual amplification of upper and lower anomalies (e.g. Resting and Midtbø, 1996). In this sense, the mechanisms are similar to those occurring in mid-latitude synoptic cyclogenesis (e.g. the cold air cyclogenesis studied by Lambert et al. (2004) was comparable to a ‘type C’ mid-latitude cyclone as defined by Deveson et al. (2002)).

However, sometimes in mesoscale polar cyclogenesis, a PV anomaly takes on a different or additional role which is more specific to high-latitude cyclogenesis. That is the direct ability to trigger CISK, without relying on an interaction with a lower-level anomaly. Grønås and Kvamsted (1995) comment that, according to Økland (1987), many polar lows are formed this way and the important effect of the upper anomaly is not to create a baroclinic development but rather a local area of deeper convection so that CISK might be initiated. Beneath a PV anomaly, the static stability is reduced with respect to surroundings at the same level (Resting and Midtbø, 1996). This effect, as pointed out by Browning (1993), is also applicable to stationary anomalies, in contrast to baroclinic instability which requires advection of positive vorticity. As the tropopause is generally low at polar latitudes (and even lower still under a dry intrusion) and the convective layer is typically deep in cold air outbreaks (Resting and Midtbø, 1996), the vertical distance between the anomaly and the convective boundary layer is small. A deep convective boundary layer is required for this mechanism (Grønås and Kvamsted, 1995). If the region of reduced static stability beneath the PV anomaly can penetrate to the top of the boundary layer, the depth of the convective boundary layer will be locally increased. An upper level anomaly of strong amplitude may then be able to penetrate through and locally lift a lower lying inversion. Convection is favoured underneath the anomaly and CISK may be initiated whilst everywhere else deep convection may be inhibited due to the strong capping inversion of the polar boundary layer.

Both the baroclinic and CISK mechanisms allow the upper level PV anomaly to have a role in organising convection, either by forcing vertical ascent (e.g. Lambert et al., 2004), or further reducing static stability. The organisation takes place on the scale of the upper anomaly, implying that the anomaly also has a role (along with other properties of the local environment) in deter-
mining the scale of the induced polar cyclone (Montgomery and Farrell, 1992, Rasmussen, 1985). Furthermore, for a baroclinic interaction for which the organised ascent becomes moist, the scale of the developing surface disturbance may be further reduced as moisture acts to shift the fastest growth rates of baroclinic instability to smaller wavelengths.

If the environmental or induced static stability favours deep penetration of the anomaly, the anomaly can induce spin up, due to the enhanced omega response and vortex stretching in a conditional neutral atmosphere (Montgomery and Farrell, 1992), and or enhance existing surface vorticity by the penetration of its own circulation (Rasmussen and Turner, 2003).

When favourable conditions are persistent, there may be an outbreak of polar lows. An upper level PV disturbance typically plays a role in triggering the first polar low, whilst the origin of subsequent formations may be more difficult to explain (Rasmussen, 1985). Ralph et al. (1994) explore several cases of multiple mesoscale cyclones which occurred within synoptic scale cyclones and find two distinct types: downstream development and pure frontal (barotropic) instability. Downstream requires an initial cyclone and so creates an age-ordered cyclone family, where as pure frontal instability can spawn quasi-simultaneous growth of multiple cyclones.

In the case of multiple disturbances, inter-level interactions may be complex. An upper level anomaly need not only be connected to the development of just one surface disturbance and more than one PV anomaly may act on the same polar low. In the first scenario, the mechanism of interaction may be different for the two (or more) surface lows (e.g. Grenäas and Kvanustø, 1995). Where two or more disturbances exist simultaneously in close proximity, they will be influenced by each other’s associated velocity fields. Ziv and Alpert (2003) discuss various combinations of interactions. Maybe the most surprising result of their research for the forecaster is that co-circulation is controlled by the intensities of the upper level anomalies. The cyclone with the stronger upper level PV anomaly will tend to remain more stationary whilst the other will cyclonically rotate around it, even if it is the stronger surface cyclone. This result emphasises the importance of knowledge of the upper level dynamics and the ability to estimate the strength of the circulation induced by the upper anomaly(ies).

**Histories of upper level precursors**

A few studies mention upper level disturbances present at cyclogenesis which can be traced through various analyses backwards in time:

- Businger (1985) notes the presence of a small sub-synoptic scale vortex at 500 hPa in his case that could be tracked back up to a day earlier.
- Moore et al. (1996) remarked that the upper-level PV anomaly in their case, which interacted with a low level disabantically induced PV anomaly to initiate a polar low, had propagated into the area from the Canadian Arctic.
- Rasmussen (1985) could trace a strong 500 hPa disturbance linked with a Barents Sea polar low in his study as a trough or closed circulation back two and a half days where it could be seen at the northern part of Novaya Zemlya. The cold core vortex was positioned in the inner cyclonic shear region of a larger scale circulation and could be identified with a ‘rather diffuse’ cloud cluster in satellite imagery, present during the pre-development phase.
- Bresch et al. (1997) carried out numerical simulations of a polar low in the Bering Sea. The control simulation showed the polar low formed near the ice edge over the Bering Sea when a lobe of anomalously large PV broke off from a migratory upper-level cold low over Siberia and advanced into an area upstream of the polar low genesis region.
- Zick (1983), Bresch et al. (1997) and Nordeng (1999) reported that mid-tropospheric mesoscale vorticity maxima rotated about upper cold-core lows before subsequently initiating a surface cyclogenesis. In the earliest of these works, Zick (1983) traced one precursor for over 30 hours. He suggested the vorticity maxima have a life cycle in which they are advected from the front of a synoptic scale trough, through the cloud head of the occluded system and out to the rear of the trough, where comma cloud development can be triggered.
Rosting and Midtbø (1996) showed an anomaly on the 290 K surface that originated from the east of Greenland and was subsequently advected southeast over a low level convergence zone to form an instant occlusion.

It is evident that precursors to polar lows exist and are often detectable many hours and sometimes days before cyclogenesis. Whereas some form by becoming separated from the stratospheric reservoir, others are recycled from a previous cyclonic development and/or may be created by orographic or diabatic processes.

**Aims**
The focus of the rest of this chapter will be to explore the origin of the upper-level precursor using the PV perspective and to consider applications of the PV perspective to forecasting. More specifically, the aims are to

- describe the favourable synoptic conditions in the PV framework using a composite analysis,
- investigate the origins and characteristics of the upper-level PV precursors in an illustrative case study,
- illustrate how the PV field relates to other features such as surface troughs,
- see if the Rossby penetration depth could be simply estimated/used operationally as a measure of the potential of an upper level anomaly to initiate surface development,
- investigate the ECMWF forecast skill in the case of a polar low.

### 7.2 Synoptic PV environment

It is well known that favourable conditions (a cold air outbreak, the resulting deep convective boundary layers and baroclinic flow) for polar low development are often achieved by the passage of a synoptic occluded low through the region. Businger (1985) and Léve et al. (1988) formed composite analyses of sea level pressure, 500 hPa heights and temperatures for days when mature polar lows were present in the Norwegian and Barents seas, and Forbes and Lottes (1985) carried out a composite analysis in the North Atlantic. Supporting the observations made in many single cases of polar lows, these three composite studies each reveal the presence of a synoptic sea level and mid-tropospheric low over Scandinavia and a ridge over Greenland. The ridge is specific to this region of polar low development. The distribution of the land and sea masses aids the generation of a strong pressure gradient between Greenland and Norway. Businger (1985) found that the time evolution of the geopotential height (with respect to the key day, the date by which a mature polar low had developed) showed ridging over Greenland and a trough over Norway even 3 days before the polar low outbreak. Forbes and Lottes (1985) note that vortices began developing about one day after a fresh cold-air mass entered the area. This gives a hint of the time required between the setting up of favourable conditions and the development of a mature low and thus adds to the importance of recognising synoptic patterns conducive to polar low development.

So far, the synoptic situation has not been climatologically studied from a PV perspective. Several observational studies however have reported the presence of a dry intrusion of stratospheric air to the west of the occluded synoptic low (e.g. Grenå and Kamstra, 1995, Nordeng and Rasmussen, 1992, Claud et al., 2004), which, according to Grenå and Kamstra (1995), typically lowers the tropopause to between 450 and 750 hPa and sometimes can be seen to curl cyclonically into the centre of the synoptic low.

In this section, the synoptic PV environment will be determined for a group of polar low cases using a isentropic composite analysis. Attention will be restricted to cases in the Norwegian and Barents Seas.

It can be expected that a region of higher PV will be found to the west of a composite synoptic scale low, as would be consistent with the observations of stratospheric intrusions described above.
7.2. SYNOPTIC PV ENVIRONMENT

and with the synoptic picture from the three composite studies. It would be of interest to find
the proportion of polar low developments matching this description. The remaining cases, where
this scenario is not applicable, will be classified to see if there are any secondary but frequently
occurring synoptic environments which can be identified as synoptic PV precursors for polar lows.
Sometimes the mesoscale PV anomalies acting as precursors are only just large enough to be re-
solved by forecast models (Browning, 1993) so it would be instructive to understand from the
synoptic PV situation where precursors are likely to form and which mechanisms are likely to be
at work.

Data

A database of polar lows developments in the Norwegian and Barents Seas was kindly supplied by
the Norwegian Weather Service and is used here to select mesocyclones of the polar low category.
The database runs from December 1999 and is still being expanded. The period chosen for the
composite study is December 1999 to May 2002, comprising three winter seasons. This period is
simply the time period common to both the database and the ERA-40 data set. In this time period,
21 polar lows were recorded in the database. The coordinates and maximum wind velocities arc
given in table 7.1 against the key date, defined as the time when the disturbance was first classified
as a fully developed polar low on the basis of satellite imagery.

The tracks of these 21 disturbances, constructed from observations in infrared satellite imagery,
are shown in figure 7.1. The beginning of a track corresponds to the first sign of a cloud feature
which can be later identified with the polar low and, similarly, the end of the track corresponds to
the last observation of a cloud feature which can be identified with the polar low. A polar low on
its key date is thus located in an intermediate section of the track. Inspection of the tracks shows
that disturbances rapidly dissipate upon land fall and that the majority of cases move to lower
latitudes within their lifetimes.

7.2.1 Geographical synoptic PV composite

Composites

The composites are constructed using ERA-40 data at the re-analysis times which most closely
match the key dates given in table 7.1.

Firstly, a geographical composite of all key dates is made for the geopotential height anomaly at
500 hPa \((Z_{500})\) and sea level pressure. The geopotential height anomaly composite is obtained by
subtracting a climatology of \(Z_{500}\) from the \(Z_{500}\) composite of key dates, following Businger (1985).
The climatology of \(Z_{500}\) is constructed for the relevant months of the years included in the study
period and is presented in figure 7.2. The \(Z_{500}\) anomaly composite and the sea level pressure
composite are shown side by side in figure 7.3.

The climatology of \(Z_{500}\) displayed in figure 7.2, compares favourably to the climatology of Businger
(1985), even though Businger’s climatology was constructed from a much larger data set. A wave
number 3 pattern is evident and the most prominent minimum lies over the Labrador Sea. The
anomaly field (7.3 (a)) reveals a pattern strikingly different from the climatological mean. A
strong minimum (120 m below the climatological mean) is positioned over Scandinavia and a
trough extends across the Arctic Ocean. A positive anomaly replaces the low of the climatology
just northwest of the Labrador Sea. The sea level pressure composite (7.3 (b)) confirms the pres­
ence of a synoptic low over northern Scandinavia, with troughs extending towards Siberia and the
Atlantic storm track. A ridge of high pressure resides over Greenland. These results echo the
findings documented in the literature, suggesting that the selection of cases is reasonably repre­
sentative of the polar lows in this region.

Geographic PV composite

Geographical PV composites created on the 290 K, 300 K, 310 K and 320 K surfaces are presented
in figure 7.4 together with selective isobars to define the sea level pressure minimum.
### Table 7.1: These data were kindly supplied by Gunnar Noer of the Norwegian Meteorological Institute in Tromsø. The data are based on AVHRR, synoptics and daily observations made in the area spanning from the Greenwich meridian to Novaya Zemlya and from 65°N to the Arctic ice edge. The list contains most but not all of the polar lows in this area. Criteria used to identify polar lows were: Wind more than 27 kts, cyclonic cloud structure present, diameter 1-500 km. Where synoptic observations were not available, the lows were identified from their appearance in AVHRR imagery. Dates and positions refer to the first image where the low is identified as a fully developed polar low, i.e., early in the life span. Maximum wind is taken from the analysis made at Tromsø covering the entire life span of the low. Horizontal lines in the table separate the three winter seasons.

<table>
<thead>
<tr>
<th>ID number</th>
<th>Key date</th>
<th>Time [UTC]</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Max. wind [kts]</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19.12.1999</td>
<td>1340</td>
<td>72</td>
<td>18</td>
<td>NE 45</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>22.01.2000</td>
<td>0250</td>
<td>72.5</td>
<td>29</td>
<td>NW 42</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>31.01.2000</td>
<td>0610</td>
<td>65</td>
<td>04</td>
<td>NW 50</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>08.03.2000</td>
<td>1900</td>
<td>69</td>
<td>04</td>
<td>NW-NE 35</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>24.03.2000</td>
<td>1230</td>
<td>72</td>
<td>21</td>
<td>NW 35-40</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>04.02.2001</td>
<td>1540</td>
<td>62.5</td>
<td>-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>05.02.2001</td>
<td>1600</td>
<td>65</td>
<td>01</td>
<td>SE 35</td>
<td>'Mike'</td>
</tr>
<tr>
<td>8</td>
<td>24.03.2001</td>
<td>0730</td>
<td>74.5</td>
<td>09</td>
<td>NE 40</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>10.04.2001</td>
<td>0650</td>
<td>76</td>
<td>02</td>
<td>N 50</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>01.11.2001</td>
<td>0200</td>
<td>71</td>
<td>19</td>
<td>N 50</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>06.11.2001</td>
<td>0100</td>
<td>68</td>
<td>06</td>
<td>NW 35</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>12.11.2001</td>
<td>0700</td>
<td>67.5</td>
<td>07</td>
<td>NW 45</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>12.01.2002</td>
<td>1200</td>
<td>73</td>
<td>44</td>
<td>N 35</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>19.01.2002</td>
<td>0400</td>
<td>70</td>
<td>47</td>
<td>N 35</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>22.01.2002</td>
<td>1100</td>
<td>75</td>
<td>28</td>
<td>NE 50</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>23.01.2002</td>
<td>1200</td>
<td>71</td>
<td>17</td>
<td>NE 35</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>19.02.2002</td>
<td>1300</td>
<td>74</td>
<td>34</td>
<td>N 55</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>23.02.2002</td>
<td>1140</td>
<td>67.5</td>
<td>07</td>
<td>NW 45</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>07.03.2002</td>
<td>1810</td>
<td>73</td>
<td>28</td>
<td>N 35</td>
<td>Sec. 7.4</td>
</tr>
<tr>
<td>20</td>
<td>05.05.2002</td>
<td>0650</td>
<td>74.5</td>
<td>12</td>
<td>N 40</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>20.05.2002</td>
<td>1436</td>
<td>73.2</td>
<td>15.3</td>
<td>NE 35</td>
<td></td>
</tr>
</tbody>
</table>

The 290 K PV composite shows that there is indeed higher PV to the west of the composite surface cyclone in the form of a streamer extending from the polar reservoir and curling cyclonically around the synoptic low. The 300 K composite also indicates higher PV to the west of the synoptic low and there is a hint of a local PV minimum positioned on the northern coast of Norway between the streamer and the synoptic cyclone. On the higher isentropes, stratospheric PV is positioned over the centre of the synoptic low in a Rossby wave-like perturbation. The lobe over Europe of the wave number 3 synoptic flow pattern appears to be strengthened with respect to the climatology. PV intrusions are typically narrow on lower (e.g. 290 K) isentropes. Considering the large variation in the track locations shown in figure 7.1, the position of the intrusions will vary substantially from case to case resulting in a large degree of cancellation between tropospheric and stratospheric PV.

### Synoptic cyclone centred PV composite

Considering the prominence of the synoptic low over Scandinavia in the sea level pressure composite and the streamer in the PV composite, it was decided to centre the composite analysis on the synoptic low. A second composite series is constructed in which the PV field for each case is rotated on the sphere such that the centre of the synoptic low is translated meridionally onto the pole. Cases without a dominant synoptic low are omitted from the composite and treated separately.

In total, 10 cases are included in the centred composite presented in figure 7.5. A suitable isentropic surface was chosen which accentuated the synoptic PV structures. Although the amplitude is
7.2. SYNOPTIC PV ENVIRONMENT

diminished in comparison to individual cases, there is clearly an intrusion of higher PV air to the west of the low, originating from the north. A range of streamer orientations is also evident. Tropospheric PV separates the high PV in the intrusion from the high PV lying directly over the synoptic low. Similarly, Grenás and Kvamstø (1995) also noticed a negative PV anomaly northwestward from the synoptic low. They suggest it to be a result of the redistribution of PV caused by the release of latent heat.

It became apparent that this typical picture can also be subdivided into two categories with different evolutions.

A Sheared streamer. The streamer is oriented parallel to the low level flow (figure 7.6 (a)). It undergoes shear instability and rolls up into isolated coherent anomalies. It is likely that several disturbances develop along the shear line. The polar low may develop either side of the streamer.

B Angled streamer. The streamer is oriented at an angle to the low level flow (figure 7.6 (b)) and is advected southward. The polar low develops ahead of the streamer, where there is positive PV advection. Both the streamer and polar low move towards the south or southwest.

It is probable that the strip of tropospheric PV seen to the north of the synoptic low's centre and to the east of the streamer in the centred composite (figure 7.5) has a significant role in type A cases. The tropospheric PV enhances the cross-streamer shear, and so acts to elongate the streamer and reduce its width. According to Browning (1993), the instability and wavelength of
Fig. 7.2: Mean 500 hPa geopotential height (contour interval: 5 dm) for all months in the composite.

(a) 500 hPa geopotential height anomaly
(b) Sea level pressure

Fig. 7.3: Key date composites of (a) the 500 hPa geopotential height anomaly (contour interval: 2 dm), where shaded regions emphasize negative anomalies less than -6 dm (dark grey) and positive anomalies greater than 2 dm (light grey), and (b) sea level pressure (contour interval: 2 hPa where shaded, 4 hPa elsewhere), where shaded regions indicate pressure less than 1000 hPa.
7.2. SYNOPTIC PV ENVIRONMENT

(Fig. 7.4: Geographical PV composite on 290 K (a), 300 K (b), 310 K (c) and 320 K (d). Composite sea level isobars at 1000, 998 and 996 hPa are overlaid.

instability depend on the half width and the magnitude of the anomaly. A separation of less than 1000 km between the regions of maximum winds on either side of the streamer is considered small. This is what appears to have been happening in the case study presented by Nielsen (1997), where PV in the 295 K - 300 K layer revealed a dry intrusion behind the cold front of a synoptic low containing several PV anomalies, one of which lay directly over the baroclinic zone at genesis time. In Nielsen's case, the polar low developed in between the dry intrusion and the synoptic low.

7.2.2 Clustered synoptic scenarios

Around half (11) of the cases were not used in the centred composite. These remaining cases have been grouped into 4 further categories based on the form of the isentropic distribution on 290 K. The clustering is currently subjective. The 'rational' is to be able to make use of isentropic PV maps to recognise the PV signatures of potential polar low precursors also when the typical pattern of a streamer to the west of a synoptic low is not satisfied. The categories are summarised in table 7.2 and described in more detail below. Illustrative examples are included in figure 7.6.

Scenario 1

This category relates to the results from the geographic composite studies here and in the li-
Fig. 7.5: 10-member PV composite on 295 K, centred on the synoptic low. Thin black contours show sea level pressure with a contour interval of 5 hPa. The thick black contour outlines the composite mean dynamical tropopause (2 PVU isoline).

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Name</th>
<th>Number of cases</th>
<th>ID numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Sheared Streamer</td>
<td>4</td>
<td>1 2 8 12</td>
</tr>
<tr>
<td>B</td>
<td>Angled Streamer</td>
<td>7</td>
<td>3 4 5 16 17 18 21</td>
</tr>
<tr>
<td>1</td>
<td>Dominant High</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>Orographic</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>Baroclinic Wave</td>
<td>5</td>
<td>9 10 13 14 20</td>
</tr>
<tr>
<td>4</td>
<td>Cut-off Anomaly</td>
<td>3</td>
<td>6 7 19</td>
</tr>
</tbody>
</table>

Table 7.2: Synoptic PV scenarios. See section 7.2.2 for further details.

A blocking high over Greenland dominates the synoptic picture. There may be a low pressure region over Scandinavia but it is weak and may be small with multiple centres. For the latter reason, it was difficult to include this case into the centred composite. Stratospheric PV is advected southwards by the high pressure region. The resulting tongue of high PV is oriented northeast-southwest, i.e. is anticyclonically curled. The polar low is generated on cyclonic side of the PV tongue and some low level vorticity may be contributed from the centre of an old occlusion. See figure 7.6 (c).

**Scenario 2**

In this category, the polar low is forced by a lower level orographically produced PV anomaly originating near Greenland. These anomalies tend to be small scale and of lower amplitude compared with those in other categories (figure 7.6 (d)). The anomalies move eastward and may be reinforced by anomalies at higher levels being advected south around a synoptic low pressure flow.
7.3. CASE STUDY: THE MIKE POLAR LOW

to the east. An example from the literature of such a case is found in Klein and Heinemann (2002). They show that the topography of east Greenland is conducive to mesoscale cyclone development. Large valleys on the coast channel and converge the flow and cyclonic vorticity is generated by vertical stretching of air columns (see their Fig. 9). The cold air transport also can significantly enhance low level baroclinicity. Klein and Heinemann note that the increased pressure gradient present when there is synoptic low between Greenland and Iceland acts to intensify katabatic flow by encouraging drainage flow from the ice sheet. The low-level PV (PV > 1.5 PVU) shown in their figure 9 is below the 275 K level. Businger (1985) also suggested that development of polar lows in this region could be enhanced by topographical forcing from the Greenland plateau.

Scenario 3

In this category, the PV distribution resembles the situation often found in synoptic scale cyclogenesis (compare figure 7.6 (e) to e.g. figure 3.1 (h)). A growing Rossby wave brings high PVsouthward and the cyclone develops ahead (to the east) of the wave under a region of positive vorticity advection (PVA). In the polar low case, the PV wave is related to a secondary baroclinic zone within the polar air mass, i.e. to the north of the main baroclinic zone. The PV anomaly propagates eastward and has a strong amplitude on 290 K compared to the other categories. It is noted by Rasmussen and Turner (2003) that polar lows occasionally form along a secondary baroclinic zone. An example from the literature is perhaps the 'most beautiful' polar low presented by Granås and Kvamsto (1995). This polar low formed from a small-scale synoptic cyclone in the Norwegian Sea and subsequently crossed eastward towards the Barents Sea near a stratospheric dry intrusion.

Scenario 4

In this last category, the PV precursor takes the form of a ‘cut-off’ anomaly (figure 7.6 (f)). The cut-offs of the relevant cases from the database formed from a mature cyclonically rolled streamer from a previous synoptic flow. The anomalies are not part of the dry intrusion to the dominant synoptic cyclone of the scene but instead are isolated in tropospheric air downstream of the streamer. This PV environment can be related to upper-level troughs trailing a synoptic cold front. These anomalies are coherent and isolated in tropospheric air. They move westward during at least part of their lifetimes and have a long history. For some cases in this category it might be that the conventional synoptic picture (mature synoptic cyclone, baroclinic flow over the ocean) is satisfied but that the precursor anomaly is not located in the vicinity of the main stratospheric intrusion.

The categories presented in this section show that there are synoptic environments conducive to polar low cyclogenesis which differ from the well known scenario. The PV perspective is instructive in understanding the differing physical mechanisms involved in cyclogenesis. To better assess the validity of the suggested categorisations and the variability within each category, a larger data set should be analysed. A more objective approach could be taken by using a quantitative analysis of the low properties to group the cases. It would also be important to determine the frequency with which polar lows occur when each of these PV environments is present. This should confirm whether or not these signals provide useful clues to the onset of polar low cyclogenesis.

7.3 Case study: The Mike polar low

Case 7 of table 7.1 was selected for further study due to its clear association with a coherent upper-level PV anomaly which could be traced back several days from the time of cyclogenesis. After having chosen this case, it was realised that the same case had also been used in a module for training forecasters about polar low forecasting techniques due to particularly good observational coverage when the storm was at its most intense. The training study\textsuperscript{4} focuses on conventional forecasting methods without consideration of potential vorticity, although reference to positive vorticity advection is briefly made. A description of the storm’s evolution follows, which is partially

\textsuperscript{4}See http://met.no/english/topics/annemk.2005/polar_low_cases/
7.3.1 Evolution of events - conventional viewpoint

The polar low, Mike, named after a weather ship on its track, developed in the Norwegian Sea on 5th February 2001, reached a diameter of 300-400 km and decayed about 18 hours later. A sequence of satellite images illustrating the development is provided in figure 7.7. The time of each
7.3. CASE STUDY: THE MIKE POLAR LOW

The synoptic situation at the surface was not the typical one for the area in which this polar low developed. Instead of a synoptic low over Scandinavia, an elongated synoptic low pressure region lay to the south over western Europe and the Atlantic and high pressure dominated over northwestern Russia, which together gave rise to persistent south-easterly winds over Norway. The flow resulted in a cold air outbreak (evident from the cloud streets seen in satellite imagery e.g. in figure 7.7 (b)) which brought very cold, dry continental air out over the relatively warm sea and acted to destabilise the vertical profile there. It is unusual that the cold air outbreak should come from Scandinavia, but the development of favourable conditions was otherwise fairly typical.

An area of slightly thicker stratiform cloud off the north coast of Norway (figure 7.7 (a)) reveals there was low level baroclinicity (a low level warm anomaly) and enhanced convection in the region. This structure, possibly the remains of an old occluded front, was partly responsible for the subsequent polar low development and can be related to a surface trough oriented NE-SW lying just off the Norwegian coast in the analysis at 00 UTC on 5th March.

At 500 hPa there was an upper level low over Scandinavia featuring two main troughs with ridging in between (see figure 7.11 (a)). One trough was oriented NE-SW along the Northern coast of Norway, parallel to and behind the surface trough, and had a large area of positive vorticity associated with it. It marked the leading edge of the cold air outbreak.

By 06 UTC on 5th March, the stratiform cloud feature corresponding with the position of the surface trough had moved south and formed a cloud cluster with some signs of cyclonic curvature (figure 7.7 (b)). Enhanced convection is likely to have taken place, due to dynamical ascent. The difference between the sea surface temperature and 500 hPa temperature exceeded 45 °C in the Norwegian Sea, a condition1 known to be conducive to polar low genesis in this region. The surface trough moved westward slowly keeping pace with the upper level feature.

By 12 UTC on 5th March, a surface low pressure minimum was positioned ahead of the 500 hPa trough axis and formed a cloud cluster with some signs of cyclonic curvature (figure 7.7 (c)). By 15 UTC, Mike reached its most negative value of -9.2 hPa. The enhanced convection had been organised into a cyclonic system with a diameter of 300-400 km (figure 7.7 (d)). Wind speeds exceeded 35 kts. The polar low was positioned just ahead but close to the axis of the 500 hPa trough, coinciding with an area of PVA. The low continued moving slowly south with little change in intensity.

The convective clouds took on a more individual appearance and the cloud signature became more spiraliform (figure 7.7 (d)-(f)). The upper level trough axis and formed a cloud cluster with some signs of cyclonic curvature (figure 7.7 (c)). By 12 UTC on 6th February, the Mike low had dissipated and by 06 UTC on 7th February (figure 7.7 (e)), cloud associated with the synoptic low had removed any remaining cyclonic structure associated with the polar low from sight.

Mike would fit into the short-wave/jet streak type of Businger and Reed (1980), characterised by a secondary vorticity maximum, PVA aloft, deep, moderate baroclinicity, modest surface fluxes and a typical comma cloud form. Yet Mike takes on more of a spiraliform structure towards the end of its life, an indication of the increasing domination of convective processes.

7.3.2 Synoptic PV environment and evolution

The dynamical pattern on isentropic surfaces is mainly associated with regions of sharp PV gradients in the region of the tropopause, i.e. jet-stream activity. The dynamics of the tropopause is of high importance for baroclinic developments and surface cyclogenesis. To examine the synoptic PV development, we therefore select a surface where there are fluctuations of the tropopause (2 PVU contour) on a synoptic scale. In this case, the 305 K level is appropriate. On the other hand, to follow the behaviour of the upper mesoscale precursors and their interactions with the surface, a
Chap!~er 7. POLAR LOWS AND STRATOSPHERIC PV ANOMALIES

Fig. 7.7: Infrared imagery showing the development of the Mike polar low. Arrows indicate the location of the Mike polar low disturbance. Images from the University of Dundee.

lower isentrope (290 K) is chosen (presented in section 7.3.3) where the precursors take the form of mesoscale stratospheric intrusions.

Figure 7.8 displays the synoptic evolution as a sequence of 305 K isentropic PV maps from 27th January 2001 (9 days before cyclogenesis) to 7th February 2001 (one day after dissipation). Isobars are drawn for sea level pressure below 1000 hPa and above 1040 hPa to draw attention to regions of low and high pressure respectively.

On 27th January (figure 7.8 (a)), the Norwegian Sea is under the influence of a mature low pressure system and a Rossby wave is growing over the west Atlantic where there is a developing synoptic cyclone. By 28th January (b), the amplitude of the Rossby wave has increased. There is a large poleward protrusion of tropospheric PV over the Atlantic, while stratospheric PV extends meridionally over the UK towards Spain. From 30th January (c) to 31st January (d), tropospheric PV pushes further north (poleward of 80°N) and connects with a tropospheric cut-off, severing the connection of the high PV region now over northeastern Europe with the stratospheric air to the west of Greenland. The wind vectors show that the flow across northern Scandinavia has correspondingly changed from northwesterly to northeasterly. Scandinavia is now under the influence of a Siberian air mass and a region of high pressure is advancing from the east. At about this time, cloud streets can be seen from satellite imagery, revealing that a cold air outbreak is forming over the Norwegian Sea (5 days before cyclogenesis). Meanwhile, high PV air is directed
7.3. CASE STUDY: THE MIKE POLAR LOW

southwestwards across Scandinavia and the stratospheric streamer extends into the Mediterranean (e). The surface conditions over the Norwegian Sea become 'ripe' for polar lows (see figure 7.9) but the high PV lies over the continent and no polar low developments are observed.

The southern part of the streamer cuts off over the Mediterranean (f) and the remainder of the streamer curls cyclonically over northeastern Europe (g). At this time, a 'elbow' of stratospheric PV is edging out from southern Norway over the North Sea towards Scotland. A surface disturbance is formed in this shear zone and develops into a mesoscale cyclone under the influence of this upper level PV structure (polar low 6 in figure 7.1). On 3rd February (h), the tip of the streamer lies over northwestern Russia at about 45°E, 65°N. This tip develops into the upper-level PV precursor of Mike and its evolution is presented in the next section. Backward trajectories were computed originating in the tip of this streamer (figure 7.10) and confirm the Siberian origin of the air mass.

Just north of the streamer tip, the streamer is very narrow and by 5th February (i), it splits from the Siberian stratospheric reservoir into a cut-off which is located over Scandinavia, the Baltic Sea, Northern Germany, Poland and Belarus. The rest of the streamer continues moving eastward towards Russia and looses its influence on the cut-off. A small synoptic cyclone can be seen approaching the UK, connected with a larger region of low pressure to the west. The stratospheric cut-off continues to rotate cyclonically (j), surrounded by tropospheric air and brings anomalously high PV over the Norwegian Sea which triggers the polar low Mike.

By 18 UTC on 5th February (k), three waves (indicated by the white arrows) of stratospheric PV can be seen moving westward in succession along the northern perimeter of the synoptic low pressure region extending over the Atlantic. Each of the three anomalies can be connected with a polar mesocyclone genesis. Polar low 6 is decaying to the south of Iceland where there is a weak (inverted) surface trough. A local minimum in sea level pressure shows the location of Mike, near the Greenwich Meridian. The third anomaly is approaching the Norwegian coast and triggers a mesoscale cyclogenesis about 6 hours later. The eventual merger of the mesoscale systems with the synoptic system brings an end to the cold air outbreak and the stratospheric cut-off is seen to be reconnected to the main stratospheric body (l).

7.3.3 Isentropic anomaly history

The 290 K isentropic surface is chosen to follow the upper level development of the mesoscale PV precursor. Although the PV structure can be detected on higher surfaces, e.g. 320 K, with larger PV values, it is much more pronounced on the lower isentropic surface because there it takes the form of a stratospheric intrusion into the troposphere. Evolution on the 290 K surface is shown in figure 7.11.

On 3rd February 12 UTC (panel (a)), a curled fragment of the stratospheric streamer can be seen lying over northern Europe and is associated with a low in the 500 hPa height. The tip of this fragment at about 57°N, 36°E, is being sequestered to form a cut-off anomaly. Six hours later (b), the cut-off has formed and the 2 PVU surface penetrates down to 500 hPa. The cut-off's longer axis is aligned with a 500 hPa trough, oriented SW-NE. Between this time and 5th February 00 UTC (panels (c) to (e)), the cut-off rotates with the trough, about the 500 hPa low, toward the Norwegian coast. In panel (e), a trough in the surface pressure (1015 hPa) can be seen to extend from the synoptic low towards and ahead of the PV precursor. By 5th February 12 UTC (g), a closed circulation forms at 500 hPa and the surface trough deepens to 1005 hPa. Three hours later, the surface disturbance is classified as a polar low. By 5th February 18 UTC (h), the PV anomaly is positioned over the polar low and there is a local minimum in surface pressure. A second PV anomaly approaches the coast and the two anomalies begin to interact. The Mike PV anomaly starts to deform into a comma-like form (i-j), possibly a signal that convective processes are beginning to dominate. By 6th February 06 UTC, the 500 hPa low moves away from the Mike low and towards the second disturbance. The form of the 995 hPa isolars indicates that there are now two surface lows. On 6th February 12 UTC, the two upper level PV anomalies appear to merge. The Mike polar low is dissipating and the second low-level disturbance becomes the dominant feature. The last panel (l) shows the synoptic low pressure region approaching from the south which eventually envelops the Mike polar low.
Fig. 7.8: Isentropic PV on 305 K in colour. Sea level pressure isolines are plotted at intervals of 5 hPa at and above 1040 hPa, and at and below 1000 hPa. Wind vectors depict the isentropic flow. The white arrows in panel k point to 3 stratospheric PV anomalies associated with polar mesocyclone developments.
7.3. CASE STUDY: THE MIKE POLAR LOW

Fig. 7.8 cont.

Fig. 7.9: Synoptic preconditioning. On 31st January 2001 (a) the Norwegian Sea is free of organised convective activity. A day later on 1st February 2001 (b) cloud streets have formed in the cold air streaming out from Norway over the Norwegian Sea. Images from the University of Dundee.
Chapter 7. POLAR LOWS AND STRATOSPHERIC PV ANOMALIES

Fig. 7.10: Backward trajectories (black), originating in the 290 K streamer tip on 3rd February 2001, 18 UTC, and traced back to 30th January 2001, 00 UTC, across Scandinavia and Siberia to the periphery of a region of stratospheric PV (shaded dark grey). The background PV field is valid at the earliest time, 30th January 2001, 00 UTC.

7.3.4 Vertical sections

Figure 7.12 displays vertical PV cross sections passing through the centre of the upper level PV anomaly. Equivalent potential temperature, relative vorticity and condensational heating are also displayed (see figure caption for details). During the sequence, the relevant anomaly moves from the right-hand side to the left-hand side of the domain.

On 3rd February 12 UTC (figure 7.12 (a)), the PV anomaly responsible for the triggering of Mike is seen at a longitude of about 37°E. At this location, downward penetrations of PV isosurfaces associated with the anomaly can be seen in the lower stratosphere, up to nearly 200 hPa, where the PV shows a local maximum above 7 PVU. The lower limit of the anomaly, taken as the 2 PVU isoline (tropopause) extends down to an altitude of about 500 hPa and remains between 500 and 550 hPa over the entire sequence. The cyclonic wind field associated with the anomaly is displayed using relative vorticity. A maximum in relative vorticity is located at about 400 hPa or 295 K, above which the isentropes bow downwards and below which the isentropes bow upwards, as is typical for an isolated upper level PV anomaly (see Hoskins et al. (1985)). Below the anomaly, at about 900 hPa, there is a low level inversion (high density of horizontally oriented isentropes), capping the polar boundary layer. Strong radiative cooling below the inversion has resulted in a low-level negative temperature anomaly and a corresponding low-level PV dipole of anomalously negative PV below anomalously positive PV. Lambert et al. (2004) similarly noticed the presence of high PV values in a very stable boundary layer ($N_{900} \simeq 1.9 \times 10^{-2} \text{Hz}$) where radiative cooling was taking place.

The vorticity fields of the upper and lower positive PV anomaly reinforce each other at mid-
7.3. CASE STUDY: THE MIKE POLAR LOW

Fig. 7.11: Isentropic PV on 290 K (colour shading) with 500 hPa geopotential heights (dashed at intervals of 5 dm) and sea level pressure (solid blue, at intervals of 5 hPa, at and below 1015 hPa). Contours of the 2 PVU surface are drawn at 450 hPa (yellow), 500 hPa (green) and 550 hPa (violet).

tropospheric levels, creating a column of vorticity through the depth of the troposphere. A cold and stable boundary layer inhibits cyclogenesis due to subsidence the subsequent decrease in convective activity. The strength of the upper anomaly’s influence is attenuated at low levels by the inversion, below which there is subsiding air and no sign of condensational heating. The intense region of convective heating and lower level vorticity seen to the far west of the section is connected with the polar low (number 6 in table 7.1) which formed between Norway and Scotland.
By 4th February 06 UTC (h), the upper anomaly has moved further north and the Norwegian Sea is captured west of 10°E in the cross section. The local environment of the upper anomaly is similar to that in panel (a) except that the boundary layer's ability to inhibit penetration of the anomaly's wind field is more evident. The depth of the convective boundary layer increases over the sea, also under the influence of another PV anomaly aloft.

On 4th February 18 UTC (i), the upper anomaly is positioned just east of the Norwegian coast. There is a shallow baroclinic zone separating stably stratified air over the continent and a region of low stability over the Norwegian Sea. Above the sea, stratification is close to being horizontal and might here be considered as an in-situ reference state into which the upper level anomaly will be advected. The top of the convective boundary layer height is about 750 hPa. Such a deep
7.3. CASE STUDY: THE MIKE POLAR LOW

Convective boundary layer is not unusual for this region (Granás and Kvarnstø, 1995).

On 5th February 00 UTC (d), the anomaly crosses the Norwegian coast and cold air from the continent is advected across the sea underneath it. The boundary layer is raised to the west and directly underneath the PV anomaly. The reversed gradient of $\theta_e$ with height, near 10°N, immediately west of the Norwegian coast, indicates a region of convective (potential) instability. Into this region, a tongue of positive vorticity descends beneath the PV anomaly.

By 5th February 06 UTC (e), the anomaly is situated over the sea, the boundary layer is clearly locally raised beneath it, compared to (e), and an area of condensational heating is seen beneath the locally deeper boundary layer together with an area of positive vorticity which shows the spin-up of the surface low. Rastu and Midltoe (1996) also point out that the stability is very low beneath the upper level PV anomaly in their case. They suggest that a large Rossby penetration depth allows the upper level PV anomaly to interact with and intensify the low-level temperature anomaly. The Rossby penetration depth will be calculated for Mike in section 7.3.5.

By 5th February 12 UTC (f), a low level warm anomaly is seen to have formed on the western edge of the low level vorticity maximum. In this case, the warm air is found to the north of the polar cyclogenesis and cold air to the south. The low level vorticity maximum increases in amplitude below the upper maximum. A similar increase is noted in vorticity by Claud et al. (2004). A column of vorticity extends through the depth of the troposphere to sea level.

By 5th February 18 UTC (g), the low level vorticity maximum reaches its maximum amplitude and is aligned with the warm anomaly and the area of condensational heating. The extremely low static stability is localised below the upper anomaly, where cyclogenesis is taking place.

By 6th February, 00 UTC (h), it is evident that the penetration of the upper level anomaly's wind field has deepened, although the lowest part of the anomaly now lies above 550 hPa. A region of low PV in the mid-troposphere has moved beneath the upper PV anomaly and acts to increase the penetration. Its presence is probably not entirely due to the creation of negative PV from diabatic heating. The area favourable for convection widens and a new region of condensational heating can be seen near the Norwegian coast. The circulation associated with the polar low Mike decreases as Mike starts to dissipate.

In this case study, the altitude of the PV intrusion does not vary greatly (as can also be seen from inspecting the contours of the 2 PVU surface presented in figure 7.11). It appears that the initial role of the PV anomaly was to locally increase the depth of the convective boundary layer, thereby increasing potential for convection in an area of the same order of scale as the PV anomaly. After cyclogenesis, a region of low PV air appears under the upper PV anomaly. It may have been simply advected into this position (the ERA-40 data shows the region of negative PV generation from condensational heating to be confined to much lower levels) but documentation in the literature of similar anomalies would seem to suggest that the anomaly is a bi-product of cyclogenesis and the convective processes leading to its formation were not adequately resolved by the ERA-40. In any case, the presence of the low PV anomaly results in further decrease of the mid-tropospheric static stability and the circulation of the upper level positive PV anomaly can penetrate further towards the surface, enhancing surface spin up.

7.3.5 Estimating the strength of upper level PV penetration

Rossby penetration depth

The Rossby penetration depth is a scale height which estimates the vertical penetration of the flow structure below (and above) the location of a PV anomaly. It is illustrated schematically in figure 7.13. The relationship can be derived by applying the invertibility principle in isentropic coordinates to a circularly symmetric PV anomaly in a horizontally uniform and statically stable reference state (Hoskins et al., 1985, Rasmussen and Turner, 2003). The assumed balance conditions are the gradient wind and hydrostatic relations. It is also required that the mass between any two isentropic surfaces be conserved. In isentropic coordinates, the Rossby penetration depth is expressed as

$$\Delta \theta \sim \frac{\rho_0 f L}{(g \theta)^{1/2}}, \quad (7.1)$$
Fig. 7.12: West-East cross sections through the upper level PV anomaly associated with the Mike polar low, showing PV (colour shading), equivalent potential temperature (black contour 5 K interval), relative vorticity (purple contour), regions of condensational heating (green contour). Longitude is given on the horizontal axis and pressure (hPa) on the vertical.
7.3. CASE STUDY: THE MIKE POLAR LOW

where $\rho$ is the density, $\theta$ is potential temperature, $f$ is the Coriolis parameter, $L$ is a horizontal length scale, $g$ is the gravitational acceleration and $\sigma$ is the inverse static stability defined as $\sigma = -\frac{1}{\rho \frac{\partial \rho}{\partial \theta}}$. The Rossby penetration depth can be transformed into physical space using the hydrostatic balance, $\Delta p = -\rho g \Delta z$, giving

$$H \sim \frac{f L}{N}. \quad (7.2)$$

This relation is appropriate when the relative isentropic vorticity is small, compared to the Coriolis parameter. This is however often not the case, with large relative vortices (up to 3$f$) having been associated with cycloonic synoptic scale flow in polar low developments (Rasmussen and Turner, 2003).

To take account of the 'background' synoptic vorticity ($\xi_{\text{av}}$), a modified scale height that replaces $f$ with $\sqrt{f(f + \xi_{\text{av}})}$ may be applied (see Rasmussen and Turner, 2003, for a derivation):

$$H^2 = \frac{f(f + \xi_{\text{av}})L^2}{N^2}. \quad (7.3)$$

The Rossby penetration depth has been applied conceptually in the literature to explain the relative influence of $\bar{\Gamma}$ anomalies on the atmosphere beneath them in different background scenarios (e.g. Røsting and Midtbø, 1996). Both observations (Grønås and Kvaamstø, 1995) and numerical sensitivity experiments (Lambert et al., 2004, Bresch et al., 1997) have illustrated that close proximity of a $\bar{\Gamma}$ anomaly to a surface disturbance is not necessarily sufficient for polar low cyclogenesis to occur, since stable boundary layers can inhibit the cyclogenetic action of the upper anomaly. In the cases studied by Grønås and Kvaamstø (1995), the Rossby Depth was mainly controlled by the vertical stability. It is noted that a $\bar{\Gamma}$ anomaly moving from a region with a stable stratification into a less stable region experiences an increase in the Rossby penetration depth (e.g. Rasmussen and Turner, 2003, Rasmussen, 1985). This is one reason why polar lows tend to be triggered only once the upper anomaly has passed from stable conditions over ice out over open water, where the air column is destabilised due to advection of cold air across the relatively warm sea.

The Rossby Depth is not usually used quantitatively as a diagnostic measure. The only exception the author has found is in Moore et al. (1996), where the Rossby depth is calculated to give credibility to their claim that a upper level $\bar{\Gamma}$ anomaly influenced a polar low's development. They take $N = 6 \times 10^{-5}$s$^{-1}$ and $L = 500$ km, giving a depth on the order of 8 km, which they say is sufficiently large to allow interaction.

Alternative measures of interaction have however been proposed, e.g. the height between the dynamical tropopause and the surface boundary layer (Grønås and Kvaamstø, 1995). Grønås and Kvaamstø (1995) studied four cases which satisfied the synoptic precursor conditions but in which only two resulted in polar low developments. The distinguishing factor was the height between the dynamical tropopause and the convective boundary layer being less than 1000 m. This distance is required to be small so that mutual interaction of the upper and lower anomalies is strong enough to trigger polar low development. Grønås and Kvaamstø propose the use of two maps to evaluate the risk of polar lows by the forecaster: one being the pressure of the tropopause and another, the height of the convective boundary layer together with surface potential temperature. Maps of boundary layer and tropopause height are not however routinely available for forecasters.

In the remaining part of this section, the strength of penetration of the upper-level $\bar{\Gamma}$ anomaly which triggers the Mike polar low will be estimated using the Rossby penetration depth over the lifetime of the anomaly.

If the synoptic isentropic relative vorticity stays approximately constant over the upper-level anomaly lifetime, it will contribute little to the relative changes in $H$ with time. Equation 7.2 is adequate for the purposes here and is chosen in preference to equation 7.3 for simplicity. The physical vertical coordinate system is chosen instead of the isentropic system so that ideas can be easily applied to forecasting.

The Mike case is particularly attractive for analysing the penetration as the anomaly is isolated and has a coherent and circular form such that the properties of the flow found by Hoskins et al.
According to relation (7.2), the Rossby penetration depth is proportional to the horizontal scale of
the anomaly and to the planetary vorticity and is inversely proportional to the static stability. It
has already been seen that the upper PV anomaly for Mike formed over a cold continental stable
boundary layer and moved out over the Norwegian Sea into a deep mixed convective boundary
layer with low static stability. As low static stability contributes to a large Rossby penetration
depth, an increase in the latter is expected as the anomaly crosses the coast.

The approach taken to estimate the three parameters $f$, $L$, and $N$, defining the Rossby penetration
depth, is outlined below.

**Coriolis parameter, $f$**

The value of the Coriolis parameter at the latitude of the anomaly centre is taken.

**Horizontal length scale, $L$**

We calculate the horizontally projected area of the anomaly defined on the 290 K surface as a cut-off
with PV values greater than 2 PVU. We define the horizontal length scale to be the diameter of
a circle having the same area as the anomaly. This estimation is only appropriate if the anomaly
is distinct from neighbouring anomalies on 290 K.

**Static stability (Brunt Väisälä frequency), $N$**

We want to estimate the Brunt Väisälä frequency of the environment experienced by the PV
anomaly in the region of its penetration without the need to remove the anomaly surgically. A
simple approach is to find a vertical pressure range which lies above regions affected by diurnal
effects of the planetary boundary layer (Ziv and Alpert, 2003), steep orography or friction and
below regions strongly influenced by the anomaly. We want to measure the environmental static
stability and not the local reduction that occurs directly below a positive PV anomaly (Hoskins
etal., 1985). The layer 850-900 hPa was selected by inspection. The value for $N$ should be rep-
resentative of the environment experienced by the anomaly and so the static stability is averaged
over a circular area of radius 300 km (a similar size to the anomaly), centred on the anomaly at
290 K. Averaging over an area also smooths out some of the small scale fluctuations.

![Rossby height or 'penetration depth' for an axi-symmetric PV anomaly, after Rasmussen and Turner (2003). The amplitude of the azimuthal wind $U$ induced by the anomaly is shown as a function of the potential temperature $\Theta$, and is a maximum $U_{max}$ at the level of the PV anomaly. The Rossby penetration depth can be transformed to physical space using the hydrostatic balance.](image)

Fig. 7.13: Rossby height or 'penetration depth' for an axi-symmetric PV anomaly, after Rasmussen
and Turner (2003). The amplitude of the azimuthal wind $U$ induced by the anomaly is shown as a
function of the potential temperature $\Theta$, and is a maximum $U_{max}$ at the level of the PV anomaly.
The Rossby penetration depth can be transformed to physical space using the hydrostatic balance.
Results

Figure 7.14 depicts the temporal evolution of the Rossby penetration depth for the upper level PV anomaly associated with the Mike polar low. On the approach of the anomaly to the coast, the Rossby penetration depth increased gradually. A sharp increase (about 1/3 of the total increase) occurs over 6 hours, just after the anomaly has crossed the coast and prior to cyclogenesis. At the point of cyclogenesis, the Rossby depth is estimated to be about 6 km. The Rossby penetration depth continues to increase as the anomaly moves out over the open ocean, until about 00 UTC on 6th February, at which point a plateau is reached (not shown).

At first the Rossby depth is too small to allow the penetration of the anomaly’s circulation through to the boundary layer, but as the Rossby depth increases, the anomaly is indeed able to have a stronger influence by lifting the inversion (see figure 7.12 (d)). Initially, the penetration is not deep enough to reach the surface. This situation changes later when surface (convective) processes act to reduce the static stability in the boundary layer. The consequential increase in penetration depth allows the anomaly’s wind field to penetrate further towards the surface and increase spin up (see figure 7.12 (g)-(h)).

The tendency of $\Pi$ can be split into contributions from its component variables $f$, $L$ and $N$ as shown in below (Eq. 7.4). In this way, the relative importance of the contributions of each variable to the temporal evolution can be assessed.

$$dH = H_f f' + H_L L' + H_N N' + \text{h.o.t.}$$

$$= \frac{L}{N} f' + L \frac{N}{N^2} dL - \frac{f L}{N^2} dN,$$

where $H_f = \frac{\partial H}{\partial f}$, $H_L = \frac{\partial H}{\partial L}$ and $H_N = \frac{\partial H}{\partial N}$.

---

Fig. 7.14: Evolution of the Rossby penetration depth for the upper level PV anomaly associated with the Mike polar low. The total Rossby penetration depth is shown by the thick contour. Contributions from $f$ (dashed), $L$ (dotted) and $N$ (dash-dotted) are also displayed, as computed by integrating the individual terms in equation 7.4 with time. The approximate times of the upper level PV anomaly’s coastal crossing and the polar low cyclogenesis are marked by dashed vertical lines.
The terms $H_f$, $H_L$, and $H_N$ are evaluated at the mid point between time steps according to equation 7.5 below. The cumulative contributions to the Rossby depth are displayed in figure 7.14. The individual evolutions of $f$, $L$ and $N$ over the same time period are shown separately in figure 7.15.

\[
H_{t+\Delta t} - H_t \approx \left( \frac{L}{N} \right)_{t+\frac{\Delta t}{2}} \left[ f_{t+\Delta t} - f_t \right] + \left( \frac{L}{N^2} \right)_{t+\frac{\Delta t}{2}} \left[ L_{t+\Delta t} - L_t \right] - \left( \frac{fL}{N^2} \right)_{t+\frac{\Delta t}{2}} \left[ N_{t+\Delta t} - N_t \right].
\] (7.5)

Inspecting the contributions of the component variables to the Rossby penetration depth in figure 7.14, it can be seen that before the coastal crossing, the contributions from $N$ and $L$ are positive and about equal. The largest contribution to the sharp increase in penetration depth is due to the destabilisation of the underlying layers as the anomaly moves over the sea. The trend in the Coriolis parameter acts in the opposite direction to reduce the penetration depth as the anomaly starts to move south; it has little influence on the trend of the penetration depth, but the large values at polar latitudes will contribute generally to larger penetration depths in the polar regions. Just after the anomaly has crossed the coast, when the first signs of cyclonic circulation appear in satellite imagery, the Rossby penetration depth has a value of about 4 km.

In order for the Rossby penetration depth to be applied to a forecasting situation, it should be estimated in a straightforward manner from standard model outputs. Simple measures of static stability have been used in the literature. Kolstad (2006) approximates $N$ as the real part of $N = \sqrt{g \frac{\Delta \ln \theta}{\Delta z}}$ in the 700-925 hPa layer and Nielsen (1997) used a measure of bulk dry static stability in the layer 900 - 400 hPa, simply calculated as $\frac{\Delta \theta}{\Delta z}$, which was used to reveal a local minimum located over a baroclinic zone. The Brunt Väisälä frequency could be estimated using the formula given by Kolstad (2006) and standard charts for $Z$ and potential temperature on 500 and 850 hPa, for example:

\[
N = \sqrt{g \frac{\ln \theta_{500} - \ln \theta_{850}}{Z_{500} - Z_{850}}},
\] (7.6)

The horizontal length scale could be estimated by eye and the Coriolis parameter calculated from the central latitude.

Determining the Rossby penetration depth at the time of cyclogenesis would suggest a suitable Rossby penetration depth threshold to be about 5 km for the region investigated and the methods used. The measured height would certainly depend on the methods used to estimate the scales of the fluctuations of the contributing variables. Results from different methods would first need to be calibrated, using a PV inversion method, for example, before a 'universal' threshold could be applied.
7.4 A potential vorticity perspective on forecasting

7.4.1 Forecasting challenges

Due to the severe weather accompanying the passage of a polar low, it is obviously desirable to accurately forecast these events.

The small scale and short lifetime of polar lows pose problems for forecasting. Polar lows are often difficult to observe by conventional means as they can slip through the low spatial and temporal resolution of the observation network unnoticed and are frequently located in sparsely populated areas or over water, where observations are few (Renfrew, 2003, Condron et al., 2006). Polar lows are at the limit of the spatial resolution of global operational forecast models, which do not, on the whole, provide a good representation of their surface development. This situation however will no doubt improve as advances in computing permit higher model resolutions (Renfrew, 2003). Operational mesoscale forecast models in polar regions on the other hand have been used to successfully forecast the genesis (once over 4 days in advance) and to detail the development sequence. However, the timing or positioning of cyclogenesis is often erroneous (e.g. Bresch et al., 1997). As mentioned in the introduction to this chapter, the majority of polar lows are triggered by flow anomalies aloft as they approach lower level baroclinic zones. Since the low level baroclinic zones are strongly related to fixed surface features and the upper level features can usually be resolved and handled successfully by forecast models (Røsting and Midtbø, 1996), in principle, it is possible for the early stages of polar low formation to be well predicted from a knowledge of the upper level dynamics (Zick, 1984, Nordeng, 1990, Bresch et al., 1997).

Later stages are more challenging to forecast. Diabatic processes become active and tend to increase the rate of development and intensity of the storm whilst decreasing its scale. Even if the storm itself is not sub-grid scale, the diabatic processes are and must be parametrised. Success of the model in forecasting subsequent stages is highly depended on the size and structures and type of the polar low in consideration (Businger and Reed, 1989).

7.4.2 Traditional forecasting methods

Despite these challenges, progress has been made by defining conditions under which polar lows develop.

A common approach taken by forecasters is to look for favourable synoptic conditions (Businger and Reed, 1989) and identify local areas where there is a significant risk of polar low development. One result of the Norwegian Polar Lows Project (Lystad, 1986) was the setting out of criteria which could be used to identify the danger areas. These were (1) cold air advection at the sea surface (2) 850-500 hPa thicknesses of less than 1.96 m (adjusted according to the sea surface temperature), (3) cyclonic or no curvature of contours at 500 and 700 hPa levels. Businger and Reed (1989) state that this method successfully predicted all polar lows that occurred, but resulted in just about as many false alarms.

Rasmussen et al. (1993) showed that a significant polar low will often develop if an upper level vortex with a 500 hPa temperature below -42 °C passes the ice edge (indicative of a potentially unstable stratification). This threshold is in current operational use as well as a similar requirement that the difference between the sea surface temperature and the temperature at 500 hPa be more than 45 °C (see teaching module on the Internet).

Such rules are very straightforward to apply but are focussed on the lower half of the troposphere. In commenting about a recent polar low forecast for November 2005 that turned out to be a false alarm, Noer (2005, personal communication), a forecaster and researcher for the Norwegian weather service, suggested that the absence of surface development was probably at least partially due to a lack of upper forcing and that better tools such as PVA are needed to diagnose the upper level forcing.
7.4.3 Prospects for use of isentropic PV in forecasting

There are several properties of PV which make the PV approach particularly attractive for diagnosing the upper level forcing of polar lows:

- PV is conserved on isentropic surfaces for adiabatic flow. At upper levels, the flow can often be treated as adiabatic, as frictional and diabatic forcing is very small. Thus upper level PV anomalies are confined to and easily traced on isentropic surfaces. Potential precursors (anomalies which are forecast to move over a known polar low genesis region) can be identified and carefully monitored days in advance of a surface signal and probably also ahead of a clear signal at 500 hPa.

- The presence of PV structures can be verified and their evolution monitored by other passive tracers. Since stratospheric intrusions can be identified by their large total column ozone amounts and low humidities (Moore et al., 1996; Renfrew, 2003), satellite water vapour imagery (Røsting and Midtbo, 1996; Santurette and Georgiev, 2005; Browning and Dicks, 2001) and, where available, radar reflectivity (Browning and Dicks, 2001) can be used to monitor the upper-level PV dynamics, verify the model PV analyses (Røsting and Midtbø, 1996), and potentially direct aircraft or ships to good locations for making in-situ measurements (Moore et al., 1996).

- Due to the sharp contrast between tropospheric and stratospheric PV values, the upper precursor, represented as an intrusion of high stratospheric PV into the tropospheric background of low PV, is often very prominent. In contrast, background pressure gradients at the surface, or geopotential gradients at 500 hPa, can be large, making the often small scale and small amplitude signal of the polar low or its precursor difficult to detect. Even at maturity, some polar lows are only detectable in the surface pressure as weak troughs rather than as surface ‘lows’ with closed isobars.

7.4.4 Influence of upper level PV on polar low development in ECMWF polar low forecasts

In this section, the nature of the ECMWF forecast model performance for case 19 of table 7.1 is presented. The polar low is short lived, surviving for less than a day before making landfall, but has a long-lived upper-level PV precursor. The polar low genesis in this case is similar to the Mike case in that it can be associated with an isolated and coherent upper level PV anomaly but differs in that there is just one polar low development. Like Mike, it falls into the fourth category (cut-off) synoptic situation presented in section 7.2.2.

Browning (1993) states that the simplicity of such a case makes it ideal for exploring the performance of dynamical components of the forecast model. He suggests that, at least in simple circumstances, an NWP model is able to produce a good forecast even of rather small upper level features if their evolution is controlled by larger scale flows which are well represented in the analysis.

Several others have used ECMWF forecast models in conjunction with polar low studies. In a comparison of the ECMWF model with a mesoscale forecast model, Nordeng (1999) found that the two models show similar large scale PV features but that the ECMWF model did not capture the very fine filament extensions at the tip of the stratospheric intrusion which were seen to wrap the polar low cyclonically in the mesoscale model. Both Nordeng (1999) and Claud et al. (2004) commented that the large scale ECMWF model forecasted a disturbance at the surface but only in the form of a weak trough. Nordeng (1999) went on to say that this behaviour is similar to a simulation using a mesoscale model with latent heating turned off. Moore et al. (1996) also found that the ECMWF data was unable to capture the full development of a Labrador Sea polar low, but nevertheless several key features associated with the development could be identified. From the above discussion, it is clear that the ERA-40 cannot be relied upon to give an accurate representation of the polar low’s surface evolution, particularly of small systems after the onset of diabatic processes. The success of the ECMWF forecast relies on the ability of the model to
correctly predict the upper level precursors of polar lows. The focus in the remainder of this section is therefore on the upper levels.

7.4.5 Case study

The ERA-40 data will be used as the ‘truth’ in this section, where upper level ECMWF forecasts for case 19 will be explored. The relevant satellite images for this case reveal the presence of a surface baroclinic leaf (see figure 7.16 (a)-(b)) and later the development of a comma-like disturbance (figure 7.16 (b)-(c)) but the temporal and areal coverage is not good for the mature stage of this case.

The stratospheric PV, which became incorporated into the precursor, extended south from the main stratospheric reservoir in the form of an intrusion to a small scale synoptic low over the Barents Sea, seen on 2nd March 2002 (not shown). A coherent anomaly formed within the streamer and broke off, circling the synoptic low cyclonically. It moved eastward across the northern coast of Finland and on 4th March at 18 UTC, the local PV maximum turned towards the north and split into two maxima, although remaining connected by a filament of stratospheric PV. At this stage, the synoptic low had filled and another synoptic low was approaching from the west. Figure 7.17 shows the PV field at 290 K on 4th March at 12 UTC (before the aforementioned splitting), on 5th March at 12 UTC and on 6th March at 12 UTC. During this time, the precursor propagated north, to the west of Novaya Zemlya over the Barents Sea.

Three ECMWF forecasts are now presented which are initiated at the times shown in figure 7.17. Contours of 2 PVU on 290 K forecast by these three runs and as given by the ERA-40 are displayed in figure 7.18 for validation dates 6th March, 18 UTC (a), 7th March, 06 UTC (b), and 7th March, 12 UTC (c).

On 6th March 18 UTC (a), all three forecasts show the PV precursor is still attached to a larger region of stratospheric PV and there is good agreement with the ERA-40, however in the longest forecast, for this date (18 hours), there is a displacement of the local PV maximum (near Novaya Zemlya) to the east and a change of orientation of the filament compared to the ERA-40 and the other forecasts. The short and intermediate forecasts indicate the anomaly moves southwards and enters the cyclonic circulation about the new synoptic low to its south.

On the 7th of March, 06 UTC (b), the PV precursor in the longest forecast is seen to be following a different evolution to the precursor in the other forecasts, taking the form of an extended filament, whilst the other forecasts show an anomaly with a more coherent, circular form.

By the 7th March, 12 UTC (c), the keydate when the lower level disturbance was classified as a polar low, the short and intermediate forecasts (24 and 48 hours) have reproduced the cut-off PV anomaly seen in the ERA-40 and match the true location well. The longest forecast for this date (72 hours) still shows an elongated PV filament with no sign of roll up into a coherent anomaly. A comparison with the ERA-40 of relative vorticity at 850 hPa shows the longest forecast produces low level vorticity associated with the upper level anomaly but well to the north of the polar low’s location (not shown). A day later, the longest forecast does produce a cut-off anomaly which propagates to the south, but well to the east of the true polar low’s location.

7.4.6 Discussion

It can be seen at 500 hPa (and also from the sea level pressure in figure 7.17 (c)), that on 6th March there is a diluent flow in the vicinity of Novaya Zemlya, with one stream to the west, around the synoptic low, and the another to the north. It is likely that small changes in the nature of the splitting of the local maxima on 4th March at 18 UTC caused the PV anomaly in the long forecast to be diverted from the true genesis site with the flow to the north rather than to the west. The two forecasts initiated after the splitting were successful in reproducing the upper level precursor observed in the ERA-40.

In this case study, the ECMWF forecasts handled the evolution of the upper level dynamics extremely well where a coherent anomaly was already present in the analysis. On the other hand,
where a splitting of an upper level PV structure occurred within the forecast period, the forecast skill was impaired, and no surface disturbance was induced in the genesis region. The ensemble forecasting technique might have an application in polar low forecasting in such situations. The ensemble output could be divided into clusters according to the evolution of the potential upper-level precursor (when present) and thereby provide an indication of the likelihood of a potential precursor propagating into an area ripe for cyclogenesis. It would be interesting to see if a cluster from an ensemble forecast, initiated at the same time as the long forecast, would have produced an upper-level evolution closer to that of the ERA-40.

### 7.5 Summary and outlook

In this chapter, the PV perspective has been applied to polar low development in the Norwegian and Barents Seas.

From a review of the literature and the cases investigated here, it is evident that the majority of polar low developments are triggered by an upper level disturbance, which often takes the form of a coherent PV anomaly. In the course of its lifetime, a PV anomaly can have more than one role in polar low cyclogenesis and can operate in conjunction with more than one surface anomaly. It may also be part (active or passive) of more than one synoptic scale system e.g. undergo recycling as Zick (1983) observed.

In the case study of the polar low Mike, an upper-level anomaly was seen to locally raise the height of the convective boundary layer, allowing deep convection to occur over a limited area of the Norwegian Sea. The combination of low stability over the sea and a mid-level negative PV anomaly permitted a deeper penetration of the upper anomaly’s wind field and enhanced surface

![Fig. 7.16: Infra-red satellite images for case 19 of table 7.1. Arrows indicate the location of the polar low disturbance. A baroclinic leaf (a-b) develops into a comma cloud (b-c). The polar low decays on landfall (d). Images obtained from the University of Dundee.](image-url)
7.5. SUMMARY AND OUTLOOK

Fig. 7.17: The ERA-40 PV field on the 290 K isentropic surface at the initialisation time for three forecasts: (a) 4th March 2002 12 UTC, (b) 5th March 2002 12 UTC, (c) 6th March 2002 12 UTC. Sea level pressure isolines are drawn at 5 hPa intervals. A black arrow points towards the PV anomaly in each panel.

Fig. 7.18: Forecasts for the 2 PVU contour on the 290 K isentropic surface, valid for (a) 6th March 2002 18 UTC, (b) 7th March 2002 06 UTC, and (c) 7th March 2002 12 UTC. Each panel shows a forecast initiated at 12 UTC on 4th March 2002 (red), 5th March 2002 (purple) and 6th March 2002 (orange) in addition to the ERA-40 re-analysis (black). Grey shading indicates regions where PV>2 PVU in the ERA-40 on 290 K. The black arrows point towards the PV anomaly in the re-analysis.

spin-up. The anomaly involved split from a larger fragment of a stratospheric PV streamer 2 days prior to polar low formation. The synoptic development, although not typical for the polar low development in the region, provided insight into the preconditioning phase. The large scale upper level flow allowed the transport of high PV air from Siberia into the development region, over 5.5 days before cyclogenesis.

A PV composite study has brought suggestions of several synoptic PV signals which could be used to recognise favourable conditions for the development of precursors. In addition to the well known occurrence of a stratospheric intrusion to the west of a surface synoptic low, the observed synoptic signals were (1) a NE-SW oriented streamer advected in the flow of a ridge of high pressure over Greenland, (2) the production of low level PV by vortex stretching as air parcels move across Greenland, (3) a large amplitude PV wave associated with a secondary baroclinic zone and (4) coherent cut-off PV structures which originate from larger synoptic PV structures associated with an earlier system.

If a polar low is triggered by an upper level disturbance, good forecasts of the development up to several days ahead is possible. Greater confidence can be placed in the upper level dynamics due to lack of diabatic effects, however, there are greater uncertainties concerning the initial formation of the upper precursor when this involves non-linear dynamics e.g. breaking of a mesoscale structure from a larger stratospheric reservoir or streamer. If the precursor is already formed before moving over a favourable region for development, confidence can be high, especially if the structure is
isolated. If, within the range of the forecast, an upper anomaly is forecast to split, the certainty of its future path is reduced.

Several have stressed the importance of careful monitoring of upper level vorticity maxima moving toward a favourable region for polar low development and detection of cloud patterns moving in phase (e.g. Zick, 1983, Røtting and Middtsø, 1996). Good upper-atmospheric data at a high enough vertical resolution is therefore crucial for the correct detection and forecasting of upper level vorticity maxima (Nordeng, 1990, Montgomery and Farrell, 1992) and the resulting polar lows.

Possible topics for future work are

- climatological study of upper level precursors and their tracks
- determining the cyclogenesis frequencies for given synoptic 'signals'
- investigating forecasts of the upper level precursor using an ensemble forecast model
- extending the study of the Rossby penetration depth tendency to further cases to see if a penetration depth criterion can be used in forecasting.
Chapter 8

Summary

In this thesis several aspects of flow anomalies in the lower-most stratosphere have been studied using the isentropic potential vorticity perspective.

An instantaneous view of isentropic PV from the ERA-40 data set revealed many localised PV maxima, some in the form of stratospheric cut-offs and streamers but many interior to the stratospheric body. These anomalies have a horizontal length scale of about 500 km and often appear as local maxima in the vertical beneath the overlying stratospheric PV reservoir.

Tracking on the 320, 330 and 340 K isentropic surfaces for 10 consecutive winter seasons permitted the general characteristics of these localised flow anomalies to be assessed. It was found that just over 25% of the isentropic anomaly population survived over 21 hours and had an exponentially decaying life time distribution. For vortices tracked at least 1 days and attaining a minimum amplitude of 1 PVU (with reference to the ambient PV), it was found that 1) the anomaly amplitude peaked above 2 PVU during the first half of the lifetime and decreased before lysis, 2) the mean background speed has a mid-life minimum, 3) a small positive trend in isotropy (compaction increases) occurs with age. These findings indicate that the isentropic anomalies observed in the ERA-40 data are dynamical distinct. It is therefore important that the amplitude and structure are accurately portrayed in the analysis.

The climatological distribution revealed that the highest event densities are located in an annular band centred around 70°N. This pattern was also found in a climatology of the most positive PV on lower middleworld isentropes for the large part the year. The band is most prominent in late spring where it is also detectable in the monthly mean PV. Linear instability analysis showed that an idealised band of the climatological dimensions is unstable to wavenumber 1 perturbations. Contour dynamics simulations illustrated that the instability of a narrower ring of elevated vorticity could go part way to explaining the small scale structure observed instantaneously.

In a last section it is seen that upper-level PV anomalies can be important in polar low cyclogenesis and provide insight into the pre-genesis phase. For one case studied, a PV anomaly contributed to development by locally raising the convective boundary layer. A composite PV analysis was used to highlight different synoptic situations on days when mature polar lows were present in the Norwegian and Barents Seas. For another case studied, ECMWF forecasts were able to capture the correct evolution of the upper-level precursor 2 days before the surface disturbance was classified as a polar low. A longer forecast was sensitive to the precursor's splitting from a larger structure.
Appendix A

Linear Instability Analysis of a Polar $\beta$-Plane Annulus

A.1 Basic state

The zonal flow in the three zonal bands takes one of the forms

$$
\frac{\bar{U}}{r} = \frac{1}{2} \cdot |(\Delta - \epsilon) - f| + \frac{1}{4} \cdot f \cdot \gamma \cdot r^2
$$

$$
\frac{\bar{U}}{r} = \frac{1}{2} \cdot |\Delta - f| + \frac{1}{4} \cdot f \cdot \gamma \cdot r^2 - \frac{1}{2} \cdot \epsilon \cdot \frac{a^2}{r^2}.
$$

A.2 Normal mode analysis of ring instability

The starting point is the nonlinear barotropic vorticity equation $\frac{\partial}{\partial t}(\xi + f) = 0$, where $\xi$ denotes the relative and $f$ the planetary vorticity. Small amplitude (linear) perturbations are governed by a linearised form of this equation which becomes:

$$
\left( \frac{\partial}{\partial t} + \bar{U} \cdot \frac{\partial}{\partial \theta} \right) \xi' + \left( \frac{\partial}{\partial r}(\xi + f) \right) \cdot \psi' = 0,
$$

where $\bar{U}/r$ and $\bar{\xi}$ refer to the zonal basic state and $(\xi', \psi')$ are the perturbation vorticity and radial flow components. These perturbations can be expressed in terms of a perturbation streamfunction $\psi'$, i.e.

$$
\psi' = \frac{\partial^2 \psi'}{\partial r^2} + \frac{1}{r} \cdot \frac{\partial \psi'}{\partial r} + \frac{1}{r^2} \cdot \frac{\partial^2 \psi'}{\partial \theta^2}
$$

$$
v' = -\frac{1}{r} \cdot \frac{\partial \psi'}{\partial \theta}.
$$

Assuming solutions of the form $\psi' = F(r) \cdot e^{i(m \theta + \sigma t)}$, the radial equation reduces to

$$
\frac{\partial^2 f'}{\partial r^2} + \frac{1}{r} \cdot \frac{\partial F}{\partial r} - \frac{m^2}{r^2} \cdot F = 0.
$$

The solutions to this differential equation (in the regions of uniform absolute vorticity) are readily seen to be $F(r) = r^\pm m$. It follows that the solution can be written as $\psi' = \psi'_1 + \psi'_2$, with
This ensures continuity of $\psi'$ at radii $r = a, b$. The additional matching conditions at the interfaces $r = a, b$ are derived as follows: Multiply the linearised barotropic vorticity equation with $r$ and integrate with respect to $r$ from $r = a - \delta$ to $r = a + \delta$ over the interface. In the limit $\delta \to 0$, the following condition results:

$$
\left( \sigma + \frac{U}{a} \cdot m \right) \left[ r \cdot \frac{\partial \psi'}{\partial r} \right]_{a^-}^{a^+} - m \cdot \psi'(a) = 0,
$$

where $[f(r)]_{a^-}^{a^+}$ is the difference $f(a + \delta) - f(a - \delta)$ in the limit $\delta \to 0$. Note that $[\xi + f]_{a^-}^{a^+} = \Delta$, i.e. the absolute vorticity step at the interface $r = a$, was used in deriving the interface condition. An analogue integral constraint can be derived for the outer interface at $r = b$.

Application of this integral condition at $r = a, b$ yields a consistency equation that can be recast as an equation for the frequency $\sigma$. 

$$
\psi'_1 = A \cdot \left( \frac{r}{a} \right)^m \cdot e^{i(m\vartheta + \sigma t)} \quad (r < a)
$$
$$
= A \cdot \left( \frac{r}{a} \right)^{-m} \cdot e^{i(m\vartheta + \sigma t)} \quad (r > a)
$$

$$
\psi'_2 = B \cdot \left( \frac{r}{b} \right)^m \cdot e^{i(m\vartheta + \sigma t)} \quad (r < b)
$$
$$
= B \cdot \left( \frac{r}{b} \right)^{-m} \cdot e^{i(m\vartheta + \sigma t)} \quad (r > b).
$$
Appendix B

Contour Dynamics

B.1 From the vorticity equation to the contour integral

The following derivation is for 2D dynamics confined to the $f$-plane.

The velocity field $u = (u, v)$ of a two-dimensional, incompressible fluid can be derived from a scalar stream function:

$$
u = -\frac{\partial \psi}{\partial x}, \quad v = \frac{\partial \psi}{\partial y}. \quad (B.1)$$

Note that the existence of a stream function entails the non-divergence and hence the incompressibility of the flow. With this stream function, the flow's vorticity is readily calculated as the two-dimensional Laplacian:

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \nabla^2 \psi. \quad (B.2)$$

Knowing the scalar $\zeta$ therefore leads by an inversion of the above Laplace equation to the stream function, and finally to the velocity vector.

If viscosity can be neglected, the fluid's evolution is completely described by the barotropic vorticity equation:

$$\frac{D\zeta}{Dt} = 0, \quad (B.3)$$

with $\frac{D}{Dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y}$.

Physically this equation expresses the fact that the vorticity $\zeta$ is conserved following the fluid's motion. It can be verified that the stream function of an arbitrary vorticity distribution can be obtained by evaluating the following area integral:

$$\psi = \frac{1}{2\pi} \int \zeta(x', y') \log(r) dx' dy' \quad (B.4)$$

with $r = \sqrt{(x - x')^2 + (y - y')^2}$.

The integration domain comprises the whole two-dimensional area $\mathbb{R}^2$ but if a $\zeta$ distribution of the form

$$\zeta(x, y) = \begin{cases} \xi_1 & \text{for } (x, y) \in B \\ 0 & \text{else} \end{cases} \quad (B.5)$$

is specified, the first of the two-dimensional integrals can be transformed into one-dimensional
contour integrals for the two wind components \( u \) and \( v \):

\[
\begin{align*}
\frac{\xi_1}{2\pi} \int_{\partial B} \log(r) dr' &= u = -\frac{\xi_1}{2\pi} \int_{\partial B} \log(r) dr', \\
\frac{\xi_1}{2\pi} \int_{\partial B} \log(r) dy' &= v = -\frac{\xi_1}{2\pi} \int_{\partial B} \log(r) dy'.
\end{align*}
\]

(B.6)

It is sufficient to know the contour line associated with a region of uniform vorticity to determine the flow associated with that specific \( \xi \) region. If the flow is known, it is possible to advect the contour line forward in time according to this flow. This is possible because the contour lines are material surfaces under the assumption of neglected viscosity. In fact, if \( x_0(t) = (x_0(t), y_0(t)) \) is a point on the contour line, the following 1st order differential equation results:

\[
\frac{d}{dt} x_0(t) = u(x_0(t)) = -\frac{\xi_1}{2\pi} \int_{\partial B} \log(|x_0(t) - x'|) dx' + u_f(x_0(t) - x').
\]

(B.7)

The aforementioned formalism can be readily extended for several homogeneous vorticity regions. The integral which describes the movement of a contour line comprises then the sum over all individual \( \xi \) contributions:

\[
u(x_0(t)) = -\frac{\xi_1}{2\pi} \sum_{k=1}^{N} Q_k \int_{\partial B_k} \log(|x_0(t) - x'_k|) dx'_k + u_f(x_0(t) - x'),
\]

(B.8)

where \( B_k \) is the boundary of a region of uniform vorticity, \( Q_k \) is the jump in vorticity across \( B_k \) and \( x_k \) is a point on \( B_k \).

**B.2 Including the planetary vorticity \( f \)**

With a small number of modifications to the above equations, the planetary vorticity can be included into the contour dynamics integration.

The barotropic vorticity equation B.3 is reformulated as a conservation law of the absolute vorticity, \( \eta \):

\[
\frac{D\eta}{Dt} = 0,
\]

(B.9)

where \( \eta = \xi + f \), and \( f \) is the planetary vorticity (Coriolis parameter).

We substitute \( \xi = \eta - f \) into equation B.4 to obtain

\[
\psi = \frac{1}{2\pi} \int \eta(x', y') \log(r) dx' dy' - \frac{1}{2\pi} \int f(x', y') \log(r) dx' dy'.
\]

(B.10)

The first integral can be solved in the same way as above, if the vorticity distribution is composed of regions of constant absolute vorticity instead of constant relative vorticity, i.e. the vorticity distribution must be of the form

\[
\eta(x, y) = \begin{cases} 
\eta_1 & \text{for } (x, y) \in B \\
0 & \text{else},
\end{cases}
\]

(B.11)

and then we can obtain

\[
\begin{align*}
\frac{\eta_1}{2\pi} \int_{\partial B} \log(r) dr' &= u_{\eta} = -\frac{\eta_1}{2\pi} \int_{\partial B} \log(r) dr', \\
\frac{\eta_1}{2\pi} \int_{\partial B} \log(r) dy' &= v_{\eta} = -\frac{\eta_1}{2\pi} \int_{\partial B} \log(r) dy'.
\end{align*}
\]

(B.12)
for the velocity components due to the absolute vorticity.

The second term of B.10 is the contribution to the stream function of the Coriolis parameter. This contribution is constant in time and therefore it needs to be evaluated and inverted just once to find the wind components $u_f$ and $v_f$. We take a polar $\beta$-plane distribution for $f(x, y)$. The fields turn out to be

$$u_f = -\frac{y}{r_f 0.5 r_c (1 - 0.25 r_c^2)}$$

$$v_f = -\frac{\eta}{2\pi} \int_{\partial B} \log(r) dy.'$$

The velocities derived from the two terms in equation B.10 add together to give the total wind field ($u = u_n + u_f$), which is used to advect the points along the contour to their new positions at the next time step.

### B.3 Numerical scheme and initialisation

The contours $R_k$ are represented by a series of computational nodes and the integral over each contour is replaced by the sum of the integrals over each segment connecting consecutive nodes. To preserve the resolution of the computation, the nodes are continually adjusted, with nodes added in regions of high curvature. Equation B.7 can be readily numerically integrated by standard techniques. In our implementation we use a 4th order Runge-Kutta scheme. This scheme is particularly efficient for the computing memory as it only requires values from $t_n$ to calculate the new values at $t_{n+1}$.

Random noise is added evenly to the entire domain to introduce very small scale perturbations to the initial set up at the start of the integration.

### B.4 Validation and sensitivity

The $\beta$-plane version of the contour dynamics tool was used to test and confirm that the results from the contour dynamics are in line with the analytical results.

Two integrations have been repeated with exactly the same initial configuration except for a different seed for the random background noise. Results are qualitatively similar, but for a thin ring (width/radius = 0.1), differing numbers of vortices (10 and 11) were obtained and there was a slight delay in the onset on instability. This shows the sensitivity of the non-linear evolution to the interactions of finite amplitude disturbances. Without added noise, the rings remain stable, attesting to the numerical stability of the algorithm.

### B.5 Contour surgery

As the contour integration proceeds, in particular if the vorticity distribution is in an unstable configuration, the distance between neighbouring points along the contour will decrease in some areas and increase in others. Regions of convergence would be very well represented whilst regions of divergence would be poorly resolved. To circumvent this problem, after a number of iterations of the algorithm, the number of points is increased (in proportion to the increase in contour length) and all points are redistributed evenly along the contour. The contour is re-parametrised by means of a cubic-spline interpolation along the distorted contour. This redistribution greatly improves the quality of the numerical integration.

Another problem arises when two non-neighbouring sections of the contour become extremely close, corresponding to the formation of a thin vorticity filament. The contour in such regions
would typically become more and more elongated, making increased demands on computational
time because of the extra points. Such thin filaments contribute little to the local velocity field
and in the real world would tend to disappear due to diffusion. If the distance between two non-
neighbouring sections becomes closer than a minimum threshold, the two segments are cut out
from the contour and the open ends are joined. The two new contours are numerically treated as
two distinct vorticity structures decoupled from each other. Entire structures may also be removed
if they are too thin.

It is also possible that the separation between two separate entities becomes so small that, in
reality, they would be said to have merged. The contour surgery is not yet capable of handling
merger. For the vorticity configurations studied, however, merger would play a role only at much
later times of the integration.

B.6 Visualisation

To visualise the evolution of the contour dynamics, the vorticity, velocity and contours are interpo-
lated onto a Cartesian grid at equal time intervals (after a set number of time steps) to produce a
sequence of images. The mean radial profile of vorticity is also computed to visualise the evolution
of the symmetric part of the flow.
References


REFERENCES


REFERENCES


Curriculum Vitae

Sarah Frances Kew

Hürststrasse 65
CH-8046 Zürich
sarah.kew@env.ethz.ch

Born on 12 October 1979
in Chatham, UK
Citizen of the United Kingdom

Education

November 2003 - May 2007:
PhD thesis at the Institute for Atmospheric and Climate Science at the Swiss Federal Institute for Technology (ETHZ) in Zürich, under the supervision of Prof. Dr. H. C. Davies and Dr. M. A. Sprenger.
Title: Structure and dynamics of distinctive flow anomalies in the lowermost stratosphere.

October 2002 - September 2003:
MSc studies in mathematics and meteorology, University of Reading, UK.
Masters dissertation, Department of Meteorology, University of Reading, UK.
Title: Development of a 3D fractal cirrus model and its use in investigating the impact of cirrus inhomogeneity on radiation.
MSc. Mathematical and Numerical Modelling of the Atmosphere and Oceans. University of Reading, UK.

October 1998 - April 2002:
MSci Physics studies, Imperial College London.
MSci project, Plasma Physics Research Centre, EPFL, Lausanne, Switzerland.
Title: Investigating current drive efficiency in a tokamak.
MSci. Physics with a Year in Europe. Imperial College.

September 1996 - July 1998:
A-Levels: Physics, Chemistry, Biology, Mathematics and
AS-Level: Further Mathematics,
Fort Pitt Grammar School, Chatham, UK.

September 1991 - July 1996:
Secondary education and GCSEs,
Fort Pitt Grammar School, Chatham, UK.
International Conferences and Workshops

- Second Thorpex International Symposium, Landshut, Germany, December 2006.
- Cyclone Workshop 2006, Monterey CA, USA, October 2006.
- EGU General Assembly 2006, Vienna, Austria, April 2006.
- First Thorpex International Symposium, Montreal, Canada, December 2004.

Publications


Investigating current drive efficiency in a tokamak, 2001, S. F. Kew, MSc Project.
Acknowledgements

I am grateful to know many wonderful people - those I have met during the course of my PhD and those who have kept in touch from 'back home'. All have been instrumental in the completion of this work, whether directly by their input of ideas and encouragement, or indirectly by helping me take time to relax.

First of all, I’d like to thank you to Huw Davies for providing the opportunity for me to come to Switzerland and work at the IAC and for his wisdom and guidance as my supervisor.

Many thanks to Chris Davis for his interest in my work, his enthusiastic suggestions and for kindly agreeing to be my external examiner, despite the very long journey involved.

I have greatly appreciated Michael Sprenger’s (often very practical) support as my second supervisor. I enjoyed many helpful and interesting discussions as well as quiet lunches with fine food. Thanks Michael!

I would like to thank my international and interlingual friends of L6 and M16.1 (Patrick, Harald, Marco, Rich and Jana) for the quiet and yet friendly office environment and their eagerness to help with any technical or scientific questions.

Thanks to all at the IAC for many memorable moments and events! Thanks to my semester students Ferenc and Urs for their work and enthusiasm and thanks to the Dynamos for all the constructive feedback and technical help.

A special thank you to Olivia for her practical help with \LaTeX, netcdf, proof-reading, translation, printing and also for her supportive understanding, especially in those last moments!

Thank you to my ‘WG’ friends Martin & Jen, Frank, Claudia, Philipp and Alla for your company and understanding and especially for feeding me! Without you things definitely wouldn’t have been the same!

A huge thank you to Daniel for his interest and countless hours given in support and encouragement and for his ongoing friendship. Thank you for all the effort together with Jen in helping to create my Middleworld Climatology website.

Many friends have supported me with encouragement and prayer. In particular I would like to thank everyone from ‘Check-in’: Martin & Mignon, Manu, José, Rahul, Thomas, Susanne and Aurélie... friends from the international ski camp: Asi & Peter, Nicky & Catherine and others (what a great break that was!)... those from ‘Vineyard All Nations’: Naomi, Caroline, Alexey, Rahul, Mathew, Jim, Susanne... and friends and family in England: Katie & Allen, Maggie & Paul, Paul & Jean, Penny, Margaret & Alec, Ken Gardiner, Lydia, Vivienne, Rosie, Amy, my grandma Kew and her friends, Jean, Joan, the Durkans, the Kews, Ken & Helen.

To my parents, thank you for everything - listening, encouraging, praying, your wise advice, interest and love. I am so grateful to have you as my parents.

To my Lord and Saviour, Jesus Christ, thank you so much for being with me and for your amazing faithfulness to your word.

Finally I would like to dedicate this work to the memory of a very special lady, my grandma, Dorothy Frances Trice.