Department of Physics and Technology

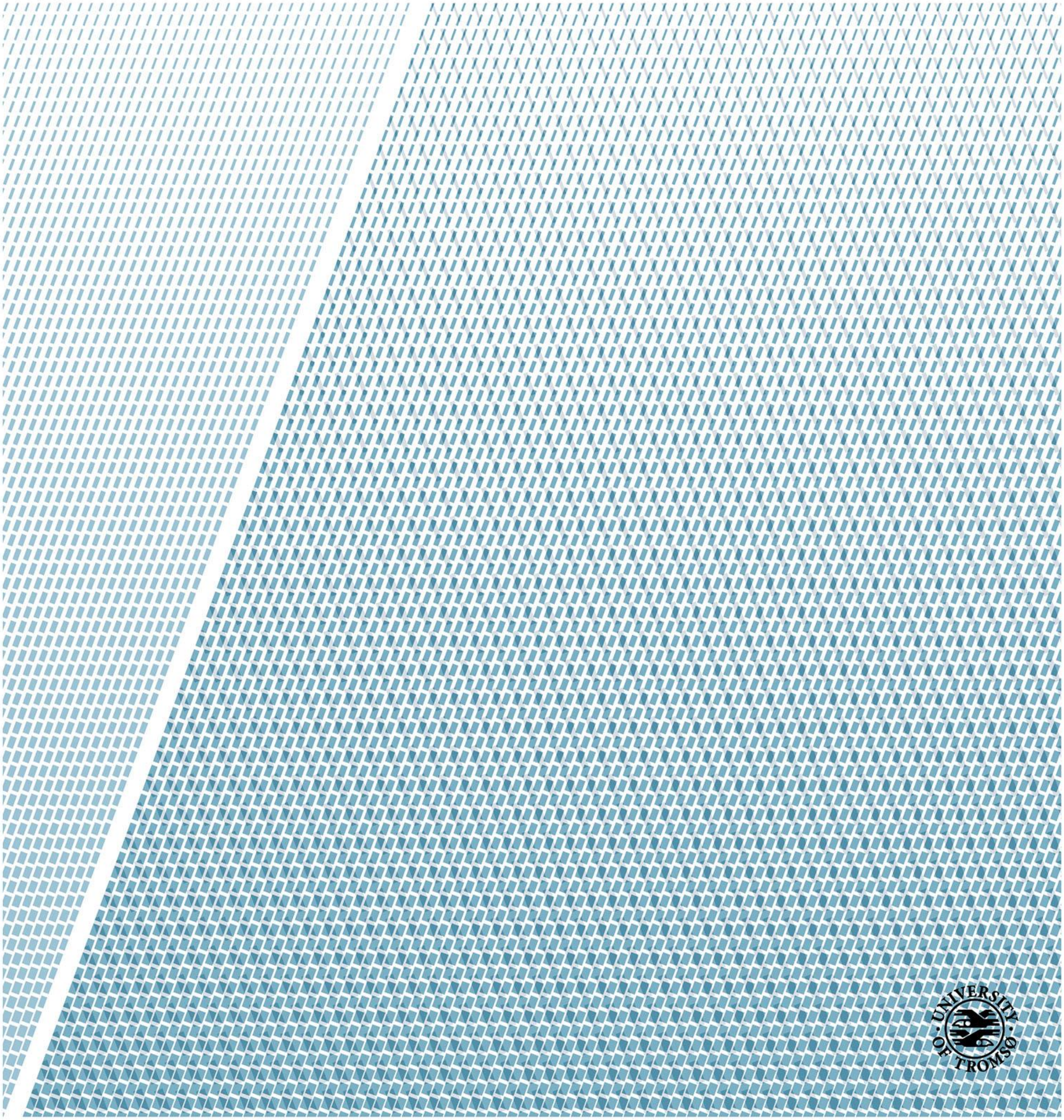
Polar low trajectories in the Nordic Seas 1999-2013: a statistical analysis using kernel density methods

—

Ylva Kirsten Kvammen

EOM-3901 Master's Thesis in Energy, Climate and Environment

June 2014



Acknowledgements

First I would like to thank my main supervisor Alfred Hanssen. I am very grateful for the opportunity I was given by you and Aker Solutions in Tromsø to work with such a relevant and exciting theme for my master's thesis. Thank you for all of the advise, motivation and fruitful discussions.

Also thank you to my co-supervisors, Gunnar Noer and Odd Erik Garcia, for much appreciated advise and help when needed.

A huge thank you should also be extended to Chantal Claud for providing me with the data set that formed the entire basis for my thesis. Allowing me to use this data set is highly appreciated.

During my work I received help with various tasks, and for that I wish to express appreciation to the following persons: Robert Jenssen, Jonas Nordhaug Myhre and Georg Elvebakk. Thank you so much for taking the time to help me.

To my family, thank you for all of your love and encouragement, and for always supporting me in any way you can, from food deliveries to tire changes!

Last, but not least, my wonderful boyfriend Jon, thank you so much for your tremendous love and support, which literally stretched to great lengths - I don't know what I would do without you!

Ylva Kvammen, June 2014

Abstract

Polar lows are mesoscale cyclones of high intensity, frequently occurring in the Nordic Seas during the winter season. They usually form in baroclinic zones, for instance in regions near ice edges, and they are often maintained by organized convection. Polar lows are characterized by severe weather conditions like heavy snowfall, strong winds, icing and large sea waves. In addition to forecasting complexity, polar lows thus represent potentially hazardous conditions for marine installations, ships and coastal communities.

In this thesis, we carried out a statistical analysis of polar low trajectories in the Nordic Seas during $1999 - 2013$ by the use of kernel density estimation as the main method. This method provides smoothed estimates of the probability density function of the data, and the degree of smoothing can be adjusted by the choice of a bandwidth parameter. Our density estimates gave an accurate representation of the space-time distribution of polar lows trajectories, and displayed the main concentration located in coastal/marine areas close to Lofoten and Vesterålen. By segmenting the data set into time intervals, the main maximum was seen to shift spatially with time, displaying the space-time distribution as highly dynamic. A temporal shift in the densities towards the Barents Sea was proposed to be connected to climate change. Polar low formation and development was found to correlate well spatially with the Norwegian Atlantic Current and its branches in the Nordic Seas.

By investigating large scale wind flow, it was found that polar lows formed mainly within a northerly flow in the Norwegian Sea, and a north-easterly flow in the Barents Sea. We connected the combination of increased sea surface temperatures and cold air outbreaks from the Arctic which leads to a decrease in static stability, which in turn promotes favourable conditions for polar low formation and development. Polar low dissipation was found to relate well with systems being steered across land or sea ice, being deprived of their main energy source. A secondary analysis was made based on curve clustering of the polar low trajectories. This revealed how trajectories vary with regards to length, propagation direction and degree of curvature. Our results confirmed previous research on these characteristics connected to the distribution of genesis locations.

Contents

Chapter 1

Introduction and motivation

Polar lows are polar mesoscale cyclones of high intensity, which form poleward of the main polar front, but are especially common in the Nordic Seas (i.e., the North Atlantic between 60 °N and 80 °N latitude and between Greenland and Novaya Zemlya) during the winter season [Heinemann and Claud, 1997]. These storm systems are small, but violent, and have impacted coastal communities and maritime activities over the centuries. They are believed to be the main reason for a number of shipwrecks in the Nordic Seas [Noer et al., 2011]. Weather conditions associated with polar lows include heavy snow showers, strong winds, large sea waves and icing, which in different ways represent potential risks for maritime traffic and industrial players within fishery and petroleum in the Norwegian and Barents Sea [Rojo et al., 2014]. Another complicating factor related to polar lows is their small temporal and spatial scales. They can develop in a few hours, and their diameters are approximately between 200 and 1000 km [Rasmussen and Turner, 2003]. This makes their detection and forecasting particularly difficult by the use of standard weather models [Rojo et al., 2014]. The combination of severe weather and forecasting complexity makes polar lows a hazard to ships, oil rigs and coastal communities in the Arctic regions [Irvine et al., 2011].

During the last three decades, several climatological studies on polar lows have been published, e.g., [Businger, 1985], [Wilhelmsen, 1985], [Ese et al., 1988], [Zahn and von Storch, 2008], [Bracegirdle and Gray, 2008], [Blechschmidt et al., 2009], and [Noer et al., 2011]. Several of these studies have focused on the climatological forcing and meteorological processes related to polar low development, and static aspects of properties like size, type, lifetime etc. [Rojo et al., 2014]. Significantly less research has been carried out on polar lows trajectories and their properties [Rojo et al., 2014], especially from a statistical, long-term point of view. Observations of polar low tracks may provide important knowledge to be applied in risk management [Rojo et al., 2014], and as a collaboration with Aker Solutions, this contributed as motivation for this thesis. The main purpose was therefore to identify how the polar low trajectories are distributed in the Nordic Seas by the use of statistical methods. By doing this, valuable information could be retrieved, for instance regarding areas of high polar low exposure, which can be of the essence in risk assessment and management in maritime activities. Put simply, the main goal of the thesis was to provide a possibility for the petroleum industry or fishermen to have a preliminary and overall idea of areas to be vary of, due to possible hazardous conditions caused by polar lows.

The statistical analysis is carried out using a data set which include the positions for polar low trajectories observed in the Nordic Seas during the years $1999 - 2003$ (produced by Rojo et al. [2014]). The main method of investigation will be kernel density estimation, a dynamic and fairly common approach of non-parametric density estimation. This method will provide information of the spatial distribution of polar low trajectories in general, i.e., indicate areas of low or high frequency with respect to tracks having formed or passed there. Because the input for these processes will be position values of trajectories (latitude, longitude) a general idea of polar low exposed areas will be provided, but not information of the tracks themselves. Therefore, a curve clustering analysis will also be applied to the data, in an effort to review the trajectories and their characteristics, and investigate whether these form the basis for any subdivisions (clusters) of the set.

The findings from the statistical analysis revealed an area outside the Lofoten and Vesterålen coast as the most heavily exposed with regards to polar lows during the period $1999-2003$. However, segmenting the data into various time intervals showed that this maximum was not constant in space and time, but shifted throughout the Nordic Seas, especially along the coast of Norway, and ended up in the Barents Sea at the end of the period. Kernel density estimation was also applied to both the genesis and dissipation points of the complete trajectory collection, respectively. The estimates resulting from the complete set and the genesis points were both found to correlate well with the Norwegian Atlantic current and outbreaks of cold air from the Arctic, which in combination cause steep temperature gradients between sea and air, resulting in high levels of instability, and favourable conditions for polar lows. The dissipation points were in most cases connected to polar lows making landfall along the Norwegian coast, or being steered across the winter sea ice edges of the Nordic Seas. Both situations caused an increase in static stability, leading to quick decay of the polar low systems. The clustering resulted in a visual representation of the full trajectories, where they were divided into clusters based mainly on propagation direction, trajectory length and degree of straightness.

The structure of the thesis is as follows: Chp. 2 will present some general theory on polar lows, e.g., classifications, processes of formation and development, and forecasting. Chp. 3 will briefly present the data set obtained from Rojo et al. [2014], and the methods they used to produce this. In Chp. 4, a detailed description of the kernel density estimation method will be given. A brief description of curve clustering analysis by the use of mixture regression models will follow. The results of both methods will be presented in Chp. 5, followed by a discussion in Chp. 6. Finally, the main findings and conclusions will be listed and summarized in Chp. 7.

Chapter 2

Theoretical background on polar lows

2.1 Introduction

Polar lows are small low-pressure systems (cyclones) of high intensity, belonging to the mesocale of atmospheric circulation. Other terms used to describe these vortices include polar air depression, Arctic instability low, mesoscale cyclone or mesocyclone and Arctic hurricane. The term polar low is often used of mesocyclones of especially high intensities. As the name implies, these systems occur at high latitudes - in the Arctic and Antarctic, poleward of the polar front [Rasmussen and Turner, 2003]. A general definition of polar lows which will be used here, provided by [Rasmussen and Turner, 2003] is the following:

> *A polar low is a small, but fairly intense maritime cyclone that forms poleward of the baroclinic zone (the polar front or other major baroclinic zone). The horizontal scale of the polar low is approximately between 200 and 1000 kilometres and surface winds near or above gale force.*

Polar lows are short-lived systems, with the time from genesis to dissipation ranging from 6 hours to 2-3 days.[1] Referring to the definition, strong surface winds are often associated with polar lows, with forces near or above gale force [Linders, 2009]. According to the Beaufort scale for wind force, winds of gale force range from about 13.9 to 28.4 m/s, in different degrees of gale.[2] The most powerful low measured since 1999 had winds reaching hurricane strength (32.5 m/s) over a 12 hour period.[1] Polar lows have an anticlockwise rotation and often form under a northerly air flow. Their propagation direction will therefore be directed southwards, and the wind on the west side of the center will be stronger than that on the east side. According to an old rule at the Meteorological Institute in Northern Norway, the wind force can vary with two to four Beaufort, corresponding to 5-10 m/s, compared to the background large scale wind field.[1] In general, the wind can be said to be strongest on the right side of the polar low center, when looking in the polar lows propagation direction.[3] In a low approaching the coast, this can lead to a treacherous

---

[1] MetLex - Meteorologisk leksikon: Polare lavtrykk `http://metlex.met.no/wiki/Polare_lavtrykk`

[2] MetLex - Meteorologisk leksikon: Beaufortskalaen `http://metlex.met.no/wiki/Beaufortskalaen`

impression, when relatively weak off-shore winds on the eastern side of the system produces rather clear weather and blue skies. However, when the low hits land, the coast will be subjected to a rapid change in wind direction and strength on the opposite side, with heavy and chaotic precipitation in the form of snow. The rapid change of the wind also affects the sea. Turbulence grows and increased presence of sea spray can in combination with low enough temperatures lead to *icing*. Sea spray icing happens when sea spray (surface water droplets whipped into the air by strong surface winds[4]) freezes to a surface, e.g., a boat, offshore installations or simply land along a coast.[5] Temperatures below 0°C will result in supercooled water droplets which instantly adhere themselves by freezing if presented with a surface.[4] Icing is not particularly common near the coast of Northern Norway, but probably more so further north in the Arctic Ocean, where temperatures are lower.[3]

Heavy precipitation in the form of snow is another weather condition following polar lows, which is also considered as the main complicating factor associated with off-shore installations like oil rigs. The sudden build-up of substantial amounts of snow can cause severe implications leading to costly delays in activity. Snowfall is often also the main issue of polar lows making landfall. The heavy snow showers with drifting snow tend to cause poor visibility ranging less than a 100 meters, which in turn causes problems in traffic and the periodic close-down of airports.[3]

Ocean wave generation is also a consequence associated with polar lows, but it is probably the least researched one. Few observations of waves around the polar low system exist, but those that do, have reported of wave heights increasing from 5 to 9 meters in less than an hour. The waves are mainly connected to the strong surface winds, which also means that the largest waves can be found in the western sector of the low.[3] Dysthe and Harbitz [1987] is one of the few examples of studies on polar lows and ocean waves. Dysthe and Harbitz [1987] pointed to a limited capacity of polar lows for wave generation due to their small scale in both space (a short fetch) and time, given the assumption that polar lows are more or less stationary. However, because polar lows systems are generally advected in a northerly large-scale flow, they stated that a wave packet can develop. This packet will be dominated by waves in a frequency band centered around waves which have a group velocity the same as the travelling speed of the low, because these waves are going to be located in the enhanced wind field (usually the west side of the system) of the low for a long time. The waves will also be confined to the sector of the low where wind direction is aligned with the propagation direction of the polar low system. Dysthe and Harbitz [1987] stated that this could lead to the polar low being accompanied by a wave packet with a sharp front. As this front passes a measuring station, a very rapid increase in wave height would be observed. They referred to this synchronising between the low and certain waves as ‘resonant fetch enhancement’ [Dysthe and Harbitz, 1987].

As the name implies, polar lows form and develop in the high-latitude polar regions. Be-

---

[3] Article by Barentswatch, `http://www.barentswatch.no/Tema/Sjotransport/Polarvar-og-istjenester/Polare-lavtrykk1/Prognose-for-polart-lavtrykk-/`

[4] MetLex - Meteorologisk leksikon: Sjøsprøyt `http://metlex.met.no/wiki/Sj%C3%B8spr%C3%B8yt`

[5] MetLex - Meteorologisk leksikon: Ising `http://metlex.met.no/wiki/Ising`

cause they arise due to an instability (discussed in more detail in Sec. 2.3), an area of steep temperature gradients between sea water and air is in many cases a prerequisite. Polar lows are thus especially common in the Norwegian Sea and the Barents Sea, where the warm water of the Norwegian Atlantic current and cold air from the Arctic meet. Along the Norwegian cost, polar lows are most common in the northern part of Norway and down to mid-Norway (Trøndelag). There have been cases of polar low formation further south in the Norwegian Sea, but this is rather rare.[1]

There are several similarities between polar lows and tropical hurricanes, which is the reason why polar lows are sometimes called Arctic or polar hurricanes. Both systems form due to temperature gradients between surface water and above-lying air, where this relatively warm water is also the main energy source for their development. Like tropical hurricanes, polar lows rapidly weaken after landfall, when they are deprived of this energy source, even though the landfall itself often can be associated with somewhat dramatic events. The appearance, when looking at satellite images, also makes it easy to relate the two systems to each other. They are both characterized by a more or less symmetrical spiral around a clear eye. Associated weather include strong winds and heavy precipitation (as snow for polar lows) from cumulonimbus clouds. While the tropical hurricanes require a threshold value of water temperatures over 27°C, polar lows need a temperature difference of 43°C between the sea surface and the air in 6 km elevation to form. However, due to the colder and drier regions of formation and development of polar lows, they are usually smaller and weaker than their tropical relative.[1]

---

[1] MetLex - Meteorologisk leksikon: Polare lavtrykk `http://metlex.met.no/wiki/Polare_lavtrykk`

2.2 Tracks and spatial distribution

The climatological study of polar lows in the Nordic Seas carried out by Noer et al. [2011] showed a fairly even distribution in the Norwegian and Barents Sea (shown in Fig. 2.1). This study applied for the period of 2000 to 2009. The majority of the polar lows investigated were found in the area north of 62 °N, south of 75 °N and east of the zero meridian. As Fig. 2.1 shows, the areas of highest concentration are associated with the northward flowing Norwegian Atlantic Current, i.e., areas of increased mean sea-surface temperatures. Within these areas some density maxima can be found, the most significant being in the area known as Tromsøflaket, around 72 °N and 18 °E. The combination of cold-air outbreaks in the Svalbard area and the high sea surface temperatures creates highly favourable conditions for polar low genesis and development in this area[Noer et al., 2011].

The polar lows in the Barents Sea is connected to cold-air outbreaks in the area between Svalbard and Novaya Zemlya. In the easternmost part of the Barents Sea however, the number of occurrences decreases. This may be connected to lower sea surface temperatures in this area, which typically has a value of around $2 - 3\,^{\circ}\mathrm{C}$ in wintertime. Another reason may be the small variations in sea surface temperature in the north-south direction, which causes northerly winds in the area to be unable to transfer as much sensible and latent heat [Noer et al., 2011]. Note that Fig. 2.1 displays only the early locations of observed polar lows, and therefore is not related to the spatial distribution of the complete trajectories.

Other observations of spatial distribution can be found in [Lystad, 1986], [Noer and Ovhed, 2003], [Kolstad, 2006] and [Blechschmidt, 2008]. Lystad [1986] found that the coastal areas of northern and western Norway were affected, and that approximately half of the cases studied crossed the area between Bear Island and the coast of northern Norway. A total of 159 cases through the years $1971 - 1985$ were studied. Noer and Ovhed [2003] (41 cases in the four seasons $1999 - 2003$) stated the high-occurrence area to be south of 75 °N and east of Jan Mayen, in the border areas of the Norwegian Atlantic Current, where sea surface temperature gradients are high. Kolstad [2006] referred to both [Lystad, 1986] and [Noer and Ovhed, 2003], and stated that both studies agreed on a maximum of polar low activity in the region between Svalbard and the northern coast of Norway. Blechschmidt [2008] studied a total of 90 polar lows between 2004 and 2005, but focused mostly on cyclone formation distribution. They found that the majority of polar lows (67%) originated in the Norwegian Sea between Iceland and Finnmark, which coincides with the findings in [Lystad, 1986] and [Noer and Ovhed, 2003]. A second maximum (22%) was found in the Barents Sea, where large sea surface temperature gradients favour development. Fewer polar lows (11%) were found to form between southern Greenland and Iceland. It was stated by Blechschmidt [2008] that polar lows in this region form in a synoptic scale westerly or northwesterly flow as air is forced to flow around or over the southernmost part of the Greenland ice cap [Rasmussen and Turner, 2003].

As presented above, most of the earlier findings related to polar low spatial distribution coincide with each other because they represent quite coarse observation areas and quantitative estimations of high-density regions. Focus has also often been on polar low genesis areas as opposed to their full trajectories. The methods applied in this thesis will provide results where more accurate conclusions with regard to spatial distribution of polar low trajectories can be drawn.

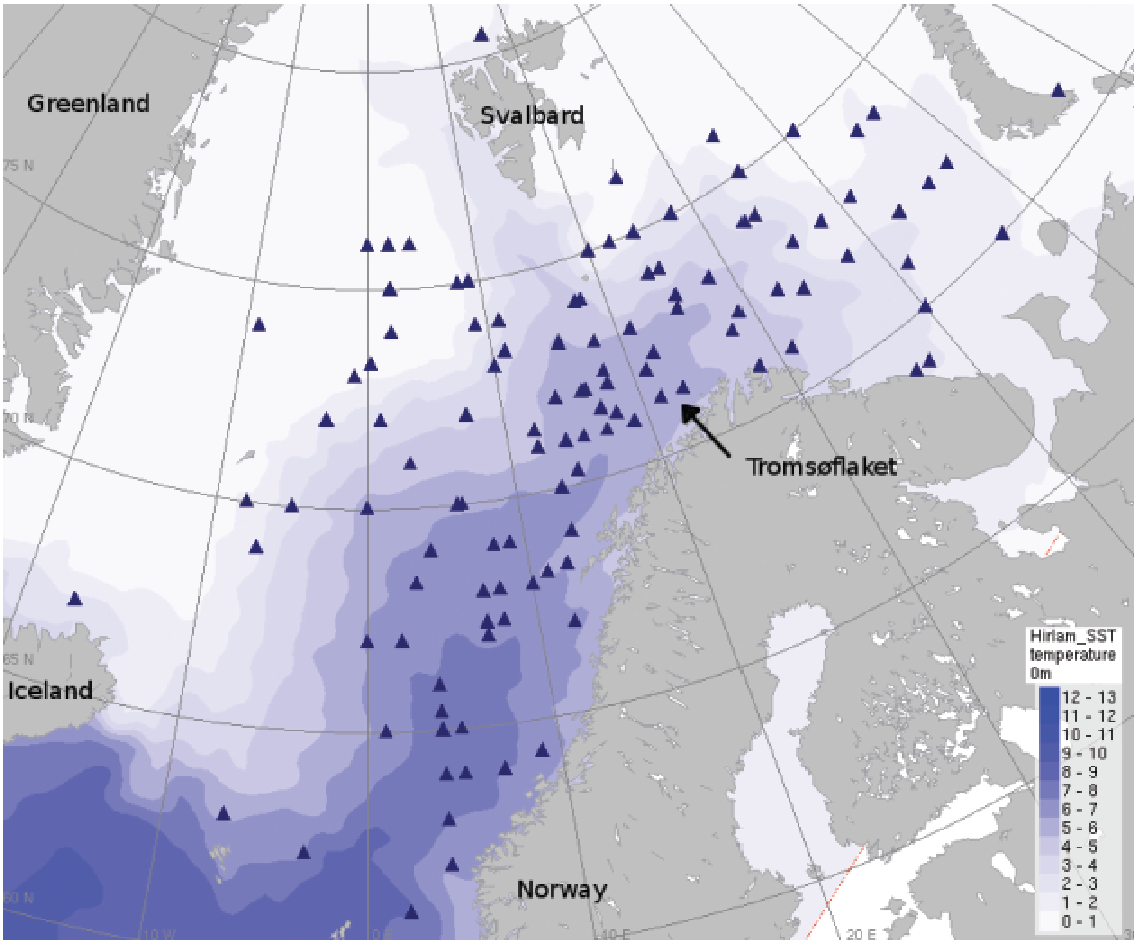


Figure 2.1: Position of polar lows 2000–2009. Locations represent the point at which the low was first identified as fully developed, i.e., early in its life span. Blue shading shows the climatological mean sea-surface temperatures for January taken from the climatological archive compiled by Engedahl et al. [1998]. Source: Noer et al. [2011]

2.3 Formation and development

There are still some insecurity related to the main forcing mechanism behind the formation of polar lows. However, polar lows are known to form following some kind of instability. The instabilities in question are then either *baroclinic* or *convective*. Formation due to baroclinic instability would produce systems similar to the midlatitude cyclones (frontal lows or wave cyclones), and convective instability would create systems with similarities to tropical cyclones. It is however often assumed that polar lows form by a combination of these two [Linders, 2009]. A more detailed discussion of the two different instabilities and their relation to polar lows follows below.

2.3.1 Baroclinic instability

Baroclinic instability is the type of instability connected to the baroclinic region of the atmosphere. This is a region where the density depends on both the pressure and the temperature. In other words, it is a region where temperature changes along the pressure surfaces. The growth of baroclinic instabilities happens by the conversion of potential energy connected to the mean horizontal temperature gradient to kinetic energy. This is governed by the ascend of warm air and descend of cold air [Rasmussen and Turner, 2003].

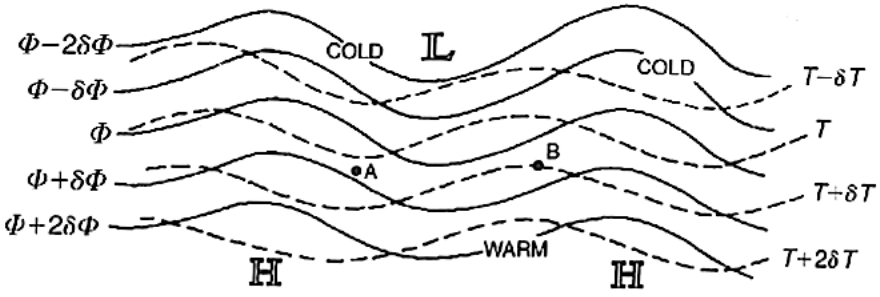


Figure 2.2: Development of a baroclinic wave in the Northern Hemisphere. Distribution of geopotential height and temperature on a constant pressure surface is shown by solid lines and broken lines, respectively. Source: Rasmussen and Turner [2003]

Fig. 2.2 shows how a baroclinic wave develops in an area where the temperature gradient is strong in the north-south direction. It is assumed that some process causes a weak and wave-like disturbance in the initially uniform zonal flow (flow along latitudes). Resulting by the meridional motions of this disturbance, the original straight east-west isotherms (broken lines in Fig. 2.2) are distorted and a wave in the temperature field forms. The isotherm distortions will develop further under influence of the horizontal temperature advection related to the the geostrophic wind field, and the temperature wave will thus amplify. The growth of the wave requires increase of the kinetic energy. This happens through the thermal circulation where cold air sinks at point A and warm air rises at point B (see Fig. 2.2). This mechanism lowers the centre of gravity of the fluid, and thus contribute to the conversion of potential energy into kinetic energy [Rasmussen and Turner, 2003]. This baroclinic development can be associated with the formation and life cycle of a midlatitude wave cyclone. In these systems, upper-level and low-level systems develop simultaneously, through a coupling of the low-level cyclone and the upper-level trough. This continuous process is part of the baroclinic instability growth [Rasmussen and Turner, 2003].

Baroclinic instability in the polar regions is considered in the effort of relating this to the formation and development of polar lows. Baroclinic instability in polar regions is usually distinguished between low-level instability in shallow layers, often along ice edges, and "deep instability" in the deeper baroclinic layers. During winter in polar regions baroclinic zones can form along the edges between the regions covered with ice and snow, and the regions which are not. These baroclinic zones have through the years been considered essential in the formation of polar lows [Rasmussen and Turner, 2003]. The precursor of polar low formation in these zones are so called Marine Cold Air Outbreaks (MCAOs). This is when cold, stable and well stratified polar and continental air masses breaks out over a relatively warmer ocean surface. In these boundary zones, small and shallow fronts known as Arctic fronts will form, that separates the cold, stable continental air from the unstable air above the warm sea surface [Kolstad, 2007]. This is why Businger and Reed [1989a] called one of their three main types of elementary polar low developments the "Arctic-front type" [Rasmussen and Turner, 2003]. The formation of an Arctic-front type polar low can in a large degree be compared to the formation of the wave cyclone, or frontal low. Businger and Reed [1989a] discussed an Arctic front just west of Spitzbergen on February 14th 1984, where two polar lows formed. Similar to conventional fronts, this boundary layer front had a low-level jet stream (≈ 30 m/s) above the ice edge and a sloping upwards of the warmer marine air [Businger and Reed, 1989a]. These Arctic-front lows can be associated with the first type of baroclinic instability in polar regions (mentioned at the start of this paragraph) in that they form mainly due to baroclinicity in the low levels [Rasmussen and Turner, 2003].

2.3.2 Thermal instability and convection

The research on polar lows through the years has made it clear that (deep) convection plays a large role in the formation and development of these systems. This has been stated in numerous of research articles regarding polar lows, one example is Businger and Reed [1989b] who stated in their definition of polar lows that their cloud masses were 'largely of convective origin'. Satellite images support this statement, a large number of images showing polar low development include the occurrence of convective clouds [Rasmussen and Turner, 2003].

An important theory proposed for the influence of convection on polar low development is the "Conditional Instability of the Second Kind" theory, abbreviated CISK. CISK was presented by Rasmussen [1977, 1979] and Økland [1977], who suggested that it provided an explanation not only for tropical hurricane mechanisms, but also for similar systems in polar regions, in spite of the threshold value of sea surface temperature being far below 26°C. Another theory suggested to explain polar lows is "Wind Induced Sensible Heat Exchange", abbreviated WISHE. Both these theories state that the polar low develops through a succession of various balanced states [Rasmussen and Turner, 2003]. Fig. 2.3 and Fig. 2.4 show the schematic diagrams of the CISK and the WISHE mechanisms, respectively.

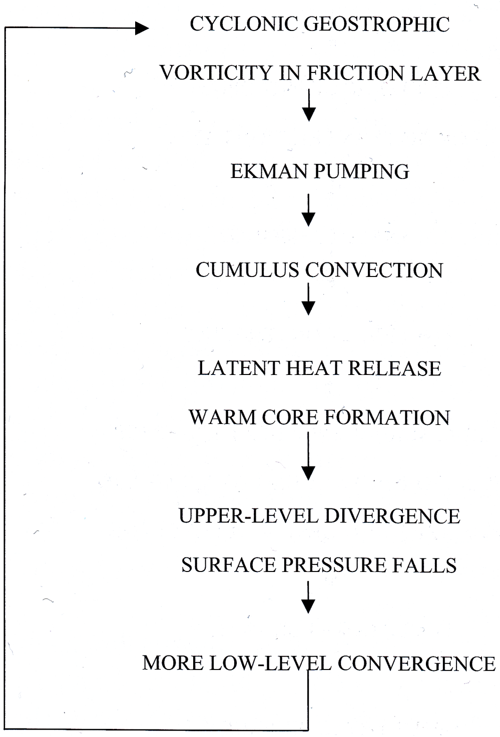


Figure 2.3: Schematic diagram of CISK. Source: Rasmussen and Turner [2003]

To begin with, it should be noted that an important condition for CISK is the presence of a reservoir of "Convective Available Potential Energy" (CAPE). By the presence of CAPE, air parcels which are lifted to their level of free convection will continue to rise. The mechanism of CISK initiates with some kind of disturbance which destabilize the surface atmosphere and triggers cyclonic vorticity in the friction layer (low-level convergence). This causes the Ekman pumping, which triggers and/or enhances cumulus convection. Latent heat release in the convection result in the formation of a warm core which induces divergence (outflow) in the upper levels. The upper-level divergence will obviously cause the pressure to drop at the surface, thus inducing more convergence in the low levels. This cycle is illustrated in Fig. 2.3. It will repeat itself until the low is weakened by some factor, for instance if it makes landfall [Rasmussen and Turner, 2003].

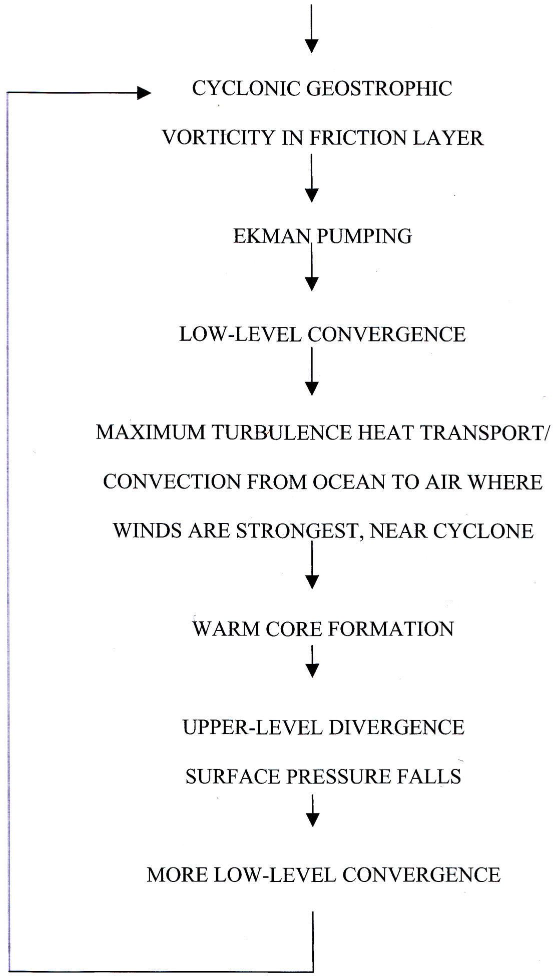


Figure 2.4: Schematic diagram of WISHE. Source: Rasmussen and Turner [2003]

As seen from Fig. 2.4 the WISHE mechanism for polar lows is quite similar to CISK. However, low-level inflow towards the centre of the early stage tropical cyclone or polar low results in increased surface fluxes from the sea surface in the region near the centre of the cyclones, where wind speeds are high. As seen from Fig. 2.4 this is the area of maximum transport of sensible and latent heat upwards, which happens through turbulence and convection. In the WISHE theory, it is assumed that the atmosphere is convectively stable, or neutral. The main source of energy is surface fluxes, and the amount of CAPE will be negligible, as opposed to the CISK theory. However, the presence of a small amount of CAPE is often assumed, and not inconsistent with the WISHE theory. As was also the case for CISK, the WISHE mechanism results in a warm core formation, followed by upper-level divergence and increased low-level convergence [Rasmussen and Turner, 2003].

The initial formation of a polar low is assumed to arise due to an "infinitesimal disturbance" which then grows due to CISK or WISHE. However, theoretical results regarding the CISK and WISHE theories have yet to explain this small disturbance, and the early stage of cyclogenesis arising because of it. Baroclinic instability has been proposed as a likely candidate, which leads to an understanding of polar lows as a two stage process: an initial baroclinic disturbance and intensification phase, followed by a maintenance phase which is related to organized convection [Nordeng, 1990; Rasmussen and Turner, 2003]. Observational results has in many cases supported this view. Satellite images has shown clear indications of deep convection in most polar lows, where many of the systems initially formed in highly baroclinic environments at low levels, i.e., regions near ice edges [Rasmussen and Turner, 2003].

Baroclinic instability, CISK and WISHE have now been presented as some the forcing mechanisms behind polar low formation and development, but even more exists. Based on this, it is not surprising that these systems appear in so many different forms. Also, the increased acceptance of combinations of these mechanisms in polar low formation has led to the understanding of a 'polar low spectrum'. At one end of the spectrum is the purely baroclinic systems, while the pure convection-driven systems represents the other end. In between there exists a variety of hybrids, e.g., the polar lows forming in two stages according to main forcing mechanism. Experience shows that the hybrid form of polar lows is by far the most common [Rasmussen and Turner, 2003].

2.4 Cloud signatures

The use of satellite data provides significant amounts of information regarding the cloud signatures of mesocyclones and polar lows. Satellite images indicate both form and composition of the clouds, which can be used to say something about the physical processes leading to their formation. Obviously, the imagery is also used to determine the lifetime of mesoscale cyclones, as the different stages represented by clouds are shown. The different forcing mechanisms leading to the formation of mesocyclones and polar lows result in different cloud signatures from which the cyclones can be identified. Visible and infra-red satellite images can thus be used to classify mesoscale vortices at high latitudes according to their characteristic cloud signatures. The two most common types observed on satellite imagery is the comma cloud and the spiraliform vortex, and they are presented below [Rasmussen and Turner, 2003].

2.4.1 The comma cloud

Comma clouds got their name from satellite images revealing the comma-shaped cloud pattern of these systems in their mature stage. Most comma clouds arises due to upper-level vorticity anomalies, like a region of upper-level PVA (positive vorticity advection) ahead of a short-wave trough. The tail of the comma cloud is then aligned with the trough axis. Comma clouds are usually found at middle latitudes close to the main baroclinic zone, but can also occur in higher latitudes, e.g., the Barents Sea. The tail of the cloud consists of (shallow) cumulus clouds, and marks the leading edge of a cold air outbreak [Rasmussen and Turner, 2003]. Fig. 2.5 shows a typical comma cloud signature over Iceland around 1000 km west of the polar front, where the comma tail was leading an outbreak of cold polar air.

In their studies of comma cloud development in the Eastern Pacific, Reed and Blier [1986b,a] suggested that an initiating element of the development was the advance of a pre-existing short-wave trough into the region. The following growth of the comma cloud resulting in the associated appearance of a surface cyclone then happened by a migration of the initial disturbance. The migration in question was southwards and over increasingly warm ocean water which created great potential for deep convection. Strong fluxes of heat and moisture from an underlying and relatively warm ocean surface was responsible for the growth of deep convection, and thus the comma cloud. This ‘migration effect’ is thought to also be important for the convective type polar low which forms at higher latitudes [Rasmussen and Turner, 2003].

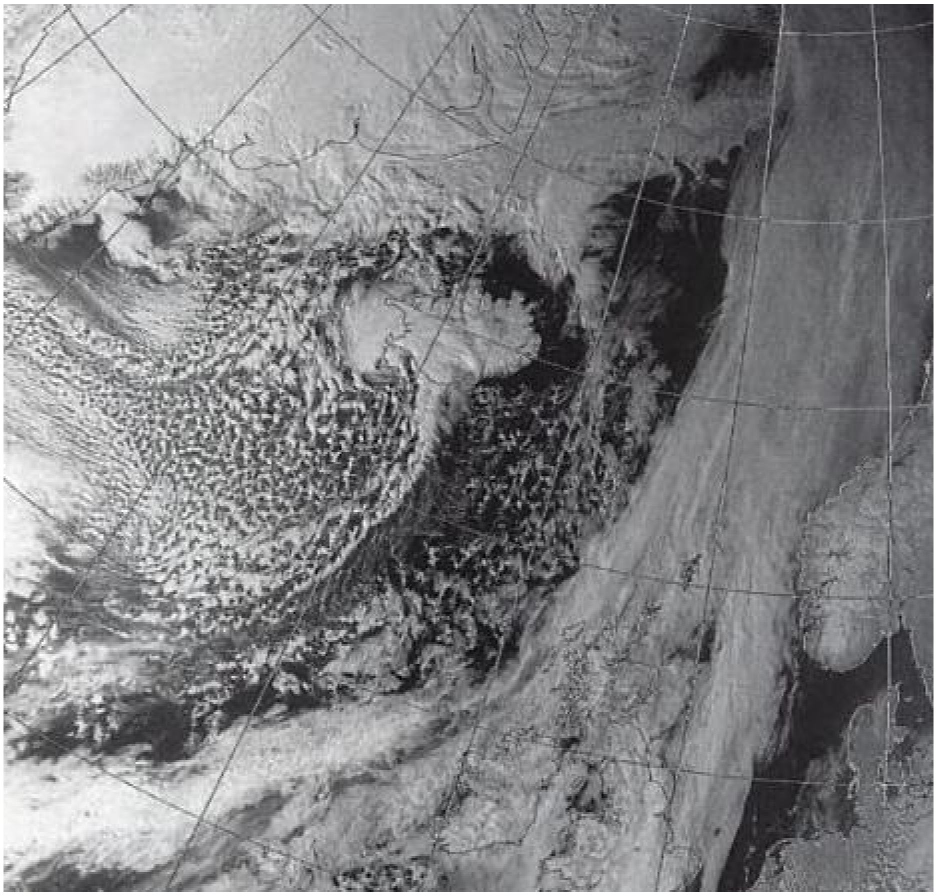


Figure 2.5: A visible wavelength satellite image showing a comma cloud over Iceland around 1000 km west of a synoptic-scale cloud band associated with the polar front. Taken by NOAA 9 at 1528 GMT 8 March 1988. Source: Rasmussen and Turner [2003]

Comma clouds are usually considered to be relatively weak surface systems, either in the form of a trough along the trailing edge of the tail, or as a weak surface low. They can however at occasions transform into polar lows by developing strong surface circulations [Rasmussen and Turner, 2003].

2.4.2 The 'spiraliform' cloud signature

The second main cloud form provided by satellite imagery is the 'spiraliform' signature. Characteristics of spiraliform systems include one or more spiral bands of convective clouds around the circulation centre, and an occasional cloud-free or nearly cloud-free area at the centre, similar to the 'eye' of a tropical cyclone. Fig. 2.6 shows a spiraliform of a marine low close to North Cape. This form is typical for many spiral-structured polar lows. The figure also shows two major spiral arms (cloud bands). The southern extending cloud band is the Svalbard boundary layer front (BLF) which is entering the low from the west and is indicated by small arrows (BLFs will be discussed briefly in Sec. 2.5). The other cloud band, which is north of the centre and in the shape of an arc marks the boundary between an outbreak of cold Arctic air flowing towards the low and a warmer air mass further south [Rasmussen and Turner, 2003].

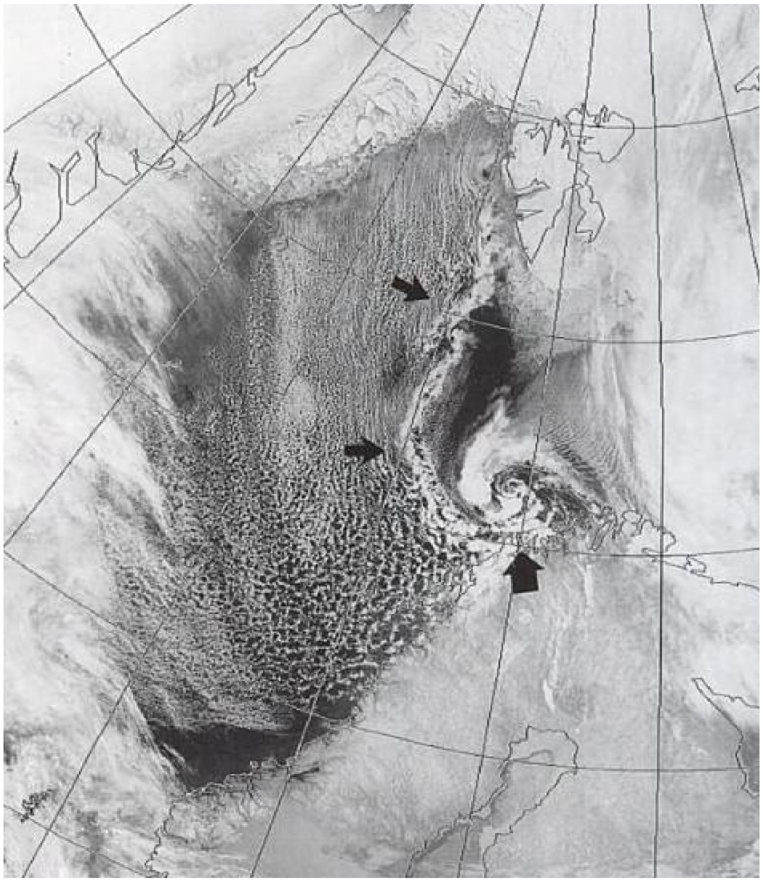


Figure 2.6: A thermal infra-red satellite image showing a spiraliform polar low over the sea close to North Cape (indicated by a large arrow). Also shown on the figure is a well defined *boundary layer front* from west of Spitsbergen and south along 10°E. Taken by NOAA 6 at 1756 GMT 26 March 1981. Source: Rasmussen and Turner [2003]

Many polar lows are triggered by cold, upper-level, short-wave troughs following tracks that lead them over an ice-free ocean. In these cases, strong PVA ahead of the trough axis causes ascent due to dynamic forcing and decreased vertical stability around the cold troughs promotes deep convection. Both of these mechanisms will contribute to the formation of a circulation at the low-level. In other cases polar lows develop near the centre of an upper-level, cold large-scale vortex. Then the upper-level PVA will be weak and the main contributing factor for polar low formation will be decreased vertical stability due to the very low upper-level temperatures near the centre. These types of polar lows are referred to as type (3), or the 'cold-low type' in Businger and Reed's classification system (discussed in Sec. 2.5). They are quite common, and will usually have a spiraliform cloud signature [Rasmussen and Turner, 2003].

Fig. 2.7 shows an example of a cold-low type of polar low development, and the associated spiraliform cloud signature. The spiral was associated with a well-developed polar low at the surface, which formed as an upper-level cold low passed over the sea around Iceland [Rasmussen and Turner, 2003].

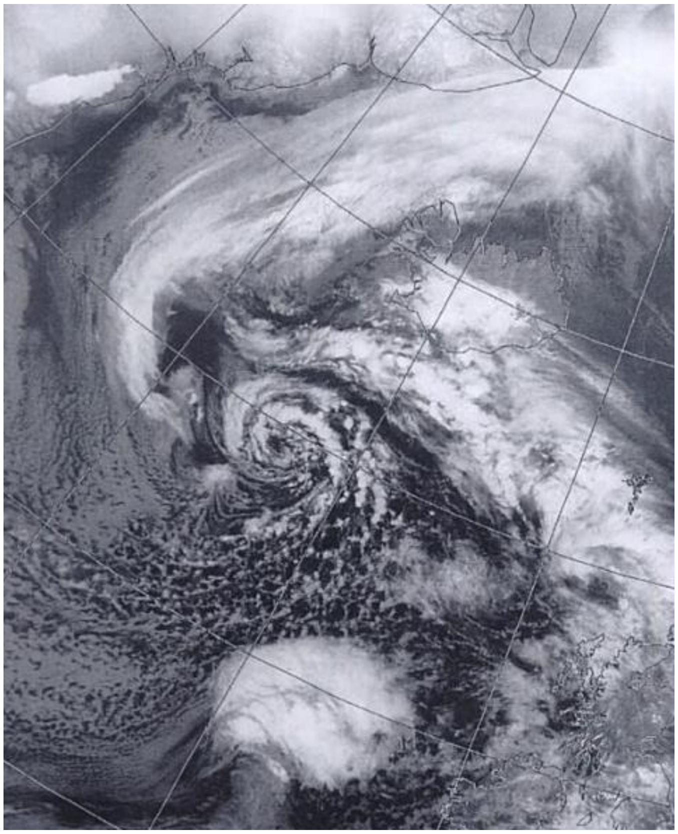


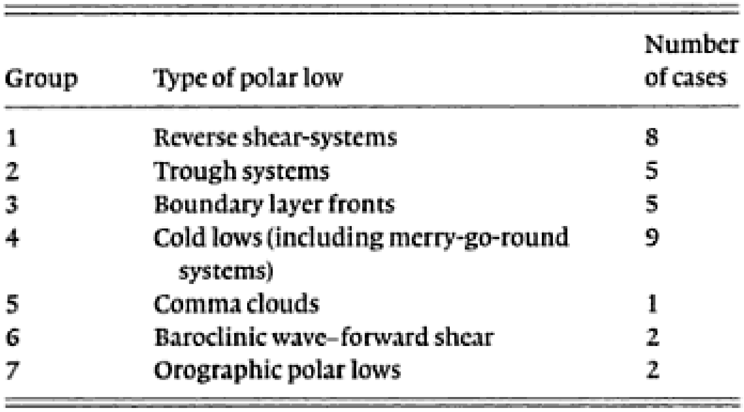
Figure 2.7: An infra-red satellite image showing a polar low with a spiraliform cloud signature south of Iceland at 1549 GMT 9 December 1993. Source: Rasmussen and Turner [2003]

2.5 Classification

Through the years of polar low research, various classifications has been used. Businger and Reed [1989a,b] presented a classification system which has since been widely used. The classification comprises three main types of polar low development based on the distinctive synoptic patterns connected to them: (1) the short wave/jet-streak type which is characterized by a secondary vorticity maximum and PVA (positive vorticity advection) aloft, deep, moderate baroclinicity, and modest surface fluxes; (2) the Arctic-front type, associated with ice boundaries and characterized by shallow baroclinicity and strong surface fluxes; and (3) the cold-low type, characterized by shallow baroclinicity, strong surface fluxes, and deep convection. Type (1) is practically the same as the comma cloud (discussed in Sec. 2.4) which is associated by PVA aloft. In fact, it is believed that upper-level PVA and the cold temperatures in the upper-level regions of the short-wave trough is not only important for the first type, but for most significant polar low developments, including Businger and Reed's type (2) polar low, the Arctic-front type [Rasmussen and Turner, 2003].

Grønås and Kvamstø [1995] suggested a modification of Businger and Reed's type (2) definition where they rather chose to call it *Arctic outbreak polar lows*. They argued the following: 'We chose to call this class *Arctic outbreak polar lows* since they do not form at the leading edge of the Arctic front. It seems evident that a mobile upper disturbance is also active.' They continued: 'The presence of an upper disturbance, using the potential vorticity concept, should therefore be included in the definition.' [Rasmussen and Turner, 2003].

A study of a number of polar lows carried out by Wilhelmsen [1985] were used in an attempt to extend and improve Businger and Reed's classification. This new classification was proposed by Rasmussen (unpublished data) and the basis of his study was 32 cases of gale-producing polar lows near Norway in the period $1978 - 1982$, identified and listed by Wilhelmsen. The resulting seven categories of polar lows are presented in Fig. 2.8 and are based on a combination of synoptic features and physical formation mechanisms, accounting for polar lows in the Nordic Seas. This scheme is quite general, meaning that practically all polar lows observed in the region could be placed within one of the categories. But due to polar low developments strong dependence on local climatological factors, this scheme would not necessarily be appropriate for other areas of polar low development [Rasmussen and Turner, 2003].



Group	Type of polar low	Number of cases
1	Reverse shear-systems	8
2	Trough systems	5
3	Boundary layer fronts	5
4	Cold lows (including merry-go-round systems)	9
5	Comma clouds	1
6	Baroclinic wave–forward shear	2
7	Orographic polar lows	2

Figure 2.8: Polar lows found in Wilhelmsen [1985]. Source: Rasmussen and Turner [2003]

Group 1: Reverse shear systems An important group of polar lows is the reverse shear systems. In contrast to the classical baroclinic systems, conditions are such that the low-level steering wind is antiparallell to the thermal wind in the adjoining layers [Duncan, 1978]. Duncan [1977] found that a necessary condition for polar lows to develop was that a vertical wind shear existed and that the thermal wind and the low level steering winds were parallel. He later noted, [Duncan, 1978], that reverse-shear conditions could also be favourable. In the 'ordinary' baroclinic instability case, relative to the motion of the system, the warm air lies to the right of the path, and the cold to the left. For the reverse shear, the thermal wind is antiparallel to the flow, and the warm air lies to the left of the prevailing flow [Kolstad, 2006]. Behind the baroclinic wave disturbance, there is warm advection and ascending motion, and ahead there is cold advection and descending motion [Businger and Reed, 1989b]. Again, the situation will be opposite in the classical case. This thermal circulation will cause the gain of kinetic energy from the reservoir of available potential energy [Rasmussen and Turner, 2003].

Reverse shear type polar lows often develop in the region between Spitsbergen and Iceland within a northerly or north-easterly low level flow. This flow is associated with the western flank of a large-scale low which is located further to the east. West of this flow, over the Fram Strait and the sea ice region along the east coast of Greenland, cold temperatures dominate, as opposed to the warmer air in the Norwegian Sea further to the east and southeast. These conditions frequently lead to low static stability and reverse shear baroclinic developments. By documentation in Wilhelmsen [1985], reverse shear baroclinic developments seem to be quite frequent also in the Barents Sea, then along the ice edge aligned close to the 75 °N latitude [Rasmussen and Turner, 2003].

Reversed shear has been proposed as a favourable condition for polar low development [Kolstad, 2006], and considered as an important subgroup of polar lows [Kolstad, 2011]. However, in a study done by Blechschmidt [2008] of 90 polar lows over the Nordic Seas in the years 2004 and 2005, only one fourth (22 %) fulfilled the criteria for reverse shear systems [Kolstad, 2011].

Group 2: Trough systems The development of this group of polar lows are associated with large-scale cyclones, and their tracks during the winter. They tend to follow a track over the northern part of the Norwegian Sea and into the Barents Sea. Polar lows may form in the 'trailing troughs' of these cyclones, which follow an east-west orientation. The trailing troughs are characterized by large values of low-level vorticity on the western flank of of the synoptic-scale lows [Rasmussen and Turner, 2003].

Upper-air charts have shown that polar lows that form within synoptic-scale surface troughs are in most cases triggered by short-wave cold troughs in upper levels. The approach of this upper-level short-wave trough (or jet streak) and upper-level PVA will result in a local increase in low-level vorticity. This again leads to the spin-up of air, and the formation of a low level vortex. Deep convection will arise due to the fact that the region is already characterized by ascending air processes, and destabilization through upper-level advection. The polar low is now formed, and will be maintained, or developed further by the release of latent heat. The polar low depicted in Fig. 2.6 is actually a trough system polar low. It formed within a trailing trough over the Norwegian/Barents Sea behind a large, synoptic-scale cyclone [Rasmussen and Turner, 2003].

Group 3: Boundary layer front type Fig. 2.6 and several other satellite images have shown the occurrence of boundary layer fronts (BLFs). They have been located along and south of the west coast of Svalbard in situations of a northerly large-scale flow. The aforementioned Arctic fronts are BLFs with their origin near the edges which separate shallow, low-level Arctic air masses from warmer, maritime polar air masses. Cloud street orientations shown in satellite images often indicate the presence of low-level horizontal wind shear across the BLF. The cross-front length of this region of shear may be a few tens of kilometres. Along the fronts, small-scale vortices may form, which due to their short lifetimes and lack of development regarding size will be considered as quite insignificant systems. However, under the influence of upper-level forcing, like in the presence of a short-wave trough, these small-scale vortices may strengthen and develop into polar lows [Rasmussen and Turner, 2003]. The BLFs, which initially form close to ice edges, may become detached from this coastline and drift away over open water, a motion driven by the synoptic-scale flow. The leading edge of this motion, i.e., the Arctic front, is often visible on satellite images as a cloud band with small vortices [Rasmussen and Turner, 2003].

Group 4: Cold low type Cold low type polar lows (type (3) in Businger and Reeds classification) are small spiral, or comma cloud patterns of convective character which forms/intensifies within the inner cores of old occlusions or cold lows. This happens without any detectable connection to upper-level short waves or low-level baroclinic features [Businger and Reed, 1989a,b]. The cold low type represents the largest of the seven polar low groups from the table in Fig. 2.8. This type of polar lows form within the cold core of large-scale cold lows, and their cloud patterns often consist of impressive spiraliform structures, easily distinguished from comma cloud patterns. Observations of several polar lows of this type have been made in the Arctic when an upper level cold low moves out over open water. Over the water, the low level circulation may intensify, which causes a vortex that develops into a polar low [Rasmussen and Turner, 2003].

Group 5: Comma clouds Comma clouds were discussed earlier, in Sec. 2.4. As expected, few of the polar lows listed in Wilhelmsen [1985] were characterized as comma clouds, seeing as most of them developed at rather northerly latitudes *away* from the main baroclinic zone. Only one (rather uncertain) comma cloud case was identified, at a rather southerly latitude [Rasmussen and Turner, 2003].

Group 6: Baroclinic waves Baroclinic instability along secondary baroclinic zones poleward of the main baroclinic zone (the polar front) has long been considered as an initiating mechanism for the formation of polar lows. These zones include shallow baroclinic zones along ice edges and/or ice/snow-covered coastlines. A purely baroclinic development will consist of stable ascent which causes the formation of cloud signatures of stratified clouds. The cloud patterns associated with baroclinic polar low developments may bear a striking similarity to those of synoptic-scale extratropical cyclones. Cloud patterns may include leaf-like structures connected to the early stages of a baroclinic polar low development, which then develops into a comma-like structure, and finally into a tight, and more spiraliform vortex [Rasmussen and Turner, 2003].

Most comma clouds are examples of baroclinic systems, even though convection may be present in the development. Two baroclinic (forward shear) polar lows were identified in Wilhelmsen [1985]. One was a rare case of a polar low developing in summer (June 1980), and the other was from October 1980, but they were still surprisingly similar. Applying for both systems was that the main baroclinic zone (the polar front) was far to the south, and their formation was associated with small-scale waves in secondary baroclinic zones. Contrary to shallow BLFs arising from ice-edges, the frontal zones were deep, thus resembling the polar front [Rasmussen and Turner, 2003].

Group 7: Orographic polar lows This last type of polar lows represents the systems that develop primarily because of influences from the orography. Two of the cases in Wilhelmsen [1985] were characterized as orographic polar lows. Examples of regions where orography play a part in polar low development include the lee of Iceland, where vortices formed during periods with a northerly flow have developed into polar lows. Another region is the sea east and northeast of Cape Farewell, the southern tip of Greenland. During outbreaks of polar/Arctic air from the North American continent, westerly or northwesterly flows of air can be forced over or around the mountains here, and vortices may form. A third region, for which an example was found in Wilhelmsen [1985], is the area around Scandinavia. Here, polar lows can form as secondary circulations associated with an eastwards moving parent circulation. They then form in the lee of the Norwegian mountains due to orographic effects. An example of this occurring in Wilhelmsen [1985] was a rather large polar low which was left over the Norwegian Sea, behind a large-scale cyclone centre which passed the Scandinavian peninsula in an eastwards direction (2-3 of January, 1978) [Rasmussen and Turner, 2003].

2.6 Forecasting and tracking of polar lows

As stated earlier, polar lows can have severe impacts on maritime installations in their path, as well as coastal installations and communities when they make landfall. Due to the large number of oil and gas platforms situated in areas like the Norwegian Sea, good forecasting methods are important, in the effort of minimizing the impact on these kinds of operations [Rasmussen and Turner, 2003]. This section will thus focus mainly on the forecasting aspects of polar lows in the Nordic Seas regions.

Forecasters use the output from numerical weather prediction (NWP) analysis and forecast systems in the means of determining the time and location of mesocyclone developments in a certain region. Obviously, the forecasting improved substantially with the utilization of satellite images and data, which are essential in the process of predicting the paths of already existing mesoscale vortices a few hours ahead. Polar lows are of the mesoscale, the scales of individual lows may vary greatly and they are very likely to appear together with other synoptic and subsynoptic-scale systems. The forecasts of polar lows must therefore involve a variety of atmospheric scales. A combination of weather forecasting methods are thus required [Rasmussen and Turner, 2003].

At the weather forecasting center in Tromsø, Northern Noway (VNN), forecasting of polar lows are constantly carried out throughout the winter season. For identifying purposes, certain facts and criteria has been determined to indicate a developed polar low [Noer and Ovhed, 2003]:

- Winds are stronger than 27 kts (13.9 m/s)
- Horizontal extent in the 200 - 600 km range
- Cyclonic appearance
- Development is associated with cold air outbreaks from the polar ice cap, large scale winds flowing northwesterly to northeasterly directions

This definition could easily apply for several other polar weather systems. Therefore, to avoid misclassification, polar lows are also always associated with having a deeply convective nature. In combination with the other criteria, this separates them from similar systems like small synoptic lows (which involves warm and cold fronts), surface troughs (which has a more linear appearance) or shallow eddies or troughs (vortices of weaker winds than those stated for polar lows) [Noer and Ovhed, 2003].

Some years ago, the methodology of polar low forecasting applied at VNN was mainly a two stage process. First a 24 hour outlook forward was made. In the case of an identified polar low, a more detailed, short-term forecast for the next 0-6 hours was carried out [Noer and Ovhed, 2003]. With the development and improvement in forecasting models the recent years, the transitions of the forecasting process are now more gradual. Usually, long term prognoses are made for the next < 48 hours, where the focus is on large-scale flow patterns, like cold air outbreaks (CAOs) and (upper-level) cold-core lows. Shorter prognoses than this involve the gradual transition to medium-scale models like HIRLAM with 8 km or 12 km resolution, or AROME with 2.5 km resolution when the forecasting is concentrated in areas close to the coast. When a polar low is identified and determined in relation to the prognoses, the forecasting process enters a ‘nowcasting’ mode which typically involves the prognoses for the next 6-12 hours. In addition to this, the ‘old’ methodology is still present in that routine evaluations of the potential development for the next 24 hours are done on a regular basis (referring to conversations with G.Noer 2014).

In the 24 hour registration, the following data are collected [Noer and Ovhed, 2003]

1. The current situation, involving troughs or polar lows, which can be compared to yesterdays situation
2. The associated current situation in the upper levels, like troughs, jet shear and the T_{500} (the temperature at 500 hPa)
3. Before the outlook for the next 24 hours is carried out, the current situation is classified according to surface phenomena
4. Surface conditions for the next 24 hours
5. Upper level conditions for the next 24 hours
6. Other comments on past or future developments

A typical polar low case registered at VNN is the one arising due to the interaction between some kind of a low level disturbance which is embedded in a CAO, with positive vorticity advection (PVA) in upper levels. The PVA is associated with a cold trough from the Arctic basin. In other polar low cases, high values of potential vorticity (PV) in areas of jet shear have lead to the formation.

The first step is to evaluate the situation in the low layers. The current situation is first investigated, directing the main attention to areas of cold air outbreaks from the polar ice cap. Through the years of forecasting, all of the following low level phenomena (embedded in CAOs) have been discovered to be responsible for polar low development: Shallow surface fronts or convergence lines, shallow eddies, open convective cells, organized troughs, and isolated baroclinic zones [Noer and Ovhed, 2003].

Next, the development for the coming 24 hours is evaluated. If the CAO persists, potential low level systems will be advected in the flow. The absence of a CAO can be interpreted as a low probability for a polar low to develop in the next 24 hours, because any development of a low level disturbance requires some time [Noer and Ovhed, 2003].

Following the surface evaluation is a consideration of upper level conditions, for instance a search for a cold trough in the Z_{500}/T_{500} prognostic fields, which may lead to a low level disturbance. A general value to look for here is a 44°C temperature difference between the sea surface temperature and the T_{500}. Judgements are made regarding the curvature and propagation speed of a potential trough. If the trough is overshooting and propagating fast, a low level vortex development will probably be short lasted, and decay once the trough has moved on. If the trough is slow moving and has a phase lock with the low level disturbance, the probability for full development to a polar low is higher. The site of polar low development will then be located in the intersecting point of the upper trough with the low level disturbance [Noer and Ovhed, 2003].

If specific polar low events are identified, short term forecasting is initiated. The movements, strength and other characteristics of the system must now be considered as close as possible, preferably in fair time before the system reaches coastal fishing areas, or other areas areas of significance [Noer and Ovhed, 2003]. One important aspect of consideration is the estimation of polar low surface winds, seeing as these are the parts of the mesocyclone system carrying the greatest potential of damage to marine operations [Rasmussen and Turner, 2003]. By the use of satellite imagery, some important visual clues in a subjective judgement of the polar low wind speed are the following [Noer and Ovhed, 2003]:

- A clear eye
- A wavelike pattern of cirrus outflow radiating out along the rim of the low indicates strong winds
- Cases of singular lows are usually stronger than when two or more centers are present

Another important, and quite recent application of forecasting and tracking of polar lows is the use of a so called Ensemble Prediction System (EPS). An EPS is a set (ensemble) of prognoses applied to a selection of equally probable defined initial states when a set of observations are given. Each of the members of this ensemble are slightly different from the others, thus providing a set of various solutions. This methodology is based on the atmosphere's own chaotic nature, and the fact that too few observations are available to give an accurate description of it.[1]

Based on the set of solutions provided by the EPS many central questions within forecasting can be answered, for instance how certain the prognosis is, the time range of its reliability, and its significance with respect to other possible outcomes. An EPS provides information of tendencies in the weather development, for instance of increasing or decreasing temperatures. The probability for a given type of weather occurring, for instance strong winds or heavy precipitation, or the track of a low-pressure center can be obtained by using an EPS. Hazardous weather conditions can also be predicted, even if the probability is low.[1]

At VNN, in the recent years, forecasting methods based on EPS are utilized for tracking purposes, to provide predictions of the most likely track of a polar low, and the probability

---

[1] Article by Barentswatch, `http://www.barentswatch.no/Tema/Sjotransport/Polarvar-og-istjenester/Polare-lavtrykk1/Prognose-for-polart-lavtrykk-/`

of associated strong winds and precipitation. A track prognosis calculated from standard prediction models are always related to an uncertainty. However, by using EPS, a set of possible tracks are given, and high probability tracks can be viewed as areas of geographical overlap of several of these. The track probability can be plotted as in Fig. 2.9, where green color indicates a moderate probability, and blue represents high probability. The plot in Fig. 2.9 is not related to time specifically, but displays a possible track during the next 42 hours from the start of the prognosis.[1]

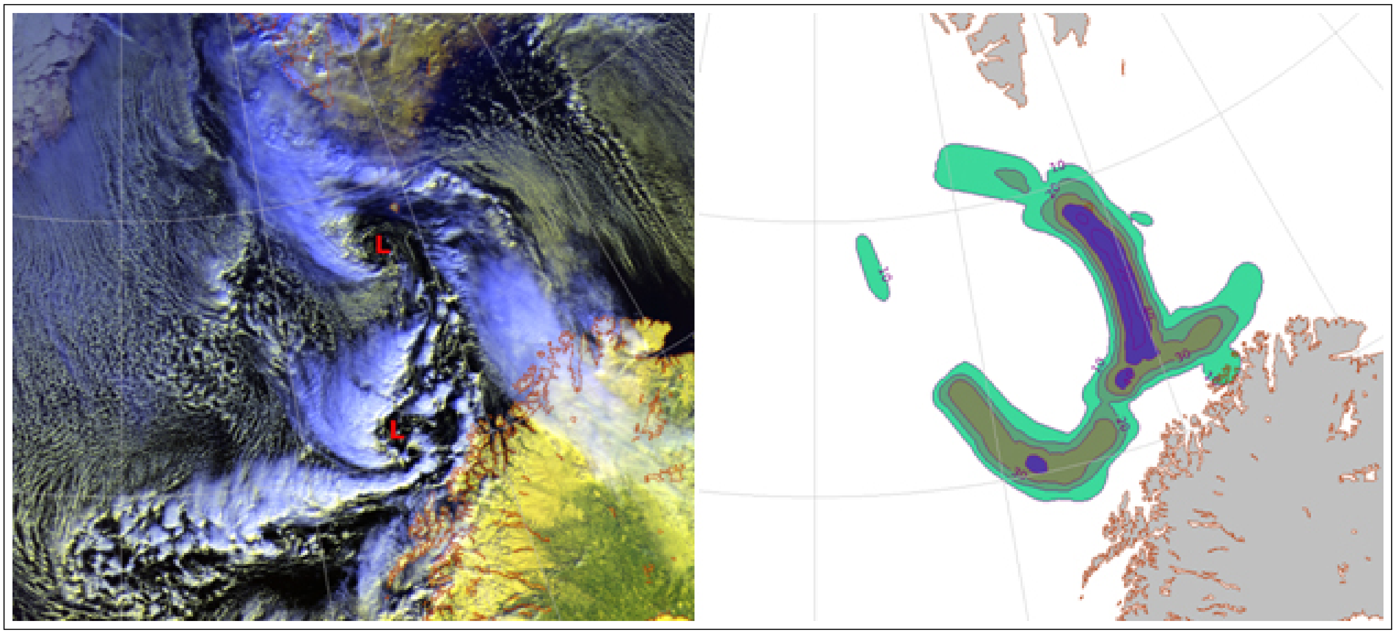


Figure 2.9: Satellite image to the left shows a center west of Bear Island, and one northwest of Troms. Both are fairly well represented in the track plot to the right. A smaller center is also present west of Lofoten, but due to its small diameter, it is not included in the track plot. Source: www.barentswatch.no (see footnote on this page)

The tracking routine is only applied to low pressure centres of diameters between 200 and 600 km, as most polar lows are within this size range. This means that smaller systems, for instance showery weather confined to local areas are filtered out to avoid fluctuations of very small scales. The same apply for large scale pressure systems, for which other types of prognoses and forecasting methods are well established. The track plots are only calculated for areas where air is highly unstable, and by combining several types of criteria, misclassifications and 'false alarm' forecasts are avoided.[1]

The EPS prognoses provide relatively good descriptions of the expected polar low track, even though the position estimates are of varying accuracy. Tracks plotted with low probability are also useful in that they can be interpreted as expected areas of unstable weather, with heavy snowfall and periods of gale force winds. Overall, the track plots offer good

---

[1] Article by Barentswatch, `http://www.barentswatch.no/Tema/Sjotransport/Polarvar-og-istjenester/Polare-lavtrykk1/Prognose-for-polart-lavtrykk-/`

indications of expected polar low trajectories, and thus withhold a significant practical value in relation to areas which may be affected.[1]

Obviously, the forecasting methods for polar lows have improved through the recent years. During the last couple of decades, significant advances in polar low research has been made due to the ability of representing these systems in numerical models. Consequently, some polar lows can now be represented explicitly in greatly improved, high-resolution NWP systems and/or through EPS prognoses. Since polar lows were first observed, great advances have been achieved in the understanding of these systems, and there is no reason to believe that the research on these systems will decline any time soon [Rasmussen and Turner, 2003].

---

[1] Article by Barentswatch, `http://www.barentswatch.no/Tema/Sjotransport/Polarvar-og-istjenester/Polare-lavtrykk1/Prognose-for-polart-lavtrykk-/`

Chapter 3

Data

The statistical analysis in this thesis is conducted on a data set developed by Rojo et al. [2014] for the article *Polar lows tracks over the Nordic Seas: a 14 winter climatological analysis*. This study presented a manual tracking of a total of 190 polar low cases, during 14 winter seasons from 1999 – 2000 to 2012 – 2013. Of these 190, 80 events were dual or multiple, and the data set thus contained coordinates for a total of 286 polar low trajectories. The following data and methods were used by Rojo et al. [2014] in determining the tracks and characteristics of the polar lows: 1) a list of polar lows, made available from the Norwegian Meteorological Institute (met.no); 2) the identification and observation of cloud signatures associated with polar lows using mainly AVHRR thermal infra-red imagery; 3) atmospheric reanalysis data [Rojo et al., 2014].

3.1 List of polar lows

The manual polar low tracking over the Nordic Seas was partly based on the annual report on polar low activity which is published by the Norwegian Meteorological Institute (met.no) [Noer et al., 2011]. This routine for registration of polar low events was introduced in 2000, by forecasters at the Norwegian Meteorological Institute. Methods for collecting data for these lists include the observation of synoptic scale circulations over the Norwegian and Barents Sea through high resolution imagery. The most recent report listed dates and positions for 190 developed polar lows in the Norwegian and Barents Sea over the time period December 1999 to June 2013 [Rojo et al., 2014].

3.2 Advanced Very High Resolution Radiometer (AVHRR) imagery observation (manual tracking)

Along with the list of polar low positions (coordinates) in their mature, developed phase, the trajectories of each event was determined by using an archive of Advanced Very High Resolution Radiometer (AVHRR)- Channel 4 (10.3-11.3 μm) images provided by the Natural Environment Research Council (NERC), Canada. For the eastern most cases of polar lows, satellite images from the Space Research Institute of Russian Academy of Sciences for Barents Sea regions were used to identify and document the tracks. The thermal infrared satellite imagery also provided possibilities for estimations of location, duration and diameter, which were made by observing cloud signatures associated with the polar low events [Rojo et al., 2014]. Fig. 3.1 shows an example of manual tracking using AVHRR images, carried out by Rojo et al. [2014].

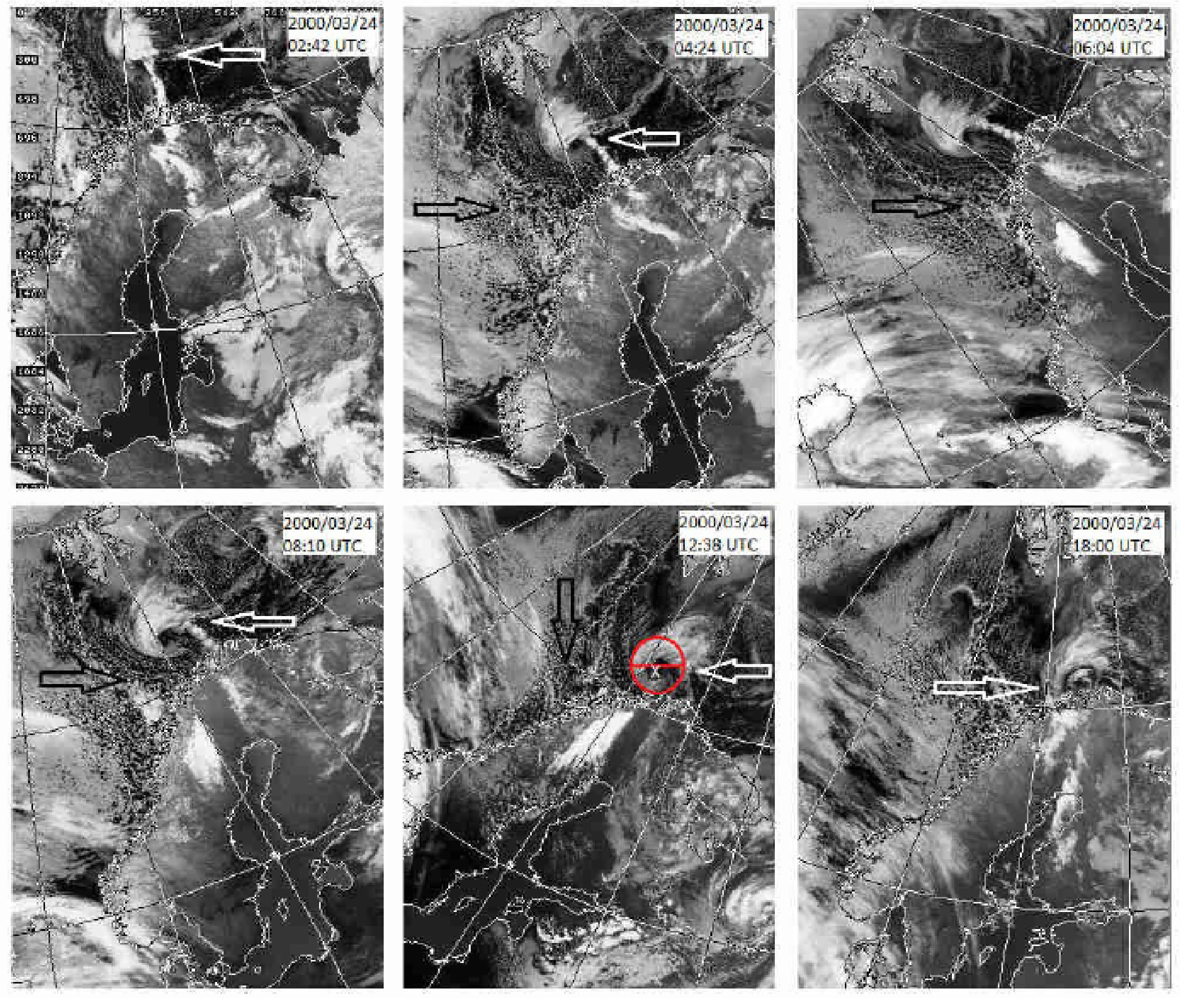


Figure 3.1: Example of a PL (polar low) manual tracking using AVHRR images. The white arrows indicate the PL, black arrows point out mesoscale convective cells (indicators of the cold air outbreak). On the first three images, a comma shaped PL forms between Svalbard and Norway. The fourth image shows a transition phase where the PL becomes spiraliform. PL reaches its mature stage close to the coast of Finnmark, still as spiraliform (image 5). The white cross indicate the centre of the PL, the red line indicate its approximative diameter, and the red circle indicates the approximative circumference. On the sixth and last image the PL dissipates over the coast of Finnmark. Source: Rojo et al. [2014]

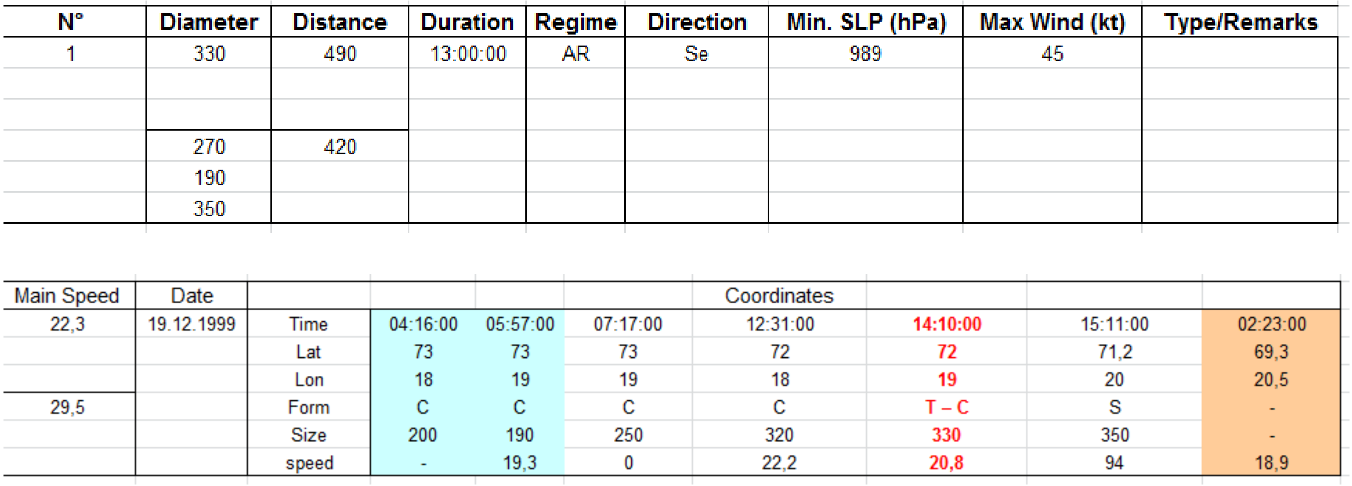
3.3 Atmospheric Re-analyses

Each manual polar low track were compared with the lower atmosphere wind field using the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis Interim (ERA-I) wind fields at 850 hPa and 950 hPa. Because the development of polar lows over the Nordic Seas are associated with the North Atlantic-Europe Weather Regime (WR) (documented by Mallet et al. [2013]) two key fields to determine WR were investigated. These two include geopotential height at 500 hPa (Z500) and the mean sea level pressure (MSLP). Mallet et al. [2013] used the daily mean values of Z500 from the NCEP/NCAR Re-analysis to determine WR. The NCEP/NCAR Re-analysis covers 17 pressure levels with a 2 ° grid resolution, over a period from 1948 to the present [Kalnay, 1996]. The WR was determined over the North Atlantic Europe domain (20 °N - 80 °N/90 °W - 30 °E) with the use of clustering techniques for Z500 daily maps for the northern hemisphere winter (November to March), over the period 1957 – 2013 [Cassou, 2008]. Four main large-scale flow patterns were identified for the winter season: Scandinavian blocking, the positive and negative phases of the North Atlantic Oscillation (NAO), and a pronounced ridge over the Atlantic (Atlantic ridge). Except for polar lows forming in September - October and April - May, and in the two last winter seasons (2011 - 2012 and 2012 - 2013), the regime for which each polar low date belonged to was included in the final dataset [Rojo et al., 2014].

Based on these data and methods, Rojo et al. [2014] were able to develop a data set containing the tracks and characteristics of each of the 190 polar low cases. The tracks include the whole of the polar lows lifetime, from genesis to dissipation. Genesis was determined as the first image showing a clearly formed cyclonic cloud structure. The dissipation corresponded to the image where cyclonic cloud structures were broken, i.e., no longer visible. The trajectories arises from the positions of each case from genesis to dissipation, where the given coordinates corresponds to the centre of the cloud vortex, i.e. the centre of the PLs eye. For each coordinate, the corresponding time is given. In cases where multiple centres occur, only data for the deepest low is measured and analysed [Rojo et al., 2014]. It should be noted that the coordinate/time measurements are not evenly sampled, time between each measurement is highly variable. The time interval in many cases ranges between 2-3 hours, but time between one measurement to the next can stretch to several hours.

One of the characteristics determined was the size of PLs. The size is determined by the diameter of the fully developed PL, meaning the circular cloud band of the PL in its mature phase. Other characteristics associated with the dynamics of the PLs was analysed and determined based on the manual PL tracking. These include duration, distance travelled, velocity of propagation, and direction [Rojo et al., 2014]. The duration corresponded to the approximate time between the first and last observation of a mesoscale vortex shown on satellite images. The duration were categorized into 9 time intervals, from less than 6 hours to more than 48 hours, with 6 hour intervals. The distance travelled by each PL were also calculated by Rojo et al. [2014]. This included the distance between the first and the last point of the PL, but also the complete distance travelled, meaning the complete distance between each point that made up the trajectory. To obtain an index of straightness, Rojo et al. [2014] divided the distance travelled by the complete distance travelled. The velocity of propagation was estimated by dividing the total distance travelled by the PL by its duration. Main direction was determined as the straight line between the point of departure and the point of arrival, i.e., the first and last point of the PLs trajectory. Directions were classified into 16 categories at 22.5 ° intervals [Rojo et al., 2014].

As an example, Fig. 3.2 displays a segment of the data set produced by Rojo et al. [2014], showing all the mentioned properties and the trajectory time and coordinates of the first polar low, which occurred on the 19th of December 1999.



N°	Diameter	Distance	Duration	Regime	Direction	Min. SLP (hPa)	Max Wind (kt)	Type/Remarks
1	330	490	13:00:00	AR	Se	989	45	
	270	420						
	190							
	350							

Main Speed	Date		Coordinates						
22,3	19.12.1999	Time	04:16:00	05:57:00	07:17:00	12:31:00	**14:10:00**	15:11:00	02:23:00
		Lat	73	73	73	72	**72**	71,2	69,3
		Lon	18	19	19	18	**19**	20	20,5
29,5		Form	C	C	C	C	**T – C**	S	-
		Size	200	190	250	320	**330**	350	-
		speed	-	19,3	0	22,2	**20,8**	94	18,9

Figure 3.2: Data for the first polar low event in the data set produced by Rojo et al. [2014].

Chapter 4

Methods of analysis

4.1 Kernel density estimation

The density estimation carried out in this thesis deals with non-parametric techniques, meaning those where the underlying probability density function (pdf) of the data is not known and has to be approximated. The oldest and most common non-parametric density estimator is the *histogram*. The method is quite simple. Given a random sample $X_1, X_2, \ldots, X_n$ from an unknown continuous distribution. An origin x_0 is chosen, and the real line is divided into so-called bins of *binwidth* h. The bins of the histogram are then the intervals

$$[x_0 + (j-1)h, x_0 + jh), \qquad j \in \mathbb{Z}. \tag{4.1}$$

The number of observations which falls into each bin j are then counted, and denoted n_j. The frequency count n_j of each bin is divided by the sample size n and by the binwidth h:

$$f_j = \frac{n_j}{nh}. \tag{4.2}$$

Finally, the histogram is constructed by placing a bar with height f_j and width h over each bin. The intervals containing most observations (the areas of highest density) will thus be recognized by the tallest bars. The collection of bars will represent an estimate, or approximation of the pdf [Härdle et al., 1995].

An example of a histogram is shown in Fig. 4.1. It shows the result of the histogram approach executed on the latitude data points of the polar lows data set developed by Rojo et al. [2014]. The sample size n in this case is then 2270, which is the total number of coordinates for all the 286 trajectories in the data set. The data points have been placed into ten bins of equal length.

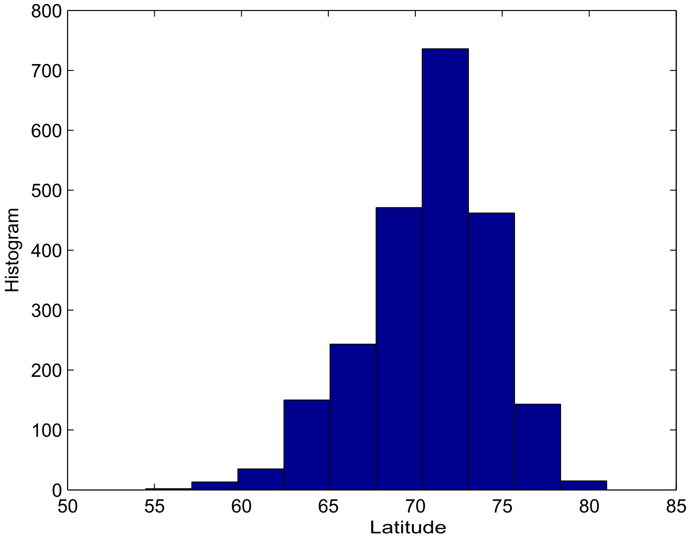


Figure 4.1: Histogram for latitude data with 10 bins of equal length.

By looking at Fig. 4.1, it becomes apparent that the choice of binwidth is what primarily controls the amount of smoothing for the histogram approach [Silverman, 1986]. To illustrate this, the histogram using latitude points is plotted again in Fig. 4.2, but now the data are distributed by using 50 bins of equal length, i.e., the binwidth is decreased.

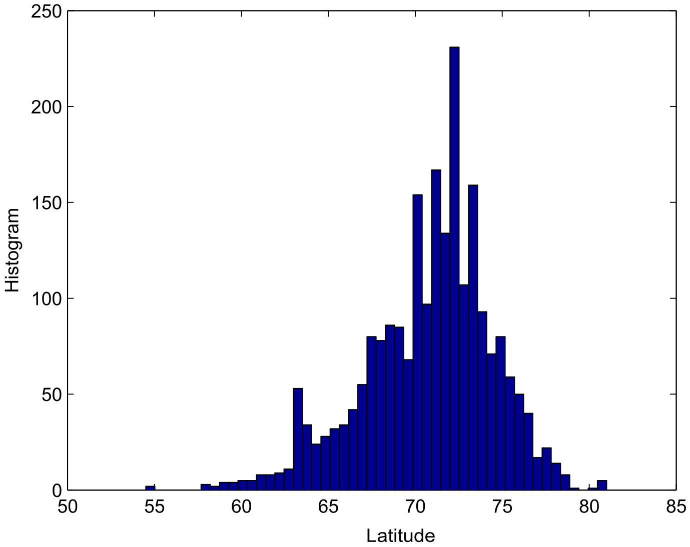


Figure 4.2: Histogram for latitude data with 50 bins of equal length.

As mentioned, the histogram provides a rather easy method for density estimation, but the approach also has some shortcomings. For instance, the histogram is not a continuous function. The discontinuities between the bins represents non-differentiable areas and the histogram has zero derivative everywhere else. These features are unfortunate if it is desirable to estimate a smooth, continuous pdf [Härdle et al., 1995].

In these cases it is preferable to use the method called *kernel density estimation* which is basically a generalization of the histogram approximation method. An intuitive and simple understanding of this method is that the bars of the histogram are replaced by continuous kernel functions, or "bumps". But instead of one bar being placed over all the data points located within the bin width (width of the bar), a bump is placed and centered over each data point. This method thus operates on a higher level of accuracy. With the histogram, all points within a certain region (bin) are assigned to the same bar, which could give a false impression, say if the majority of the points were clinched together at the outer edge of the bin. The density estimate at a given x can be found by vertically summing over the bumps [Härdle et al., 1995]. Fig. 4.3 shows a kernel density estimate for an artificial data set which illustrates this understanding. The red crosses represent the data points, and the blue "bumps" the kernel functions. The black function represent the density estimate, a vertical summation of the kernel functions.

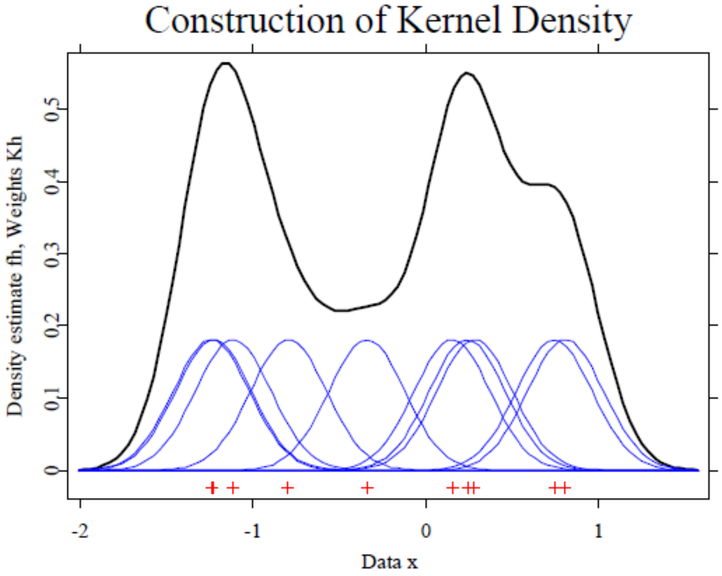


Figure 4.3: Kernel density estimation as a sum of bumps. Source: Härdle et al. [1995].

The kernel density estimate $\widehat{f}_h(x)$ at point x is defined by

$$\widehat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right), \tag{4.3}$$

where n is the sample size of the random sample $X_1, X_2, \ldots, X_n$. $K(\bullet)$ is the kernel function, and h is the so-called bandwidth of the kernel function [Härdle et al., 1995]. The kernel function is a smooth function that represents the basic shape of the probability mass assigned to each location where a data point occurred. Usually, non-negative, symmetrical density functions are chosen as kernel functions. The bandwidth, or smoothing parameter, corresponds to the scale over which the mass is 'spread out'. It controls the kernel width [Rumpf et al., 2010]. To illustrate the relations between the histogram and the kernel method, a univariate (1D) kernel density estimate is made based on the latitude data used for the histograms in Fig. 4.1 and Fig. 4.2. The result is shown in Fig. 4.4. The y-axis now displays an estimate of the probability density function, as opposed to a frequency count in the histogram y-axis. The curve in Fig. 4.4 is obviously a scaled and smoothed version of the histogram.

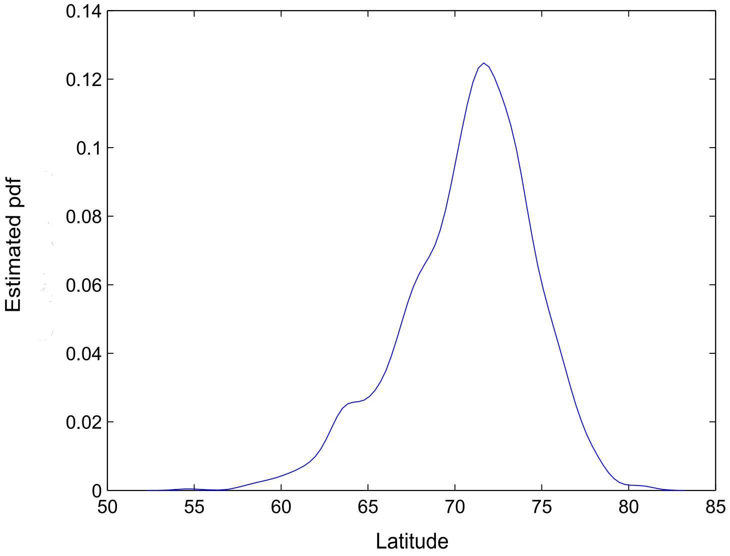


Figure 4.4: 1D kernel density estimate for latitude data. Gaussian kernel function and bandwidth $h = 0.7366$ (calculated as a function of the number of latitude points) is used.

The estimator above (Eq. 4.3) only applies for one-dimensional data. For estimation of multivariate densities, the d-dimensional random vector $\mathbf{X} = (X_1, \ldots, X_d)^T$ is considered. $X_1, \ldots, X_d$ are one-dimensional random variables. A sample of size n in this setting means that there are n observations for each of the d random variables, $X_1, \ldots, X_d$. The ith observation of each of the d random variables is the vector $\mathbf{X}_i$ [Härdle et al., 1995]:

$$\mathbf{X}_i = \begin{pmatrix} X_{i1} \\ \vdots \\ X_{id} \end{pmatrix}, \qquad i = 1, \ldots, n. \tag{4.4}$$

Estimating the probability density of $\mathbf{X} = (X_1, \ldots, X_d)^T$, is the equivalent of the joint pdf f of the random variables $X_1, \ldots, X_d$ [Härdle et al., 1995]

$$f(\mathbf{x}) = f(x_1, \ldots, x_d). \tag{4.5}$$

The kernel density estimator can thus be adapted to the d-dimensional case by writing

$$\begin{aligned}\widehat{f}_h(\mathbf{x}) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h^d}K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h}\right)\\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h^d}K\left(\frac{x_1-X_{i1}}{h},\dots,\frac{x_d-X_{id}}{h}\right),\end{aligned} \tag{4.6}$$

where K is a multivariate kernel function operating on d arguments, and the bandwidth h is assumed to be the same for each component. Not assuming equal bandwidth results in a vector of bandwidths $\mathbf{h} = (h_1, ..., h_d)^T$ and the estimator becomes [Härdle et al., 1995]

$$\widehat{f}_{\mathbf{h}}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_1\dots h_d}K\left(\frac{x_1-X_{i1}}{h_1},\dots,\frac{x_d-X_{id}}{h_d}\right). \tag{4.7}$$

Different bandwidths will be used if the spread of the data points is much larger in one of the coordinate directions than the others [Silverman, 1986].

The analysis in this thesis will involve two-dimensional data, where each data point consists of a latitude and a longitude representing the current position of a polar low. This will thus correspond to the kernel density estimator for $d = 2$. Considering the two-dimensional case where $\mathbf{X} = (X_1, X_2)^T$ the estimator becomes [Härdle et al., 1995]

$$\begin{aligned}\widehat{f}_{\mathbf{h}}(\mathbf{x}) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_1}\frac{1}{h_2}K\left(\frac{x_1-X_{i1}}{h_1},\frac{x_2-X_{i2}}{h_2}\right)\\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_1}\frac{1}{h_2}K\left(\frac{x_1-X_{i1}}{h_1}\right)K\left(\frac{x_2-X_{i2}}{h_2}\right).\end{aligned} \tag{4.8}$$

Each of the n observations are of the form (X_{i1}, X_{i2}). The two components represent the value the random variables X_1 and X_2 takes on the ith observation [Härdle et al., 1995].

Common types of kernel functions include uniform (box), triangle, Epanechnikov, Gaussian (standard normal), quartic (biweight), tricube (triweight) and cosine [Haikun et al., 2009]. Plots of the four first are presented in Fig. 4.5

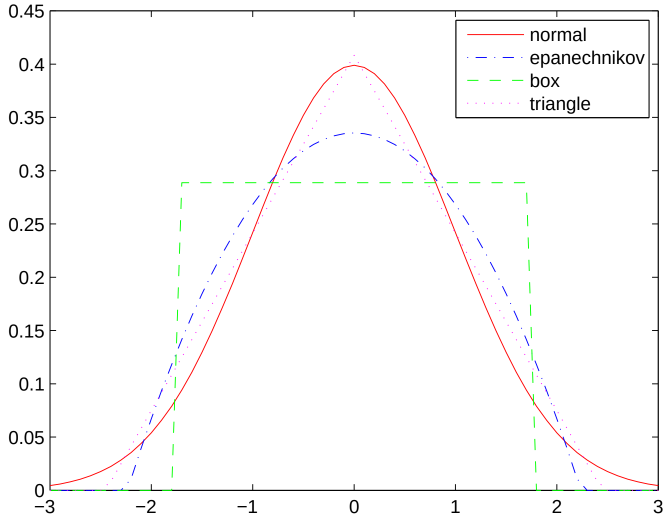


Figure 4.5: Plot of the Gaussian (normal), Epanechnikov, uniform (box) and triangle kernel functions.

Of the kernel functions displayed in Fig. 4.5 it can be noted that every kernel function except the Gaussian represent bounded functions with finite support. Only the Gaussian kernel function has infinite support. Practical implications of this will be discussed shortly, and shown in Sec. 5.1.1. Related to practical use, the Epanechnikov and Gaussian kernel are perhaps the most popular [Rumpf et al., 2010]. All of the mentioned kernel functions are listed in the table below (obtained from Härdle et al. [1995])

Kernel	$K(u)$
Uniform	$\frac{1}{2}\,\mathrm{I}(\lvert u\rvert \le 1)$
Triangle	$(1-\lvert u\rvert)\,\mathrm{I}(\lvert u\rvert \le 1)$
Epanechnikov	$\frac{3}{4}(1-u^2)\,\mathrm{I}(\lvert u\rvert \le 1)$
Quartic (Biweight)	$\frac{15}{16}(1-u^2)^2\,\mathrm{I}(\lvert u\rvert \le 1)$
Triweight	$\frac{35}{32}(1-u^2)^3\,\mathrm{I}(\lvert u\rvert \le 1)$
Gaussian	$\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}u^2)$
Cosine	$\frac{\pi}{4}\cos(\frac{\pi}{2}u)\,\mathrm{I}(\lvert u\rvert \le 1)$

where

$$K_h(\bullet) = \frac{1}{h}K(\bullet/h). \tag{4.9}$$

$K(\bullet)$ is a kernel function like one of those given in the table, and h is the bandwidth. The indicator function I denotes which points the kernel function should apply for. If the point is outside of the given interval, the kernel function takes the value zero [Härdle et al., 1995]:

$$\mathrm{I}(|u|\leq 1) = \begin{cases} 1; & |u|\leq 1 \\ 0; & |u|> 1 \end{cases}$$

The two-dimensional kernel density estimator in Eq. (4.8) with K taken as the Gaussian kernel function becomes [Härdle et al., 1995]

$$\begin{aligned} \widehat{f}_{\mathbf{h}}(\mathbf{x}) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_1}\frac{1}{h_2}K\left(\frac{x_1-X_{i1}}{h_1}\right)K\left(\frac{x_2-X_{i2}}{h_2}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_1}\frac{1}{h_2}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x_1-X_{i1}}{h_1}\right)^2\right\}\cdot\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x_2-X_{i2}}{h_2}\right)^2\right\} \\ &= \frac{1}{2\pi n h_1 h_2}\sum_{i=1}^{n}\exp\left\{-\frac{1}{2}\left(\left(\frac{x_1-X_{i1}}{h_1}\right)^2+\left(\frac{x_2-X_{i2}}{h_2}\right)^2\right)\right\}. \end{aligned} \tag{4.10}$$

Assuming equal bandwidths for each component, $h_1 = h_2 = h$, Eq. (4.10) is rewritten to

$$\widehat{f}_{\mathbf{h}}(\mathbf{x}) = \frac{1}{2\pi n h^2}\sum_{i=1}^{n}\exp\left(-\frac{(x_1-X_{i1})^2+(x_2-X_{i2})^2}{2h^2}\right). \tag{4.11}$$

The two-dimensional kernel density estimator in Eq. (4.8) with K taken as the Epanechnikov kernel becomes [Härdle et al., 1995]

$$\begin{aligned} \widehat{f}_{\mathbf{h}}(\mathbf{x}) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_1}\frac{1}{h_2}K\left(\frac{x_1-X_{i1}}{h_1}\right)K\left(\frac{x_2-X_{i2}}{h_2}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_1}\frac{1}{h_2}\cdot\frac{3}{4}\left\{1-\left(\frac{x_1-X_{i1}}{h_1}\right)^2\right\}\mathrm{I}\left(\left|\frac{x_1-X_{i1}}{h_1}\right|\leq 1\right) \\ &\cdot\frac{3}{4}\left\{1-\left(\frac{x_2-X_{i2}}{h_2}\right)^2\right\}\mathrm{I}\left(\left|\frac{x_2-X_{i2}}{h_2}\right|\leq 1\right). \end{aligned} \tag{4.12}$$

As can be seen from Eq. (4.12) observation i will contribute to the sum only if X_{i1} falls into the interval $[x_1-h_1, x_1+h_1)$ and X_{i2} falls into the interval $[x_2-h_2, x_2+h_2)$. If one or both of the components fall outside of the(ir) respective intervals then the corresponding indicator function(s) takes the value 0 and the the observation is not included in the frequency count [Härdle et al., 1995]. Concerning the density analysis in this thesis, the Gaussian and the Epanechnikov kernels will mainly be considered.

Finally, it is useful to emphasize the practical difference between kernels with finite support and those without. For the univariate case, note that for kernels with finite support, for instance the Epanechnikov kernel, the observations located in a cube around x are used to estimate the density at the point x. For the multivariate case, observations from a circle around x are used to estimate the probability density function at x. The points outside of the cube or the circle are excluded from the calculations. The Gaussian kernel, which does not have finite support, never takes the value zero. This means that every observations is considered when estimating the density function at x. However, observations close to x are weighted higher than those further away [Härdle et al., 1995]. A practical result of finite support kernels are thus density estimates of higher degree of detail, while the infinite support Gaussian kernel tend to result in more smoothed estimates. This will be seen in practice in Sec. 5.1.1.

4.1.1 Optimal bandwidth and statistical properties

Both theoretical and practical research suggests that the choice of kernel function does not contribute significantly in terms of the quality of the kernel estimator [Haikun et al., 2009; Rumpf et al., 2010]. Choice of bandwidth however is usually considered more important. Too large bandwidth h will result in too much bias and an “oversmoothing”. Too small values of h will cause too much variance and a “wiggly” and too rough appearance of the estimate (“undersmoothing”). Both under and oversmoothing can substantially reduce the precision of the estimate [Haikun et al., 2009].

The optimal bandwidth is found by calculation using expressions for estimation accuracy. The MISE (Mean Integrated Square Error) provides a global measure of estimation accuracy between the estimated density and the true density. The idea is to choose h so that the MISE becomes as small as possible. MISE is the integrated expression of the MSE (Mean Squared Error), which is the local measure of estimation accuracy, the discrepancy of the density estimator $\hat{f}$ from the true density f, when considering estimation at a single point. MSE is defined by

$$\mathrm{MSE}_x(\hat{f}) = E\{\hat{f}(x) - f(x)\}^2, \tag{4.13}$$

which is equal to the sum of the squared bias and the variance at x [Silverman, 1986]

$$\mathrm{MSE}_x(\hat{f}) = \{E\hat{f}(x) - f(x)\}^2 + \mathrm{var}\hat{f}(x), \tag{4.14}$$

where the bias and variance of an estimator are denoted by [Silverman, 1986]

$$\begin{aligned} \mathrm{bias}(\hat{f}) &= E\{\hat{f}(x)\} - f(x) \\ \mathrm{var}(\hat{f}) &= E\{\hat{f}^2(x)\} - (E\{\hat{f}(x)\})^2. \end{aligned} \tag{4.15}$$

The bias of the kernel density estimator is given by (derivation not shown here)

$$\mathrm{bias}\{\widehat{f}_h(x)\} = \frac{h^2}{2} f''(x)\mu_2(K) + O(h^2), \qquad \text{as } h \to 0, \tag{4.16}$$

where the variable $s = (u - x)/h$ is used, $\mu_2(K) = \int s^2 K(s)\,\mathrm{d}s$, and the bias is proportional to h^2 [Härdle et al., 1995]. Furthermore, $O(h^2)$ represents terms that converge to zero faster than h^2 as h approaches zero.

The variance is given by (derivation not shown here)

$$\mathrm{var}\{\widehat{f}_h(x)\} = \frac{1}{nh}\|K\|_2^2 f(x) + O\left(\frac{1}{nh}\right), \qquad \text{as } h \to \infty, \tag{4.17}$$

where $\|K\|_2^2$ is shorthand for $\int K^2(s)\,\mathrm{d}s$, the so-called squared L_2 norm of K. Note that the kernel density estimator is nearly proportional to nh^{-1} [Härdle et al., 1995].

Looking at Eq. (4.16) and Eq. (4.17) a trade-off between variance and bias becomes apparent. Preferably, both the bias and the variance should be kept small, but increasing h will lower the variance while raising the bias, and decreasing h will result in the opposite. Minimizing the MSE, i.e., the sum of the squared bias and the variance thus represents a compromise between oversmoothing (large bias, small variance) and undersmoothing (small bias, large variance) [Härdle et al., 1995].

Revisiting Eq. (4.14) for the MSE, Eq.(4.16) and Eq. (4.17) now yield

$$\mathrm{MSE}\{\widehat{f}_h(x)\} = \frac{h^4}{4} f''(x)^2 \mu_2(K)^2 + \frac{1}{nh}\|K\|_2^2 f(x) + O(h^4) + O\left(\frac{1}{nh}\right). \tag{4.18}$$

From Eq. (4.18) it can be seen that the MSE of the kernel density estimator goes to zero as $h \to 0$ and $nh \to \infty$, confirming that the kernel density estimator is consistent. However, the expression depends on f and f'', which is unfortunate because both of these functions are unknown in practice. After deriving the expression for h that minimizes the MSE, $h_{opt}(x)$, these expressions are still included and a need for suitable substitutes arise. Also, $h_{opt}(x)$ depends on x, and therefore only represents a local bandwidth [Härdle et al., 1995]. As mentioned earlier, by using the MISE instead of the MSE, the estimation accuracy is expanded to a global measure. The MISE optimal bandwidth is therefore rather focused on [Härdle et al., 1995].

The MISE for the kernel density estimator is given by

$$\begin{aligned}
\mathrm{MISE}(\widehat{f}_h) &= \int \mathrm{MSE}\{\widehat{f}_h(x)\}\,\mathrm{d}x \\
&= \frac{1}{nh}\|K\|_2^2 \int f(x)\,\mathrm{d}x + \frac{h^4}{4}\{\mu_2(K)\}^2 \int \{f''(x)\}^2\,\mathrm{d}x + O\left(\frac{1}{nh}\right) + O(h^4) \\
&= \frac{1}{nh}\|K\|_2^2 + \frac{h^4}{4}\{\mu_2(K)\}^2 \|f''\|_2^2 + O\left(\frac{1}{nh}\right) + O(h^4) \\
&\text{as } h \to 0, nh \to \infty.
\end{aligned} \tag{4.19}$$

If higher order terms (the last two terms) are ignored the formula for the AMISE (Approximate Mean Integrated Squared Error) can be given as

$$\mathrm{AMISE}(\widehat{f}_h) = \frac{1}{nh}\|K\|_2^2 + \frac{h^4}{4}\{\mu_2(K)\}^2 \|f''\|_2^2. \tag{4.20}$$

Differentiating the AMISE with respect to h and solving for the first order of h gives the AMISE-optimal bandwidth

$$h_{opt} = \left(\frac{\|K\|_2^2}{\|f''\|_2^2 \{\mu_2(K)\}^2 n} \right)^{1/5}, \tag{4.21}$$

or, written out [Härdle et al., 1995]

$$h_{opt} = \left\{ \int s^2 K(s)\,\mathrm{d}s \right\}^{-2/5} \left\{ \int K^2(s)\,\mathrm{d}s \right\}^{1/5} \left\{ \int f''(x)^2\,\mathrm{d}x \right\}^{-1/5} n^{-1/5}. \tag{4.22}$$

Obviously, this result is still rather disappointing, since h_{opt} still depends on the unknown density which is being estimated. However, h_{opt} is used to investigate the statistical properties of the kernel density estimator a bit further. The expression for h_{opt} is inserted into Eq. (4.20) which yields

$$\mathrm{AMISE}(\widehat{f}_{h_{opt}}) = \frac{5}{4} \left(\|K\|_2^2 \right)^{4/5} \left(\mu_2(K) \|f''\|_2 \right)^{2/5}, \tag{4.23}$$

where it is indicated that the terms preceding $n^{-4/5}$ are constant with respect to n. If the sample size then gets larger and larger, the AMISE is converging at the rate $n^{-4/5}$ [Härdle et al., 1995].

4.1.2 Smoothing parameter selection

Several methods for selecting the smoothing parameter, or the bandwidth, have been developed. A natural method for choosing the bandwidth is the subjective choice, where several curves of varying bandwidths are plotted out. The curves are then compared to the a priori expectations of the density, and the estimate which coincides the most is chosen. In many cases, this method is sufficient, and the examination of the different plots will give a good insight in the data, by for instance offering several possible explanations for its distribution [Silverman, 1986].

Other methods for bandwidth selection include more automatic approaches. Two of the methods most frequently used are the plug-in method and the method of cross-validation. The focus here will be on one representative of each of the methods, the "quick & dirty" plug-in method introduced by Silverman [1986], and the least squares cross-validation method [Härdle et al., 1995].

Silverman's rule of thumb Plug-in methods are based on the principle of replacing an unknown parameter in an expression with an estimate [Härdle et al., 1995]. Eq. (4.21) for the optimal bandwidth can be used as an example in this case. The expression involves the unknown quantity $\|f''\|_2^2$. It is then assumed that the unknown density f belongs to the normal distribution with mean μ and variance σ^2, and where ϕ is set to be the pdf of the standard normal distribution

$$\begin{aligned} \|f''\|_2^2 &= \int f''(x)^2 \,\mathrm{d}x \\ &= \sigma^{-5} \int \phi''(x)^2 \,\mathrm{d}x \\ &= \sigma^{-5} \frac{3}{8\sqrt{\pi}} \approx 0.212\sigma^{-5}. \end{aligned} \tag{4.24}$$

Substituting this in Eq. (4.21) for the optimal bandwidth and setting K as the Gaussian kernel generates the following "Silverman's rule-of-thumb" bandwidth $\widehat{h}_{rot}$ after calculation [Silverman, 1986]

$$\widehat{h}_{rot} = \left(\frac{4}{3}\right)^{1/5} \sigma n^{-1/5} \approx 1.06\sigma n^{-1/5}. \tag{4.25}$$

In practice it is obviously not known if the data are in fact normally distributed. If they are, $\widehat{h}_{rot}$ will provide the optimal bandwidth. If not, $\widehat{h}_{rot}$ gives a bandwidth close to the optimum of the distribution if the data are not too different from the normal distribution, which is referred to as the reference distribution. This method provides fairly good result for unimodal density distributions, but has a tendency to oversmooth and hide features for multimodal data. This is obviously regarded as a drawback for the use of rule-of-thumb bandwidths [Härdle et al., 1995].

In one or two dimensions it is simpler to choose an adequate bandwidth by using the subjective method mentioned above, where several plots of different bandwidths are made and the best one is chosen. Even though, for analytical purposes, the bandwidth selectors should be applied to the data in this thesis, and then a generalization to the multivariate case is required.

Extension to the multivariate case For the multivariate case, the bandwidth h and the variance σ are extended to diagonal matrices, the bandwidth matrix $\mathbf{H} = \operatorname{diag}(h_1, \ldots, h_d)$ and the covariance matrix $\mathbf{\Sigma} = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_d^2)$, respectively. The rule-of-thumb for general $\mathbf{H}$ and $\mathbf{\Sigma}$ is derived as a generalization of Scott's rule (which is derived from the Silverman rule-of-thumb, see Scott [1992])

$$\widehat{h}_j = n^{-1/(d+4)} \widehat{\sigma}_j. \tag{4.26}$$

The bandwidth matrix $\mathbf{H}$ is chosen to be proportional to $\mathbf{\Sigma}^{1/2}$, and the multivariate generalization of Scott's rule is thus [Härdle et al., 1995]

$$\hat{\mathbf{H}} = n^{-1/(d+4)} \widehat{\mathbf{\Sigma}}^{1/2}. \tag{4.27}$$

Cross validation The focus here will be on least squares cross-validation, which is a completely automatic method for bandwidth selection. The method is based on a very simple idea, and was first suggested by Rudemo [1982] and Bowman [1984]. The ISE (Integrated Square Error), another version of the error measure between $\widehat{f}$ and f is now considered:

$$\mathrm{ISE}(h) = \mathrm{ISE}\{\widehat{f}_h\} = \int \{\widehat{f}_h - f\}(x)^2 \,\mathrm{d}x = \int \{\widehat{f}_h(x) - f(x)\}^2 \,\mathrm{d}x. \tag{4.28}$$

By comparing Eq. (4.28) to the expression for MISE (Eq. (4.19)) it can be established that the MISE is the expected value of the ISE (as the name also implies). The goal in this bandwidth selection method is to choose an h so that the value for ISE is minimized [Härdle et al., 1995]. Eq. (4.28) is rewritten

$$\mathrm{ISE}(h) = \int \widehat{f}_h^2(x)\,\mathrm{d}x - 2\int \{\widehat{f}_h f\}(x)\,\mathrm{d}x + \int f^2(x)\,\mathrm{d}x. \tag{4.29}$$

The last term does not depend on h and can be ignored in relation to the minimization. The first term can be calculated from the data. The middle term however, depends on h and involves the unknown quantity f. It can be established that $\int \{\widehat{f}_h f\}(x)\,\mathrm{d}x$ is the expected value of $\widehat{f}_h(X)$, where the expectation is calculated with respect to an independent random variable X. This expected value can be estimated by

$$\widehat{E\{\widehat{f}_h(X)\}} = \frac{1}{n}\sum_{i=1}^{n} \widehat{f}_{h,-i}(X_i), \tag{4.30}$$

where

$$\begin{aligned} \widehat{f}_{h,-i}(x) &= \frac{1}{n-1}\sum_{j=1, i\neq j}^{n} K_h(x - X_j) \\ &= \frac{1}{(n-1)h}\sum_{j=1, i\neq j}^{n} K\left(\frac{x - X_j}{h}\right). \end{aligned} \tag{4.31}$$

Here, $\widehat{f}_{h,-i}(x)$ is the leave-one-out estimator, meaning that the ith observation is not used in the calculation of $\widehat{f}_{h,-i}(X_i)$. The leave-one-out estimator ensures that the observations used for calculating $\widehat{f}_{h,-i}(\bullet)$ are independent of X_i, the observation at which $\widehat{f}_{h,-i}(x)$ is estimated in Eq. (4.30) [Härdle et al., 1995].

The criterion function which should be minimized with respect to h is thus

$$\mathrm{ISE}(h) = \int \widehat{f}_h^2(x)\,\mathrm{d}x - 2E\{\widehat{f}_h(X)\} + \int f^2(x)\,\mathrm{d}x. \tag{4.32}$$

As mentioned above, the last term does not depend on h and can therefore be ignored. It is brought to the left side of the equation

$$\mathrm{ISE}(h) - \int f^2(x)\,\mathrm{d}x = \int \widehat{f}_h^2(x)\,\mathrm{d}x - 2E\{\widehat{f}_h(X)\}. \tag{4.33}$$

Then the expressions for $E\{\widehat{f}_h(X)\}$ (Eq. (4.30) and Eq. (4.31)) above are plugged in to generate the *cross-validation criterion*

$$\mathrm{CV}(h) = \int \widehat{f}_h^2(x)\,\mathrm{d}x - \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1,i\neq j}^{n} K_h(X_i - X_j). \tag{4.34}$$

Now it only remains to express $\int \widehat{f}_h^2(x)\,\mathrm{d}x$ using sums rather than intergrals. It can be shown that [Härdle, 1991, p.230]

$$\int \widehat{f}_h^2(x)\,\mathrm{d}x = \frac{1}{n^2 h} \sum_{i=1}^{n} \sum_{j=1}^{n} K * K\left(\frac{X_j - X_i}{h}\right), \tag{4.35}$$

where $K * K(u)$ is the convolution of K, i.e., $K * K(u) = \int K(u-v)K(v)\,\mathrm{d}v$. Inserting this expression into Eq. (4.34) yields the following function to minimize with respect to h [Härdle et al., 1995]

$$\mathrm{CV}(h) = \frac{1}{n^2 h} \sum_{i=1}^{n} \sum_{j=1}^{n} K * K\left(\frac{X_j - X_i}{h}\right) - \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1,i\neq j}^{n} K_h(X_i - X_j). \tag{4.36}$$

As supposed to the rule-of-thumb the least square cross-validation provides a bandwidth selection method where the need for assumptions regarding density distribution is eliminated [Haikun et al., 2009].

Extension to the multivariate case The multivariate extension of the cross-validation method is straightforward and yields the multivariate cross-validation criterion as a perfect generalization of the CV in the one-dimensional case:

$$\begin{aligned} CV(\mathbf{H}) = {} & \frac{1}{n^2 \det(\mathbf{H})} \sum_{i=1}^{n} \sum_{j=1}^{n} K * K\{\mathbf{H}^{-1}(\mathbf{X}_j - \mathbf{X}_i)\} \\ & - \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1,i\neq j}^{n} K_{\mathbf{H}}(\mathbf{X}_j - \mathbf{X}_i). \end{aligned} \tag{4.37}$$

A few methods of bandwidth selection have now been presented, but several more exist. In this thesis the subjective method, the rule-of thumb, and the least squares cross-validation will be applied to the polar low trajectory data set. The results will be presented and discussed in Chp. 5.

4.2 Curve clustering analysis

Clustering is a form of classification of data where one want to investigate whether a collection of data can be divided into groups, or clusters, based on similarities. Clustering is an example of an *unsupervised* method of pattern recognition, meaning that the training data used in the classification does not correspond to a set of target values, e.g., class or group labels [Bishop, 2007].

Clustering analysis applied in this thesis will thus represent a method of dividing the polar low trajectories of the data set into groups according to similar qualities. These qualities can for instance be trajectory length, propagation direction and degree of bend of the trajectory curve. By using clustering techniques, large sets of trajectories can be decomposed into subgroups of individual spatio-temporal characteristics. These groups can then be analysed further, e.g., in terms of relationship to larger-scale atmospheric circulation. This was part of the motivation for the curve clustering performed in Gaffney et al. [2007]. The clustering analysis in this thesis is adopted from the methodology used in Camargo et al. [2007] and Gaffney et al. [2007], which again is based on methods from Gaffney [2004]. Camargo et al. [2007] and Gaffney et al. [2007] used a new probabilistic clustering technique, based on a regression mixture model to describe tropical cyclone (TC) tracks in the western North Pacific and extratropical cyclone (ETC) tracks over the North Atlantic, respectively.

Camargo et al. [2007] assumed that each trajectory, or cyclone track, was generated by one of K different regression models, where each of them had its own shape parameters. The clustering was divided into two parts: (i) Given the TC tracks, learn the parameters of all K models, and (ii) determine which of the K models are most likely of generating each track. Each track could then be assigned to the mixture component, i.e., the cluster, with the highest probability of generating the track, given the model. This cluster would thus have the highest posterior probability given the track. The model parameters were learned by defining an expectation-maximization (EM) algorithm, which is fairly uncomplicated in terms of implementation and use. Below follows the description of this method, obtained from the methodology section in Gaffney et al. [2007] and the appendix in Camargo et al. [2007].

4.2.1 Using mixtures of regression models for cluster analysis of trajectories

In the standard mixture model framework, the probability density function (pdf) for a d-dimensional vector $\mathbf{x}$, as a function of model parameters ϕ, is modelled by the mixture density

$$p(\mathbf{x} \mid \phi) = \sum_{k}^{K} \alpha_k p_k(\mathbf{x} \mid \theta_k), \tag{4.38}$$

where $\phi = \{\alpha_1, \ldots, \alpha_K; \theta_1, \ldots, \theta_K\}$, α_k is the kth component weight, and p_k is the kth component density with parameter vector θ_k; e.g., K Gaussian densities each with a d-

dimensional mean vector and a $d \times d$ covariance matrix [Gaffney et al., 2007]. The mixture weights α_k sum to one and are nonnegative. This way, a finite mixture model is a pdf which consists of a weighted average of component density functions [McLachlan and Peel, 2004].

The mixture model framework is applied to data clustering in the following manner: A data set of n vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ which is assumed to be a random sample from the underlying mixture model is considered. Each of the data vectors $\mathbf{x}_i$ are generated by one of the K components, but the generating component is not known. The parameters for each density component $p_k(\mathbf{x} \mid \theta_k)$, and the corresponding weights α_k, can be estimated from the data by using the Expectation-Maximization (EM) algorithm, which is a common used technique for maximum-likelihood parameter estimation with mixture models [McLachlan and Krishnan, 2007]. Relating to clustering, the estimated component models, $p_k(\mathbf{x} \mid \theta_k)$, $1 \leq k \leq K$, are represented as K clusters, where each cluster is defined by a pdf in the d-dimensional input space $\mathbf{x}$ [Gaffney et al., 2007].

Furthermore, by using Bayes rule and Eq. (4.38), the probability that $\mathbf{x}$ was generated by the kth cluster (or component) can be calculated by using Eq. (4.39). In Eq. (4.39), the membership probabilities, w_{ik}, represent the a posteriori uncertainty about which cluster each data vector $\mathbf{x}_i$ (denoted $\mathbf{z}_i$ in Eq. (4.39)) originated from:

$$\mathrm{w}_{ik} = \frac{\alpha_k f_k(\mathbf{z}_i \mid \mathbf{T}_i \boldsymbol{\beta}_k, \Sigma_k)}{\sum_j^K \alpha_j f_j(\mathbf{z}_i \mid \mathbf{T}_i \boldsymbol{\beta}_j, \Sigma_j)}. \tag{4.39}$$

The new parameters in Eq. (4.39) are related to the clustering of trajectories, and will be explained shortly. These w_{ik} are the estimated cluster memberships formed by the ratio of the likelihood of trajectory i under cluster k, to the sum-total likelihood of trajectory i under all clusters. They give the probabilities that the ith trajectory was generated from cluster k. The cluster memberships are binary values, meaning that the ith trajectory was either generated by the kth cluster, or it was not [Gaffney et al., 2007]. The clustering of the original data $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ can then be performed by assigning each vector $\mathbf{x}_i$ to the cluster f_k with the highest membership probability, i.e., the cluster it was most likely generated from [Gaffney et al., 2007].

To further relate the theory above to the clustering of trajectories, we consider a set (e.g., a cell array) of n cyclone trajectories $\mathbf{Z} = \{\mathbf{z}_1, ..., \mathbf{z}_n\}$. Each element $\mathbf{z}_i$ is an $n_i \times 2$ matrix containing the latitude and longitude measurements for trajectory i, where n_i denotes the number of measurements. Corresponding to this set is a set of associated measurement times $\mathbf{T} = \{\mathbf{t}_1, ..., \mathbf{t}_n\}$. Each element $\mathbf{t}_i$ is an $n_i \times 1$ vector of discrete time indices $\{0, 1, ..., n_i - 1\}$. This cell is optional to use, and it can also hold other independent variables which are of interest [Camargo et al., 2007].

Longitude and latitude is modelled with a polynomial regression model of order $p = 2$, with time $\mathbf{t}_i$ as the independent variable. Under the assumption that trajectory i was generated by cluster k Camargo et al. [2007] had

$$\mathbf{z}_i = \mathbf{T}_i \boldsymbol{\beta}_k + \epsilon_i, \qquad \epsilon_i \sim N\left(0, \boldsymbol{\Sigma}_k\right), \tag{4.40}$$

where $\mathbf{T}_i$ is the $n_i \times (p+1)$ Vandermonde regression matrix associated with the vector $\mathbf{t}_i$; $\boldsymbol{\beta}_k$ is a $(p+1) \times 2$ matrix of regression coefficients for cluster k, containing the longitude coefficients in the first column and the latitude coefficients in the second. Furthermore, ϵ_i is an $n_i \times 2$ matrix of multivariate Gaussian noise (normal errors), with zero mean and a 2×2 covariance matrix $\mathbf{\Sigma}_k$. The assumption of a normal noise model is computationally straightforward, and a common choice for additive noise in regression models [Gaffney et al., 2007]. The covariance matrix $\mathbf{\Sigma}_k$ contains diagonal elements σ_{k1}^2 and σ_{k2}^2, the noise variances for each longitude and latitude observation, respectively. The cross covariance is set to 0.

The formula for the conditional density (conditioned on membership in the kth cluster) for the ith cyclone, is thus given by

$$\begin{aligned} p(\mathbf{z}_i \mid \mathbf{t}_i, \theta_k) &= f(\mathbf{z}_i \mid \mathbf{T}_i\boldsymbol{\beta}_k, \mathbf{\Sigma}_k) \\ &= (2\pi)^{-n_i}|\mathbf{\Sigma}_k|^{-n_i/2}\exp\left\{-\frac{1}{2}\mathrm{tr}[(\mathbf{z}_i - \mathbf{T}_i\boldsymbol{\beta}_k)\mathbf{\Sigma}_k^{-1}(\mathbf{z}_i - \mathbf{T}_i\boldsymbol{\beta}_k)']\right\}, \end{aligned} \tag{4.41}$$

where the parameter set $\theta_k = \{\boldsymbol{\beta}_k, \mathbf{\Sigma}_k\}$ contains the regression coefficients $\boldsymbol{\beta}_k$ and the noise covariance matrix $\mathbf{\Sigma}_k$ [Gaffney et al., 2007].

We can derive regression mixtures for the cyclones by substituting the unconditional density components $p_k(\mathbf{x} \mid \theta_k)$ in Eq. (4.38) with the conditional regression density components $p_k(\mathbf{z} \mid \mathbf{t}, \theta_k)$, defined in Eq. (4.41). This results in the following unconditional (on k) regression mixture model with K clusters:

$$p(\mathbf{z}_i \mid \mathbf{t}_i, \phi) = \sum_k^K \alpha_k p_k(\mathbf{z}_i \mid \mathbf{t}_i, \theta_k) = \sum_k^K \alpha_k f_k(\mathbf{z}_i \mid \mathbf{T}_i\boldsymbol{\beta}_k, \mathbf{\Sigma}_k), \tag{4.42}$$

where α_k is the probability of a randomly selected track belonging to cluster k, where $\Sigma_k \alpha_k = 1$, and ϕ represents the overall set of mixture parameters ($\boldsymbol{\beta}_k, \Sigma_k$ and $\alpha_k; 1 \leq k \leq K$) [Camargo et al., 2007]. Each trajectory is assumed to be generated by one of the K different regression models, where each model has its individual "shape" parameters $\theta_k = \{\boldsymbol{\beta}_k, \Sigma_k\}$ [Gaffney et al., 2007].

Recalling the sets $\mathbf{Z}$ and $\mathbf{T}$, the full probability density of $\mathbf{Z}$ given $\mathbf{T}$ is

$$p(\mathbf{Z} \mid \mathbf{T}, \phi) = \prod_i^n \sum_k^K \alpha_k f_k(\mathbf{z}_i \mid \mathbf{T}_i\boldsymbol{\beta}_k, \mathbf{\Sigma}_k), \tag{4.43}$$

which is also known as the conditional likelihood [Gaffney et al., 2007].

The product form of Eq. (4.43) arises due to the assumption of conditional independence of the $\mathbf{z}_i$'s given both the $\mathbf{t}_i$'s and the mixture model, meaning that cyclones are assumed not to influence each other. From a meteorological point of view regarding the polar low events this is not necessarily true, because the dual and multiple polar lows often occur at the same time. However, when performing clustering analysis of trajectories based on their shape, this assumption is acceptable [Gaffney et al., 2007].

As mentioned above, to perform the clustering, the likelihood expression in Eq. (4.43) is maximized to determine estimates of the parameters ϕ given data. The expectation-maximization (EM) algorithm represents an iterative method of finding local maxima of the likelihood. It can thus be determined for each trajectory, given a learned model, which of the K models are most likely to have generated it, hence to which model (cluster) it should be assigned to [Camargo et al., 2007]. More detail of this iterative maximum likelihood procedure can be found in Appendix A in Gaffney et al. [2007].

In this thesis, the method explained above was applied to polar low trajectories instead of tropical or extratropical cyclone trajectories. The clustering was performed using a curve clustering toolbox in Matlab called CCToolbox 1.0, made by S. Gaffney.[1]

---

[1] CCToolbox documentation: `http://www.datalab.uci.edu/resources/CCT/doc/`

Chapter 5

Results

5.1 Kernel density estimation

The dataset of polar low trajectories provides an array consisting of a total of 2270 data points (for 286 trajectories), where each point represents a coordinate (latitude, longitude) of a certain position of a polar low trajectory. In this array, the trajectories are aligned in chronological order, from the first trajectory in 1999 to the last in 2013. This set will be referred to as the ‘complete trajectory data’ from here on. It will form the basis for the kernel density estimation processes. Fig. 5.1 shows a scatter plot of all the 2270 data points on a map background.

It is not very easy to draw conclusions from this plot regarding the density of polar low occurrences, but some areas seem to be more exposed than others. By eye measure, the concentration seem to be the highest in an area between 65 °N - 72 °N and 0 °E - 20 °E. This actually correlates well with Fig. 2.1, obtained from the climatological study by Noer et al. [2011] which showed the locations of early development for polar lows during the years 2000 - 2009. An especially high concentration can be observed outside the Lofoten area. There may be several reasons for the high exposure in this area, for instance climatological factors favouring formation and development, but also its proximity to the Tromsøflaket region where developed polar lows travel south and southwestwards into the area (see Fig. 2.1). Some maxima also seem to exist in the Barents Sea, but these are less apparent.

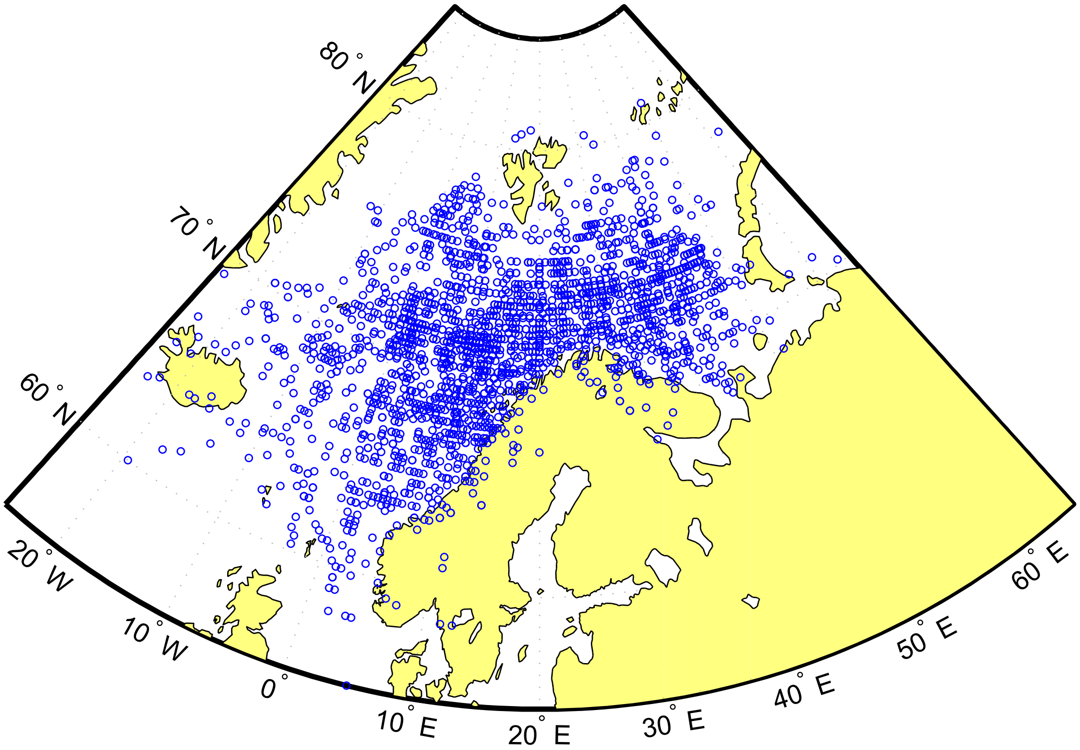
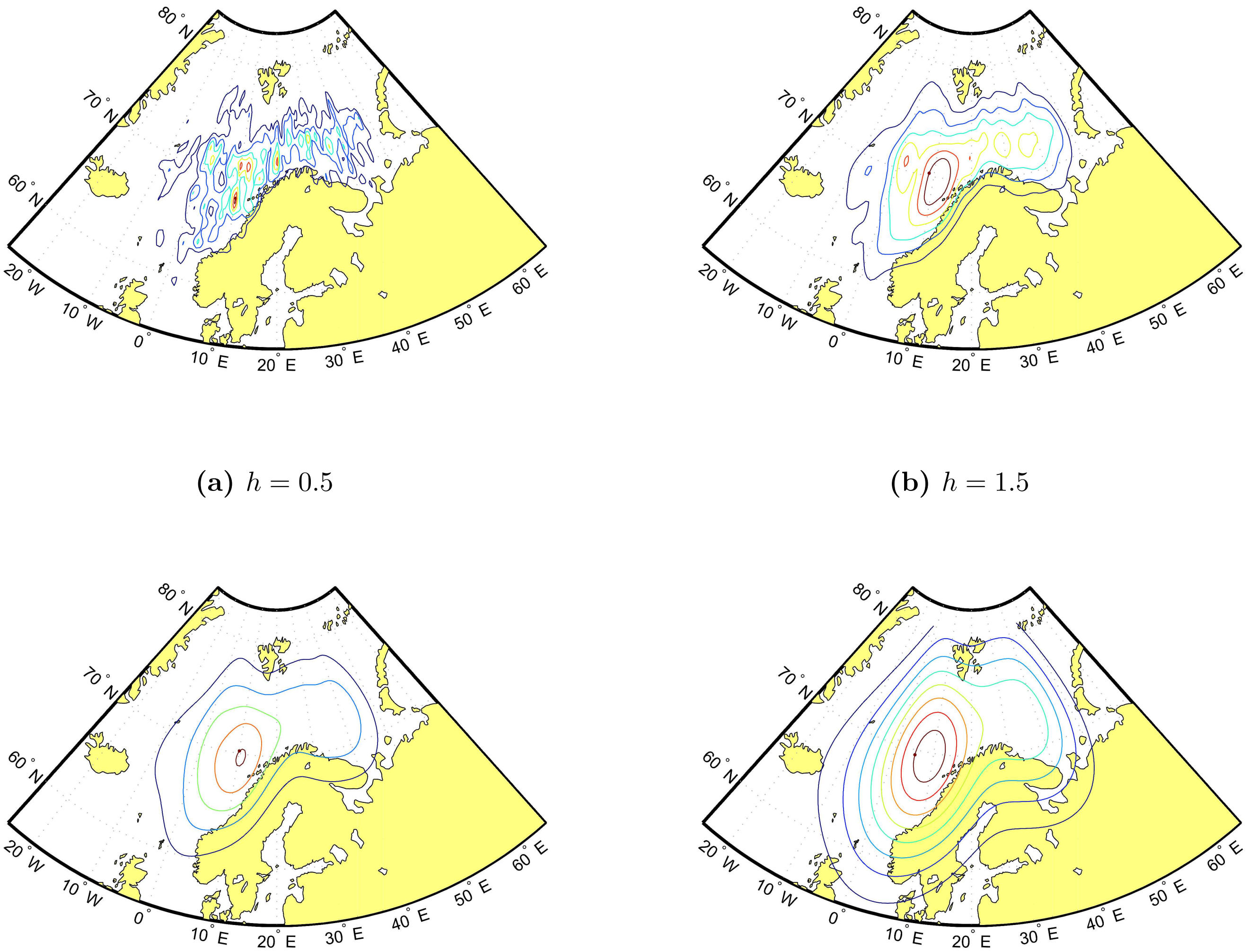


Figure 5.1: Scatter plot of the coordinates of polar low trajectories over the Nordic Seas, winter seasons 1999-2013.

5.1.1 Kernels and bandwidth selection

The concentrations become more clear after the kernel density estimations. Both Gaussian and Epanechnikov kernels were used, and a variety of bandwidths were tested for each, according to the subjective method of bandwidth selection (see Sec. 4.1.2). Fig. 5.2 shows plots of the density estimates of the complete trajectory data with a Gaussian kernel for four different bandwidths.

Looking at the four plots in Fig. 5.2 it can quickly be established that the bandwidths chosen in Fig. 5.2a and Fig. 5.2d are not suitable for these data, as they tend to cause strong under and oversmoothing, respectively. The very small bandwidth in Fig. 5.2a produces “wiggly” density estimates showing a few, very small maxima and the plot is in general difficult to retrieve information from. The opposite happens for bandwidth $h = 5$ (Fig. 5.2d), where the estimates are strongly oversmoothed. Only one quite large maximum is displayed - the data have been estimated as a unimodal distribution. The large bandwidth also causes an erroneous impression that polar lows in many cases have occurred far beyond the coastal boundaries, which is obviously not the case.



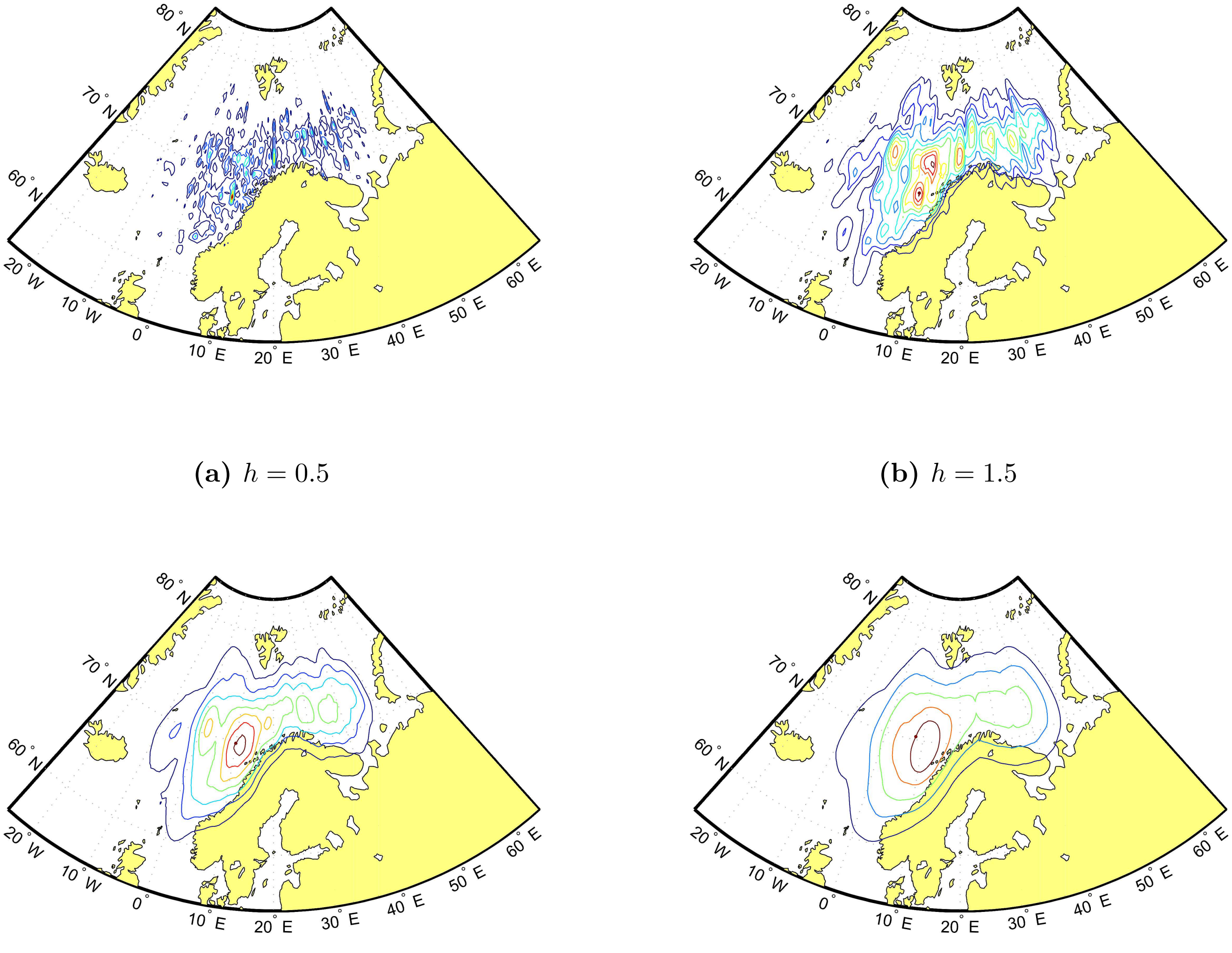
(a) $h = 0.5$ **(b)** $h = 1.5$

(c) $h = 3$ **(d)** $h = 5$

Figure 5.2: Kernel density estimates of the complete trajectory data with Gaussian kernel function for four different bandwidths.

Fig. 5.2c seems more appropriate in terms of bandwidth, but the plot still appears somewhat oversmoothed. Fig. 5.2b represents the plot with the most suitable bandwidth used of the four. The density contours follow the coastal boundaries and topography well and features which were "smoothed away" in the two last figures (Fig. 5.2c and Fig. 5.2d) become apparent. In comparison to Fig. 5.2c and Fig. 5.2d which seemed more and more unimodal, in that there were only one dominant maxima, Fig. 5.2b seems to illustrate a more bimodal distribution. Areas of high densities exist in the Barents Sea, in addition to a large area in the Norwegian Sea outside Lofoten and Vesterålen. This fits well with the observations which were made earlier for the scatter plot (Fig. 5.1).

Using the same collection of bandwidths, the densities were also estimated using the complete trajectory data and an Epanechnikov kernel function. The plots are shown in Fig. 5.3.



(a) $h = 0.5$

(b) $h = 1.5$

(c) $h = 3$

(d) $h = 5$

Figure 5.3: Kernel density estimates of the complete trajectory data with Epanechnikov kernel function for four different bandwidths.

An overall difference in the plots can immediately be observed. As for the Gaussian kernel, bandwidth $h = 0.5$ (Fig. 5.3a) is obviously way too small, but for the Epanechnikov kernel it undersmooths even more. Fig. 5.3b also illustrates the use of a bandwidth which is too small, but the general features start to appear. Fig. 5.3c is highly comparable to the plot of the densities with a Gaussian kernel and $h = 1.5$ (Fig. 5.2b), and in similarity with this represents the most appropriate choice of bandwidth of the four. The largest bandwidth used (Fig. 5.3d) oversmooths also in this case, but not nearly as much it did for the Gaussian kernel (Fig. 5.2d).

Judging from these results, it seems as the Gaussian kernel requires a lower bandwidth than the Epanechnikov kernel to produce similar estimates. This can be related to the finite support of the Epanechnikov kernel, discussed in Sec. 4.1. Because the Gaussian kernel does not have finite support, it needs smaller bandwidths to focus the weighting (and the degree of detail) in a local area. The Epanechnikov kernel does this automatically, as it excludes the points located outside a local area.

The bandwidth values of $h = 1.5$ and $h = 3$ seem to be in the range of the most suitable bandwidths for Gaussian and Epanechnikov kernel, respectively. Because of its finite support, the Epanechnikov kernel produces more detailed results than the Gaussian for the same bandwidths. The Epanechnikov kernel function will therefore mainly be used from here on out.

The results presented above represent the subjective method for bandwidth selection, and provided a good insight regarding a suitable range of bandwidth value. As stated in Sec. 4.1.2, two other common methods of bandwidth selection are the Silverman's rule-of-thumb and the Least-squares cross-validation. Both of these methods have been implemented to find optimal bandwidth and to compare the findings with those of the subjective method. The complete trajectory data set were used in both implementations.

The bandwidth using plug-in rule-of-thumb was found by generating a simple code of Eq. (4.27) in Matlab. The output bandwidth matrix is reproduced below:

$$\hat{\mathbf{H}} = \begin{bmatrix} 1.0254 & 1.4454 \\ 1.4454 & 4.4462 \end{bmatrix}.$$

The Least-squares cross-validation (LSCV) was carried out by using a built-in function in R (a software environment for statistical computing) created for this purpose. The result is shown below:

$$\hat{\mathbf{H}} = \begin{bmatrix} 1.077566 \times 10^{-11} & 6.154842 \times 10^{-5} \\ 6.154842 \times 10^{-5} & 5.801234 \times 10^{2} \end{bmatrix}.$$

These matrices give the bandwidths for two-dimensional data, where the values of interest, h_1 and h_2, are located in the diagonal. However, because the kernel density estimates in this thesis are carried out consistently using the same bandwidth for both dimensions, these results rather have to be interpreted as a range, where the optimal bandwidth lies somewhere between h_1 and h_2.

Judging by these results, it is safe to say that the plug-in method provides the preferable bandwidth matrix. The values of h_1 and h_2 (1.0254 and 4.4462, respectively) are in the same range as the bandwidths which gave decent results in the subjective method.

The matrix resulting from the least-squares cross-validation gives a very small value for h_1 and a very large value for h_2. These values indicate that Least-squares cross-validation is not a suitable method for bandwidth selection is this case, for these particular data. The reasons for this are not clear. However, one possibility might be the fact that some latitude or longitude values appear more than once in the data set, which may be undesirable associated with leave-one-out methods. Accompanied by the output matrix was a warning message stating that LSCV is not 'well behaved for duplicated data'. Another possible undesirable quality of the data associated with LSCV may be the low level of accuracy, or resolution, of the latitude/longitude measurements, which cause a smaller spread of the data points. For instance, data values which potentially could be expressed as 73.0976 and 71.5674 are expressed as 73 and 71.5 instead. The LSCV was therefore run again after adding a very small and random high-resolution value to each position data point in the set, to ensure that no points were equal, and increase the resolution of points. This significantly changed the output:

$$\hat{\mathbf{H}} = \begin{bmatrix} 5.433913 \times 10^{-5} & 0.003209305 \\ 0.003209305 & 21.232078476 \end{bmatrix}.$$

However, it is a possibility that this change is mostly the result of computational differences, following these new values. Actually, this might as well be the case, as the same process was run again, now by adding a slightly larger value to all the data points:

$$\hat{\mathbf{H}} = \begin{bmatrix} 0.01008589 & 0.01999632 \\ 0.01999632 & 9.86410467 \end{bmatrix}.$$

Based on the two last matrices, it seems as the value of h_1 and h_2 increases and decreases, respectively, with increasing values of data points. Also, the values of h_1 and h_2 are still quite small and large, respectively, compared to those given by the plug-in method, and those deemed suitable from the subjective bandwidth selection.

Based on the observations above, the subjective method is considered the most preferable method for bandwidth selection in this thesis. As stated by Silverman [1986], subjectively choosing the smoothing parameter is probably sufficient, and also desirable if the main purpose of the density estimation is to explore the data to suggest possible models and hypotheses. This can be said to be the purpose in this thesis, and examining several plots of differently smoothed data will give a much better insight into the data than considering only one automatically produced plot [Silverman, 1986]. Also, when using the same bandwidth for both dimensions, the subjective method is far more adjustable and easy to relate to than the automatic ones, in that the automatic methods are constrained to univariate *or* multivariate data. The subjective method with same bandwidth for both dimensions represents a mix between methods for univariate and multivariate data, in that the data used are bivariate, but the bandwidth is given as a scalar, not a vector or a matrix.

5.1.2 Kernel density estimates - General observations

After gaining an insight in suitable bandwidths to apply, some general observations of the polar low densities were made. These results were obtained by kernel density estimation using the complete trajectory data. The results thus indicate high density areas, regarding both polar formation and passing developed lows. To explore some general observations of the kernel density estimates, lets take a closer look at a plot where Epanechnikov kernel and bandwidth $h = 2.6$ are used (Fig. 5.4).

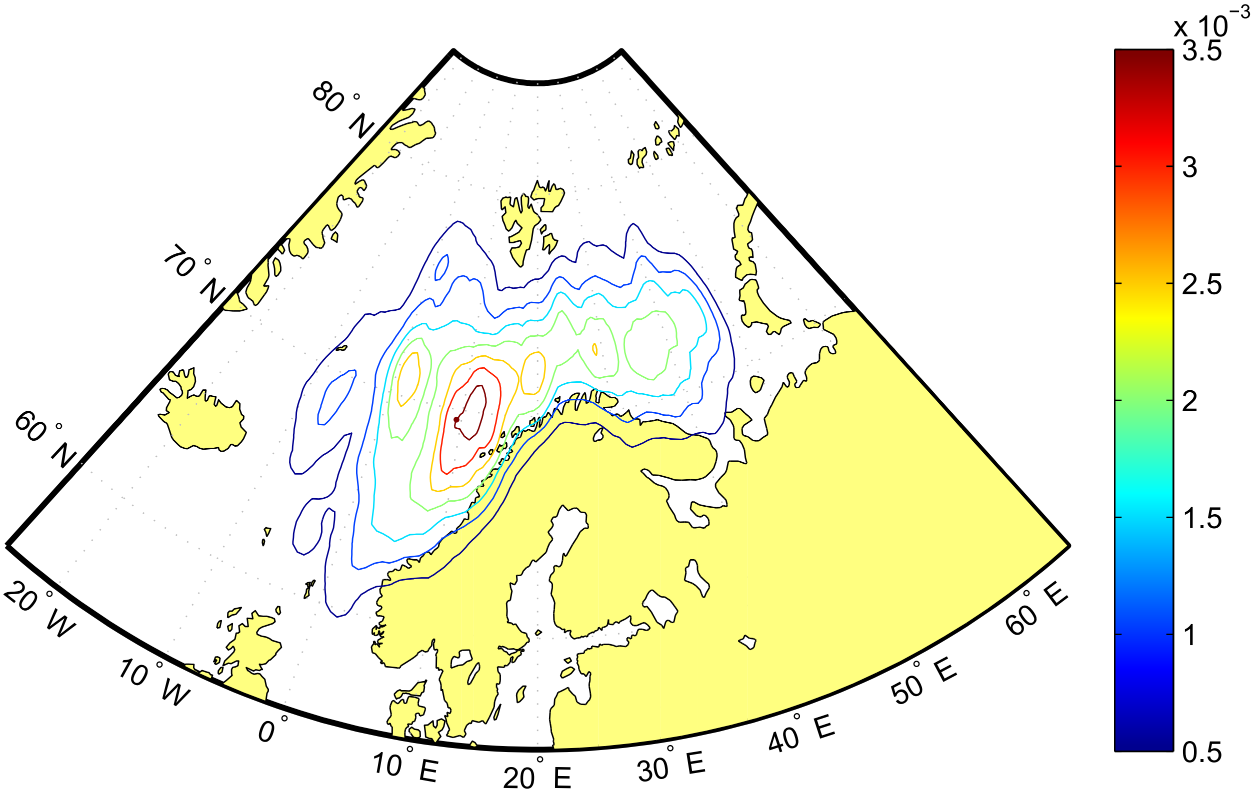


Figure 5.4: Kernel density estimates of the complete trajectory data with Epanechnikov kernel and bandwidth $h = 2.6$. The estimated density values of the contours are represented by a colorbar.

As mentioned earlier, the main maximum exists just outside of the Lofoten and Vesterålen area. The highest concentration area (deep red contour) stretches from approximately 68 °N - 72 °N and 10 °E - 14 °E. Two other high-density (but lower than the first one) areas are also present in the Norwegian Sea, one further away from the coast (approximately 69 °N - 72 °N and 0 °E - 4 °E) and another in the previous mentioned Tromsøflaket (71 °N - 73 °N and 17 °E - 21 °E). Both areas are marked by an orange contour. A fourth, very small area of same density (also orange contour) is located in the Barents Sea (approximately

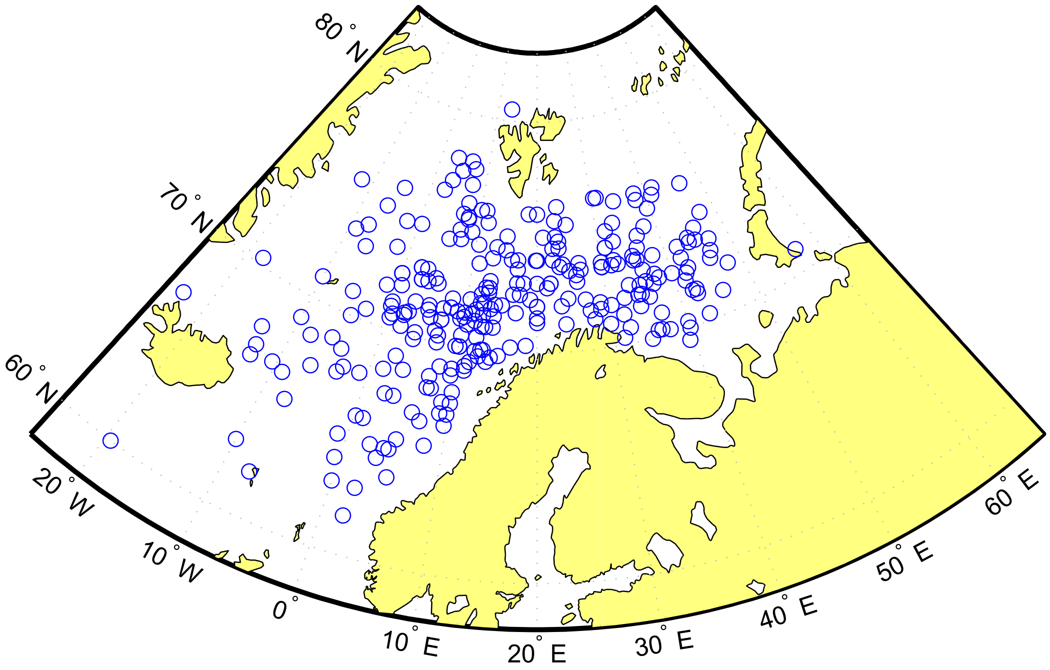
71.5 °N - 73.5 °N and 28 °E - 30 °E). For larger bandwidths like $h = 3$ (Fig. 5.3c) this small maximum is not visible, but rather a part of a larger area of lower densities. Further east in the Barents Sea, the density drops even more, but there is still an area here representing a local maximum (green contour).

Beyond the mentioned areas, a general (coarse) impression is that the density decreases with distance from the Norwegian coastline. Exceptions here are two regions, one located approximately between 67.5 °N - 70.5 °N and 8 °W - 5 °W, and another between 75 °N - 77 °N and 0 °E - 3.5 °E. These are areas where the density increases slightly for then to drop again.

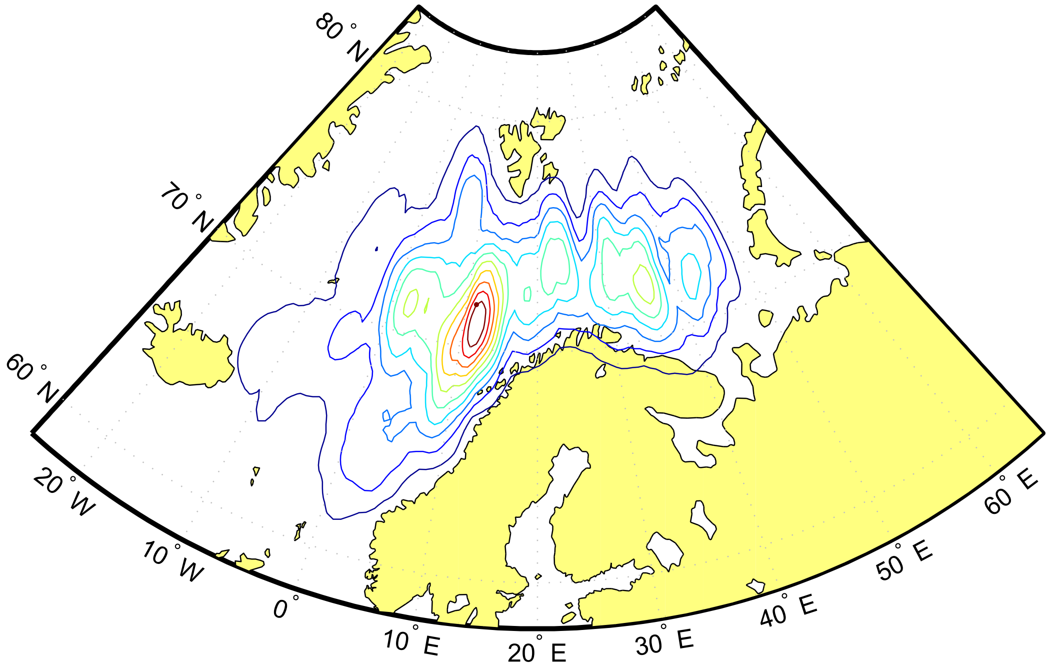
5.1.3 Kernel density estimates - Polar lows genesis

The previous section focused on density estimations using the complete trajectory data set, and thus resulted in the investigations of high-frequency areas of polar lows (PLs) in general through the winter seasons of $1999 - 2013$. However, it is also interesting to determine the areas where PLs are most likely to form, i.e., the areas of cyclogenesis. The kernel density estimation method was therefore also applied by using only the genesis points of each trajectory in the data set. The result is shown in Fig. 5.5, where Fig. 5.5a shows a scatter plot of all the genesis points and Fig. 5.5b shows the density estimates.

Once again, an area quite close to Lofoten and Vesterålen protrudes as the main maximum. Although quite similar to Fig. 5.4 or Fig. 5.3c, closer inspection of Fig. 5.5b reveals that the main maximum is now located a few degrees further north, the highest density contour (deep red) is located in an area between 70 - 72 °N and 10 - 13 °E. Other than that, some local maxima exist in the Barents Sea, in addition to a very small area located approximately between 71 - 72 °N and 2 - 3 °E.



(a) Scatter plot of the genesis points.



(b) Kernel density estimates of the genesis points. Epanechnikov kernel and $h = 3$ are used.

Figure 5.5: Area distribution of polar low genesis over the Nordic Seas from winter 1999/2000 to winter 2012/2013. Upper panel: scatter plot, lower panel: kernel density estimate.

5.1.4 Kernel density estimates - Polar lows dissipation

As mentioned earlier, polar lows tend to decay and dissipate when they hit land and are deprived of their main source of energy.[1] The event is nonetheless often associated with dramatic weather conditions. It was therefore interesting to find out which parts of the Nordic Seas coastlines are heavily exposed with regards to PL dissipation. Also, not all polar lows dissipate over land, some decay over ocean areas. A kernel density estimation using dissipation points would also say something about this distribution.

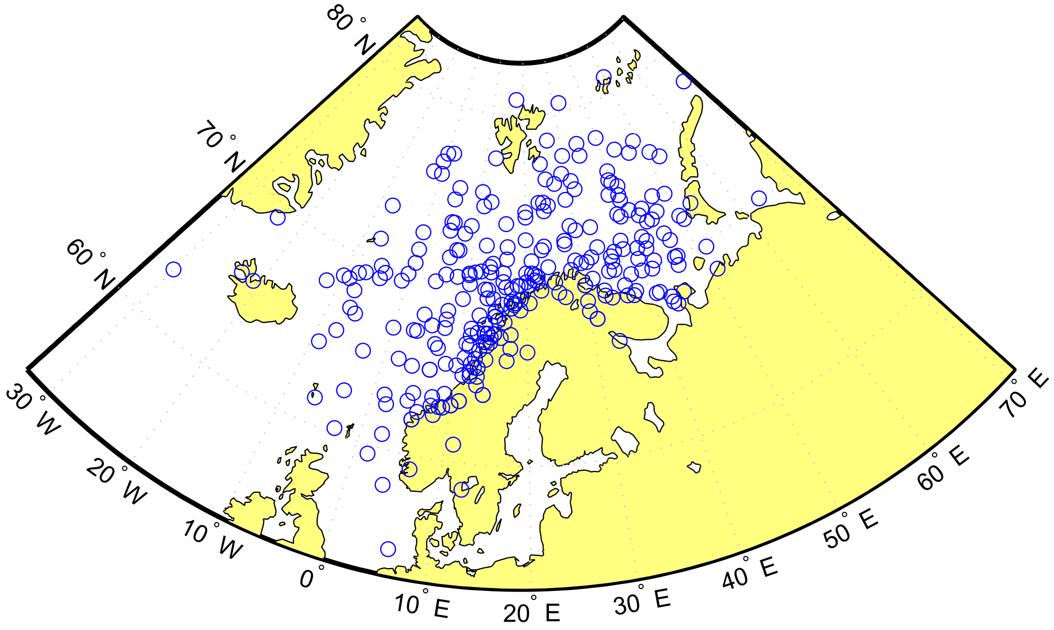
Like for the genesis points, Fig. 5.6 consists of a scatter plot of the dissipation coordinates for each trajectory (Fig. 5.6a), and a plot showing the estimated densities based on these points (Fig. 5.6b). From the scatter plot in Fig. 5.6a it is obvious that along the Norwegian coastline, Mid and Northern Norway have experienced the most polar low landfalls during the winter seasons of $1999 - 2013$. Many of the cases have also dissipated over the ocean, but these are more randomly spread throughout the Nordic Seas. What is surprising with this plot is that some of the polar lows have reached quite far into the land before dissipating completely.

The density estimates in Fig. 5.6b provide more detail and display the coastal area just south of Lofoten (Bodø, Helgeland) as having the highest densities (red contour). This region is bounded by approximately 66 - 68 °N and 10 - 15 °E. Though decreasing in density outwards from this high-density center, the contours stretches along large parts of the Norwegian coastline, especially Mid and Northern Norway, which was predicted from the scatter plot (Fig. 5.6a).

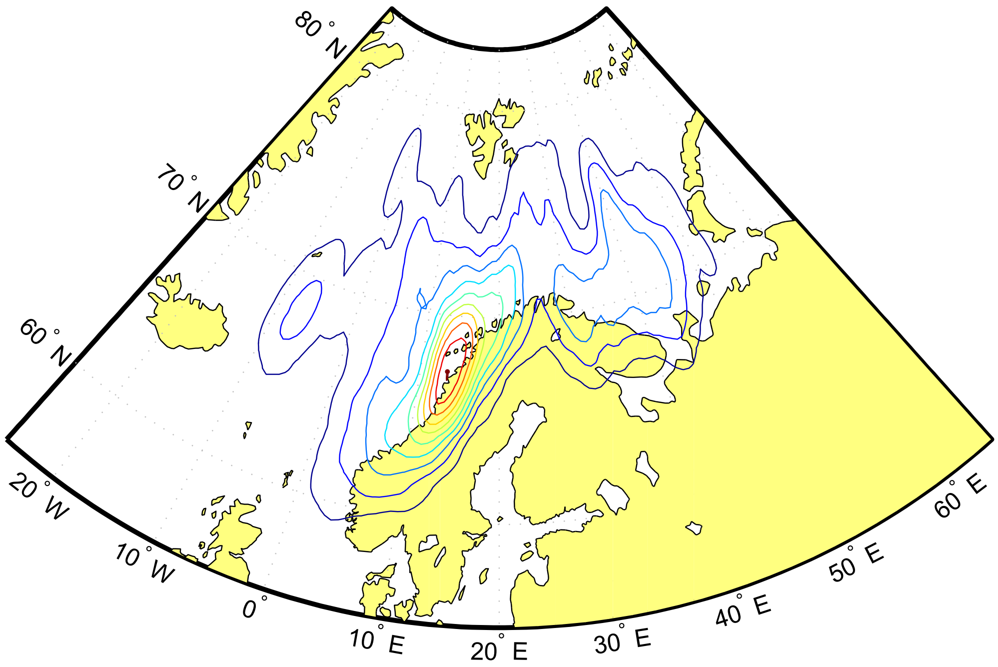
Other than this there are two other areas of increased local density, one in the Barents Sea and one northeast of Iceland, but because the data points in these areas very few or dispersed compared to the maximum region along the coast, the density values are very low.

---

[1] MetLex - Meteorologisk leksikon: Polare lavtrykk `http://metlex.met.no/wiki/Polare_lavtrykk`



(a) Scatter plot of the dissipation points.



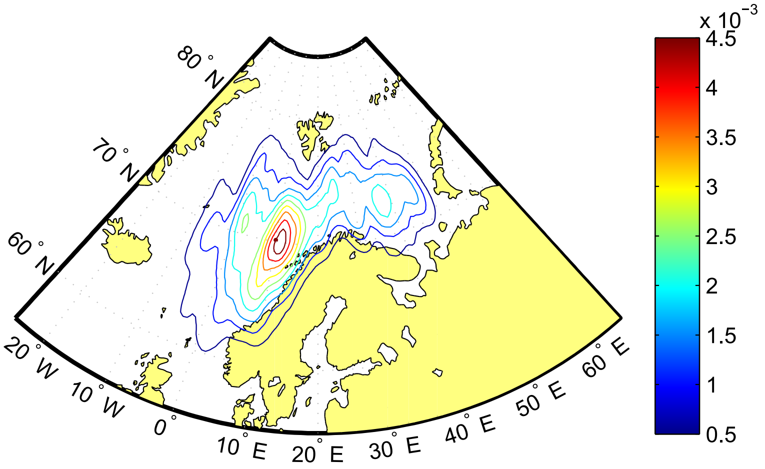
(b) Kernel density estimates of the dissipation points. Epanechnikov kernel and $h = 3.3$ are used.

Figure 5.6: Area distribution of polar low dissipation over the Nordic Seas from winter 1999/2000 to winter 2012/2013. Upper panel: scatter plot, lower panel: kernel density estimate.

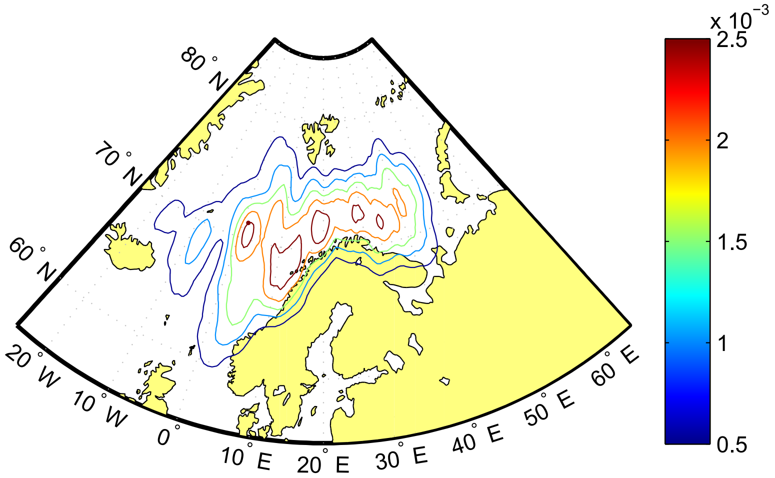
5.1.5 Time evolution of trajectory densities

In an effort to say something about the time evolution of the densities the complete trajectory data set was segmented into several numbers of intervals which were then used in the estimations. This means that the 2270 data points of the 286 polar low trajectories (the complete trajectory data) were divided into sections containing an equal number of data points, in chronological order. For instance, the array of 2270 data points were first split in half, to result in two smaller arrays of 1135 points, then divided by three, to produce three arrays of approximately 756 points, and so forth. The intervals were then correlated to time by checking which time window applied to each. The method of kernel density estimation was applied for each interval of each segmentation. This process was done for 2, 3, 4, 5 and 6 sections, respectively. When the total number of data points were not divisible by the number of intervals in question, the data set was divided into approximately equally sized intervals.

It should be mentioned that these plots provide a somewhat uneven relation to time because the number of PL events per year (and number of observations for each event) varies greatly throughout the data set. It is therefore not given that interval 1 and 2 corresponds to the first and second half of the time period 1999 – 2013, respectively. Therefore, the month and year of each interval is indicated on the figures. Also, because each interval is the result of an equipartition, thus containing an equal number of data points, it should be noted that the same bandwidth had to be used for both, to be able to compare them to each other. This obviously implies for each of the segmentations.
Fig. 5.7 illustrates the density estimates for both parts when the data set was split in half.



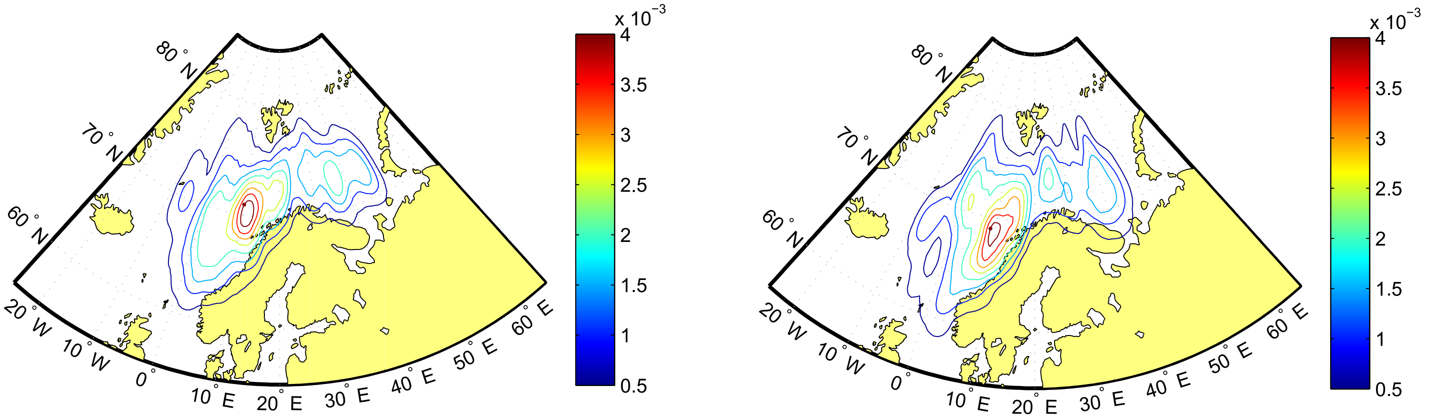
(a) Interval 1:Dec 99 (19.12.99) - Jan 09 (07.01.09)



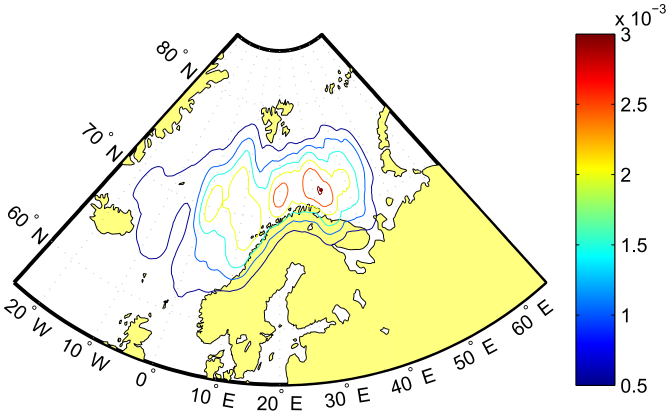
(b) Interval 2: Jan 09 - Apr 13 (09.04.13)

Figure 5.7: Kernel density estimates for two equally sized intervals of the data set. Epanechnikov kernel and bandwidth $h = 3$ are used. The intervals are defined by month and year, e.g., Jan 09 refers to January 2009.

An interesting observation to be made from Fig. 5.7 is that the maximum of the estimates actually expands and shifts spatially from one interval to the other, and thus undergoes quite dramatic changes through time. In the first part (Fig. 5.7a) the previously "known" maximum in the Lofoten area is clear. In the second part (Fig. 5.7b) however, there seem to be a presence of several smaller maxima, spread along the coast from the Lofoten area and northeastwards into the Barents Sea. Fig. 5.7a is very similar to the density estimate plot when using the complete trajectory data (Fig. 5.4), indicating that the main changes in spatial distribution take place after January 2009.



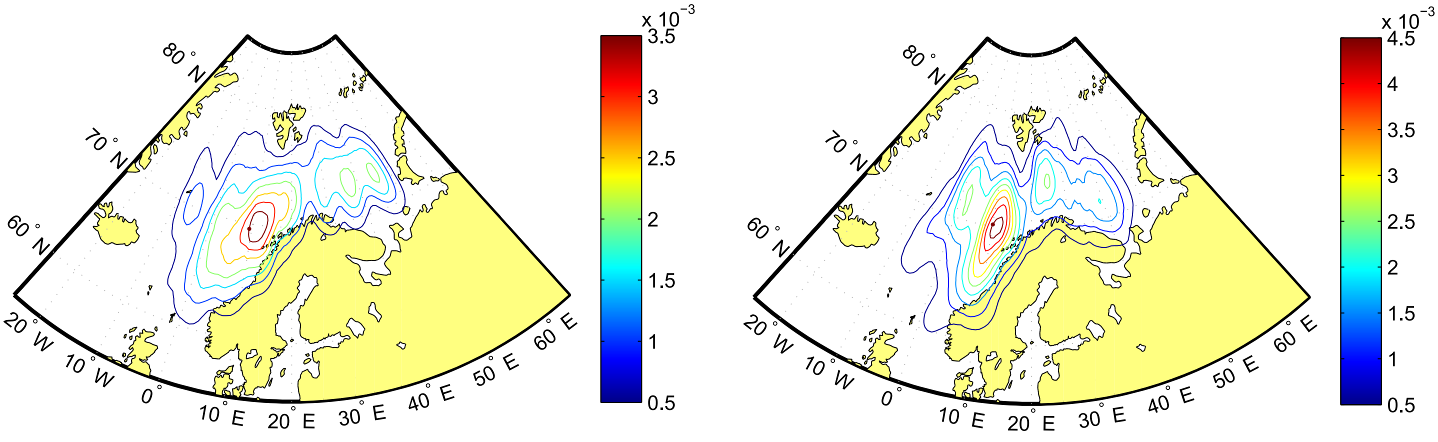
(a) Interval 1: Dec 99 (19.12.99) - Mar 06 (22.03.06) **(b)** Interval 2: Oct 06 (29.10.06) - Dec 10 (06.12.10)



(c) Interval 3: Dec 10 - Apr 13 (09.04.13)

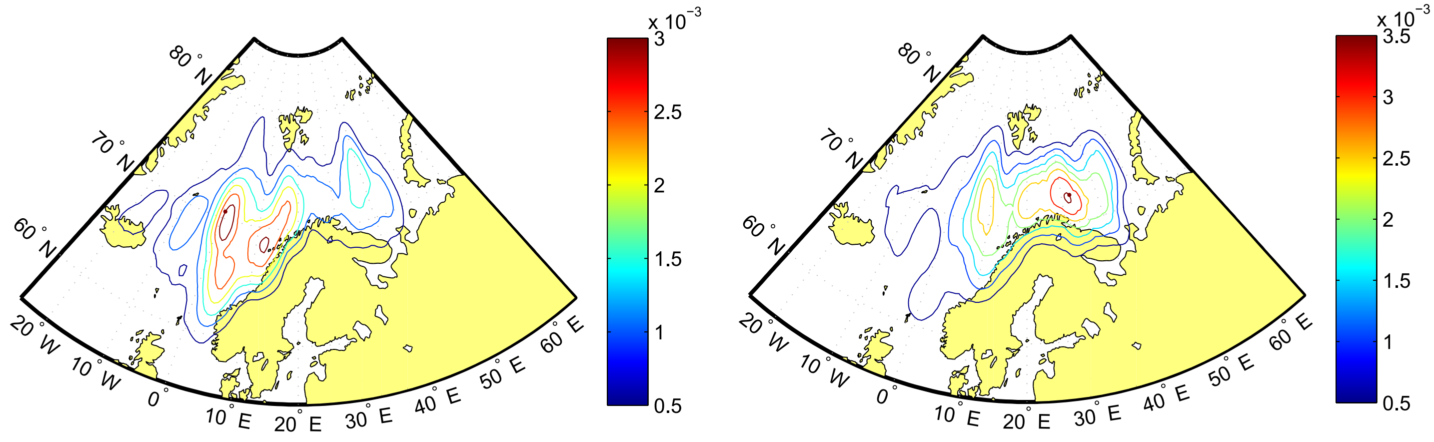
Figure 5.8: Kernel density estimates for three equally sized intervals of the data set. Epanechnikov kernel and bandwidth $h = 3.5$ are used.

The estimates for three equally sized sections of the data set (Fig. 5.8) present the same development, but in a little more detail. The main maximum can now be seen to move a little southwestwards along the Norwegian coast from interval 1 (Fig. 5.8a) to interval 2 (Fig. 5.8b), before being located in the Barents Sea in the last time window (Fig. 5.8c).



(a) Interval 1: Dec 99 (19.12.99) - Feb 05 (26.02.05)

(b) Interval 2: Feb 05 - Jan 09 (07.01.99)

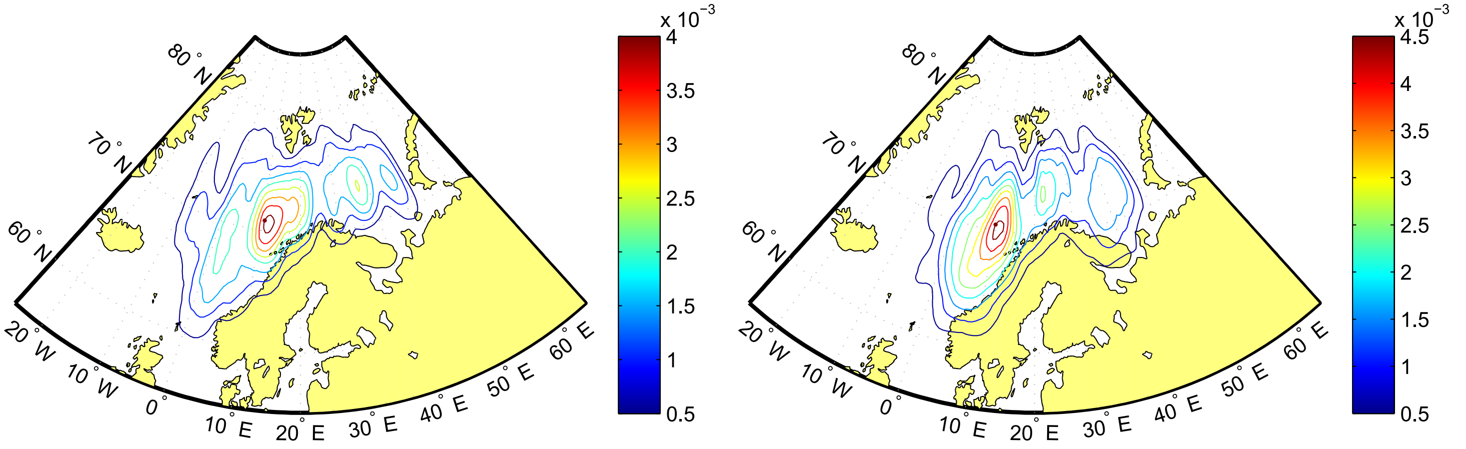


(c) Interval 3: Jan 09 - Mar 11 (24.03.11)

(d) Interval 4: Mar 11 - Apr 13 (09.04.13)

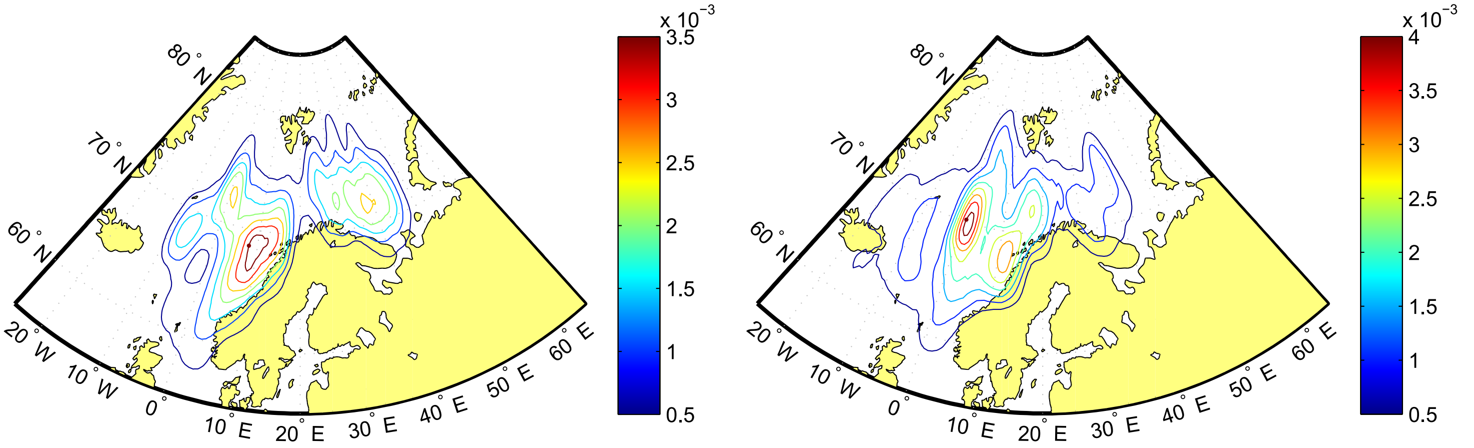
Figure 5.9: Kernel density estimates for four equally sized intervals of the data set. Epanechnikov kernel and bandwidth $h = 3.7$ are used.

Fig. 5.9 shows the estimates when four intervals are used. The densities in the two first intervals (Fig. 5.9a and Fig. 5.9b) are quite similar, except for the maximum being more "stretched out" and concentrated in the second one. In the third interval (Fig. 5.9c), three maxima are visible, now covering a larger area. Interval four (Fig. 5.9d) again shows the highest densities mainly being present in the Barents Sea.



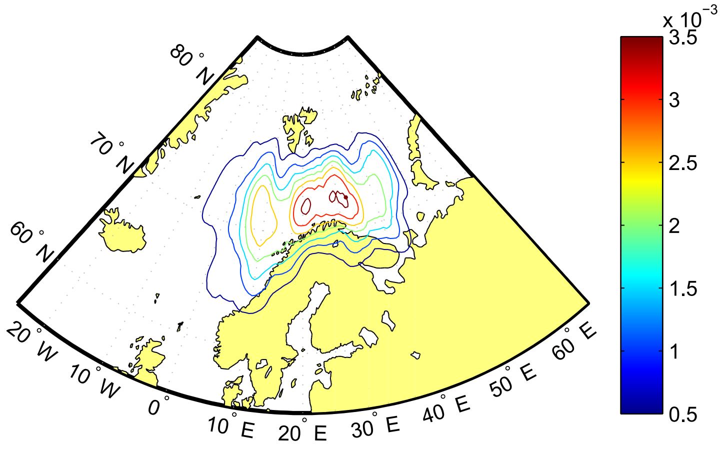
(a) Interval 1: Dec 99 (19.12.99) - Mar 04 (27.03.04)

(b) Interval 2: Mar 04 - Feb 08 (29.02.08)



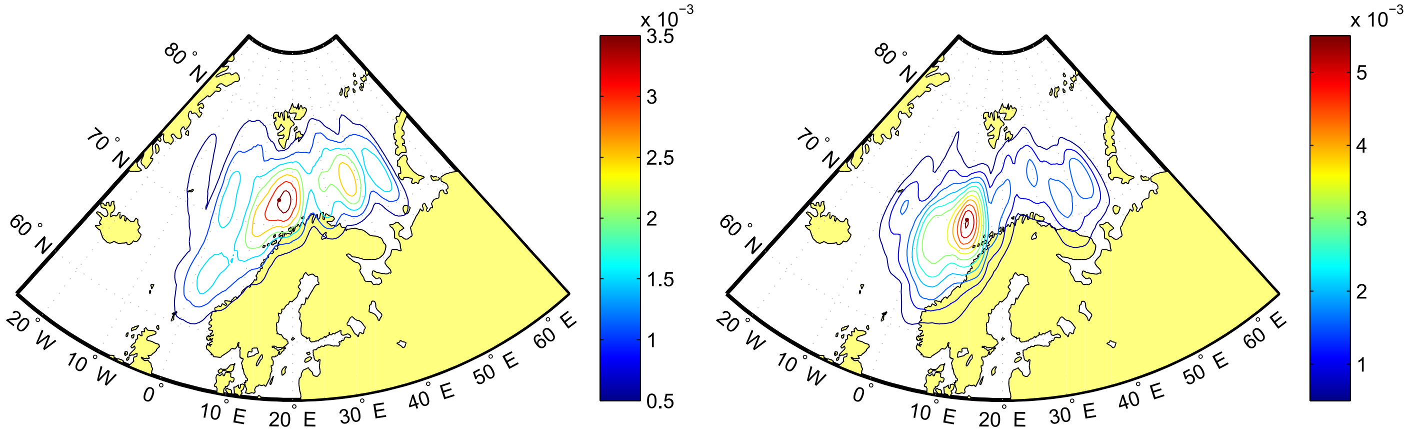
(c) Interval 3: Feb 08 - Mar 10 (14.03.10)

(d) Interval 4: Mar 10 - Mar 12 (17.03.12)

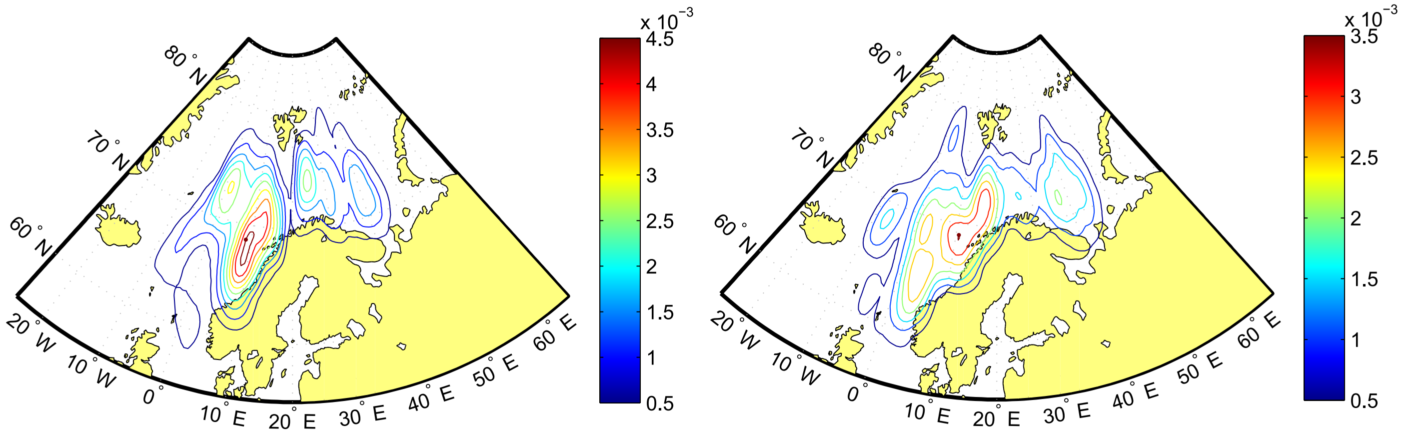


(e) Interval 5: Mar 12 - Apr 13 (09.04.13)

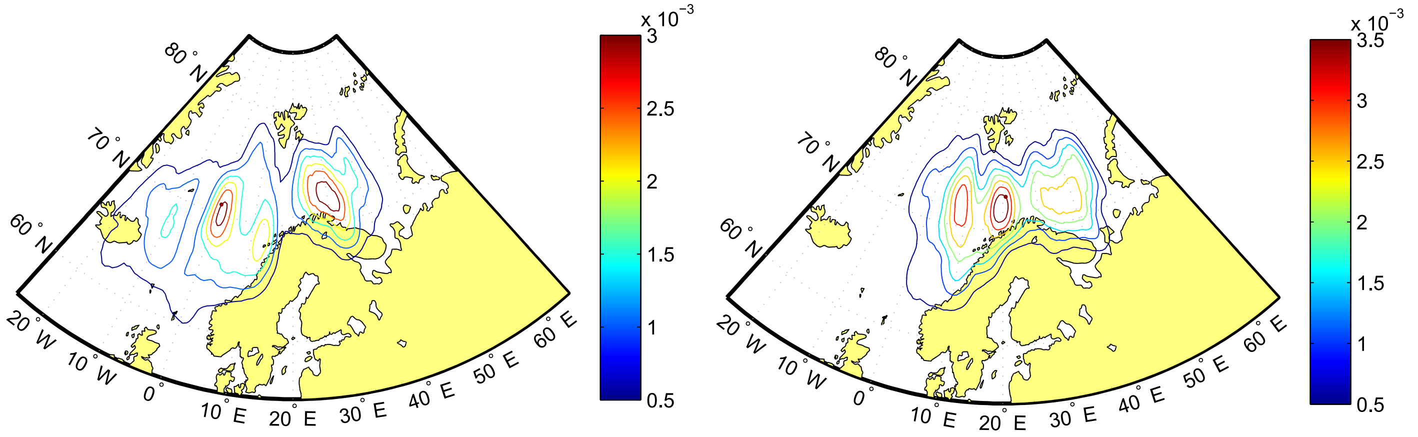
Figure 5.10: Kernel density estimates for five equally sized intervals of the data set. Epanechnikov kernel and bandwidth $h = 3.7$ are used.



(a) Interval 1: Dec 99 (19.12.99) - Dec 03 (26.12.03) **(b)** Interval 2: Dec 03 - Mar 06 (22.03.06)



(c) Interval 3: Mar 06 - Jan 09 (07.01.09) **(d)** Interval 4: Jan 09 - Dec 10 (09.12.10)



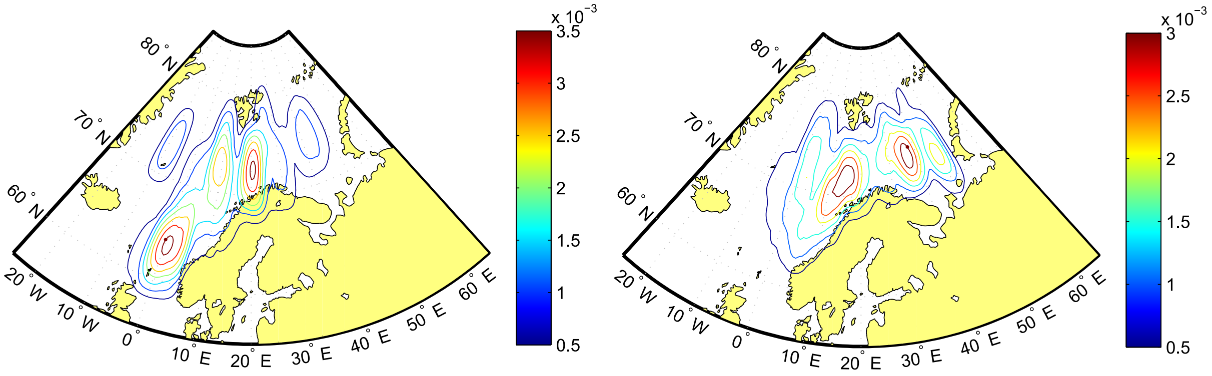
(e) Interval 5: Dec 10 - Oct 12 (09.04.13) **(f)** Interval 6: Oct 12 - Apr 13 (09.04.13)

Figure 5.11: Kernel density estimates for six equally sized intervals of the data set. Epanechnikov kernel and bandwidth $h = 3.8$ are used.

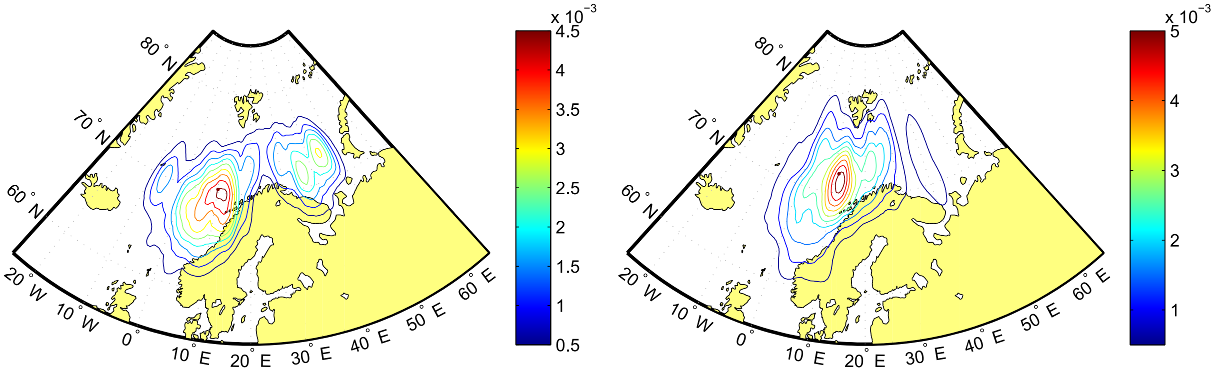
As the number of intervals increase to five (Fig. 5.10) and six (Fig. 5.11) a more and more detailed evolution becomes evident. Fig. 5.10 seem to follow a development similar to the preceding plots, where the maximum "ends" in the Barents Sea. In Fig. 5.11 however, the development differs. The plots of the three first intervals (Fig. 5.11a, Fig. 5.11b and Fig. 5.11c) are quite similar to those of Fig. 5.10, with high densities being concentrated in, or close to the Lofoten area. The fourth plot (Fig. 5.11d) also displays the greatest concentrations in this area, though somewhat more spread out along the coast. The fifth plot (Fig. 5.11e) shows an area of high density in the Barents Sea, but also one located quite far out in the Norwegian Sea, between Norway and Iceland. This maxima can also be spotted in Fig. 5.9 and Fig. 5.10. Looking at the scatter plot in Fig. 5.1, this area (around 70 °N and 0 °E) can actually be identified as an area of high occurrence of polar low activity. In the sixth plot (Fig. 5.11f) in Fig. 5.11 the maximum is located in the Tromsøflaket region, with lower densities in the Barents Sea.

To be able to say something about the evolution more specifically related to time, kernel density estimates were also calculated for two-year intervals of the data set. The results can be seen in Fig. 5.12. Because of highly varying numbers of PL cases in the two-year intervals, the plots in Fig. 5.12 exhibit a larger degree of variability than some of the previous figures, where each time interval consisted of an equal number of PL data points. Because of the occasionally large differences in number of data points, the bandwidth for each interval were considered carefully and adjusted individually to give smooth, yet informative plots.

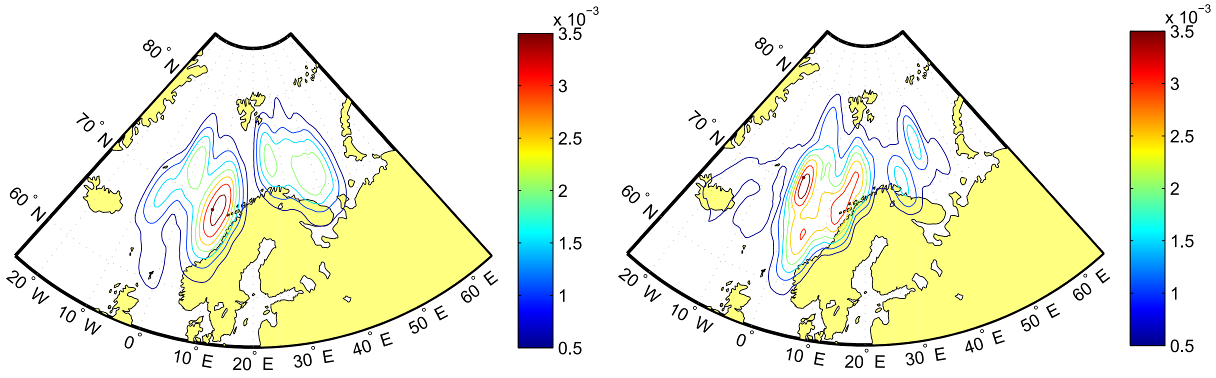
Many of he previously observed maxima are also present in Fig. 5.12, like the one outside of Lofoten and Vesterålen, the one in the Tromsøflaket area, and the one between Norway and Iceland (around 70 °N and 0 °E). One new maximum can be seen in Fig. 5.12a, which is the first one to appear this far to the south (around 65 °N and 0 °E). It should be mentioned that $1999 - 2001$ was the two-year period when fewest PLs occurred, meaning that this maximum is based on very few data points compared to the other intervals. Due to this, the maximum should not be assumed to have a significant meaning in relation to PL spatial activity. Similar to Fig. 5.8, Fig. 5.9 and Fig. 5.10 the last plot in Fig. 5.12, (Fig. 5.12g) displays the main maximum located in the Barents Sea, yet again suggesting a spatial time evolution of PL activity.



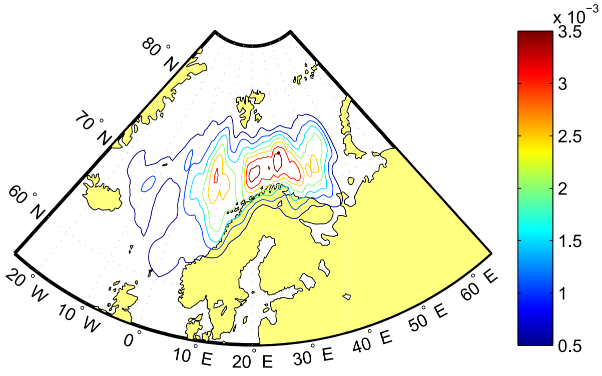
(a) Year 1999-2001, $h = 5$ **(b)** Year 2001-2003, $h = 4$



(c) Year 2003-2005, $h = 3.5$ **(d)** Year 2005-2007, $h = 4.25$



(e) Year 2007-2009, $h = 4.5$ **(f)** Year 2009-2011, $h = 3.5$



(g) Year 2011-2013, $h = 3$

Figure 5.12: Kernel density estimates for two-year intervals of the data set. Epanechnikov kernel is used, and bandwidths are fitted individually for each interval by the subjective method.

5.2 Clustering analysis

The discussion now moves away from densities based on singular observation points, and turns the focus over to the full trajectories of the polar lows. To begin with, a plot of all the trajectories (for each of the 286 events) is shown in Fig. 5.13

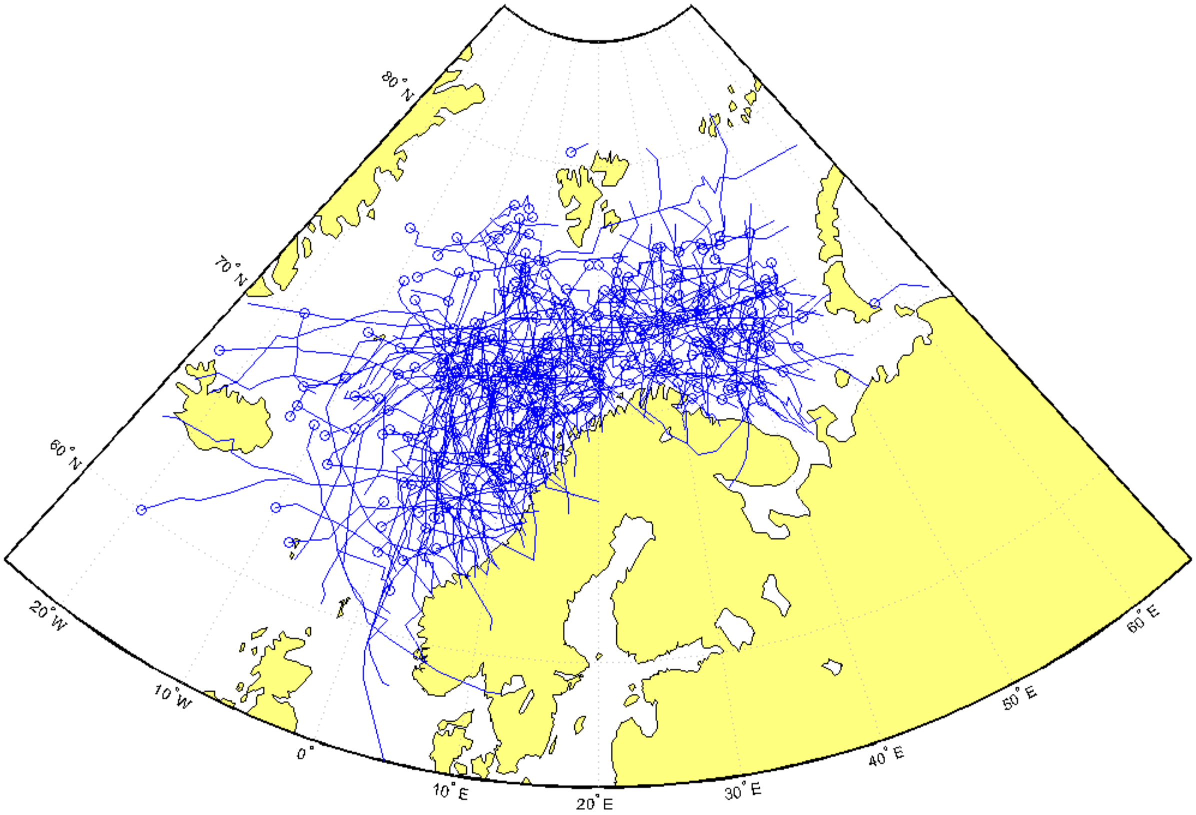
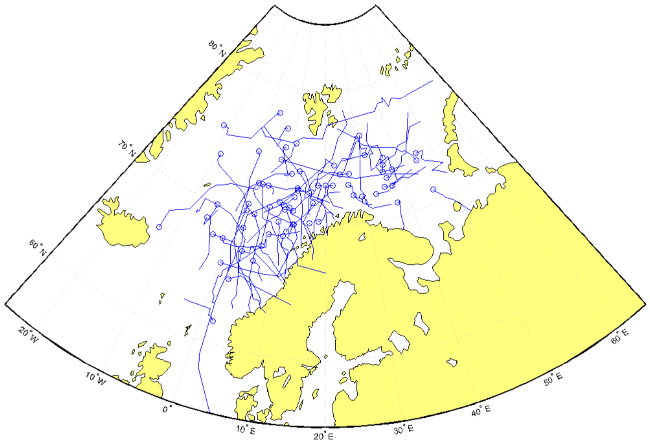


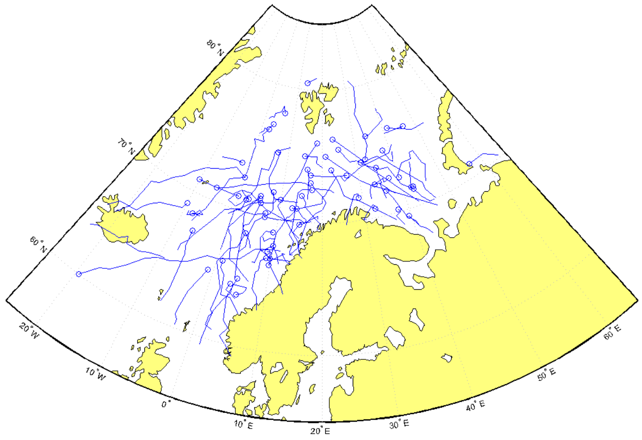
Figure 5.13: Polar low trajectories in the Nordic Seas between 1999-2013. The first point of each trajectory is marked by a circle.

As seen from the figure, the collection of trajectories appear to be rather chaotic. Overlapping tracks makes visualizing difficult. The tracks which can be visualized clearly seem to be random in shape and direction, and no subgroups are protruding, at least not judging by the naked eye. For visual purposes, the number of trajectories were divided into four approximately equal segments and plotted the same way as in Fig. 5.13.



(a) Polar low trajectories 99-13, interval 1

(b) Polar low trajectories 99-13, interval 2



(c) Polar low trajectories 99-13, interval 3

(d) Polar low trajectories 99-13, interval 4

Figure 5.14: Polar low trajectories during 1999-2013, divided into four segments, each containig approximately 70 trajectories. The first point of each track is marked by a circle.

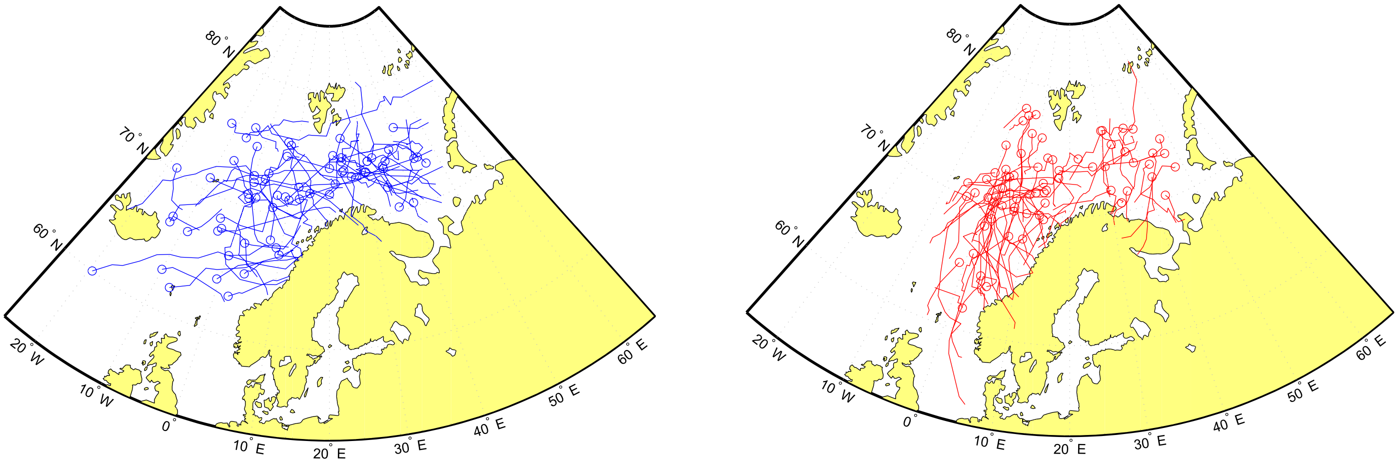
The trajectories are now more evident, and certain characteristics can be identified. There is great variety in their shape. Some tracks are almost completely straight, others are slightly bent, and some contain great curves and bends in their structure. Also, the direction of the tracks differs from case to case. Many of them have an overall, or net, southwards movement, meaning that their point of dissipation is further south than the point of origin. The southwards motion is then usually combined with an east or westward component. However, some of the tracks have a net displacement in the northwards direction.

All these (286) trajectories were used as the input for the Curve Clustering Toolbox.[1] The clustering process was carried out by including the observations for each trajectory in the complete trajectory data set (latitude-longitude coordinates, referred to as the cell array **Z** in section 4.2.1), and a cell array of date indexes (referred to as the cell array **T** in section 4.2.1). Each trajectory corresponded to a date index which started on 1 for the first trajectory and counted upwards for each new date. Dual or multiple PLs occurring on the same date were therefore given the same index.

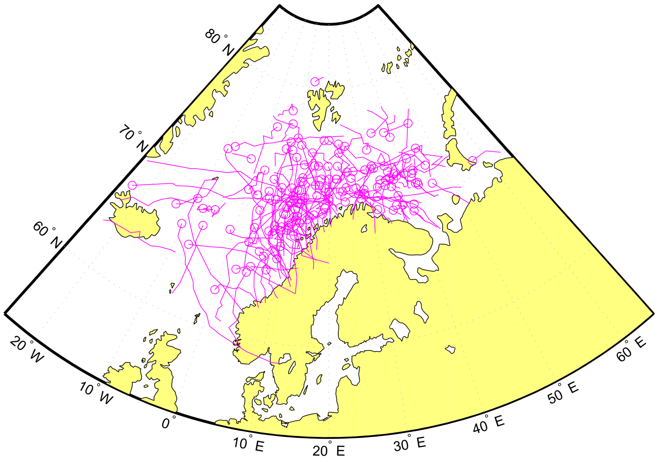
The clustering was run with a model option that excluded initial starting position of trajectories as a possible source of variation. This option forced all PLs to begin at the relative position of (0,0), hence removing the genesis position as a source of variation in the clustering. The clustering then becomes more concentrated on cyclone track shape than spatial distribution.[1] Use of the opposite option, 'nozero', were also tested, but this confirmed what was stated above, i.e., the clustering focused more on genesis location and spatial distribution. The 'zero' option was therefore determined as preferable. The model was first run by dividing the trajectories into 2, 3, 4, 5 and 6 clusters, respectively. The number of clusters are predefined. After investigating the results it was decided that the clustering with 3, 5 and 6 predefined clusters provided the most informative and easily interpreted plots. These results were therefore included, and can be seen in Fig. 5.15 (3 clusters), Fig. 5.16 (5 clusters), and Fig. 5.17 (6 clusters).

---

[1] CCToolbox documentation: `http://www.datalab.uci.edu/resources/CCT/doc/`

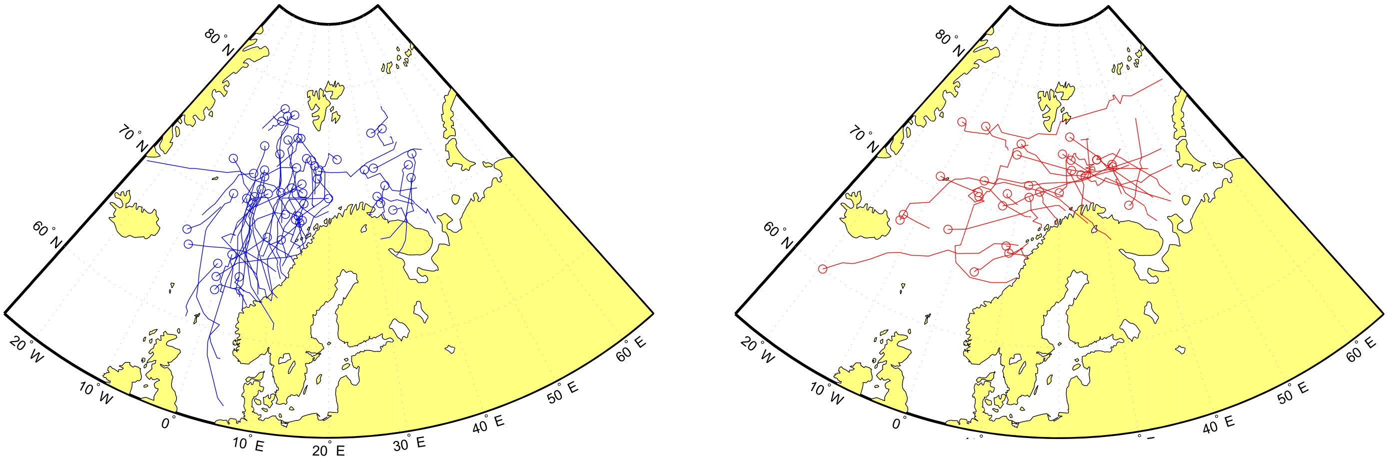


(a) Cluster 1 **(b)** Cluster 2

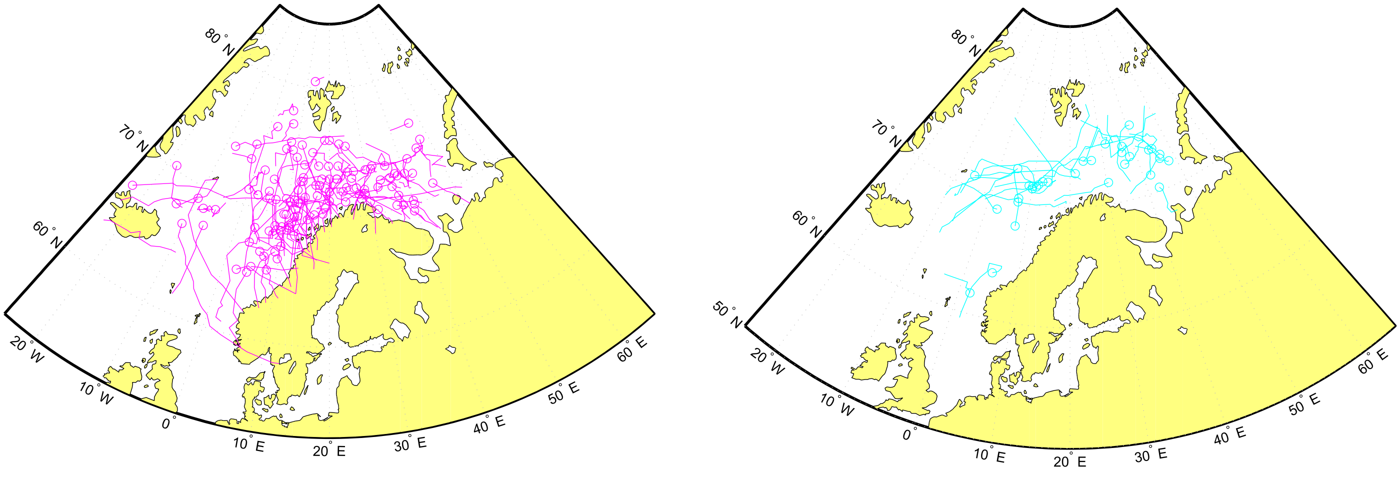


(c) Cluster 3

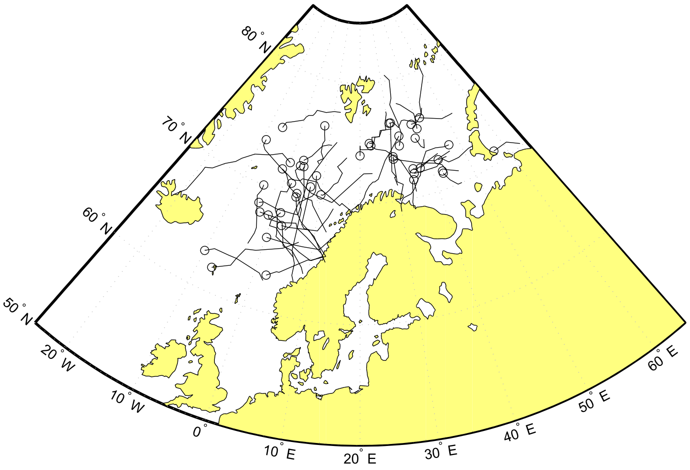
Figure 5.15: Clustering analysis with 3 clusters.



(a) Cluster 1 **(b)** Cluster 2

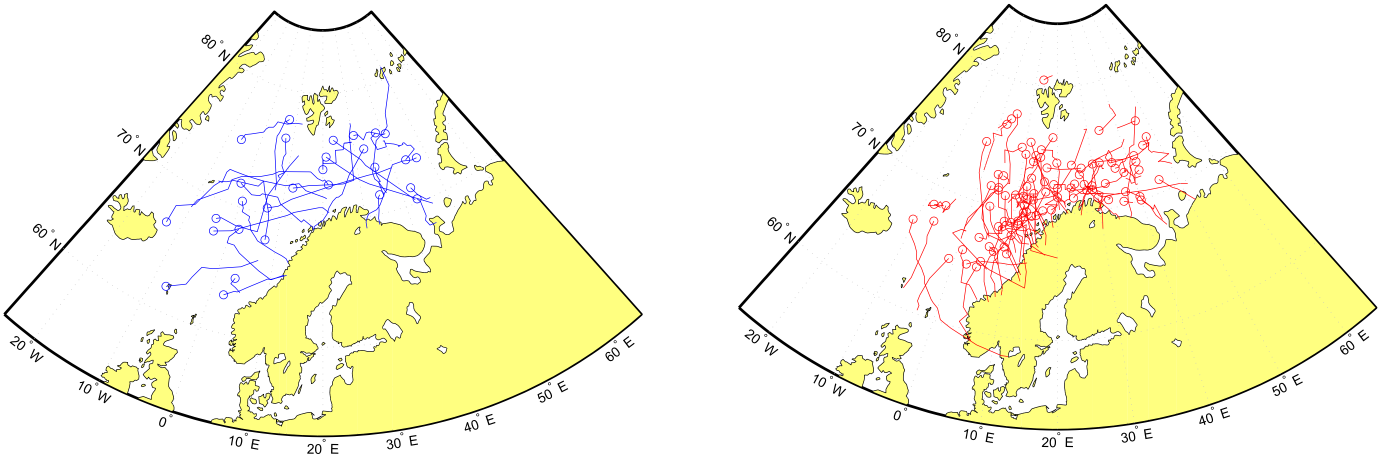


(c) Cluster 3 **(d)** Cluster 4

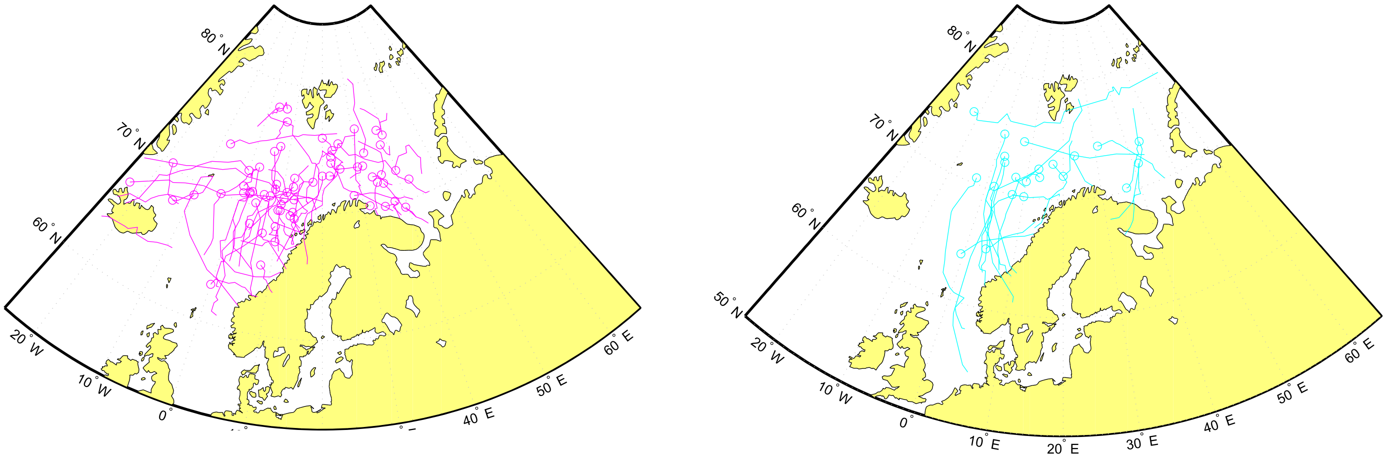


(e) Cluster 5

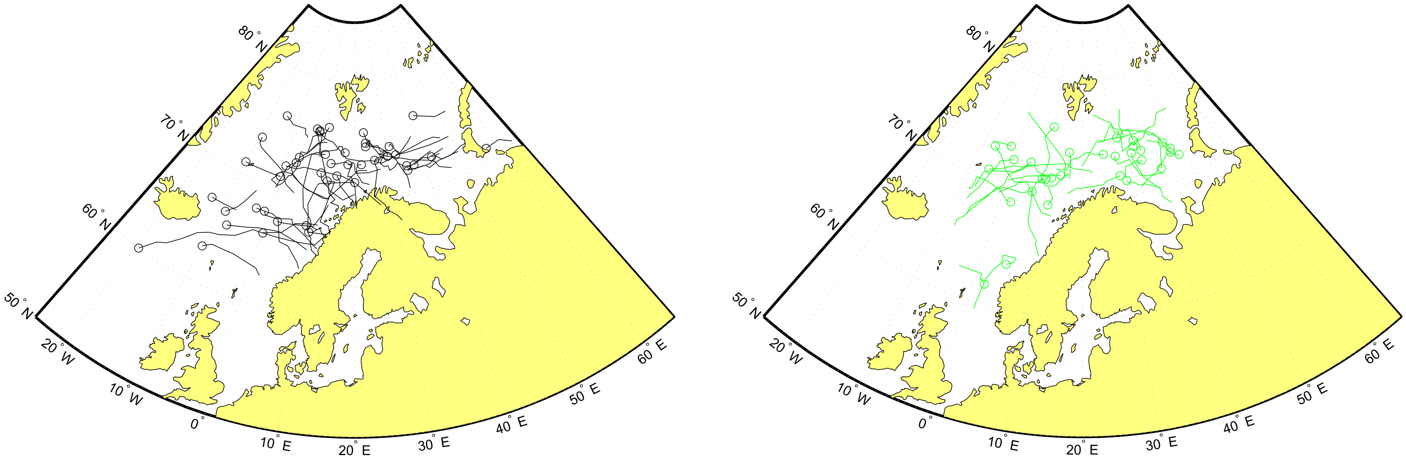
Figure 5.16: Clustering analysis with 5 clusters.



(a) Cluster 1 **(b)** Cluster 2



(c) Cluster 3 **(d)** Cluster 4



(e) Cluster 5 **(f)** Cluster 6

Figure 5.17: Clustering analysis with 6 clusters.

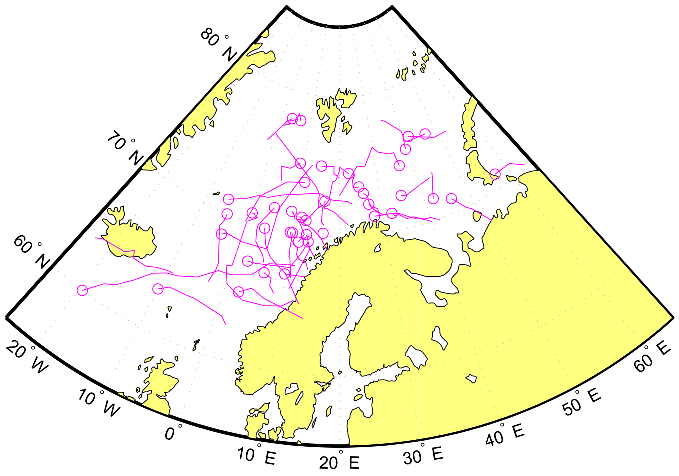
As seen from the plots, curve clustering can provide interesting information regarding the shape characteristics of trajectories. The clustering results will be further discussed in Chp. 6.

5.3 Analysis of polar lows and weather regime

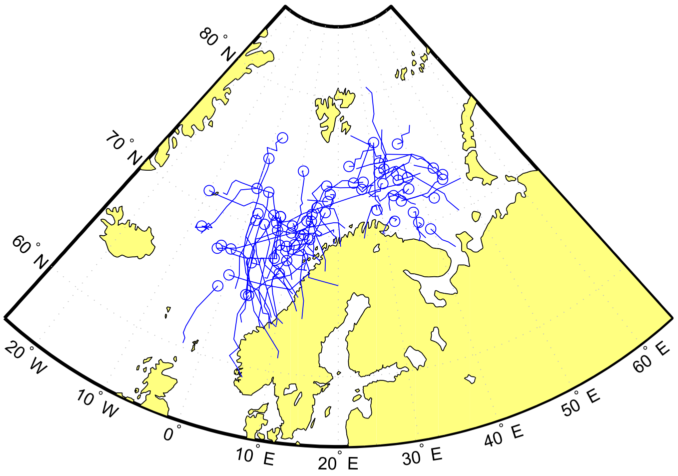
In addition to the density estimates and the clustering analysis, the polar lows were analysed according to their belonging weather regime, which represent one of the other properties of the polar lows determined in the data set by Rojo et al. [2014]. The weather regime analysis was carried out mainly just by plotting, to gain a visual representation which could then be associated with the findings of Rojo et al. [2014].

Fig. 5.18 displays the polar low trajectories of the winter seasons $1999 - 2013$, plotted according to which winter weather regime they belonged to/formed under. Fig. 5.18a displays the polar low trajectories belonging to the positive phase of the North Atlantic Oscillation (NAO+), Fig. 5.18b those belonging to the negative phase of the North Atlantic Oscillation (NAO-), Fig. 5.18c shows the trajectories of the Atlantic Ridge (AR) and Fig. 5.18d shows the trajectories of the Scandinavian Blocking (SB). Trajectories formed outside of the winter season November-March, i.e., in September-October or April-May were excluded by Rojo et al. [2014], and likewise here. The PL trajectories of the last two years of the data set, $2011 - 2012$ and $2012 - 2013$, were not assigned to a weather regime by Rojo et al. [2014]. These are therefore also excluded from the plots in Fig. 5.18.

As seen from Fig. 5.18, the trajectories in every subfigure are spread reasonably well throughout the Nordic Seas. In other words, location of genesis or spatial distribution do not seem to be immediately related to WR. However, PL trajectory characteristics seem to vary according to which WR they formed and developed under. The most significant characteristics appear for the NAO- regime (Fig. 5.18b) and the SB regime (Fig. 5.18d), where the trajectories are dominated by a net southwards and net eastwards travel direction, respectively. It turns out that large-scale wind flow of WRs are closely connected to characteristics in travel direction of Nordic Seas PLs. This will be further discussed in Sec. 6.3.



(a) NAO+ **(b)** NAO-



(c) AR **(d)** SB

Figure 5.18: Polar low trajectories in the Nordic Seas during the winter seasons 1999-2013 plotted according to the weather regime they belong to. The first point of each trajectory is marked by a circle.

Chapter 6

Discussion

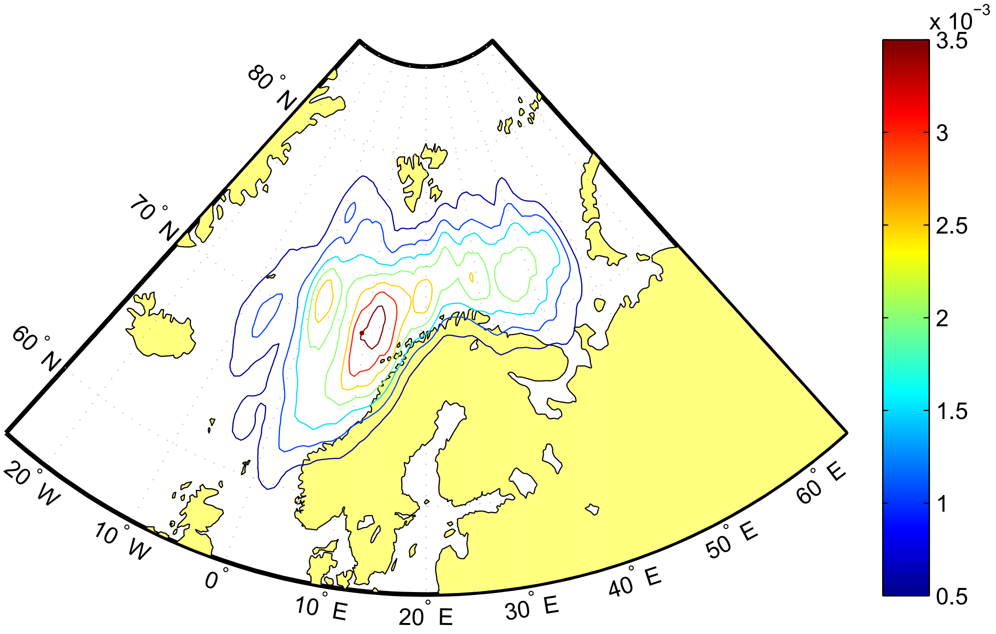
The discussion part aims to give some explanations, or suggestions for the spatial and temporal distribution results presented in Ch. 5, which were obtained by using the methods presented in Ch. 4. This will be based on previous research on polar lows, and the climatological factors connected to them, both on local and larger scales.

6.1 Kernel density estimation

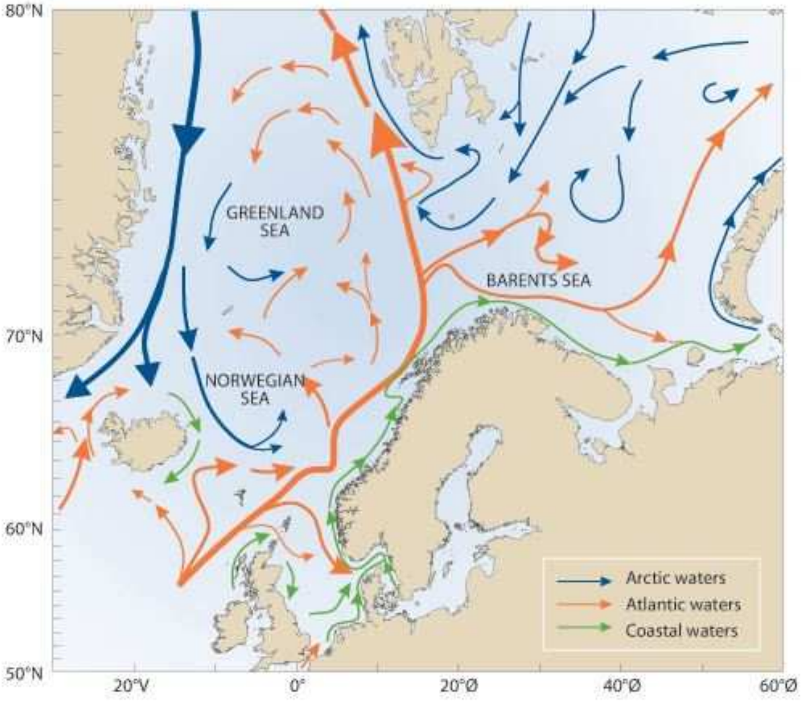
6.1.1 General observations

The general observations of polar low densities are considered first. These results were obtained using the complete trajectory data (for all 286 trajectories) and thus indicate where polar lows not only form, but also where developed lows pass frequently. These areas most likely represent regions where conditions are right for polar low genesis, but also for maintaining the energy supply needed for developing polar lows.

To try to correlate the estimated densities with something that can help explain their distribution, an intuitive suggestion is to compare the densities with the distribution of ocean currents in the Nordic Seas (Fig. 6.1). The favourable conditions for PL formation arise over these relatively warm ocean areas during so-called cold air outbreaks (CAOs), or marine cold air outbreaks (MCAOs) [Bracegirdle and Gray, 2008; Kolstad, 2011]. MCAOs are large-scale advections of cold, polar air masses over warmer, ice-free waters. MCAOs are especially common in the North Atlantic ocean during the winter season over the North Atlantic Current (NAC) and its relatively warm sea surface temperatures (SSTs). The Norwegian Atlantic Current (NwAC), which is a branch of the North Atlantic Current, brings warm, saline water along the coast of Norway and into the Arctic regions. This results in unusually warm wintertime water temperatures, for instance are temperatures of 7°C not uncommon just south of Svalbard [Noer et al., 2011].



(a) Kernel density estimates with Epanechnikov kernel and $h = 2.6$, revisited.



(b) Ocean currents in the Nordic Seas. Source: Institute of Marine Research, Norway.[1]

Figure 6.1: 2D kernel density estimates vs. ocean currents in the Nordic Seas.

---

[1] Figure obtained from: `http://www.regjeringen.no/en/dep/kld/documents-and-publications/government-propositions-and-reports-/reports-to-the-storting-white-papers-2/20012002/Report-No-12-2001-2002-to-the-Storting/2/1/1.html?id=452048`

The large air-sea temperature gradients which arise during MCAOs over warmer waters are likely to promote polar low formation and development [Kolstad et al., 2009; Linders and Sætra, 2010; Sætra et al., 2008].

A first impression from comparing Fig. 6.1a and Fig. 6.1b is that the high density ocean areas fit well with areas of increased mean sea surface temperature [Noer et al., 2011]. The area of highest frequency outside Vesterålen and Lofoten fits well with the small bend in the Norwegian Atlantic Current in this area. Other areas of high density also seem to follow the smaller branches of the Norwegian Atlantic Current in the Barents Sea and in the Norwegian Sea, between Norway and Iceland (around 70 °N and 0 °E).

The relation between PLs and cold air outbreaks can be discussed using results from a study carried out by Mallet et al. [2013]. Mallet et al. [2013] made observations of synoptic environment conditions, and how they influence PL development in the Norwegian and Barents Sea, by using the dataset from Noer et al. [2011], among others. This was performed by determining significant anomaly patterns over large areas centered over PL genesis zones, for instance anomalies related to near-surface wind, suggesting cold air outbreaks and stratospheric intrusions.

Analyses were conducted for PLs (from Noer's list) that formed over the Norwegian Sea (87 cases) and those that formed over the Barents Sea (47 cases), during the years 2000 to 2009. The results for the near-surface wind anomalies for PLs key days, i.e., outbreak days in Noer's list are reproduced in Fig. 6.2. They are divided according to formation region, the Norwegian Sea (left hand side) or the Barents Sea (right hand side). It turned out that the PL outbreak days corresponded to sudden intensifications of climatological anomalies, especially related to near-surface wind (CAOs) [Mallet et al., 2013].

The near-surface wind anomaly pattern for Barents Sea PL cases (Fig. 6.2, right hand side) displays an easterly component (i.e., a wind directed westwards) on the eastern side of the PL formation region. This indicates that most PLs forming in the Barents Sea do so during cold air outbreaks from the area between Spitsbergen and Novaya Zemlya [Mallet et al., 2013]. The wind anomaly for the Norwegian Sea PLs (Fig. 6.2, left hand side) is northerly, stating that PLs in this region form during cold air outbreaks from the Fram Strait [Mallet et al., 2013]. Mallet et al. [2013] could thus conclude that PLs develop within a northerly flow in the Norwegian Sea, and a northeasterly flow in the Barents Sea. Similar results were also stated in Ese et al. [1988] and Noer et al. [2011].

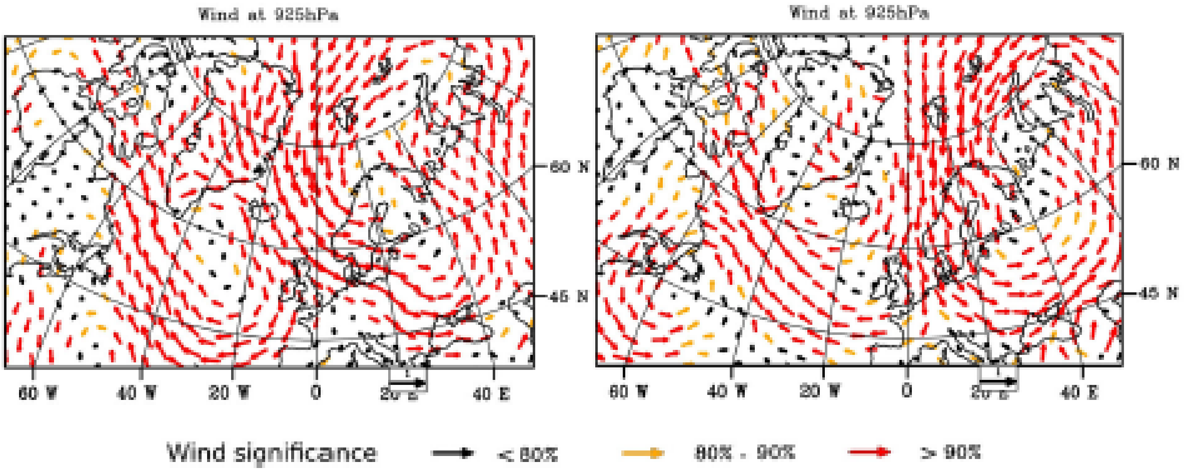


Figure 6.2: Standardized anomalies of wind at 925 hPa in cold season months (Oct-March) for PLs key days from Noer's (date) list, according to the formation location: the Norwegian Sea on the left side and the Barents Sea on the right side. The significant levels are indicated by the color of the arrows (wind). Source: Mallet et al. [2013].

The density estimates can also be related to other PL datasets in the Nordic Seas to check for similarities in spatial distribution. One possibility is the collection of identified PLs between 1978 and 1982 in Wilhelmsen [1985]. This study presented 33 PL cases, most of them constrained to the Norwegian Sea. It should be noted that the study was based on meteorological weather maps, which again were based on conventional observations. Due to the lack of observations over ocean areas, it is likely that some polar lows forming and developing over these regions were missed. Also, because Wilhelmsen had limited access to Russian observations, a number of PLs in the eastern part of the Barents Sea may have been missed [Noer et al., 2011]. Fig. 6.3 is reproduced from Wilhelmsen [1985] and shows the trajectories of the 33 PL cases identified between 1978 and 1982.

Wilhelmsen's figure actually paints another picture than that of the kernel density estimates in Fig. 6.1a. It appears that the highest frequencies of polar low activity, both by forming systems and passing tracks, are located in the previous mentioned region known as Tromsøflaket, around 72 °N and 18 – 22 °E. The densest number of PL genesis points is found in this area, and a large number of them travel a short distance before they dissipate over the coast of Northern Norway. The coastal areas of Lofoten and Vesterålen, which protruded as highly affected areas from the density estimates in Fig. 6.1a are now characterized by a rather limited PL activity. The reasons for this shift in maxima are not clear. However, as the main high-density area is still well aligned with the NwAC and its branches, the shift could just represent the dynamic nature of the polar low systems.

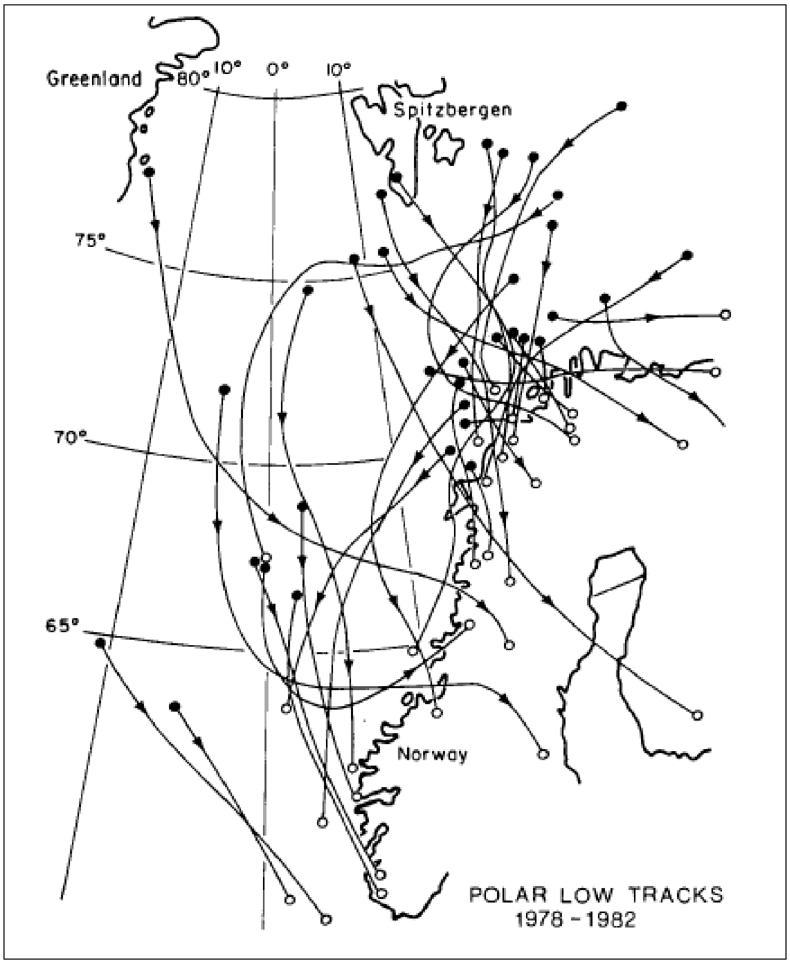


Figure 6.3: Tracks of 33 polar lows in the years 1978-1982. Source: Wilhelmsen [1985].

This certainly became clear in the time evolution plots in Sec. 5.1.5, where spatial distribution and maxima were characterized by a significant variability with time. Actually, Fig. 5.11a, which shows the density estimates for a time interval between Dec 1999 and Dec 2003 (also five years), is quite similar to Fig. 6.3 in terms of location of the PL activity main maximum. Even local areas of high density are comparable in both figures, for instance an area located around 65 °N and 3 °E. This might serve as an explanation of the change in density pattern from Wilhelmsen's set to the findings obtained when using the data set produced by Rojo et al. [2014]: the PL distribution is characterized by a natural variability in both space and time. Also, the Wilhelmsen set was based on only 5 years and 33 PL cases, thus representing only a fraction of the 14 years and 286 PL cases in Rojo et al. [2014]. It therefore seems unreasonable to assume that these two sets would provide us with the same observations regarding PL spatial distribution.

We have now presented some climatological factors and processes which provide favourable conditions for polar low formation and development, and related these to the Nordic Seas, and the PLs forming and developing here. By comparing our results to another PL study we have pointed to a natural variability in space and time, which PLs can be assumed to have.

6.1.2 Polar lows genesis

Local weather conditions like warm ocean currents and MCAOs discussed in the previous section are obviously related to polar low genesis, as these provide favourable conditions for formation. Rojo et al. [2014] made observations regarding this. They stated that the PL genesis density is larger along the warm currents of the Norwegian and Barents Seas. The maximum is located at the overlap of the Norwegian Atlantic Current (NwAC) and the West Spitsbergen Current (WSC) [Rojo et al., 2014].

The kernel density estimates of the genesis points in Fig. 5.5b can be compared to the display of Nordic Seas ocean currents in Fig. 6.1b in Sec.6.1.1. The high-density area of the estimates in Fig. 5.5b correlates well with the path of the NwAC, and the overlap between the NwAC and the WSC is located in the immediate proximity of the maximum density contour which was located around $70 - 72\,°$N and $10 - 13\,°$E. The impression given by Rojo et al. [2014] is that there should be a complete and more accurate spatial coincidence between the mentioned maximum and the ocean current overlap. However, the lack of this coincidence should not necessarily raise a cause for doubt in Rojo et al. [2014]'s statement. Given the dynamic system that is an ocean current, a plot is bound to operate with some degree of spatial error, and will most likely vary from representation to representation. Also, their statement were based on a histogram approach of the genesis spatial distribution, which degree of accuracy is not as high as for the kernel density estimation (which will be discussed further below). In addition to the area of global maximum, there are some local maxima in the Barents Sea which can also be connected to the NwAC by its branches in this region (see Fig. 5.5b and Fig. 6.1b). The main conclusion regarding the genesis locations is that they share a close spatial relation to the relatively warm SST which arises because of the NwAC and its branches, thus confirming the statement given by Rojo et al. [2014].

The genesis of the polar lows were also examined in Rojo et al. [2014], in a process similar to the histogram approach. The region of interest was divided into areas of 2.5° in latitude and 5° in longitude. The number of polar lows that formed in each area was then counted and the result was displayed in a figure which is reproduced here, Fig. 6.4.

As seen in Fig. 6.4, the area where most polar lows formed (14 cases) during the winter seasons $1999-2013$ is constrained by 70 - 72.5 °N and 10 - 15 °E. It is interesting to compare the performances of the histogram method used in Rojo et al. [2014] and the kernel density estimation method used in this thesis. For this purpose, another contour plot of the genesis estimates was made (Fig. 6.5), now with latitude and longitude grids chosen equally to the ones used in Rojo et al. [2014]'s histogram. This was to make visualization and comparison easier and more accurate.

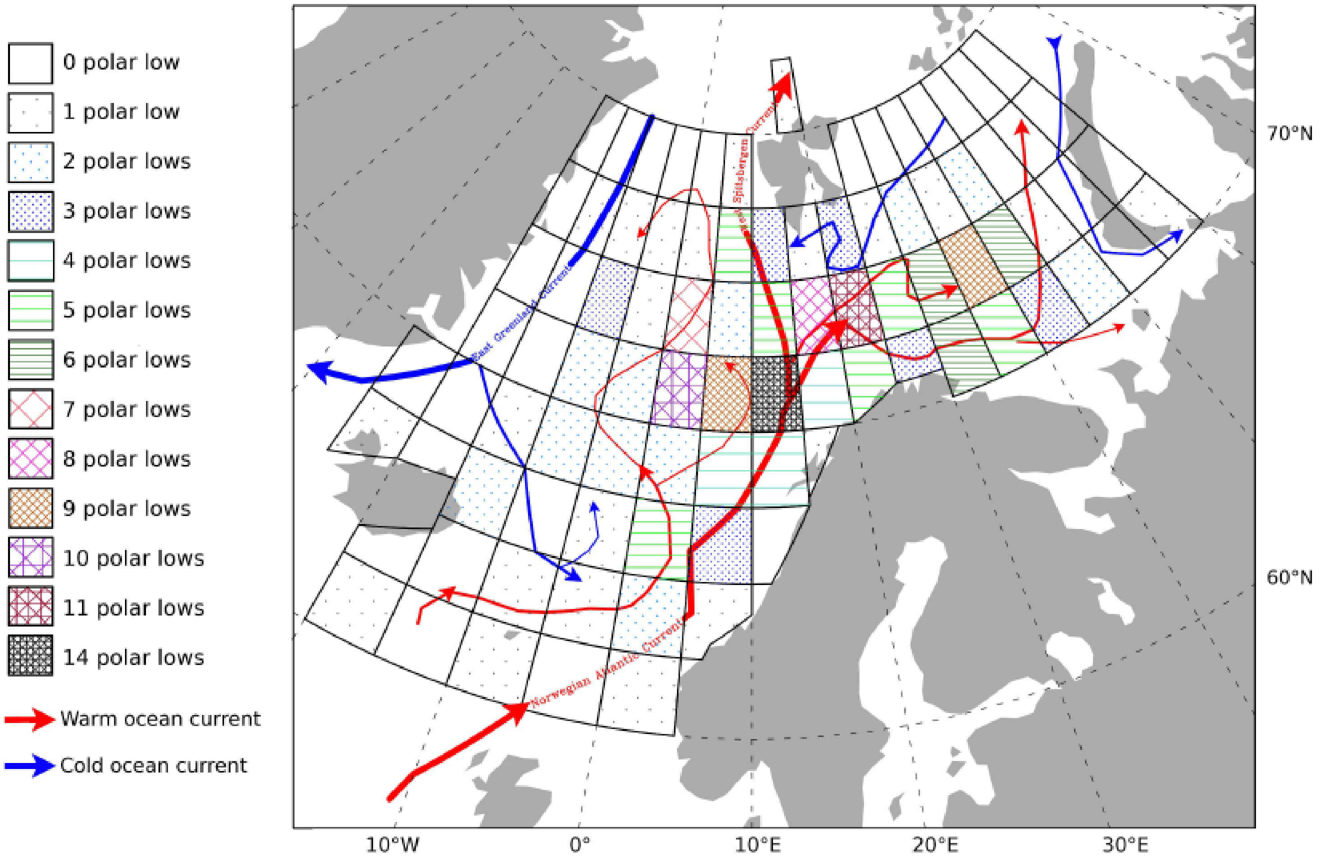


Figure 6.4: Area distribution of polar low genesis over the Nordic Seas from winter 1999/2000 to winter 2012/2013. Source: Rojo et al. [2014]

Comparing the quadrants, or boxes of Fig. 6.4 and Fig. 6.5 one can easily correlate several of them to each other. For instance the contour of highest density (deep red) in Fig. 6.5 is also constrained by 70 - 72.5 °N and 10 - 15 °E, like the 14 polar low quadrant in Fig. 6.4. The same applies for the quadrant 70 - 72.5 °N and 0 - 5 °E which holds a smaller contour of lower density, and relates to 10 polar lows in Fig. 6.4. However, when it comes to density contours which can not be contained within box boundaries, the differences in the methods become more clear. An example here is the four quadrants between 70 - 75 °N and 30 - 40 °E. The contour of the highest densities (green) in this region lies in the junction between the grids. This makes it difficult to compare with Fig. 6.4, where the region is divided into four boxes and marked according to how many polar lows occurred in each box.

Also, the spatial distribution in Fig. 6.5 gives a more accurate image of where *within the box* the frequencies were highest, whereas Fig. 6.4 fills the boxes out to its boundaries.

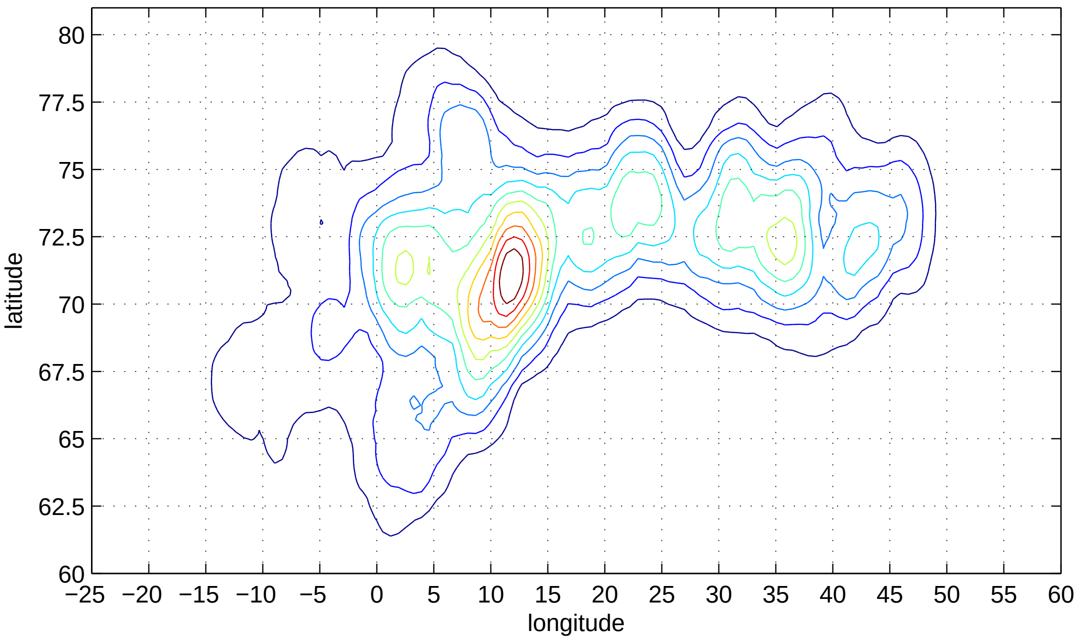


Figure 6.5: Kernel density estimates of the genesis points. Epanechnikov kernel and $h = 3$ are used. Latitude and longitude grids are chosen equally to those in the histogram approach by Rojo et al. [2014] (Fig. 6.4).

For instance, the previously mentioned contour located between 70 - 72.5 °N and 0 - 5 °E only takes up a small part of the box, as opposed to in Fig. 6.4, where the box is filled out, indicating that the entire region is a high-frequency area.

The overall impression from comparing the two figures Fig. 6.4 and Fig. 6.5 and thus the two methods of histogram and kernel density estimation is that there are advantages and disadvantages connected to both. The histogram provides a raw and simple way to display the coarse regions of where polar lows are most likely to form. The kernel density estimation gives a smoother and more dynamic picture of the exposed areas. One can see more clearly how the the densities vary spatially throughout the area of interest, which instantly provides a higher level of spatial resolution of the genesis distribution. Because it is smoothed it also has the ability to provide a higher level of statistical accuracy.

Of, course, it should be remembered that the two figures/methods provide us with different information regarding the polar low genesis areas. The histogram (Fig. 6.4) is direct and gives the exact number of how many polar lows were formed in the respective areas. The kernel method however, provides the probability density estimate based on these numbers.

To conclude, the choice of method thus depend on what is the main interest. It can however be determined that the kernel method for density estimation is in fact a smoother, and more accurate version of the histogram approach, with the capability of providing spatially resolved density estimates.

6.1.3 Polar lows dissipation

The kernel density estimates of polar low dissipation points (Fig. 5.6b) displayed a maximum stretched along the mid and northern coastline of Norway, with the highest densities located just south of Lofoten. A large part of the dissipation points contributing to this maximum are probably PLs forming in relatively close proximity to the western coast of Norway or the coast of Finmark (see Fig. 5.5b). The combination of high genesis density in this area, their proximity to the coast and a general southwards displacement (for 70% of the PLs according to [Rojo et al., 2014]) makes the coastline south of Lofoten exposed with respect to PL dissipation.

However, this maximum is not to be considered as the only area of significance. As seen from the scatter plot of genesis points (Fig. 5.6b), a substantial fraction of the dissipation points occur over the ocean, but because these are spread and not as densely distributed as along the coast, they are obviously not as visible in the density plot. The results actually showed that a larger fraction of the PLs dissipated over ocean areas than land in the Nordic Seas. Of 286 PLs, about 117, or approximately 40% of the cases studied dissipated over land.

As Rasmussen and Turner [2003] stated, most PLs, or at least phases of them, are associated with convection which can happen because of relatively warm sea surfaces and high humidity over the sea. If the system makes landfall, meaning that it is steered across a coastline, it will be bound over a cold and often snow/ice covered surface. This will result in a cooling and drying of the lowest layer above the surface, which weakens the convection until it eventually dies out. This process is associated with a rapid decay of the polar low system until it dissipates completely, which happens within a short period of time [Rasmussen and Turner, 2003].

The same will obviously happen for PLs being steered across ice-covered areas of the sea. The contours of the dissipation points over maritime areas can therefore be compared to a plot of typical winter sea-ice cover during the time interval of the observations. Fig. 6.6 is a Google Earth image displaying the sea ice extent in the Nordic Seas as it appeared in March 2005. This time is chosen randomly as a representative of the time interval $1999 - 2013$ for which the dissipation points apply.

Comparing Fig. 6.6 to the contours of the dissipation points (Fig. 5.6b), one can actually correlate several spatial features. An overall correlation between the ice edges (Fig. 6.6) and the extension of the density contours (Fig. 5.6b) can be observed. The contours do not extend further than approximately 10 °W (which is also apparent from the scatter plot in Fig. 5.6a, where very few points are located beyond this longitude). Further south than approximately 75 °N, the ice edge (Fig. 6.6) is located west of this longitude. It is therefore reasonable to suggest that many of the dissipation points in this area (south of 75 °N and close to the 10 °W meridian) are results of PLs suddenly being located above sea ice connected to the east coast of Greenland. Further north of 75 °N, and west of Svalbard, the ice extends to the prime meridian (0 °E), and the area located around this longitude is visible as a small ‘tip’ in both Fig. 6.6 and Fig. 5.6b, again representing a spatial correlation between sea ice and PL dissipation.

The most easily correlated spatial sea ice feature from Fig. 6.6 is probably the protruding ice tip below Svalbard (around the 20 °E longitude). This is clearly visible in Fig. 5.6b as a depression of the contour in the same area. Also the scatter plot (Fig. 5.6a) shows an almost absence of points within this depression area, but several points along it, probably representing PLs which encountered the sea ice and decayed.

In the Barents Sea, the sea ice edge and the contours seem to follow each other quite well, with a number of PLs presumably dissipating over the ice connected to Novaya Zemlya and the Russian coast, east of the Kola Peninsula.

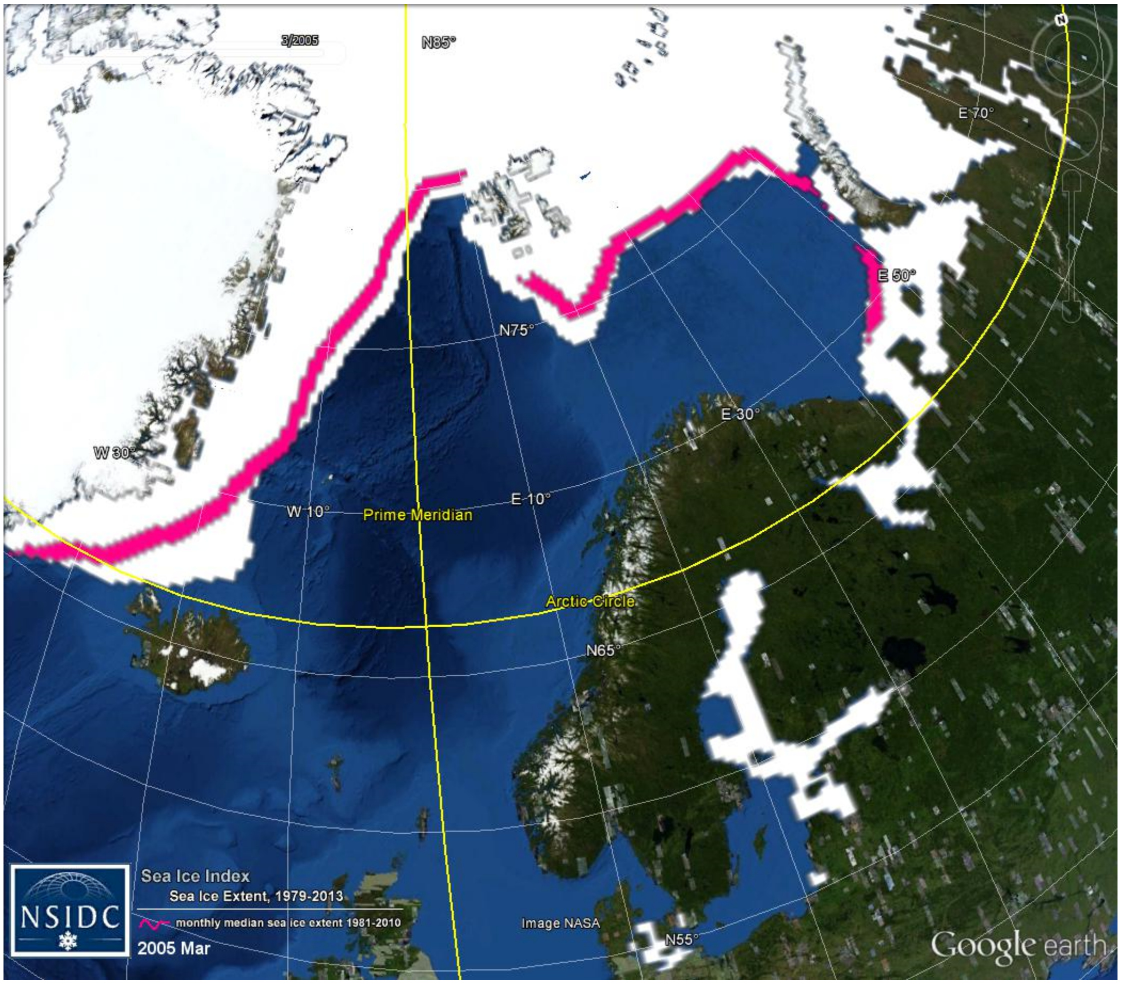


Figure 6.6: Sea ice extent in the Nordic Seas March 2005. The pink line indicates the monthly median sea ice extent for the March months (maximum sea ice extents) of 1981-2010. Source: National Snow & Ice Data Center, obtained from `https://nsidc.org/data/google_earth/`.

Associated with the sea ice discussion, it should be noted that the sea ice cover in certain regions of the Nordic Seas is characterized by a great variability in time, causing occasional dramatic changes from one year to the next. This variability is not clear in

Fig. 6.6. Therefore, another figure displaying sea ice extent is included. Fig. 6.7 shows the yearly maximums (April) of sea ice edges in the Barents Sea, as they appeared in the years 2001 through 2011. The line marking the ice edge is plotted with different colors according to year, making the yearly variability evident.

From Fig. 6.7, it can be observed that the ice tip protruding from the southern tip of Svalbard is relatively stable from year to year, no major variations in ice extent are visible. The dissipation discussed above associated with this area can therefore still be assumed to be related to sea ice. In the eastern part of the Barents Sea the situation is different. Stronger annual variations are evident, and the year 2003 (dark blue line) represents an extreme. From April 2002 (green line) to April 2003, the ice extended several degrees southwards, for then to retreat considerably again, to April 2004 (yellow line).

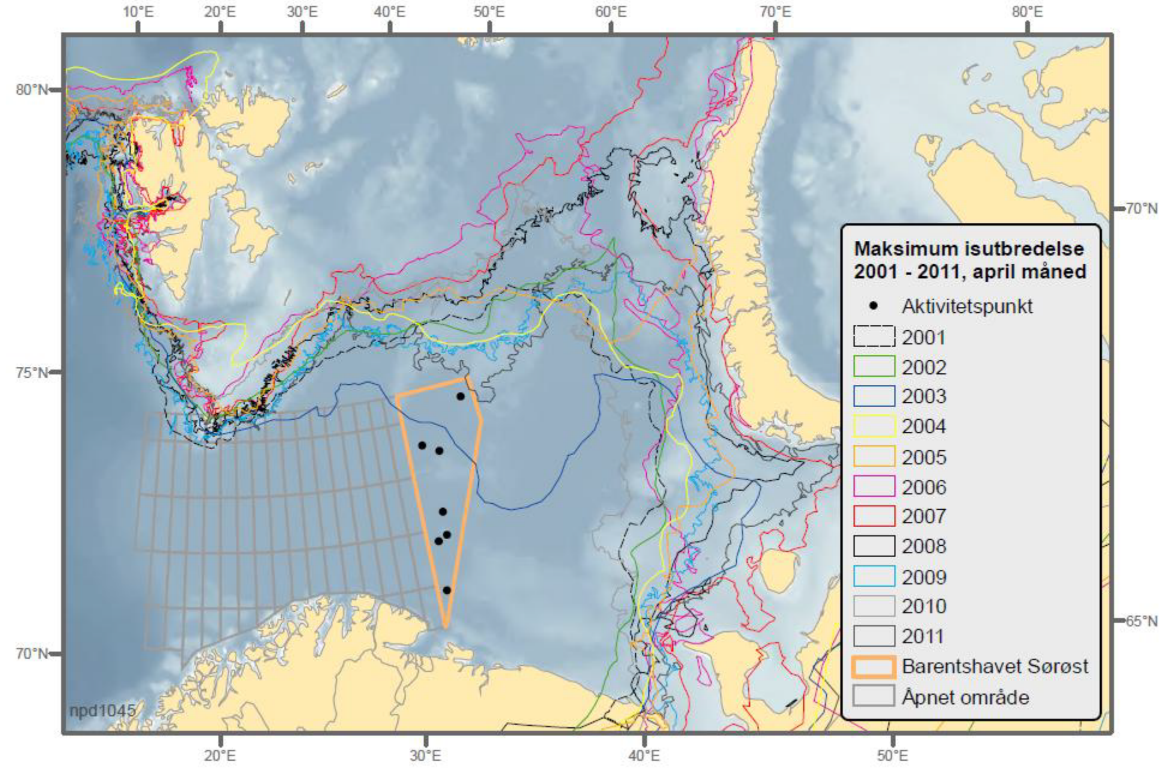


Figure 6.7: Maximum sea ice extent in the Barents Sea for the month of April, in the period 2001 – 2011. Source: Iden et al. [2012].

Relating the 2003 extreme to the dissipation points turned out to be difficult. No clear indications of correlations between the 2003 ice cover and the PL dissipation are detectable in Fig. 5.6b. This is not very surprising, due to the fact that this extreme is constrained to only one year, i.e., if there are clear relations between the 2003 extreme and the dissipation distribution, they probably “drown” in the rest of the observations from when ice cover were not as extensive. However, the two-year time evolution plot can be useful here, as it presents the PL spatial distributions on a higher resolved temporal scale. Revisiting Fig. 5.12, the plots for years 2001 – 2003 (Fig. 5.12b) and year 2003 – 2005 (Fig. 5.12c), and the transition between them can be observed. The transition from Fig. 5.12b to Fig. 5.12c does not display very obvious indications of the 2003 extreme. However, density contours are in general stretching further north in the Barents Sea in Fig. 5.12b, while they retreat several degrees southwards in Fig. 5.12c. Although not unmistakable, this may be an indication of the increase in sea ice cover.

Another possible reason for dissipation points over ocean areas (not connected to sea ice) is simply the steering of PL systems away from areas of favourable conditions regarding intensification and development. For instance, PLs forming along the Norwegian Atlantic Current and its branches which then are steered away from these areas of increased SST will experience a lower air-sea temperature gradient and thus an increased level of static stability. Similar to the PLs entering a coastline, this will lead to weakened convection. A continued situation under these conditions will eventually cause the PL to decay and quickly disintegrate.

After discussing the kernel density estimates for PL genesis, dissipation and the general spatial distribution of PL trajectories, a decent assessment of the PL activity in the Nordic Seas has been obtained. Actually, by viewing the three density plots for genesis, overall spatial distribution and dissipation (Fig. 5.5b, Fig. 5.4 and Fig. 5.6b, respectively) together, an intuitive timeline of PL spatial maximum becomes apparent. The PLs form around 72 °N over the ocean, northwest of, and in relative proximity to the coastal areas of Vesterålen. They intensify and travel southwards, approaching the coastal regions of Vesterålen and Lofoten. In their last phase, they hit the coast of Nordland, weaken in strength, and eventually dissipate completely.

6.1.4 Time evolution

The kernel density estimates from different time intervals provided some interesting results on how high-density polar low areas develop spatially over time. From the impression that polar lows were quite concentrated in a certain region when using the complete trajectory data, the division into various numbers of time segments allowed for a different picture, where the maximum shifted from different locations in the Norwegian and Barents Sea.

As an initial remark, it should be mentioned that these variations should not be interpreted as a general pattern in the time development of polar low activity. For that the data set is still too sparse. The results represent the development from relatively few years, and might as well represent a random, natural variability. To draw conclusions regarding statistical trends one would probably need a much longer data set. Nonetheless, the results are very interesting, and the statistical analysis provides the possibility to draw conclusions on shorter time scales. Suggestions providing possible explanations for the development can also be discussed.

In most of the figures showing the time evolution, one could easily get the impression that the main maximum of polar low activity is one the move northeastwards, from the Norwegian Sea, and into the Barents Sea (Fig. 5.8c, Fig. 5.9d, Fig. 5.10e, Fig. 5.12g). Recent years research on polar low activity have included relations to climate change, and these connections can be worth looking into in this thesis. Zahn and von Storch [2010, 2012] and Woollings et al. [2012] represent some of the research within this field. Zahn and von Storch [2010] used a series of regional climate model simulations to downscale a set of global climate change scenarios from the International Panel of Climate Change (IPCC). Formation of polar lows systems in the North Atlantic were simulated and the individual cases counted. In projections for the end of the twenty-first century, they found that the frequency of polar lows decreased significantly, but also that their mean genesis region shifted to the north, as a response to increased atmospheric greenhouse gas (GHG) concentration [Zahn and von Storch, 2010]. This change can be a consequence of increase in the North Atlantic sea surface temperature and mid-troposphere temperature, where the latter is found to rise faster. This causes the temperature gradient between sea surface and air to decrease, and the resulting vertical atmospheric static stability to increase [Woollings et al., 2012; Zahn and von Storch, 2010, 2012]. Because polar low systems form and intensify in environments characterized by instabilities, stronger stability will act as a dampening factor for polar low formation. Zahn and von Storch [2010] viewed these results as a rare example of a decrease in extreme weather connected to climate change, as opposed to the opposite which is more common.

The downscaled IPCC scenarios using various regional climate models and the related polar low density distribution are shown in Fig. 6.8 (adopted from Zahn and von Storch [2010]). The downscaling was performed on two periods of transient simulations conducted for the Fourth Assessment Report (AR4) of the IPCC with the ECHAM5/MPI-OM model. The first period which was downscaled was a 30-year control period driven by GHG concentrations valid for $1960 - 1989$, which served as a reference. It is seen in Fig. 6.8 as C20. The second period was also 30 years, and was driven by three of the most commonly used IPCC-AR4 future GHG concentration scenarios. These can be seen in Fig. 6.8 as B1, A1B and A2 [Zahn and von Storch, 2010].

In addition to simulations based on future climate, Zahn and von Storch [2010] also did a downscaling analysis of the period $1948 - 2006$ using a regional climate model which was processing NCEP/NCAR re-analysis data. This simulation was called REA. Comparison between the REA and empirical polar low data from certain available time periods proved that downscaling large-scale information with a regional climate model generated realistic statistics of polar lows [Zahn and von Storch, 2010].

The number of polar lows per polar low season found in C20 were considerably lower than the number found in the REA simulation, due to an increase in vertical stability. The estimate of future change in polar low frequency was derived by comparing the three IPCC scenarios B1, A1B and A2 with the control simulation C20. All three scenarios present a reduction in the number of polar low events per season during the years $2070 - 2099$ [Zahn and von Storch, 2010].

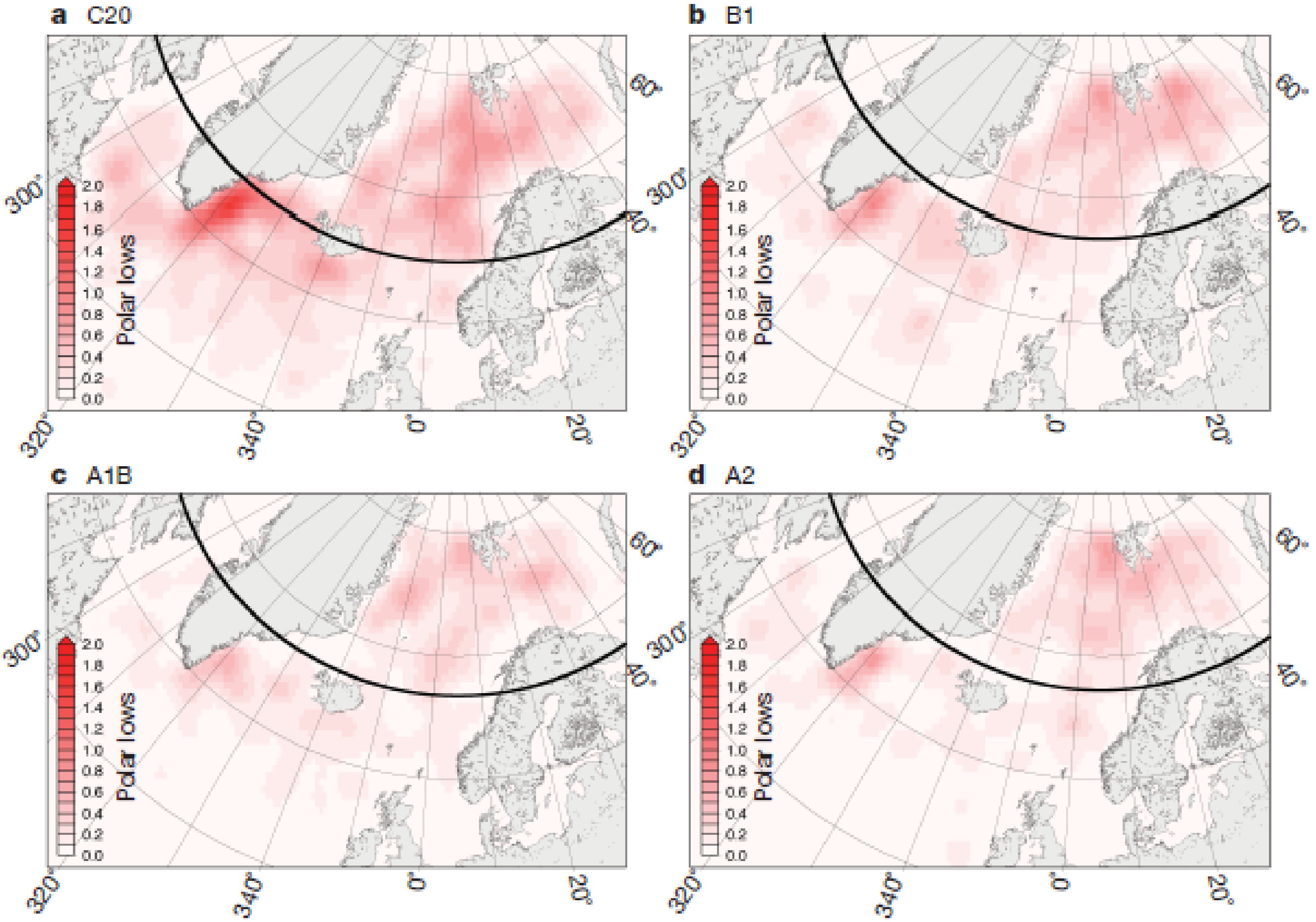


Figure 6.8: Polar low density distribution as a response to a warmer climate. IPCC scenarios: a) C20; b) B1; c) A1B; d) A2. The colour scale measures modelled polar lows per 2500 km^2. The thick black line in each panel indicates the mean latitude of polar low genesis. Source: Zahn and von Storch [2010].

In addition to the frequency of polar low events, the warmed climate also affects their spatial distribution (Fig. 6.8). A projected northward shift becomes apparent both from Fig. 6.8 and the values of the mean latitude of polar low genesis which is 64.9 °N in C20 and 66.8 °N, 66.8 °N and 67.3 °N in B1, A1B and A2, respectively. As mentioned earlier, the dynamical forcing mechanisms causing this weakening and northward shift are related to the prognosed increase in tropospheric temperatures, which happens faster than the comparatively moderate warming of the North Atlantic. This tendency causes increased vertical stability, which means less favourable conditions for polar low formation. To indicate this stability Zahn and von Storch [2010] did calculations on the temperature difference of the area in question, and time-averaged ice-free SST and 500-hPa air temperatures over the maritime northern North Atlantic of the simulation area during October-March, the main polar low months. The lower the difference this was, the higher stability it indicated.

Based on the calculations, Zahn and von Storch [2010] could state it obvious that the mean vertical stability were predicted to rise with an increase in GHG concentration and climate warming [Zahn and von Storch, 2010].

The scenarios also projected a northward shift in the Arctic ice edge in the North Atlantic, which presently is known as an area of shallow baroclinic zones which provide excellent initial conditions for polar lows. Following this shift is a northward shift of cold air outbreaks, and again, polar low formation [Zahn and von Storch, 2010]. Zahn and von Storch [2012] and Woollings et al. [2012] reached the same conclusions regarding increase of vertical static stability and decrease in polar low activity.

Could the northeastward shifting nature of the PLs investigated in this thesis be a precursor of the development projected in Zahn and von Storch [2010, 2012]? Or do the results just indicate natural and coincidental shifts in maximum areas? As stated earlier, the data set is currently too sparse to conclude on trends, and assumptions should therefore not be made.

The time evolution was also commented on in Rojo et al. [2014]: "In the recent years, the number of polar lows has increased especially in late winter and in two specific areas: to the west of Spitzbergen and in the Barents Sea" [Rojo et al., 2014]. This was explained by the retreat of sea ice in winter and its late formation in autumn in recent years, which causes an extension in the area where polar lows can form and develop [Rojo et al., 2014]. Of course, this too is assumed to be related to climate change and increasing temperatures.

6.2 Clustering analysis

The results of the clustering analysis in Fig. 5.15, Fig. 5.16 and Fig. 5.17 displayed some interesting and informative groups of the trajectory data set. Starting with Fig. 5.15, this figure shows the results of the clustering process when using 3 predetermined clusters. Some polar low characteristics are clearly separated between the clusters. In cluster 1 (Fig. 5.15a) the tracks propagate almost exclusively to the east, or with eastern direction as their main component. Cluster 2 (Fig. 5.15b) reveals a similar exclusivity, practically all trajectories are characterized by a southward directed displacement, with small components in the west, or in a few cases, east direction. The last cluster (cluster 3, Fig. 5.15c) is more chaotic. The cluster contains many short tracks in close proximity to the Norwegian and Kola Peninsula coastline, but still has some long tracks forming in the west ($0 - 20\,°$W).

Fig. 5.16 displays the clustering result when 5 clusters were used. By looking at this figure it seems as the clustering in addition to direction of propagation can focus on similarities within trajectory length, and degree of straightness, or curvature. Cluster 1 (Fig. 5.16a) is quite mixed regarding both straightness and length. Most of the trajectories have a net displacement southwards, even though some northwards cases exist too. Cluster 2 (Fig. 5.16b) is probably the cluster from which most qualities can be easily observed. The trajectories are mainly directed eastwards, most of them have quite a high degree of straightness and the travel distances are long. The trajectories in cluster 3 (Fig. 5.16c) are probably the most curved and erratic with regards to displacement direction. They are mixed with respect to length with long trajectories forming mainly between 0 °W and 20 °W, and short ones forming closer to the coastline of Norway and the Kola Peninsula and in the Barents Sea. Cluster 4 (Fig. 5.16d) seems to consist of short, and relatively straight trajectories, but the propagation displacements seem to sprawl in several different directions. Only one of the trajectories in cluster 4 makes landfall. In cluster 5 (Fig. 5.16e) the situation is similar to that of cluster 1 (Fig. 5.16a). A variety of trajectories are displayed, regarding length, direction and straightness. However, there are not many of the very long tracks.

Fig. 5.17 presented the results of the last clustering process, with 6 predetermined clusters. The first cluster (Fig. 5.17a) consists mostly of tracks of similar length, and few extremes regarding length (i.e., very short or very long) can be seen. Most of them have a net displacement to the east, and the overall degree of straightness is not very high. The trajectories of cluster 2 (Fig. 5.17b) are concentrated along the coast, and most of them have a southwards displacement. They are however highly erratic in both length and degree of straightness. A large number of the cluster 2 tracks dissipate over land, and most of these tend to be quite short. Cluster 3 trajectories (Fig. 5.17c) are probably the most curved of all the 6 clusters. Other than that, their length and direction varies greatly. Cluster 4 (Fig. 5.17d) has a large number of tracks with a very long travel distance. Most of them have a large degree of straightness and travel direction is either mainly southwards or mainly eastwards. The tracks belonging to cluster 5 (Fig. 5.17e) are mostly short, straight and directed in the eastward direction. Cluster 6 (Fig. 5.17f) also displays quite curved and short tracks. Many of them have a net displacement to the west, with a few cases of eastward directed systems in the Barents Sea. Just one of the tracks dissipates over land.

The clustering plots discussed above can be related to some results stated in Rojo et al. [2014]. Rojo et al. [2014] found that area of genesis in a large degree influences several polar low characteristics. PLs forming in the western part of the Nordic Seas ($0-20$ °W) tend to live longer and travel longer distances than PLs forming in the Barents Sea. This can be observed in Fig. 5.15a and Fig. 5.15c, for instance, where the longest tracks in the clusters are those forming west of the zero meridian. In comparison, the PLs of the same cluster, but with genesis in the Barents Sea are shorter, confirming the statement of Rojo et al. [2014]. Had the clustering process been performed with the ‘nozero’ option, this would have been more clear, because the clustering is then based strongly on genesis region and spatial distribution. In this case however, we settle with viewing the dissimilarities of track length within the clusters.

Rojo et al. [2014] also stated that PLs forming between $0-10$ °E and $70-72.5$ °N are also characterized by longer lifetime and distance. Fig. 5.17d might serve as an example of this. Although not very many, the cluster in this plot contain long trajectories forming in this particular region, many of them eventually dissipating over the coast of Noway. The cluster also include the special case of an extremely long PL trajectory which formed between $70 - 72.5$ °N and travelled all the way through the Norwegian Sea before it dissipated in the North Sea, close to the coastal border between Germany and the Netherlands.

Close to the coast of Norway and the Kola Peninsula the shortest PLs form, with regard to both lifetime and distance travelled. This is because of their proximity to the coast and PLs tendency to decay quickly when they hit land [Rojo et al., 2014]. Clusters with these qualities (short travel distance and formation close to land) can be seen for instance in Fig. 5.16c and Fig. 5.17b. In these clusters, the high concentration areas tend to follow the coastline of Northern Norway and the Kola Peninsula, and are characterized by quite short trajectories. Also mutual for the clusters is that a large fraction of the trajectories decay when they hit the closely located coastal areas.

Some of the clusters can also be used to confirm, and shed new light on the results obtained by kernel density estimation, discussed in Sec. 6.1 As the high density contours were well aligned with the Norwegian coastline in Fig. 6.1a, the most numerous clusters also tend to be. Areas of large overlapping of tracks in these clusters correlate well with the density maximum of polar low activity seen in Fig. 6.1a. The area outside Lofoten/Vesterålen can thus be confirmed as a highly exposed area. Examples of these clusters can be seen in the following figures: Fig. 5.15c, Fig. 5.16c, and Fig. 5.17b. The densest areas of these clusters are characterized by trajectories forming rather close to the coast, which during their lifetime travel past or dissipate in the coastal areas of Lofoten and Vesterålen (in the approximate area between $67 - 72\,°$N and $8 - 15\,°$E). Because of their genesis location being close to the coast, their travel distances are mostly short.

The clustering analysis has thus shed new light on, and expanded the knowledge obtained from the kernel density estimates regarding the high density area of polar low activity in the Nordic Seas. The results have confirmed this area as a region of high PL activity, regarding both passing and decaying PLs. It has also provided the ability to determine characteristics of the trajectories themselves in this area. For instance that they form in close proximity to the coastline, and that their travel distances therefore mostly are short.

6.3 Analysis of polar lows and weather regime

From Fig. 5.18 some observations could be made regarding polar low trajectories and their corresponding weather regime. Connections between PL trajectories and WR were determined by Rojo et al. [2014]. For instance, it was observed that the zonal flow regimes (NAO+ and NAO-) are more favourable for multiple PL development than the meridional flow regimes (AR and SB). NAO- and AR are however the most favourable regimes for PL formation in general (Fig. 5.18b and Fig. 5.18c) [Rojo et al., 2014].

Regarding the trajectories direction of movement, it was stated by Rojo et al. [2014] that this tend to be controlled by the large-scale flow in the lowest layers. Significant differences are observed both for the negative phase of the NAO and the SB regime, compared to the other weather regimes. The NAO- regime seem to favour displacement mainly in the westward and southwestward direction (Fig. 5.18b). Because NAO- also tend to favour the formation of multiple PLs, this is also the WR for which the highest number of polar lows form under [Rojo et al., 2014]. Fewest polar lows form under the SB regime (Fig. 5.18d), but most of them share the same characteristics regarding direction, their displacements are almost exclusively directed eastwards. The eastward motion leads to them having a greater impact on the Norwegian coast compared to PLs belonging to other WRs [Rojo et al., 2014].

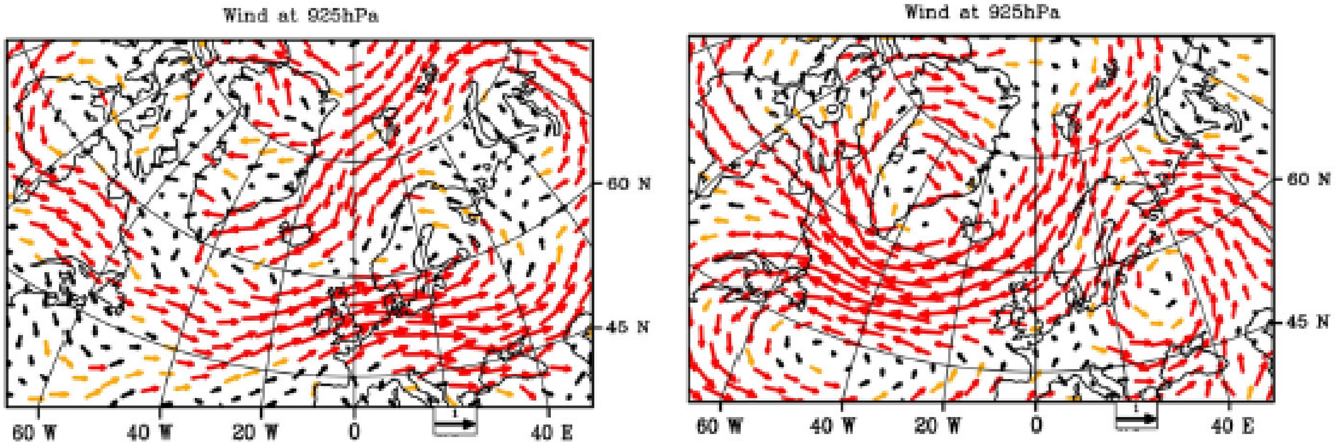
Mallet et al. [2013] documented the association between PLs and the North Atlantic-Europe WR on the development of PLs over the Nordic Seas during the years $1999-2011$. They showed the atmospheric conditions prevailing on PL days in the Norwegian and Barents Sea (based on the data set by Noer et al. [2011]) for each regime. Among others, the standardized anomalies of wind at 925 hPa were displayed for PL formation days classified as AR, NAO-, NAO+ and SB. These are displayed in Fig. 6.9.

For NAO+ cases the wind anomaly represents a northeasterly flow along the ice edge (Fig. 6.9a). For NAO- cases (Fig. 6.9b), the low-level wind also has a northeasterly anomaly. The wind is however stronger in the Norwegian and Barents Sea, especially close to the Norwegian coastline and the Kola Pensinsula coast. AR cases are related to a wind anomaly flowing southwards, with a southeastwards component south of approximately 65 °N (Fig. 6.9c). The wind anomaly for SB cases represent the most different wind conditions compared to the other regimes. The low-level wind displays a westerly flow anomaly south of 75 °N and a northerly flow to the north of this latitude (Fig. 6.9d) [Mallet et al., 2013].

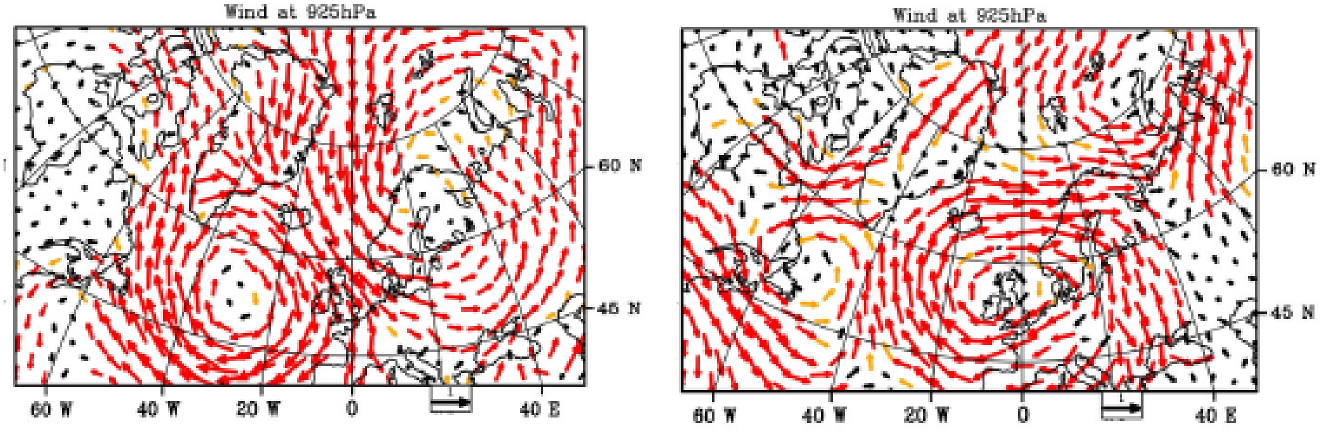
Both the NAO- PLs and the SB PLs in Fig. 5.18b and Fig. 5.18d, respectively, can be closely related to their respective displays of low-level wind anomaly (Fig. 6.9b and Fig. 6.9d). By comparing these figures, the observations made by Rojo et al. [2014] are yet again confirmed, as a large portion of the PLs in Fig. 5.18b follow the large scale northeasterly flow. In Fig. 5.18d the PL activity is even more influenced by the unusual westerly flow of the low level wind. The PLs forming furthest to the north in Fig. 5.18d can also be seen to be affected by the large scale northerly flow occurring north of 75 °N, or they are affected of both the northerly and westerly flow (Fig. 6.9d) during their lifespan.

For the NAO+ PLs (Fig. 5.18a), the large-scale flow is not as apparent as for NAO- and SB, but some systems are seen to follow the northeasterly flow direction (Fig. 6.9a). The wind is not as significant in the entire Nordic Seas areas as in for instance the NAO- regime, and this might be a reason for the weakened influence. There are also some PLs in Fig. 5.18a which are directed almost completely eastwards. A possible explanation for this is that these are PLs which formed and developed during a weather regime change, for instance between NAO+ and SB or opposite.

The southwards flow in AR regimes is apparent in many of the PL trajectories in Fig. 5.18c. Between 60 °N and 70 °N, one can also observe the southeastwards component mentioned above, which result in the landfall of many PLs along the Norwegian coast in this latitudinal area.



(a) NAO+ **(b)** NAO-



(c) AR **(d)** SB

Wind significance → < 80% → 80% - 90% → > 90%

Figure 6.9: Composite standardized anomalies of wind at 925 hPa in cold season months (Oct-March) for key days classified as NAO+, NAO-, AR and SB (Noer et al. [2011]'s list). The significance levels of the wind are indicated by the color of the arrows. Source: Mallet et al. [2013].

We have now seen how the large scale wind flow during various weather regimes influences the main directions of PL trajectories. Mallet et al. [2013] also investigated other synoptic scale climatological anomalies of the WRs and how these to a various degree provided favourable conditions for PL formation and development.

Relating PLs to synoptic scale WRs will introduce important elements in PL forecasting methodology, as the main challenge of PL forecasting is often connected to their small scale and detection difficulty in standard forecasting models. Detailed knowledge of significant large-scale anomaly patterns over large areas centered over PL genesis zones (which are easier to detect and forecast), will in turn improve the detection and forecasting possibilities of PLs [Mallet et al., 2013].

Chapter 7

Conclusions

In this thesis, polar low trajectories during the winter seasons of $1999 - 2013$ in the Nordic Seas were analysed by the use of various statistical methods. The most important method used was the kernel density estimation, which was applied to various segments of the available data. It was estimated by using the complete trajectory data set (latitude, longitude position for every trajectory during $1999 - 2013$) that a coastal area outside Lofoten and Vesterålen experienced the highest densities of polar low activity, with areas of lower densities in the Barents Sea. The high-density areas were suggested to correlate with the Norwegian Atlantic Current and its branches in the Nordic Seas. It was also stated that these polar lows formed during so-called Marine Cold Air Outbreaks, within a northerly flow in the Norwegian Sea, and a northeasterly flow in the Barents Sea, which was based on previous research (e.g., from Mallet et al. [2013]).

Kernel estimates were also made based only on the genesis points and the dissipation points of the polar lows data set, respectively. For the polar lows genesis, the main concentration was located a few degrees north of the maximum when using the complete data set, which correlated well with the spatial distribution of the Norwegian Atlantic Current. The genesis estimates were also compared to a histogram approach carried out using the same data points in Rojo et al. [2014], pointing to the kernel density estimation as a more accurate, smoother and more dynamic method of non-parametric density estimation.

The estimates of the dissipation points displayed a very dense area along the coast of Norway, especially south of Lofoten. Although more dispersed, many of the polar low trajectories also dissipated over what appeared to be open ocean areas. These areas were explored with regard to sea ice extent, and density contours (indicating occurrence of dissipation points) were found to align well with the sea ice edges in several locations. It could be stated that polar lows tend to dissipate shortly after encountering land areas and sea ice alike, due to the loss of the important air-sea temperature gradient, which in large part is responsible for their development.

Kernel density estimation were also carried out on various time intervals of the data set, to investigate how the polar low spatial distribution developed with time. By dividing the data points into several equally sized sections (with regards to data points), and also by two-year intervals, it became evident that the heavily exposed areas were not constant in time and space, but shifted throughout the Nordic Seas with time. In several of the time evolution results, the main maximum appeared to "end" in the Barents Sea, which were discussed as a possible indication of research results (e.g., from Zahn and von Storch [2010]) stating that the distribution of polar lows will move northwards as a result of climate change. However, nothing of this sort could be assumed, as the data set time range is currently too short.

To our knowledge, this is the first time this type of statistical method have been applied to polar low data, resulting in a new and very clear representation of the polar low spatial distribution in the Nordic Seas. The observations revealed the locations of exposed areas, and visually confirmed polar lows as highly dynamic systems.

In addition to kernel density estimation, a curve clustering analysis of the polar low trajectories was carried out. While the kernel density estimates provided results on spatial distribution, the clustering process focused on shape characteristics of the polar low trajectories themselves. The graphical representation of the clustering results displayed the trajectories divided into several groups, based on significant differences in travel direction, travel displacement and degree of straightness. These observations could be associated to previous internal results of the thesis, and to previous external research results. For instance, the coastal area outside Lofoten/Vesterålen could be confirmed as a highly exposed area, by observing the most numerous clusters in this area. Results obtained by Rojo et al. [2014] regarding characteristics and genesis location could also be observed from several clusters. For instance, long trajectories forming in the western part of the Nordic Seas, and short trajectories forming close to the coastline of Norway and the Kola Peninsula.

The polar low cases between 1999 – 2013 were also investigated in relation to the large scale weather regime they corresponded to. A visual representation of the trajectories forming under the negative and positive North Atlantic Oscillation, the Atlantic Ridge, and the Scandinavian Blocking regime, respectively, showed how the trajectory properties varied from regime to regime. The trajectories were then compared to the typical large scale wind flow of the various regimes, and in many cases correlation between trajectory direction and wind flow could be observed. An increased knowledge of polar lows associations to large scale weather patterns were concluded to represent an improvement in the polar low forecasting methodology. Because large scale weather conditions are easier to detect and forecast via standard forecasting models, the ability to relate these conditions to probabilities in space and time for polar low formation and development would play a role in improving polar low forecasting.

The findings in this thesis represent highly relevant information, which can be of both practical and theoretical use. For risk management purposes within the petroleum industry the observations can provide a basic idea of areas characterized by higher probability of encountering polar lows. This information can also be valuable for players within the fishing industry, seeing as polar lows may present especially hazardous conditions for fishermen and vessels operating in the Nordic Seas. Forecasters may also make use of the results. Not only do they provide a clear visual representation of the polar low activity, they can also serve as indications of which areas to put extra focus on during detection and forecasting processes.

As mentioned earlier, the data set used in this thesis (provided by Rojo et al. [2014]) is still to sparse to assume statistical trends regarding polar lows and their distribution. However, the data set is to our knowledge the currently longest set of polar low data in the Nordic Seas, which makes it an important contribution to polar low research. The data, and the findings in this thesis, for which these data formed the basis, may be combined with, and compared to future research on polar lows to possibly indicate trends, e.g., the predictions of Zahn and von Storch [2010].

Although extensive, it should be mentioned that this thesis represents only a part of the possible analysis which can be carried out using the Rojo et al. [2014] data set. The main focus in this thesis was on the trajectories, and analysis was performed mainly by the use of the trajectory coordinates. Due to time constraints, only the weather regime data served as an additional polar low property to analyse. However, the data set provided by Rojo et al. [2014] contains several other polar lows properties, e.g., size (diameter), distance travelled, duration, direction, propagation speed and associated wind velocities. These could form the basis for further analysis. For instance, the intensity of polar lows could be investigated by applying histogram or kernel density methods on wind speeds to explore whether certain areas are characterized by especially intense polar lows. The wind speeds of individual trajectories could also be explored for potential patterns, e.g., some sort of periodicity from genesis to dissipation. The same applies for the sizes of polar lows (diameter of the fully developed circular cloud band), where one could investigate if there exist some pattern in the growth or spatial decrease of polar lows during their lifetime. Properties like diameter, travel distance, lifetime and propagation speed were analysed according to genesis area in Rojo et al. [2014], but it could be interesting to also relate these to the full trajectories, for instance to study the shortest and longest trajectories (based on both duration and distance) separately. The trajectory coordinates could also be put to further use, for instance by calculating the average direction of each trajectory, then averaging over all to make visual representations of average travel direction for all trajectories. As mentioned, the data set provide a range of possibilities regarding polar low research in addition to the observations made in this thesis.

Bibliography

C.M. Bishop. *Pattern Recognition and Machine Learning.* Springer, 2007.

A.-M. Blechschmidt. A 2-year climatology of polar low events over the Nordic Seas from satellite remote sensing. *Geophysical Research Letters*, 35(9), 2008.

A.-M. Blechschmidt, S. Bakan, and H. Graßl. Large-scale atmospheric circulation patterns during polar low events over the Nordic Seas. *Journal of Geophysical Research: Atmospheres*, 114(D6):n/a–n/a, 2009. ISSN 2156-2202. doi: 10.1029/2008JD010865. URL `http://dx.doi.org/10.1029/2008JD010865`.

A.W. Bowman. An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, 71(2):353–360, 1984. doi: 10.1093/biomet/71.2.353.

T.J. Bracegirdle and S.L. Gray. An objective climatology of the dynamical forcing of polar lows in the Nordic Seas. *International Journal of Climatology*, 28(14):1903–1919, 2008. ISSN 1097-0088. doi: 10.1002/joc.1686. URL `http://dx.doi.org/10.1002/joc.1686`.

S. Businger. The synoptic climatology of polar low outbreaks. *Tellus A*, 37(5):419–432, 1985.

S. Businger and R.J. Reed. Cyclogenesis in cold air masses. *Weather and forecasting*, 4, 1989a.

S. Businger and R.J. Reed. Polar lows. *Polar and arctic lows*, pages 3–45, 1989b.

S.J. Camargo, A.W. Robertson, S.J. Gaffney, P. Smyth, and M. Ghil. Cluster Analysis of Typhoon Tracks. Part I: General Properties. *Journal of Climate*, 20(14), 2007.

C. Cassou. Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. *Nature*, 2008. doi: 10.1038/nature07286.

C.N. Duncan. A numerical investigation of polar lows. *Quarterly Journal of the Royal Meteorological Society*, 103(436):255–267, 1977.

C.N. Duncan. Baroclinic instability in a reversed shear-flow. *Meteorological Magazine*, 107 (1266):17, 1978.

K.B. Dysthe and A. Harbitz. Big waves from polar lows? *Tellus A*, 39(5):500–508, 1987.

H. Engedahl, G. Eriksrød, and B. Ådlandsvik. A climatological oceanoghaphic archive covering the Nordic Seas and the Arctic ocean with adjacent waters. *ACROSS-Atmospherie Circulation Related to Oscillations in Sea-ice and Salinity*, page 13, 1998.

T. Ese, I. Kanestrøm, and K. Pedersen. Climatology of polar lows over the Norwegian and Barents Seas. *Tellus A*, 40A(3):248–255, 1988. ISSN 1600-0870. doi: 10.1111/j.1600-0870.1988.tb00345.x. URL `http://dx.doi.org/10.1111/j.1600-0870.1988.tb00345.x`.

S.J. Gaffney. *Probabilistic Curve-Aligned Clustering and Prediction with Regression Mixture Models*. PhD thesis, University of California, Irvine, 2004.

S.J. Gaffney, A.W. Robertson, P. Smyth, S.J. Camargo, and M. Ghil. Probabilistic clustering of extratropical cyclones using regression mixture models. *Climate Dynamics*, 29(4):423–440, 2007. ISSN 0930-7575. doi: 10.1007/s00382-007-0235-z. URL `http://dx.doi.org/10.1007/s00382-007-0235-z`.

S. Grønås and N. G. Kvamstø. Numerical simulations of the synoptic conditions and development of Arctic outbreak polar lows. *Tellus A*, 47(5):797–814, 1995.

Z. Haikun, W. Liguang, and Z. Weican. Kernel density estimation applied to tropical cyclones genesis in Northwestern Pacific. 3:89–92, July 2009. doi: 10.1109/ESIAT.2009.238.

G. Heinemann and C. Claud. Report of a workshop on "theoretical and observational studies of polar lows" of the European Geophysical Society Lows Working Group. *Bulletin of the American Meteorological Society*, 78(11):2643–2658, 11 1997. URL `http://search.proquest.com/docview/232628318?accountid=17260`.

W. Härdle. *Smoothing techniques: with implementation in S*. Springer, 1991.

W. Härdle, M. Müller, S. Sperlich, and A. Werwatz. *Nonparametric and Semiparametric Models*. Springer, 1995.

K.A. Iden, M. Reistad, O.J. Aarnes, R. Gangstø, G. Noer, and N.E. Hughes. Kunnskap om vind, bølger, temperatur, isutbredelse, siktforhold mv. - Barentshavet SØ. Technical report, Meteorologisk Institutt, 2012.

E. A. Irvine, S. L. Gray, and J. Methven. Targeted observations of a polar low in the Norwegian Sea. *Quarterly Journal of the Royal Meteorological Society*, 137(660):1688–1699, 2011. ISSN 1477-870X. doi: 10.1002/qj.914. URL `http://dx.doi.org/10.1002/qj.914`.

E. Kalnay. The NCEP/NCAR 40-year Reanalysis Project. *Bulletin of the American Meteorological Society*, 77(3):437–470, 1996.

H. Økland. On the intensification of small-scale cyclones formed in very cold air masses heated by the ocean. *Institute report series*, 26, 1977.

E. W. Kolstad. A new climatology of favourable conditions for reverse-shear polar lows. *Tellus A*, 58(3):344–354, 2006.

E.W. Kolstad. *Extreme winds in the Nordic Seas: polar lows and Arctic fronts in a changing climate*. PhD thesis, University of Bergen, Norway, 2007.

E.W. Kolstad. A global climatology of favourable conditions for polar lows. *Quarterly Journal of the Royal Meteorological Society*, 137(660):1749–1761, 2011. ISSN 1477-870X. doi: 10.1002/qj.888. URL `http://dx.doi.org/10.1002/qj.888`.

E.W. Kolstad, T.J. Bracegirdle, and I.A. Seierstad. Marine cold-air outbreaks in the North Atlantic: temporal distribution and associations with large-scale atmospheric circulation. *Climate Dynamics*, 33(2-3):187–197, 2009. ISSN 0930-7575. doi: 10.1007/s00382-008-0431-5. URL `http://dx.doi.org/10.1007/s00382-008-0431-5`.

T. Linders. *Polar low interaction with the ocean*. PhD thesis, University of Oslo, 2009.

T. Linders and Ø. Sætra. Can CAPE maintain polar lows? *Journal of the atmospheric sciences*, 67(8), 2010.

M. Lystad. Polar lows in the Norwegian, Greenland and Barents Sea - Final report. Technical report, The Norwegian Meteorological Institute, Oslo, 1986.

P-E. Mallet, C. Claud, C. Cassou, G. Noer, and K. Kodera. Polar lows over the Nordic and Labrador Seas: Synoptic circulation patterns and associations with North Atlantic-Europe wintertime weather regimes. *Journal of Geophysical Research: Atmospheres*, 118 (6):2455–2472, 2013. ISSN 2169-8996. doi: 10.1002/jgrd.50246. URL `http://dx.doi.org/10.1002/jgrd.50246`.

G. McLachlan and T. Krishnan. *The EM algorithm and extensions*, volume 382. John Wiley & Sons, 2007.

G. McLachlan and D. Peel. *Finite mixture models*. John Wiley & Sons, 2004.

G. Noer and M. Ovhed. Forecasting of polar lows in the Norwegian and the Barents Sea. *Proc. Ninth meeting of the EGS Polar Lows Working Group*, 2003. URL `http://www.uni-trier.de/index.php?id=28161#c60628`.

G. Noer, Ø. Saetra, T. Lien, and Y. Gusdal. A climatological study of polar lows in the Nordic Seas. *Quarterly Journal of the Royal Meteorological Society*, 137(660):1762–1772, 2011. ISSN 1477-870X. doi: 10.1002/qj.846. URL `http://dx.doi.org/10.1002/qj.846`.

T. E. Nordeng. A model-based diagnostic study of the development and maintenance mechanism of two polar lows. *Tellus A*, 42(1):92–108, 1990. ISSN 1600-0870. doi: 10.1034/j.1600-0870.1990.00009.x. URL `http://dx.doi.org/10.1034/j.1600-0870.1990.00009.x`.

E. Rasmussen. The polar low as a CISK phenomena. 1977.

E. Rasmussen. The polar low as an extratropical CISK disturbance. *Quarterly Journal of the Royal Meteorological Society*, 105(445):531–549, 1979.

E.A. Rasmussen and J. Turner. *Polar Lows - Mesoscale Weather Systems in the Polar Regions*. Cambridge University Press, 2003.

R. J. Reed and W. Blier. A further study of comma cloud development in the eastern Pacific. *Monthly weather review*, 114(9):1696–1708, 1986b.

R.J. Reed and W. Blier. A case study of comma cloud development in the eastern Pacific. *Monthly weather review*, 114(9):1681–1695, 1986a.

M. Rojo, C. Claud, P-E. Mallet, G. Noer, A.M. Carleton, and M. Vicomte. Polar lows tracks over the Nordic Seas: a 14 winter climatological analysis. *Accepted for publication in Tellus, 2014*, 2014.

M. Rudemo. Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, 9(2):pp. 65–78, 1982. ISSN 03036898. URL `http://www.jstor.org/stable/4615859`.

J. Rumpf, H.l Weind, E. Faust, and V. Schmidt. Structural variation in genesis and landfall locations of North Atlantic tropical cyclones related to SST. *Tellus A*, 62(3), 2010.

D.W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization.* John Wiley & Sons, 1992.

B. W. Silverman. *Density Estimation for Statistics and Data Analysis.* Chapman & Hall, 1986.

Ø. Sætra, T. Linders, and JB. Debernard. Can polar lows lead to a warming of the ocean surface? *Tellus A*, 60(1):141–153, 2008.

K. Wilhelmsen. Climatological study of gale-producing polar lows near Norway. *Tellus A*, 37A(5):451–459, 1985. ISSN 1600-0870. doi: 10.1111/j.1600-0870.1985.tb00443.x. URL `http://dx.doi.org/10.1111/j.1600-0870.1985.tb00443.x`.

T. Woollings, B. Harvey, M. Zahn, and L. Shaffrey. On the role of the ocean in projected atmospheric stability changes in the Atlantic polar low region. *Geophysical Research Letters*, 39(24):n/a–n/a, 2012. ISSN 1944-8007. doi: 10.1029/2012GL054016. URL `http://dx.doi.org/10.1029/2012GL054016`.

M. Zahn and H. von Storch. A long-term climatology of North Atlantic polar lows. *Geophysical Research Letters*, 35(22):n/a–n/a, 2008. ISSN 1944-8007. doi: 10.1029/2008GL035769. URL `http://dx.doi.org/10.1029/2008GL035769`.

M. Zahn and H. von Storch. Decreased frequency of North Atlantic polar lows associated with future climate warming. *Nature*, 467(7313):309–312, 2010.

M. Zahn and H. von Storch. Investigation of past and future polar low frequency in the North Atlantic. *Geophysical Monograph Series*, 196:99–110, 2012.