A Dynamical Perspective on Predictions of Severe European Cyclones: The Role of Large-Scale Conditions

Jennifer Sarah Ruth Owen

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Earth and Environment
April 2014

Declaration of Authorship

The candidate confirms that the work submitted is her own and that appropriate credit has been given within the thesis where reference has been made to the work of others

Acknowledgements

First, I would like to thank my supervisors, Peter Knippertz and Doug Parker. Your knowledge, encouragement and patience were both beneficial to the work and inspiring to me. I am grateful to Tomek Trzeciak, for the storm tracker and for fortitude in the face of a complex problem.

Many thanks to the AXA Research Fund for providing the financial support for this work, embedded in the overall project 'A Seamless Approach to Assessing Model Uncertainties in Climate Projections of Severe European Windstorms'. Thanks to the University of Leeds, for providing the infrastructure for this thesis, and to to my Research Support Group of Piers Forster, Alan Haywood and John Marsham for keeping things on track. Acknowledgements also go to: NCAS, for inviting me to attend the Climate Modelling Summer School in September 2011; the DIAMET project, for allowing me to take part in their fieldwork in December 2011; and to NCAR, for welcoming me at the Advanced Study Program Summer Colloquium in June 2012. I am thankful for these opportunities to discuss ideas with a diversity of people and the fresh perspective they provided.

I am grateful to the ECMWF for providing the vast majority of the data. Thanks also to the MetOffice for the data they provided and to Malcolm Brooks for getting it. Many thanks to the Freie Unviversität Berlin for the Wetterkarten, and particularly to Volker Emert. Many thanks to Richard Rigby for Linux support, and to Ben Parkes for LaTeX (and moral) support. Thank you to Rob Wilbraham for providing results from his BSc dissertation. Thanks to Lindsay Bennett for all of the comments on my Data & Methods Chapter. Countless colleagues provided academic discussion over a brew or beer over the years, and I am grateful for each and every one of these; special mentions go to Tim Baker, Jen Catto, Sophie Cowie, Matt Hawcroft, Tim Hewson, Alex Roberts, Victoria Sinclair, and Joaquim Pinto.

Numerous people have provided encouragement over the years. Thank you to all who provided the much-needed hybrid of academic and emotional support. I am thankful to The Knitters near and far for keeping things in perspective, and all of the other nutters who've provided some light relief over the years. Particular thanks to Steph, Hannah, Irina, Emma and Michael for coping with my academic eccentricities while living under the same roof as me, to Christie for believing in me even when I didn't believe in myself, and to Jo for knowing when I needed a fieldtrip.

Thank you to Dr. Jon Owen, for going well beyond required parental duties. Thanks also to Dr. Janet Humphreys, for endless tea and encouragement. I am grateful that you both understand the PhD experience like only a comrade can. Finally, I would like to thank Martin Pirret, for supporting me throughout and still choosing to marry me.

Abstract

Severe windstorms regularly inflict damage throughout Europe. This research examines a set of 31 midlatitude cyclones, investigating each in depth but encompassing a wide variety of such storms. The set is selected using the Storm Severity Index, and categorised using two methods. The first method is based on the relationship between the storm and the jet stream. Four groups emerge: storms that move along the edge of the jet; that cross the jet stream early; that cross later; or have a split jet. The second method is based on the Pressure Tendency Equation. This establishes whether horizontal temperature advection or diabatic processes dominate during storm development. These two approaches are linked: storms in the first two jet groups tend to be driven by horizontal temperature advection, and the other two groups by diabatic processes.

This work then studies the storms' forecast quality and spread using ECMWF data. It finds that storm intensity tends to be under-forecast, the forecast storms move too slowly and are too far south. Forecast quality improves and spread decreases earlier in storms that cross the jet early, compared to those that cross later, suggesting a link between jet interaction and forecast error. Storms where horizontal temperature advection dominates are on average less well forecast than their diabatic counterparts, but diabatically driven storms tend to have greater forecast spread.

Finally, this study proposes metrics for storm-prone situations, examining the configuration of the atmosphere prior to the development of the storms. These describe a variety of key factors for cyclogenesis, such as baroclinicity, barotropicity and moist stability. Of the 31 storms, 29 are associated with a value greater than the 98th percentile of one or more metrics. There is a large overlap between storms where baroclinicity is strong and those where horizontal temperature advection dominates the deepening, confirming that the two approaches are dynamically linked. This relationship between the storms, the dynamics, and the metrics will allow future work to identify sources of uncertainty in modelling severe European windstorms.

Contents

Declaration of Authorship **1**

Acknowledgements **2**

Abstract **3**

List of Figures **6**

List of Tables **13**

Abbreviations **16**

1 Introduction **1**
- 1.1 Motivation . . . 2
- 1.2 Context . . . 4
- 1.3 Aims . . . 7
- 1.4 Scope . . . 7
- 1.5 Summary . . . 10

2 Literature Review **12**
- 2.1 Introduction . . . 12
- 2.2 Theory of Midlatitude Cyclones . . . 12
- 2.3 Data Sets for Investigating Historic Cyclones . . . 27
- 2.4 Cyclones in Climate Models . . . 32
- 2.5 Reasons for Discord . . . 36
- 2.6 Summary . . . 46

3 Data and Methods **48**
- 3.1 Data . . . 49
- 3.2 Selecting Storms . . . 52
- 3.3 Tracking . . . 55
- 3.4 Determining and Assessing Storm Types . . . 58
- 3.5 Comparing Classification Methods . . . 63
- 3.6 Measuring Forecast Quality . . . 65
- 3.7 Storm-Prone Situations . . . 72

3.8 Summary 73

4 The Storms **75**
4.1 Measuring Storm Severity 75
4.2 Reasons for Rejecting High SSI Days 79
4.3 Selected Storms 86
4.4 Storm Categories 88
4.5 Pressure Tendency Equation 102
4.6 Comparing Categories 107
4.7 Summary 109

5 Storms in Numerical Weather Prediction Models **111**
5.1 Opening Remarks 111
5.2 Review of Methods 113
5.3 Deterministic Forecasts 114
5.4 A Simple Metric for Forecast Error 134
5.5 Ensemble Forecasts 144
5.6 Summary 162

6 Storm-Prone Situations **165**
6.1 Theory 166
6.2 Method 173
6.3 Comparison of Growth Rates 178
6.4 Storms in the Growth Rates 184
6.5 Examples of Storms and their Storm-Prone Situation 191
6.6 Null Cases 203
6.7 Storm-Prone Situations and Intensity 208
6.8 Storm-Prone Situations and Storm Type 210
6.9 Storm-Prone Situations and Storm Predictability 219
6.10 Summary 221

7 Discussion and Conclusions **225**
7.1 Storms Selection 226
7.2 Storm Categorisation 227
7.3 Storm Predictability 228
7.4 Identifying Storm-Prone Situations 229
7.5 Storm-Prone Situations and Storm Categories 232
7.6 Future Work 233
7.7 Summary 235

Appendix A: Additional Ensemble Forecast Plots **252**

Appendix B: Tables of Variables **262**

References **237**

List of Figures

1.1 Hawkins and Sutton (2009) (Figure 2) 6

2.1 Schematic of the Norwegian Cyclone Model. Adapted from Figure 15a in Schultz et al. (1998), which was based on the work of Bjerknes and Solberg (1922). . . . 13
2.2 Schematic of the Shapiro-Keyser Model. Adapted from Figure 15a in Schultz et al. (1998), which was based on the work of (Shapiro and Keyser, 1990). 15
2.3 Schematic of the warm conveyor belt (red) and cold conveyor belt (blue) in a typical midlatitude cyclone: (a) before occlusion and (b) after occlusion. Figure 9 from Schultz and Vaughan (2011). 16
2.4 Isentropic flow (solid arrows) in a baroclinic wave (surface pressure and fronts also shown). Figure 1 in Thorncroft et al. (1993) 17
2.5 Schematic illustrating how rearranging a horizontal density gradient into a vertical one lowers the centre of gravity (black dot). The colder, denser fluid is shaded grey. Based on Figure 8.3 in Martin (2006) 18
2.6 The Lorenz Energy Cycle, illustrating the flow of energy caused by atmospheric processes, such as extra-tropical cyclones. The stores are mean-state potential energy (Pm), mean-state kinetic energy (Km), eddy potential energy (Pe) and eddy kinetic energy (Ke). 19
2.7 Showing the location of regions of divergence and convergence, associated with the jet stream. C indicates regions of convergence; D indicates divergence. 22
2.8 Northern Hemisphere 500hPa geopotential anomaly for December to February, averaged over 1991 to 2010. 26
2.9 Figures from Raible et al. (2008) (their Figure 2), showing (a) total cyclone centre density in ERA-40, for storms that with a minimum lifetime of 72h for winters (DJF) 1961-1990 and (b) the difference in the same measure between ERA-40 and NCEP-NCAR reanalyses. Black contours show 5% significance level. Cyclone centre density is the number of cyclones per unit time (in this case, a season), per unit area ($1000 km^2$) 30
2.10 Track density (in cyclone days per winter) of intense cyclones for (left) NCEP reanalysis (1958-1998) and for (right) a simulation with ECHAM5/MPI-OM1 for recent climate conditions (1960-2000). The two data sets have similar horizontal resolutions. From Pinto et al. (2009), Figures 4a and 13a. 33
2.11 Main storm tracks in the Northern Hemisphere at 850hPa (solid) and 250hPa (dashed). From Hoskins and Hodges, 2002 (Fig. 14). 35
2.12 Northern Hemisphere 500hPa geopotential anomaly for December to February for 1999-2000 (left) and 2010-2011 (right), compared to 1981-2010 climatology, to illustrate the NAO positive and negative phases respectively. Made using NCEP/NCAR reanalysis data at http://www.esrl.noaa.gov/ 40

2.13 Willison et al. (2013), Figure 4: Minimum of MSLP for first 32 hours of WRF simluation of a case study midlatitude cyclone. Blue = 120km resolution, orange = 20km resolution, think line=control run, thin line = perturbed physics ensemble members. Two dry runs are also included, red=120km, green = 20km resolution. 42
2.14 Track density ($10^6\ km^2\ month^{-1}$) for the cyclones tracked in winter (DJF) in the Northern Hemisphere. From Greeves et al. (2007), Figure 1. 44

3.1 Map illustrating the area over which SSI was summed [40^oN to 60^oN, 10^oW to 20^oE]. Overplotted are the SSI values for 1200h on 2009-01-24, the day with the highest SSI in the time period, which was related to storm Klaus. 53
3.2 Diagram illustrating how the smoothing criterion is assessed. P_1, P_2 and P_3 are the first, second and third points on the track under consideration, and P_m is the midpoint between P_1 and P_3. The red arrow indicates the distance calculated as part of the smoothing calculation. 56
3.3 An example of sectional plotting for storm Klaus for total wind speed ($\sqrt{u^2 + v^2}$ in ms^{-1}) at 300hPa. 59
3.4 Schematic of how the PTE works, based on Fink et al. (2012), Figure 1. See text in Section 3.4.2 for explanation. 61
3.5 For storms Jennifer, Xynthia and Klaus, testing the sensitivity of different distance limits (in degrees) when matching forecast to analysis tracks. Analysis track is in black. Coloured tracks represent the best match for a variety of lead times. 68

4.1 (a) Mean sea level pressure (hPa), (b) wind speed (ms^{-1}) and (c) SSI for 1200UTC 30th March 1995. 80
4.2 Polar low Petra at 1200UTC on 22nd Febuary 1999: (a) wind speed at 300hPa [ms^{-1}], (b) θ_e at 850hPa [K], (c) cyclone core pressure [hPa] against time [h]. (a) and (b) are plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4. 82
4.3 (a) Mean sea level pressure (hPa), (b) wind speed (ms^{-1}) and (c) SSI for 1200UTC on 5th March 1990. 84
4.4 (a) Mean sea level pressure (hPa), (b) wind speed (ms^{-1}) and (c) SSI for 0000UTC on 18th February 1997. 85
4.5 Schematic showing the typical track of an 'edge' storm (black) relative to the jet stream (red). 89
4.6 Storm Agnes, an example of an 'edge' storm from January 1993 (see Table 4.2 for exact dates): (a) wind speed at 300hPa [ms^{-1}], (b) divergence of the wind field at 300hPa [$10^-4, s^{-1}$]. Plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4. 90
4.7 Storm Agnes, an example of an 'edge' storm from January 1993: (a) θ_e at 850hPa [K], (b) cyclone core pressure [hPa] against time [h]. (a) is plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4. 91
4.8 Schematic showing the typical track of a 'cross-late' storm (black) relative to the jet stream (red). 92
4.9 Storm Jennifer, an example of a 'cross late' storm from January 2002: (a) wind speed at 300hPa [ms^{-1}], (b) divergence of the wind field at 300hPa [$10^-4, s^{-1}$]. Plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4. 93

4.10 Storm Jennifer, an example of a 'cross late' storm from January 2002: (a) θ_e at 850hPa [K], (b) cyclone core pressure [hPa] against time [h]. (a) is plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4. 94
4.11 Schematic showing the typical track of a 'cross-early' storm (black) relative to the jet stream (red). 95
4.12 Storm Kyrill, an example of a 'cross early' storm from January 2007: (a) wind speed at 300hPa [ms^{-1}], (b) divergence of the wind field at 300hPa [$10^{-}4, s^{-1}$]. Plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4. 96
4.13 Storm Kyrill, an example of a 'cross early' storm from January 2007: (a) θ_e at 850hPa [K], (b) cyclone core pressure [hPa] against time [h]. (a) is plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4. 97
4.14 Schematic showing the typical track of a split-jet storm (black) relative to the jet stream (red). 98
4.15 Storm Xynthia, an example of a 'split' jet storm from February 2010: (a) wind speed at 300hPa [ms^{-1}], (b) divergence of the wind field at 300hPa [$10^{-}4, s^{-1}$]. Plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4. 100
4.16 Storm Xynthia, an example of a 'split' jet storm from February 2010: (a) θ_e at 850hPa [K], (b) cyclone core pressure [hPa] against time [h]. (a) is plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4. 101
4.17 The components of the PTE for (a) Agnes [edge jet stream type], (b) Lothar [cross late], (c) Daria [cross early], and (d) Xynthia [split]. 103
4.18 Components of the PTE, for storms (a) Wiebke (b) Emma. 103
4.19 The components of the virtual temperature term in the PTE for (a) Agnes [edge jet stream type], (b) Lothar [cross late], (c) Daria [cross early], and (d) Xynthia [split]. 105

5.1 Pressure (analysis - forecast) against lead time for the selected storms, for (a) operational and (b) control forecasts. 116
5.2 Latitude (purple) and longitude (green) differences (analysis - forecast) against lead time for the selected storms, for (a) operational and (b) control forecasts. . . 118
5.3 Residuals (observed - regression) for: (a) pressure in operational forecasts; (b) latitude in operational forecasts; (c) longitude in operational forecasts; (d) pressure in control forecasts; (e) latitude in control forecasts; (f) longitude in control forecasts. Please note the different axis scales for each variable. 121
5.4 Pressure (analysis - forecast) against lead time for the selected storms divided by jet stream type (edge = purple +, split = green $\times$, cross early = blue $*$, cross late = red $\blacklozenge$), for (a) operational and (b) control forecasts. 122
5.5 Latitude (analysis - forecast) against lead time for the selected storms divided by jet stream type (edge = purple +, split = green $\times$, cross early = blue $*$, cross late = red $\blacklozenge$), for (a) operational and (b) control forecasts. 124
5.6 Longitude (analysis - forecast) against lead time for the selected storms divided by jet stream type (edge = purple +, split = green $\times$, cross early = blue $*$, cross late = red $\blacklozenge$), for (a) operational and (b) control forecasts. 125

5.7 Pressure (analysis - forecast) against lead time for the selected storms divided by PTE category (horiz = green ×, diab = purple +), for (a) operational and (b) control forecasts. 127
5.8 Latitude (analysis - forecast) against lead time for the selected storms divided by PTE category (horiz = green ×, diab = purple +), for (a) operational and (b) control forecasts. 128
5.9 Longitude (analysis - forecast) against lead time for the selected storms divided by PTE category (horiz = green ×, diab = purple +), for (a) operational and (b) control forecasts. 130
5.10 Storms divided by native model resolution for: (a) pressure in operational forecasts; (b) pressure in control forecasts. The numbers in the legend refer to the T or T_L native resolution of the forecast model. 131
5.11 Storms divided by native model resolution for: (a) latitude in operational forecasts; (b) latitude in control forecasts. The numbers in the legend refer to the T or T_L native resolution of the forecast model. 132
5.12 Storms divided by native model resolution for: (a) longitude in operational forecasts; (b) longitude in control forecasts. The numbers in the legend refer to the T or T_L native resolution of the forecast model. 133
5.13 Box plots of $\Delta(p)$ for each jet stream type. $\Delta(p)$ is the interpolated forecast error in pressure at 24 hours lead time, based on linear regression for each storm. . . . 139
5.14 Box plots of (a) $\Delta(lon)$ and (b) $\Delta(lat)$ for each jet stream type. 141
5.15 Box plots of $\Delta(p)$ for each PTE category. $\Delta(p)$is the interpolated forecast error in pressure at 24 hours, based on linear regression for each storm. 143
5.16 Box plots of (a) $\Delta(lon)$ and (b) $\Delta(lat)$ for each PTE category. $\Delta(lon)$ and $\Delta(lat)$ are the interpolated forecast error in longitude and latitude (respectively) at 24 hours, based on linear regression for each storm. 144
5.17 Differences between the ensemble forecasts and reanalysis for all storms, for (a) pressure (in hPa), (b) latitude, (c) longitudes (both in degrees) 147
5.18 With lead times binned by day, box plots showing the differences between the ensemble forecast and reanalysis for (a) pressure (in hPa), (b) latitude, (c) longitude (both in degrees), for all storms. 149
5.19 With lead times binned by day, box plots showing the absolute values of the differences between the ensemble forecast and reanalysis for (a) pressure (in hPa), (b) latitude, (c) longitude (both in degrees), for all storms. Line shows the root-mean-square error. 150
5.20 Differences between the ensemble forecasts and reanalysis pressure, with the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split. . . 153
5.21 Differences between the ensemble forecasts and reanalysis latitude, with the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split. . . 154
5.22 Differences between the ensemble forecasts and reanalysis longitude, with the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split. 155
5.23 Differences between the ensemble forecasts and reanalysis pressure, with the storms divided by the terms that dominate the PTE: (a) 'horiz', (b) 'diab'. Please note that the y-acis scales' differ, in order to better present the data. 156
5.24 Differences between the ensemble forecasts and reanalysis latitude, with the storms divided by the terms that dominate the PTE: (a) 'horiz', (b) 'diab'. 157

5.25 Differences between the ensemble forecasts and reanalysis longitude, with the storms divided by the terms that dominate the PTE: (a) 'horiz', (b) 'diab'. 158
5.26 Differences between the ensemble forecasts and reanalysis pressure, with the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639. . 159
5.27 Differences between the ensemble forecasts and reanalysis latitude, with the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639. . 160
5.28 Differences between the ensemble forecasts and reanalysis longitude, with the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639. 161

6.1 (Emanuel et al., 1987, Figure 3a), showing how maximum growth rate (σ) and the wavelength of that growth rate (λ) vary with the ratio r (Equation 6.11). Note that $r = 1$ is at the left and $r = 0$ is at the right. 171
6.2 Figure showing the area over which the four different versions of growth rate are averaged. 174
6.3 Time series illustrating the sensitivity of the QGD growth rate [s^{-1}] to different combinations of pressure levels used in the calculation. 175
6.4 Time series illustrating the sensitivity of the QGD growth rate [s^{-1}] to different areas, over which the average is performed. 176
6.5 Time series illustrating the sensitivity of the SGM growth rate [s^{-1}] to different combinations of θ_e and θ_{es} on the vertical levels selected in Section 6.2.1.1. For example, 'Tes - Te' refers to the difference between θ_{es} at 300hPa and θ_e at 900hPa. 177
6.6 Frequency density plots for the different versions of baroclinic growth rate [s^{-1}] for the period 1970/1980-2010/2011: (a) QGD against QGM, (b) SGM against QGM. 180
6.7 Frequency density plots for the different versions of baroclinic growth rate [s^{-1}] for the period 1970/1980-2010/2011: (a) SGD against QGD, (b) SGM against SGD. 181
6.8 Frequency density plot of QGD and BT. 182
6.9 Time series for 2010-10-01 to 2011-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate [all in s^{-1}]. For (a), QGD, QGM and SGD are all plotted using the scale on the left of the plot, but SGM uses the scale on the right. 183
6.10 Maps of two of the four baroclinic growth rates [s^{-1}] for 0000 on 2009-01-24: (a) QGD, (b) QGM. The box marks the area over which the values are averaged. Please note that the colour scales are different in these plots, because the raw values are different (see Figure 6.9). The scales are determined using a percentiles approach. 185
6.11 Maps of two of the four baroclinic growth rates [s^{-1}] for 0000 on 2009-01-24: (a) SGD, and (b) SGM. The box marks the area over which the values are averaged. Please note that the colour scales are different in these four plots, because the raw values are different (see Figure 6.9). The scales are determined using a percentiles approach. 186
6.12 Map of the barotropic growth rate [s^{-1}] for 0000 on 2009-01-24. 187
6.13 Time series for 2008-10-01 to 2009-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate. The lines indicate when storms passed through Europe: Klaus on 2009-01-24; Quinten on 2009-02-10. 192

6.14 Time series for 1989-10-01 to 1990-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate. The lines indicate when storms passed through Europe: Daria on 1990-01-25; Nana on 1990-02-11; Vivian on 1990-02-27; Wiebke on 1990-03-01. 195
6.15 Time series for 2004-10-01 to 2005-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate. The line indicates when storm Gero passed through Europe, on 2005-01-12. 196
6.16 Time series for 1994-10-01 to 1995-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate. The line indicates when storm Urania passed through Europe on 1995-01-23. 198
6.17 Time series for 1999-10-01 to 2000-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate. The lines indicate when storms passed through Europe: Anatol on 1999-12-03; Franz on 1999-12-12; Lothar on 1999-12-26; Martin on 1999-12-27; Kerstin on 2000-01-29. 199
6.18 Surface weather chart for 1999-12-11, illustrating storm Franz. Contours of mean sea level pressure, diagnosis of fronts and station plots also shown. 200
6.19 Weather chart on the 500hPa isobaric surface for 1999-12-11. Solid contours of geopotential height of the 500hPa surface, dashed contours of 1000 to 500hPa thickness. Data from radiosonde launches also shown. 201
6.20 Two of the random null case dates, showing how high pressure can block midlatitude cyclones: (a) the Met Office surface pressure chart for 1999-11-21, (b) QGD for 1999-11-21, (c) the Met Office surface pressure chart for 2008-02-08, (d) QGD for 2008-02-08 204
6.21 Two of the random null case dates, both showing an occluded midlatitude cyclones: (a) the Met Office surface pressure chart for 2002-11-24, (b) QGD for 2002-11-24, (c) the Met Office surface pressure chart for 2010-02-06, (d) QGD for 2010-02-06. 205
6.22 Two of the random null case dates, both showing Europe being affected by a mature cyclone: (a) the Met Office surface pressure chart for 2001-11-30, (b) QGD for 2001-11-30, (c) the Met Office surface pressure chart for 2006-11-30, (d) QGD for 2006-11-30. 207
6.23 Time series of the four versions of baroclinic growth rate for (a) for 1994-10-01 to 1995-03-31 (the black vertical line indicates when storm Urania passed through Europe on 1995-01-23), (b) 1999-10-01 to 2000-03-31 (Anatol on 1999-12-03; Franz on 1999-12-12; Lothar on 1999-12-26; Martin on 1999-12-27; Kerstin on 2000-01-29), and (c) 2009-10-01 to 2010-03-31 (Xynthia on 2010-02-28). 218

1 Differences between the ensemble forecasts and reanalysis pressure with lead times binned by day, the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split. 253
2 Differences between the ensemble forecasts and reanalysis latitude with lead times binned by day, the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split. 254
3 Differences between the ensemble forecasts and reanalysis longitude with lead times binned by day, the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split. 255

4 Differences between the ensemble forecasts and reanalysis pressure with lead times binned by day, the storms divided by the terms that dominate the PTE: (a) 'horiz', (b) 'diab'. 256
5 Differences between the ensemble forecasts and reanalysis latitude with lead times binned by day, the storms divided by the terms that dominate the PTE: (a) 'horiz', (b) 'diab'. 257
6 Differences between the ensemble forecasts and reanalysis longitude with lead times binned by day, the storms divided by the terms that dominate the PTE: (a) 'horiz', (b) 'diab'. 258
7 Differences between the ensemble forecasts and reanalysis pressure with lead times binned by day, and the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639. 259
8 Differences between the ensemble forecasts and reanalysis latitude with lead times binned by day, and the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639. 260
9 Differences between the ensemble forecasts and reanalysis longitude, with the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639. 261

List of Tables

1.1 Selected storms with dates and details of damage, based on data from Centre for Research on the Epidemiology of Diseases (2012). 2

2.1 Summary table of studies considering cyclones under climate change in a GCM . 37

3.1 Resolution of the ECMWF forecast models. The values in italics are the approximate conversion from spectral resolution to kilometres, at $60^{o}N$. 51
3.2 Tables showing examples of contingency tables for fixed row and column totals, and the probabilities of obtaining them or a more extreme version. 64
3.3 A general contingency table. C values refer to the column totals; R to the row totals; E to the expected value in each cell; and T to the grand total. 65

4.1 Dates and times of the top SSI values, alongside either the name of the associated North Atlantic storm or the reason for rejection. 76
4.2 Selected storms with dates and maximum SSI. The names of the storms are obtained from Free University of Berlin charts, for all but one storm. The value of SSI quoted is the maximum reached, during the time the storm was passing over Europe. Details of fatalities and estimated damage are provided based on data from Centre for Research on the Epidemiology of Diseases (2012), where available. The values of destruction are given for both the value at the time, and the value at 1st January 2011 corrected for inflation (Office for National Statistics, 2011) to ease direct comparison. 87
4.3 Storms with percentage contribution to deepening from selected terms of the PTE, ranked by horizontal temperature advection terms (horiz). 106
4.4 The list of storms, described by the two methods of categorisation and ordered by intesity (SSI). 108
4.5 Table showing counts of storms in each jet category (columns) and PTE type (rows). 109
4.6 Table showing the differences between the observed and expected values for each cell in Table 4.5. Shading illustrates whether the value is positive (pink) or negative (blue). 109

5.1 For all storms, correlation coefficients (R) for differences against lead time, and tests for its statistical significance (t), both to 3 significant figures. As discussed in Section 3.6.4, t is compared to a reference table to test significance, based on the number of points included in the correlation (Operational = 120, Control = 82). . 117
5.2 Analysis of the variance for the linear regression in the two types of deterministic forecast. Numbers are quoted to 3 significant figures. 119
5.3 Number of storms at each resolution. 132

5.4 The selected storms (ordered by SSI), with the 'standardised' forecast error at 24 hours, calculated using linear regression, for longitude, latitude and pressure. . . 134
5.5 The correlation (using Pearson's correlation coefficient, R) between the different 'standardised' forecast error at 24 hours; the test of that correlation using a t-test (t); and the significance (Sig?) of the correlation, obtained by comparing t to a reference value. 135
5.6 Table showing the results of two correlation tests (Pearson, R; Spearman's Rank, R_S) of $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$ against intensity, and their significance. 136
5.7 The number of storms of each jet stream type (n), included in the 'standardised' forecast error analysis, and the p-value of the Kolmogorov-Smirnov (K-S) test for normal distribution in each category and whether that states it is normally distributed ($p<0.05$). 137
5.8 Summary statistics for each of the four jet stream types: median, inter-quartile range (IQR), mean average and standard deviation (St. Dev). 139
5.9 The median and inter-quartile range (IQR) in $\Delta(lon)$ and $\Delta(lat)$ for each of the four jet stream types. $\Delta(lon)$ and $\Delta(lat)$ are the interpolated forecast error in longitude and latitude (respectively) at 24 hours, based on linear regression for each storm. 140
5.10 The median and inter-quartile range (IQR) for each of the PTE categories. 142
5.11 Table showing the results of two correlation tests (Pearson, R; Spearman's Rank, R_S) of $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$ against resolution, and their significance. 144

6.1 List of abbreviations used throughout this chapter, for the five different versions of growth rate. 173
6.2 Test of correlation between variables 178
6.3 The list of storms (ordered by date), and the percentile of the growth rate with which they are associated 189
6.4 Assessing the number of storms identified by different combinations of the SPS metrics. 190
6.5 Comparing the instances of the presence (✓) or absence (×) of an SPS and the presence or absence of a selected storm. 191
6.6 The list of storms, described by the two methods of categorisation and ordered by intensity (SSI), along with the percentile values of each growth rate with which each storm is associated, shaded as per Table 6.3. 209
6.7 Tables showing the number of storms associated with each of the four baroclinic growth rates at the different percentile thresholds, for each jet stream category. 98th refers to the number of storms associated with a value at or above the 98th percentile; 90th refers to those between the 90th and 98th percentiles; 80th to those between the 80th and 90th percentiles; Fail to the number of storms associated with a value less than the 80th percentile. 211
6.8 Tables showing the percentage of storms associated with each of the four baroclinic growth rates at the different percentile thresholds, using the raw numbers from Table 6.7. The jet stream types are abbreviated: CE = cross early, CL = cross late, ED = edge, SP = split. 211
6.9 For the barotropic growth rate, (a) the raw numbers and (b) the percentages of the storms in each jet stream category identified in each percentile threshold, as per Tables 6.7 and 6.8. The jet stream types are abbreviated: CE = cross early, CL = cross late, ED = edge, SP = split. 212

6.10 Difference between the number of storms observed and expected to be associated with each SPS metric at the 98th percentile threshold, for each jet stream category. Shading illustrates whether the value is positive (pink) or negative (blue). 213
6.11 Tables showing the number of storms associated with each of the four baroclinic growth rates at the different percentile thresholds, for each PTE type. 98th refers to the number of storms associated with a value at or above the 98th percentile; 90th refers to those between the 90th and 98th percentiles; 80th to those between the 80th and 90th percentiles; Fail to the number of storms associated with a value less than the 80th percentile. 214
6.12 Tables showing the percentage of storms associated with each of the four baroclinic growth rates at the different percentile thresholds, using the raw numbers from Table 6.11. 215
6.13 For the barotropic growth rate, (a) the raw numbers and (b) the percentages of the storms in each PTE type identified in each percentile threshold, as per Tables 6.11 and 6.12. 215
6.14 Difference between the number of storms observed and expected to be associated with each SPS metric at the 98th percentile threshold, for each PTE type. Shading illustrates whether the value is positive (pink) or negative (blue). 216
6.15 Testing correlation between the SPS metrics and predictability metric for longitude, using Pearson's correlation coefficient (R), its test for significance (t), Spearman's rank correlation coefficient (R_S), and whether the two tests are significant or not. 219
6.16 Testing correlation between the SPS metrics and predictability metric for latitude, using Pearson's correlation coefficient (R), its test for significance (t), Spearman's rank correlation coefficient (R_S), and whether the two tests are significant or not. 220
6.17 Testing correlation between the SPS metrics and predictability metric for pressure, using Pearson's correlation coefficient (R), its test for significance (t), Spearman's rank correlation coefficient (R_S), and whether the two tests are significant or not. 220

Abbreviations

Abbreviations	Expansion
CFSR	Climate Forecast Systems Reanalysis
ECMWF	European Centre for Medium-Range Weather Forecasting
GCM	Global Climate Model
GEFS	Global Ensemble Forecasting System
GFS	Global Forecasting System
MSLP	Mean Sea Level Pressure
NCEP	National Centers for Environmental Prediction
NWP	Numerical Weather Prediction
PTE	Pressure Tendency Equation
RCM	Regional Climate Model
SPS	Storm-Prone Situation
SSI	Storm Severity Index

Chapter 1

Introduction

This thesis examines the midlatitude cyclones that may inflict damage on Europe. It is motivated by previous studies, where modelling indicates that there are limits to our understanding of such storms. This thesis will concentrate on short-term forecasts made by Numerical Weather Prediction (NWP) models, and the large-scale atmospheric conditions that affect the development of midlatitude cyclones that may impact upon their forecast quality and spread.

The objective of this thesis is to identify potential sources of uncertainty in modelling severe midlatitude cyclones, by indicating which types of storm are poorly forecast and why. First, it is necessary to select a set of such cyclones with the potential to inflict damage. This set should be large enough to encompass a large part of the spectrum of these severe storms, while remaining sufficiently small that each storm can be examined in depth. These storms will be categorised based on the mechanisms that govern their deepening, in order to search for potential sources of uncertainty. This work also investigates the 'storm-prone situation' (SPS); that is, the configuration of the atmosphere prior to the development of these severe storms. These will be linked to the deepening mechanisms of the storms. The NWP forecasts of the selected storms will also be assessed, and related to the storm categories and SPSs.

Before specifying the aims of this study (Section 1.3) and discussing the scope of the current work in more depth (Section 1.4), a discussion of the study's motivation (Section 1.1) and a brief review on the context (Section 1.2) will ensue.

1.1 Motivation

European windstorms are a high-impact weather phenomenon. While the passage of midlatitude cyclones is a near-daily occurrence over Europe in winter, only a small percentage are sufficiently intense to inflict damage. Globally, midlatitude cyclones are the second most damaging weather phenomenon, after tropical cyclones (Munich Re, 2009). Any improvement in the quality of their forecasts would mean mitigating action could be taken, and losses both human and economic could be reduced. Table 1.1 shows examples of damaging European wind storms. The table also illustrates that these wind storms regularly inflict dozens of human losses and economic losses on the order of billions.

Storm Name	Date	People Killed	Estimated Damage (Million US$)
Daria	1990-01-25	85	6 860
Vivian	1990-02-26	50	3 230
Wiebke	1990-02-27	67	2 260
Udine	1991-01-05	48	909
Verena	1993-01-14	6	385
Anatol	1999-12-04	27	2 963
Lothar	1999-12-26	137	11 350
Martin	1999-12-27	14	4 100
Jennifer	2002-01-28	17	150
Jeanette	2002-10-28	38	2 531
Gero	2005-01-12	7	50
Kyrill	2007-01-18	46	9 010
Emma	2008-03-01	13	1 800
Klaus	2009-01-24	28	5 100
Xynthia	2010-02-28	64	6 074

TABLE 1.1: Selected storms with dates and details of damage, based on data from Centre for Research on the Epidemiology of Diseases (2012).

As well as being important to society, midlatitude cyclones are scientifically interesting. Although research into them stretches back many decades (e.g. Bjerknes and Solberg, 1922), some questions remain, such as the relative importance of different processes that drive deepening; for instance, baroclinic instability and diabatic processes. These processes are especially important, because deepening is the point at which it can be established whether a storm will have the potential to inflict damage and, if so, where. Improving the understanding of the processes that govern

the deepening phase of European windstorms will improve their modelling on time scales from climate modelling to everyday weather forecasts.

While some studies have assessed short-term forecast quality, others have used climate models. The approach that each takes has been different: climate modelling concentrates on large-scale averages over long time periods, whereas NWP concentrates on short-term predictions of a range of variables. Climate modelling is governed by boundary conditions and NWP is governed by initial conditions. However, the limitations of one model inform model development for all time scales.

Furthermore, the climate modelling community has become more interested in high-impact, extreme weather events under climate change (e.g. Meehl et al., 2007). This led naturally to the idea of 'seamless' modelling, where in order to improve the representation of key processes, the limitations of one model are fed into other models that have different applications or run on a different timescale (Palmer et al., 2008); for example, identifying deficiencies in a model used for daily weather forecasting could be used to improve a seasonal or climate model. While this is a recent idea, it already has practical applications including at the UK Met Office where the same Unified Model is used for simulations for a wide range of simulations from weather to climate. This is because there is a wide variety of earth processes that act on all timescales. Palmer et al. (2008) illustrate this from the effects that anthropogenically-induced changes in greenhouse gases have through the following stages:

1. on radiative forcing (timescales of approximately 1 day)
2. on the atmosphere (10 days)
3. on ocean-atmosphere-land interactions (100 days)
4. on cryosphere and biogeochemical interactions (1000 days and longer)
5. on climate effects (decades and longer).

It is clear that these timescales all interact. For example, while radiative forcing affects the cryosphere by causing sea-ice to melt, the very act of white sea ice melting to reveal dark ocean affects the Earth's albedo and so influences the radiative forcing. Palmer et al. (2008) argue that, although determining the accuracy of climate models cannot occur for decades, determining the accuracy of the processes within them depends on the time scale over which a given process

operates. The verification of 'fast', short-timescale processes is possible with a much shorter observational record, than their slower counterparts. Current work, such as the Transpose-AMIP project (Williams et al., 2013) are utilising this seamless idea, in order to verify climate models by running them at high resolution, allowing the models to be improved, and this informs modelling on a wide range of time scales. .

The current work is part of a wider project entitled 'A Seamless Approach to Assessing Model Uncertainties in Climate Projections of Severe European Windstorms (SEAMSEW)', funded by the AXA Research Fund. That work aims to apply these ideas to midlatitude cyclones, by investigating historic severe windstorms in climate models, with the intention to describe the reasons behind the models' discord regarding storms' future projections. That work will take a seamless approach, running a climate model at the temporal and spatial resolution typical of NWP models but for a very short period of time, and initialised using reanalysis data. This means that SEAMSEW can use results from the high-resolution climate model simulations of the midlatitude cyclones to separate the sources of uncertainty and so inform model development. Given its striking similarity to the way operational forecasts are generated, therefore, this thesis' analysis of NWP will act as a benchmark for these simulations, in order to ascertain the pertinent limitations of climate models.

Overall, this work is motivated by a need to narrow down the sources of uncertainty in the modelling of midlatitude cyclones on the range of time scales from weather to climate. Although this work will concentrate on short NWP forecasts, analysing their shortfalls could resonate through model simulations across the spectrum of time scales. Most importantly, improving the forecasts of these damaging weather phenomena leads to mitigating action being taken that reduces future damage.

1.2 Context

Previous studies have assessed the forecast quality, in terms of both accuracy and spread, of midlatitude cyclones. They tend to take one of two approaches: either analysing case studies, or looking at the picture over weeks or months. Examples of the former approach include: Wernli et al. (2002), who investigated storm Lothar; Fink et al. (2009) (storm Kyrill); and Liberato et al. (2011) (storm Klaus). These papers concentrate on a single or a small group of extreme events that had a large impact. An example of the latter approach is carried out by Froude et al., across several papers (Froude, 2009, 2010; Froude et al., 2007a,b). That work automatically tracks cyclones in

data that cover periods lasting from weeks to six months, providing an important, broad picture of the forecast quality of midlatitude cyclones. However, with such a large number of tracks, it is not possible to unpick the reasons why the forecasts are of limited quality, or analyse the impact of every storm.

For cyclones in climate models, the broad picture from the literature becomes clear: while climate models are much improved and in much better agreement in recent years, there is disagreement about how North Atlantic cyclones will evolve under climate change (Chapter 2, Table 2.1). Evolution in terms of their spatial, temporal or physical properties (e.g. core pressure or associated wind strength) has a large degree of uncertainty. A change in the regions that the storms affect most often is expected, but this is not certain (e.g. Stratton, 2004). Cyclones could become more frequent, or less frequent but with a greater frequency of intense cyclones (e.g. McDonald, 2011). The reasons for the discord will be discussed in Chapter 2, but the overall picture from the climate modelling community indicates that midlatitude cyclones are not sufficiently well understood or sufficiently well represented at the resolution of climate models.

Even in measures of the climate that are better predicted (e.g. global average temperature), there is a degree of uncertainty in climate projections. Identifying the sources of uncertainty means they can be reduced; however, as the atmosphere is stochastic, the uncertainty will never be eliminated (Lorenz, 1969). As discussed by Hawkins and Sutton (2009), there are three broad sources of uncertainty:

- **Scenario uncertainty** occurs because future emissions of greenhouse gases are not known. Therefore, different scenarios are used, each with different amounts of greenhouse gas emitted over different time periods.
- **Internal variability** describes the natural variation the atmosphere exhibits on all time scales.
- **Model error** describes all the limitations of using a computer model. These include:
  - basic state: the drift that models can have if the boundary conditions (such as sea surface temperatures, Keeley et al. (2013)) or the systems to which the atmosphere is coupled (such as ocean or sea ice models) are not correct,

– dynamical core: which set of fundamental equations are used to determine the properties of an air parcel (e.g. Eulerian or semi-Lagrangian) can affect the results (Section 2.5.1),

– model physics: the parameterised equations used to describe subgridscale physical processes, such as diabatic processes,

– resolution: the effect of discretising the atmosphere onto a grid (Section 2.5.3).

These three sources of uncertainty are illustrated in Figure 1.1, which shows the temporal evolution of these three factors, along with their sensitivity to the area and time over which they are averaged. Internal variability's effect on the overall uncertainty in the projection remains constant with time, because it is related to the chaotic nature of the atmosphere. Both scenario uncertainty and model error grow with time; the former because future emissions are more uncertain, and the latter because the errors are cumulative.

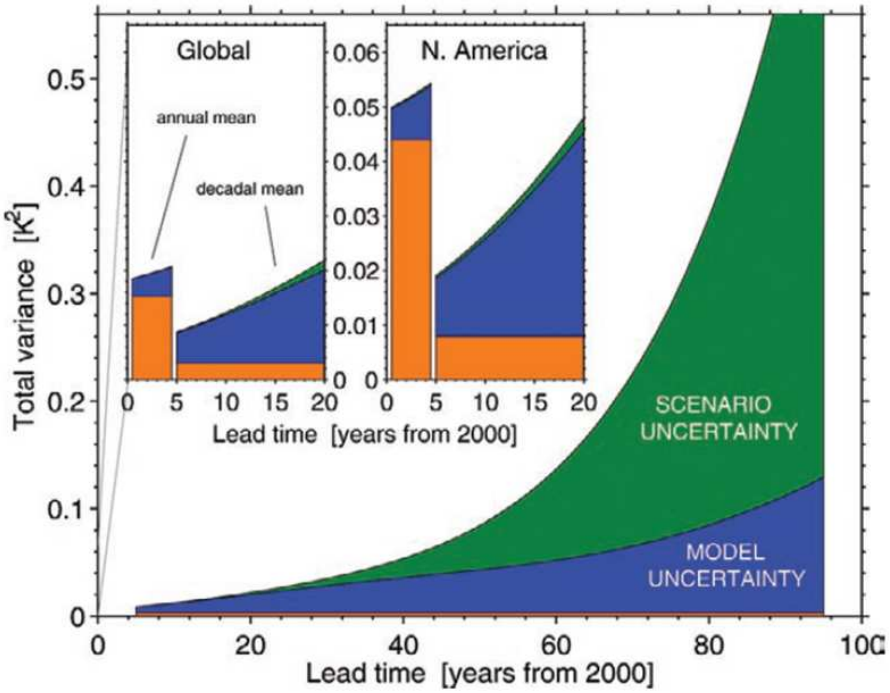


FIGURE 1.1: From Hawkins and Sutton (2009) (Figure 2): 'the relative importance of the three sources of uncertainty changes significantly with region, forecast lead time, and the amount of any temporal meaning applied. Main panel: Total variance for the global mean, decadal mean surface air temperature predictions, split into the three sources of uncertainty. Insets: As in the main panel, but only for lead times less than 20 years for (left) the global mean and (right) a North American mean. The orange regions represent the internal variability component. For lead times shorter than 5 years, we plot the results using annual mean data.'

The effect internal variability has on projections for the late 21st century is small, compared to the other sources. However, there is still merit in investigating its effects, because for the near-term climate, it contributes a large fraction of the total uncertainty. Unfortunately, the uncertainty in future emissions cannot be mitigated, because it depends on human factors. However, reducing the model uncertainty is feasible and would have a large impact on modelling of the atmosphere at all time scales. The SEAMSEW project intends to contribute towards reducing such errors for midlatitude cyclones. Initial work with short climate model runs (Ma et al., 2014) indicates that such runs show the sources of error, which propagate and become larger problems in longer climate model runs, notably those in clouds and precipitation. This means that identifying errors in NWP models could inform climate model developers, as well as guiding future improvements to short-term forecasts. This thesis will contribute towards identifying these errors.

1.3 Aims

This work aims to assess the NWP forecast quality of a set of historical windstorms, and search for links between predictability, the dynamical type of storm, and the large-scale atmospheric situation at the storm's inception. This will further the understanding of the processes that govern the development of severe windstorms and of the factors that influence their predictability. The research questions this work asks are:

1. What are the processes that govern the deepening phase of the selected storms?
2. How well are the selected storms forecast, and is forecast quality or spread connected to the processes that govern deepening?
3. To what extent are the selected storms, storm-prone situations and, and the processes that govern their deepening related?

1.4 Scope

This work will take a novel middle route between the case-study and more statistical approaches that have previously been used to assess the forecasts of severe storms, by examining a set of midlatitude cyclones. This will consist of approximately 30 storms. This size is a compromise

between a desire to capture as much of the spectrum of midlatitude cyclones as possible, and a need to be able to analyse each in enough depth as to identify the drivers of each storm's development. These storms will be categorised, based on the key factors that drive deepening. Their forecasts will then be assessed, in terms of both quality and spread. The next steps will be to examine the large-scale state of the atmosphere, at the time of each storm's initiation. This should lead to the identification of one or more metrics for a 'storm-prone situation', that describe the state of the atmosphere preceding a severe midlatitude cyclone. The final stages of the work will draw together the key themes of intense storms, their forecasts and their storm-prone situations, and identify links between them, in order to examine the degree to which the large-scale forcing controls storm development and forecast quality.

The vast majority of severe storms over Europe (95%) occur between October and March (Alexander and Tett, 2005), so this work will only consider winter storms. Furthermore, winter weather is expected to be affected more strongly under climate change than the other seasons (Deser et al., 2010), because the resulting changes in the net surface energy budget are delayed compared to the loss of sea ice, so the climate response to the maximum sea ice loss in late summer will be some months later. Therefore, this study will be limited to severe winter storms over the North Atlantic and Europe. This choice of location is made because the area is regularly affected by storms, and in Europe they pass over a densely populated area, giving a large potential to inflict damage.

The first step is to establish the set of storms. A metric must be selected that describes the risks that midlatitude cyclones pose. The two major risks from such cyclones are precipitation and wind. While precipitation can cause substantial damage through flooding and landslides, a flood model would be required alongside the NWP model, to include factors such as surface run-off and soil moisture. Therefore, for simplicity, flooding will not be considered in the present work. Another hazard associated with midlatitude cyclones is a storm surge, where low atmospheric pressure and strong winds cause coastal flooding. However. this would also require a separate model to quantify the hazard, to include tidal and wave information. Therefore, the current study will concentrate on the wind risk. There are many metrics that could be used to quantify wind risk. It is important that the selected metric is related to the damage a storm can inflict, but this can be complicated by socio-economic factors such as the affluence and preparedness of the areas over which the storm passes. Therefore, the metric selected will be based on meteorological quantities alone. Further discussion on the selection of the metric can be found in Section 3.2. Only severe storms will be

selected: i.e. those with an unusually high value of the chosen metric. Once selected, the storms will be tracked (Section 3.3), so that the evolution of each storm can be compared.

Next, the storms will be categorised, based on the processes that govern deepening (Research Question 1). From Bader et al. (1995, Chapter 5), it is clear that a wide spectrum of cyclones pass over Europe, so generalising results from case studies can prove problematic because what might be true for a subset of cyclones may not be true for all cyclones. Finding a method to categorise such storms is important because identifying types of cyclones would allow aspects of each group to be investigated and comparisons made between categories. Although the best way of categorising them remains an open question, this work will concentrate on the processes that cause a cyclone to become more intense. The processes that govern this deepening of cyclones act on a wide range of temporal and spatial scales. The large-scale forcing of cyclones includes factors such as baroclinicity (Section 2.2.4.1), which are relatively well understood, but how they affect individual cyclones' deepening is not. Smaller-scale, diabatic processes are particularly poorly understood, because they act on sub-grid scales and so are difficult to model (Section 2.2.4.4). These processes are of particular interest to current research; notably, the DIAMET project (DIAbatic influences on Mesoscale structures in ExTratropical storms) is investigating these through direct measurement and modelling (Vaughan et al., 2014).

As noted in Section 1.1, the deepening phase is used because this is key for midlatitude cyclones. The methods used for categorisation are discussed in Section 3.4. The results for the selection and categorisation of the storms are discussed in Chapter 4. Dividing the storms into categories based on processes that affect the deepening phase will facilitate the identification of sources of uncertainty. If a particular process is poorly modelled, then a poor forecast of a storm will result. However, there are caveats with this. For example, modelling errors can compensate for each other, or the initial conditions that are fed into the model can be uncertain. These will be discussed further in Sections 2.3.3 and 2.5.1, respectively.

Once the storms have been chosen and categorised, the next portion of the work concentrates on their NWP forecasts (Research Question 2). It is not within the scope of this work to examine any output from the NWP-style climate model runs, and so the comparison with this analysis is future work. The current work will examine the NWP output, again by tracking the storms. It will assess the forecast quality and spread of these storms, by using deterministic and ensemble forecasts. This analysis will be performed for both storm intensity and position, in order to ascertain if one is forecast better than the other. Forecasts will also be assessed in terms of the groupings made

in terms of storm type, which used key factors for cyclogenesis. These results are presented in Chapter 5.

The final part of this work investigates storm-prone situations (SPSs), the large-scale arrangement of the atmosphere just before a severe storm develops. This work will also allow sources of uncertainty to be established. If a forecast contains a strong SPS but a storm does not develop, that indicates a different problem to a forecast without a strong SPS at all. The first steps are to identify potential metrics for SPSs, again concentrating on processes that contribute to cyclogenesis. Once candidate metrics have been identified, then the interaction between the SPSs and the storms is examined, to determine whether they could prove useful in identifying sources of uncertainty in model simulations. Again, SPSs are analysed in terms of the categories of storms, to identify whether one category of storm is more likely to be related to one metric than another. Finally, the link between the storm-prone situation and storm predictability is examined, to ascertain whether the presence or magnitude of the SPS associated with a storm affects the quality or spread of the forecast. All of the results pertaining to SPSs are presented in Chapter 6.

1.5 Summary

The current project will select a a set of severe, midlatitude cyclones, assess their forecast quality, and examine their relationship with the large-scale situation. Previous work in this field has investigated forecast quality in terms of individual case studies at one end of the spectrum, and in terms of all the cyclones in a given season at the other end. The middle approach taken by the current work is novel. The climate modelling of midlatitude cyclones indicates that there are limitations to our understanding of such storms, because they disagree as to how their location, frequency or intensity will evolve under a changing climate. This motivates the current work, which endeavours to understand the reasons behind why some storms are better forecast than others, in order to identify the potential sources of uncertainty in modelling on all timescales. Furthermore, the assessment of forecast quality will act as a benchmark for simulations made as part of the SEAMSEW project, using climate models to simulate the set of storms at high NWP-level resolutions.

The first step of this investigation is to select the storms, based on a measure that represents their potential to inflict damage, before examining their evolution. This will facilitate the grouping of storms, using factors that drive cyclogenesis. These results are discussed in Chapter 4. The next stages are to assess forecast quality, in terms of both storm location and intensity, and any

relationship with the grouping of storms will be investigated (Chapter 5). The final portion of the work studies SPSs, to ascertain whether the atmosphere is configured in a particular way prior to severe storms' development. This is also done in terms of both the grouping and predictability of the storms (Chapter 6). Finally, a summary of key results and a discussion of the findings comprises Chapter 7. Before the discussion of the results, a review of the literature can be found in Chapter 2, before a summary of the relevant data and methods in Chapter 3.

Chapter 2

Literature Review

2.1 Introduction

European weather is governed by the passage of extra-tropical cyclones and anti-cyclones. Anti-cyclones tend to cause calm conditions, whereas cyclones and associated fronts generate wind and precipitation. Not only do cyclones influence people's daily lives, but in rare cases, cause death and destruction (Table 1.1). There are many processes that affect cyclones' development, and since some of them are small-scale and have to be parametrised, models can struggle to accurately simulate a storm's path or intensity. This may be particularly true of climate models, compared to numerical weather prediction (NWP) models, due to the lower resolution and the related increased need for processes to be parametrised. Reliable predictions for extra-tropical cyclones' path and intensity would prove crucial over Europe, on time scales from short-term forecasts to climate projections for the coming century. Improved forecasts would assist socio-economic planning, for example when formulating building codes, undertaking large-scale infrastructure construction, planning food security, or deciding whether to take action to protect against adverse weather.

2.2 Theory of Midlatitude Cyclones

The cyclones that affect Europe usually undergo growth and development over the North Atlantic, and are then guided towards Europe by large-scale flow in the atmosphere. This large scale flow is strongly related to the Gulf Stream in the ocean, which brings warm water from the Caribbean up

and across the Atlantic. This induces a strong gradient in sea-surface temperatures, and so a similar gradient develops in atmospheric temperatures. The relationship between atmospheric flow and temperature gradients will be discussed in Section 2.2.4.1. Over the course of a cyclone's lifecycle, it will grow (known as cyclogenesis) and dissipate (cyclolysis). Some methods of describing cyclones will now be discussed, followed by a description of cyclogenetic processes.

2.2.1 Norwegian Cyclone Model

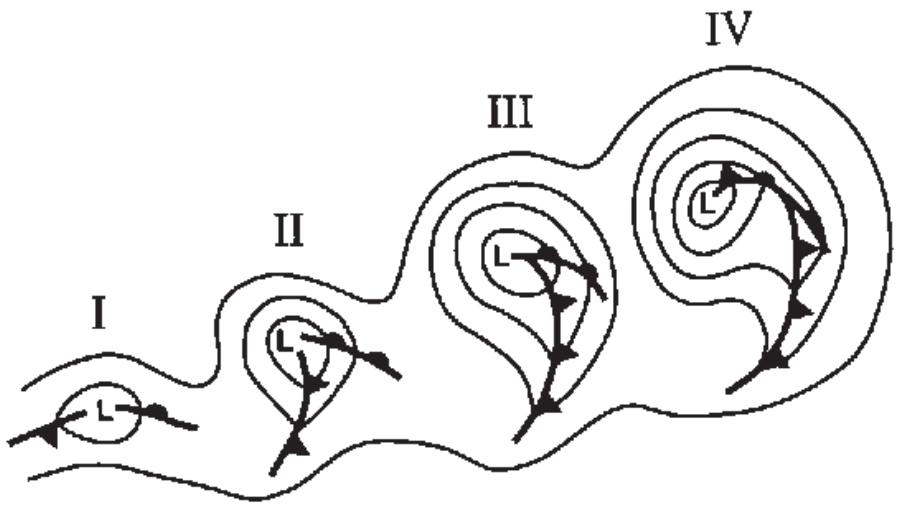


FIGURE 2.1: Schematic of the Norwegian Cyclone Model. Adapted from Figure 15a in Schultz et al. (1998), which was based on the work of Bjerknes and Solberg (1922).

The lifecycle of midlatitude cyclones was first comprehensively described by Bjerknes and Solberg (1922), who developed the Norwegian Cyclone Model, as shown in Figure 2.1. The Norwegian Cyclone Model portrays the Polar Front, a boundary between warm, moist sub-tropical air and cold, dry polar air that encircles the globe, with a clear horizontal temperature difference. Perturbations to this develop into low pressure centres with closed pressure contours (Figure 2.1 I), and then mature into a classical cyclone with encircling winds and fronts (Figure 2.1 II). The warm and cold air are delimited by the fronts, which are associated with overturning, clouds and precipitation (Wallace and Hobbs, 2006). In between the two fronts is the remnants of the tropical air known as the warm sector. Ahead of the warm front and behind the cold front is the remaining colder, polar air. As the cyclone intensifies, the cold front lengthens but the warm front remains

short (Figure 2.1 III). The cold front moves faster than the warm front, and so catches up with the warm front, forming an occluded front (Figure 2.1 IV). This is now a fully matured cyclone, and dissipation begins.

Though it was innovative at the time and remains a useful descriptive tool, limitations to this model have subsequently been identified. These include the lack of discussion of the vertical structure of cyclones or the influence of the jet stream, due to a lack of observations at the time. A further limitation of the Norwegian model is its representation of fronts; they are not sharp boundaries between air with different properties such as temperature or humidity, but rather are regions of steep across-front gradients in these properties (Martin, 2006). Furthermore, cyclones do not always occlude, or do occlude but continue to intensify rather than dissipate (Schultz and Vaughan, 2011). For this reason, further explanation of midlatitude cyclones' development is required.

2.2.2 Shapiro-Keyser Model

With the further observation of cyclones facilitated by the advent of satellites and measurements made by aircraft, it became clear that not all cyclones behave as described by the Norwegian cyclone model, and so the Shapiro-Keyser model was developed (Shapiro and Keyser, 1990), illustrated in Figure 2.2. The initial stage is also a perturbation in the region of the strong temperature gradient or front. As the cyclone begins to grow, differences to the Norwegian model emerge; here the warm front lengthens and the cold front remains relatively short. The cold front becomes perpendicular to the warm front, and the intersection is a weak point so the fronts fracture apart rather than undergoing occlusion. The warm front wraps around the cyclone centre to form a bent-back front, and cuts off an area of warm air at the cyclone core. Therefore, the warm air can be said to be secluded, as opposed to the occlusion that occurs in the Norwegian model. However, as with the occlusion process described by the Norwegian Cyclone Model, during seclusion the centre of the low and the warm sector are decoupled, so the cyclone begins to dissipate (Schultz and Vaughan, 2011).

The Norwegian and Shapiro-Keyser models can be considered as two members of the wide range of cyclones that can be observed (Schultz et al., 1998). Hart (2003) argues that cyclones are best described as members of a continuous spectrum of cyclones. Despite this, the processes that cause midlatitude cyclones to develop and dissipate are universal; differing relative strengths of these

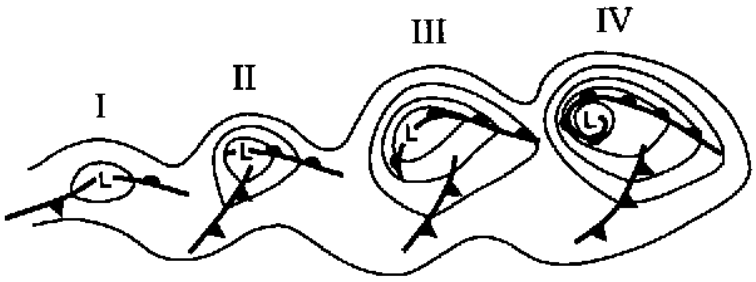


FIGURE 2.2: Schematic of the Shapiro-Keyser Model. Adapted from Figure 15a in Schultz et al. (1998), which was based on the work of (Shapiro and Keyser, 1990).

processes lead to the wide spectrum of possible cyclones. The cyclogenetic processes involved in any cyclone's lifecycle will be discussed in Section 2.2.4. First, a discussion of the airflow within typical cyclones will ensue.

2.2.3 Structure of a Cyclone

The vertical structure of a cyclone has been studied extensively since the time of the Norwegian Cyclone model. The ability to make direct measurements away from the surface using satellites, radiosondes and aircraft measurements has developed, meaning insight has been gained into the movement of air aloft in a cyclone (see Bader at al. (1995) for a review). This has led to the observation of large areas of air flowing around the cyclone, along isentropic surfaces (i.e. surfaces with constant potential temperature and therefore constant entropy), causing uplift and potentially cloud formation. There are two known as 'conveyor belts': one warm (WCB) and one cold (CCB) (Figure 2.3a). They are usually found in midlatitude cyclones, but their relative strengths vary between cyclones and over a cyclone's lifetime. The WCB draws warm, moist air up from the warm sector between the fronts, lifting it to form the mass of cloud ahead of the warm front. The CCB moves alongside the warm front and under the WCB then quickly ascends, inducing a band of cloud to the west of the cloud made by the WCB (Browning and Roberts, 1994).

After the cyclone occludes and/or dissipation begins, the WCB splits into two branches, which turn cyclonically and anticyclonically (Figure 2.3b). Thorncroft et al. (1993) proposed that the way a cyclone dissipates depends on which branches of the WCB and CCB are stronger (Figure 2.4). In the cyclonic wave breaking case, branches B and C are stronger, whereas in the anticyclonic breaking case, A and D dominate. Whether the cyclonic or anticyclonic flow dominates is strongly

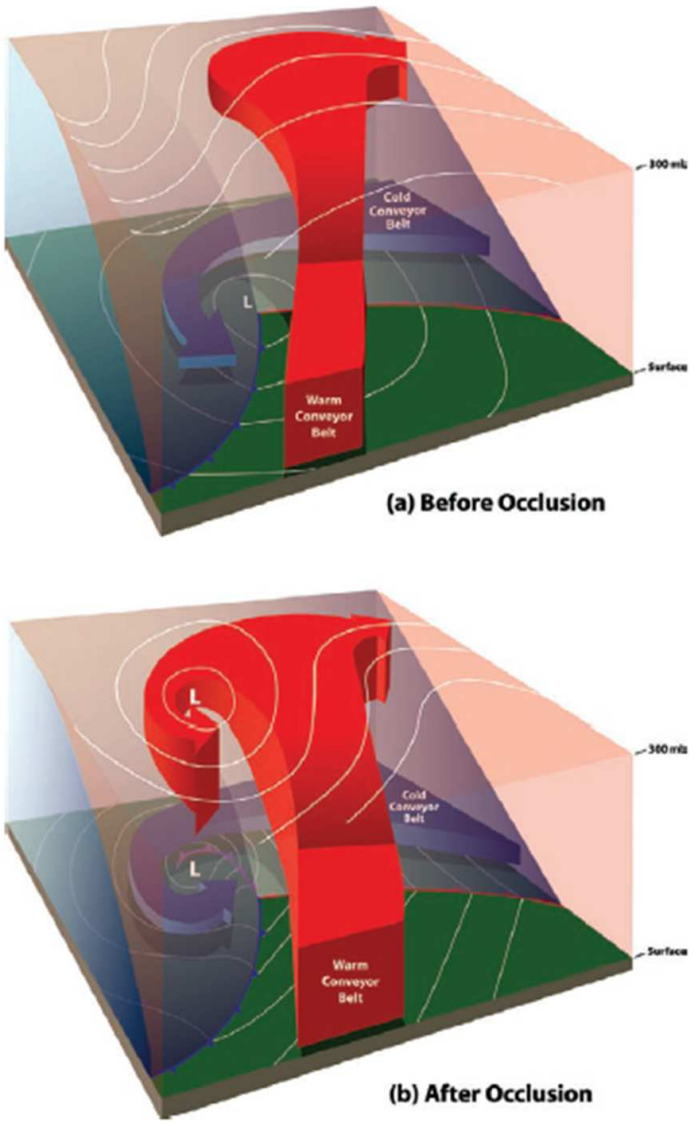


FIGURE 2.3: Schematic of the warm conveyor belt (red) and cold conveyor belt (blue) in a typical midlatitude cyclone: (a) before occlusion and (b) after occlusion. Figure 9 from Schultz and Vaughan (2011).

dependent on the background horizontal wind shear (Davies et al., 1991). Idealised simulations of these cyclones indicate the direction of wave breaking can affect key properties of a cyclone, including its track and strength.

These regions of uplift are important in the development of a cyclone, because they contribute towards cloud development, and so explain the patterns observed on satellite images. Given that both the WCB and CCB move air upwards, their strengths compared to those of downdraughts

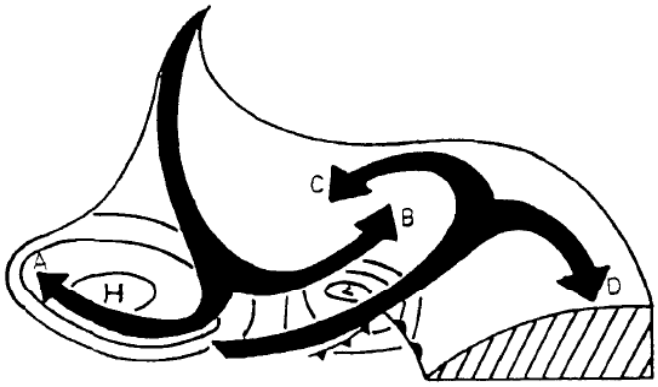


FIGURE 2.4: Isentropic flow (solid arrows) in a baroclinic wave (surface pressure and fronts also shown). Figure 1 in Thorncroft et al. (1993)

affect how efficiently air is moved from lower to upper levels in the cyclone, and therefore affect the core pressure.

2.2.4 Cyclogenetic Processes

The processes that cause extra-tropical cyclones to deepen will now be discussed. A baroclinic atmosphere (Section 2.2.4.1) provides the potential for energy to be released from overturning, and for temperature advection that can favour cyclonic deepening. The interaction with the jet stream (Section 2.2.4.2) is another important factor to discuss. Furthermore, diabatic processes can also cause cyclogenesis, for reasons that will be discussed in Section 2.2.4.4. However, it is important to note that, although all of these processes encourage cyclogenesis, they are not individually sufficient to generate a midlatitude cyclone. An initial perturbation to the temperature or pressure field is necessary, which these processes then encourage to grow into a mature midlatitude cyclone.

2.2.4.1 Baroclinic and Barotropic Instabiltiy

In this work, two types of atmospheric instability need to be considered; barotropic and baroclinic. A barotropic atmosphere is where density depends solely on pressure, whereas in a baroclinic atmosphere density depends on pressure and temperature. Barotropic instability is related to the horizontal shear of the jet, and perturbations grow by converting kinetic energy from the jet into cyclonic circulations. Therefore, barotropic instability relates the conversion of large-scale kinetic energy to smaller-scale motion. Baroclinic instability differs because the potential energy, which

can be extracted from converting a baroclinic atmosphere into a barotropic one, is also available to convert to the kinetic energy of the cyclone.

During occlusion, the air in the warm sector is lifted above the colder, denser polar air. The initial horizontal temperature gradient into a vertical temperature gradient, with warm air overlying the colder air (Figure 2.5), lowering the centre of gravity of the system and reducing its gravitational potential energy.

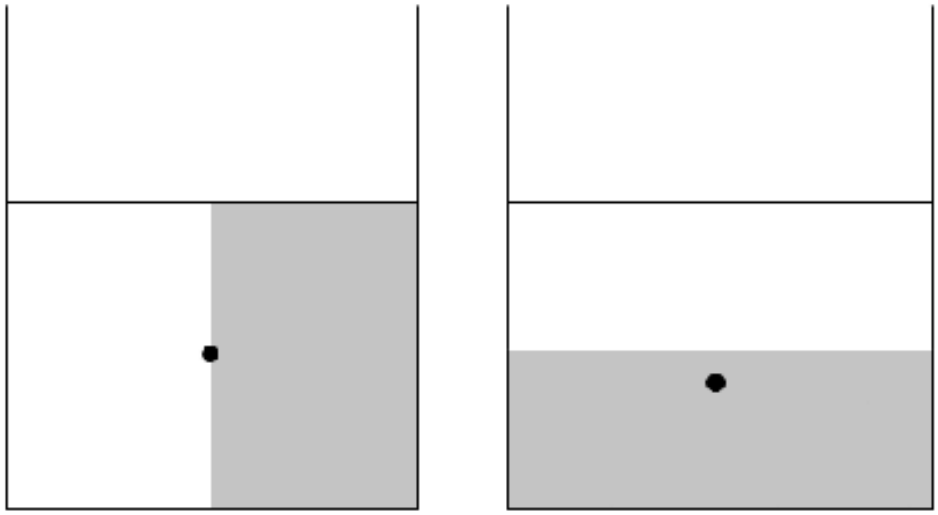


FIGURE 2.5: Schematic illustrating how rearranging a horizontal density gradient into a vertical one lowers the centre of gravity (black dot). The colder, denser fluid is shaded grey. Based on Figure 8.3 in Martin (2006)

Lowering the centre of gravity releases the Available Potential Energy (APE) from the horizontal temperature gradient into kinetic energy (Holton, 2004, Section 8.1), which works to strengthen the winds and conveyor belts within the cyclone, and encourages deepening. Baroclinicity is the key factor in cyclone development, as a storm will not develop without it. Additionally, barotropic instability can be released, by converting kinetic energy from the large-scale jet stream into kinetic energy of the synoptic-scale eddy.

The difference between baroclinic and barotropic instability is well illustrated by the Lorenz Energy Cycle (Figure 2.6). Lorenz postulated that, across the globe, energy transfer in the atmosphere is between four reservoirs: potential energy of the mean flow (Pm), kinetic energy of the mean flow (Km), potential energy of eddies (Pe), kinetic energy of the eddies (Ke). The source of the potential energy is differential heating (the sun is more intense at the Equator than at the Poles, leading

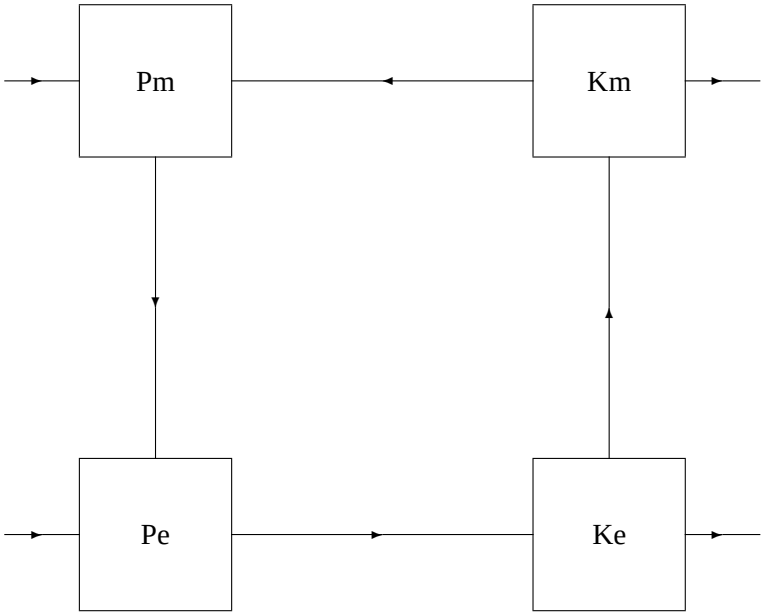


FIGURE 2.6: The Lorenz Energy Cycle, illustrating the flow of energy caused by atmospheric processes, such as extra-tropical cyclones. The stores are mean-state potential energy (Pm), mean-state kinetic energy (Km), eddy potential energy (Pe) and eddy kinetic energy (Ke).

to a temperature gradient); the energy is converted to kinetic energy by circulations forced by the temperature gradient; the remover of kinetic energy from the atmosphere is friction. The transfer of energy is caused by a variety of processes, two of which are relevant to midlatitude cyclones. Baroclinic instability is released during cyclogenesis, transferring energy from Pm to Pe, then on to Ke. Barotropic instability is released by converting Km into Ke (Holton, 2004, Section 10.4). It is from these two types of instability that midlatitude cyclones draw their kinetic energy, leading to cyclogenesis.

2.2.4.2 Jet Stream

There is a strong link between the jet stream and midlatitude cyclones. Once a perturbation appears on the jet stream, it will tend to grow because the jet stream is unstable. A mature perturbation causes flow to accelerate, inducing ageostrophic motion. This in turn causes divergence aloft, which means mass can be evacuated from the column of air, and so the surface pressure lowers.

Therefore, how much a cyclone will deepen depends partly on whether it passes through regions of upper level divergence.

The acceleration of the flow that causes the ageostrophic motion can be linear or circular. If it is linear, regions of particularly fast-moving air form, known as jet streaks; each will have divergence at the right entrance and left exit regions, due to ageostrophic circulation around the axis of the jet stream (Murray and Daniels, 1952). Therefore, a storm that crosses the jet stream and passes through these two regions usually undergoes strong deepening, though it is a nonlinear interaction (Gilet et al., 2009). The locations of the regions of divergence and convergence for straight and curved jets are shown in Figure 2.7. It is the regions of divergence aloft that provide favourable conditions for cyclogenesis. Alternatively, this can also be explained through the idea of relative vorticity (ζ), given by:

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \tag{2.1}$$

where u and v are the zonal and meridional components of the wind speed, and x and y are the zonal and meridional spatial co-ordinates. The relative vorticity describes the spin of a fluid at a point, and is related to the air's circulation per unit area, when viewed from Earth. The absolute vorticity (ζ_a) is the circulation per unit area when viewed from an inertial, non-rotating frame of reference, and is given by:

$$\zeta_a = \zeta + f \tag{2.2}$$

where f is the planetary vorticity or Coriolis parameter, quantified by:

$$f = 2\Omega sin\phi \tag{2.3}$$

where Ω is the rate of rotation of the Earth, and ϕ is the latitude. Relative vorticity — often referred to as vorticity — is positive for cyclonic flow (where an air parcel would turn anticlockwise) and negative for anticyclonic (clockwise) flow. In situations of horizontal wind shear, one side of an air parcel will be associated with higher wind speeds than the other, which imparts a turning motion on the air parcel. This is known as shear vorticity. The result is that on the northward side of a jet streak there is positive vorticity and on the southern side there is negative vorticity.

However, it is the advection of vorticity by the jet stream that leads to convergence and divergence aloft (Lynch and Cassano, 2006, Section 7.4). Positive vorticity advection (PVA) is associated with divergence aloft, and negative vorticity advection (NVA) is associated with convergence aloft. This is because, as an air parcel increases its vorticity, its circulation is conserved so its radius decreases and vertical extent increases. Given that a jet streak has positive vorticity on the northward side and negative vorticity on the southern side, PVA occurs in the right entrance and left exit regions, thereby providing favourable conditions for cyclogenesis.

A curved jet stream results in centripetal acceleration acting on the flow. The resulting ageostrophic motion is perpendicular and to the left of the acceleration vector (Holton, 2004, Equation 6.56), meaning there is convergence behind the trough and divergence ahead of the trough (Figure 2.7). From a vorticity perspective, it is necessary to consider curvature vorticity, or the spin imparted on an air parcel due to curvature of the flow. Vorticity will be most strongly positive in the base of the trough, and most negative at the peak of the ridge. Therefore, behind the trough there is NVA and associated convergence aloft, but ahead of the trough there is PVA and divergence aloft (Holton, 2004, Section 6.3.2).

In summary, the presence of accelerations within the flow that generate agostrophic motion is key for cyclogenesis, because through divergence aloft, mass is evacuated from the cyclone's core, causing the core pressure to decrease. Furthermore, there is an important link to be made with vorticity generation. If a column of air is stretched upwards so that its radius decreases, then through the conservation of angular momentum it will tend to 'spin up', which also causes divergence aloft.

2.2.4.3 Charney and Eady

The ideas of baroclinicity and vorticity advection are key in both the initiation of a mid-latitude cyclone, and in its subsequent development. If the jet stream advects cyclonic vorticity into a region, it will encourage divergence aloft and strengthen upward motion (Martin, 2006, Section 6.2). This means that, for a deepening system, the maximum in relative vorticity associated with a trough will tend to slope backwards with height; the surface trough is below the region of strongest PVA, ahead of the upper trough (Charney, 1947; James, 1995, Section 5.2).

Idealised versions of baroclinic waves were described by Charney (1947) and Eady (1949), using mathematical models. Eady (1949) developed a two-layer mathematical model to examine the

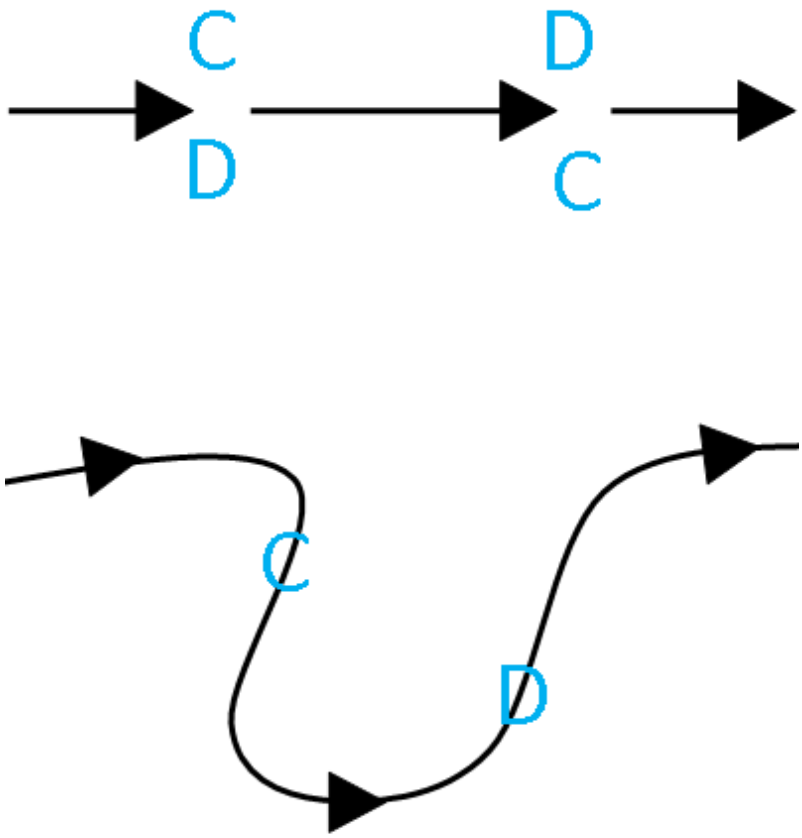


FIGURE 2.7: Showing the location of regions of divergence and convergence, associated with the jet stream. C indicates regions of convergence; D indicates divergence.

growth of baroclinic waves, as an approximation to the midlatitude atmosphere. The idealised model had fixed upper and lower boundaries, a fixed channel width, and a dividing line between the two layers assumed to be flat. The flow within the channel was zonal and varied linearly with height from the lower boundary; faster motion aloft, and always flowing in the same direction. The model in Charney (1947) was similar, but a little more flexible as it lacked the fixed the upper boundary or dividing line. However, this complicates the problem mathematically, and so the Eady set-up is used more commonly.

Eady (1949) observed that a small perturbation in the thermal field tended to grow into a large-scale wave structure in his two-layer model. Although initially many different wavelengths were triggered by the perturbation, one grew exponentially, and so evolved into the predominant wave that characterises size, structure and growth rate. The fastest-growing resulting wave was seen in both the temperature and pressure fields, so it was a baroclinic wave, and its amplitude is differential (varies with height). The wave was an idealised version of a midlatitude wavetrain, with alternating low and high pressure centres, known as ridges and troughs respectively. Therefore, it contained an idealised version of a midlatitude cyclone. The idealised version was described analytically by Charney and Eady, in this case using the partially-developed set of quasi-geostropic equations. These described an atmosphere in hydrostatic and geostrophic balance, which is a good approximation for midlatitude synoptic-scale flow, away from the Earth's surface (Martin, 2006, Section 5.4). Eady (1949) agreed with Charney (1947) that the trough slopes backwards with height, but also found that waves in the temperature field — that is, alternating regions of warm and cold air — tends to slope forwards with height. This causes overturning, and so the release of kinetic energy from the potential energy of a horizontal temperature gradient.

Furthermore, a baroclinic atmosphere has a horizontal temperature gradient that is not parallel to the pressure gradient, and so provides the potential for differential temperature advection. Ahead of the low pressure, warm advection occurs, causing sensible heating of the air behind the warm front. This air is further warmed by the latent heating caused by cloud formation. Warm advection causes upward motion, so the air column stretches and 'spins up' cyclonic vorticity, causing surface pressure to fall. Behind the low pressure, cold advection occurs causing 'spin down' and rising surface pressure. These respectively pull and push the low pressure further forward, meaning that the cyclone propagates. The uplift also causes more cloud to form, so latent heating is increased and the warm advection is strengthened, causing a positive feedback with more uplift. Also, the latent heat release in the cloud ahead of the warm front causes the ridge ahead of the system to intensify, meaning the upper level pattern amplifies as its wavelength shortens. The upper-level trough generates stronger PVA, which also causes stronger ascent and strengthens the surface cyclone. This idea is known as self-development, as the cyclone reinforces itself (Martin, 2006, Section 8.5).

2.2.4.4 Diabatic Processes

Further to these large-scale processes, diabatic processes can contribute to a cyclone's development. Adiabatic processes are those where any change in an air parcel's temperature is due solely to its expansion or contraction, due to vertical motion in the atmosphere. Diabatic processes include the release of latent heat and sensible heat, as well as interactions with the land and sea surface. As clouds form, the phase change of water provides a source of diabatic heating, converting latent heat energy into kinetic energy and affecting local vertical temperature gradients and therefore stability. These can combine to provide stronger updraughts, a stronger WCB (Chagnon et al., 2012), and more intense fronts, than if the atmosphere were dry and clouds did not form.

However, diabatic processes occur on small scales, making it necessary to parameterise them in NWP and climate models (Section 2.5.1). Further to the effects on the small scale stability, latent heat release can affect the large-scale structure of the cyclone. It can affect uplift speed in the WCB, and so decrease the core pressure of the cyclone further and intensify the cyclone. Chang et al. (1982) found that diabatic processes affect large-scale structure of the cyclone, as the tilt of the cyclone is different when latent heating is modelled.

2.2.5 Convection

Convection occurs when the air near the surface is warm compared to the air aloft, so if the air lower down is perturbed it will tend to rise rise. This air is said to be unstable. If the air has sufficient moisture, this can lead to the formation of convective clouds. Within a cyclone, there are convective features. Convective cells are often found in the cold air behind the cold front, as the air near surface is warmed by the land or sea over which it passes but the air above it is colder, which can give showery weather (Wallace and Hobbs, 2006). Also, there can be embedded convection in split cold fronts, which occur when the WCB slopes forward (Bader et al., 1995), and can generate heavy precipitation and are a significant hazard to aviation. While the uplift of air is not a diabatic process, it is still parametrised because it is often on a scale much smaller than the model grid. The highest resolution data that will be used in this project has grid squares of around $8km$, which is significantly larger than the average convective cell. Furthermore, when clouds form, there is latent heat release, which is a diabatic process and is parametrised. While convection is an important facet of midlatitude cyclones, it will only be considered implicitly as part of this work, because it is parametrised within the forecasting model of choice.

2.2.6 Planetary Rossby Waves

Although it remains a useful concept in cyclone development, the presence of the Polar Front is not necessary for frontal development. One reason for this is that cyclones and anticyclones tend to alternate along the path of the jet stream, because one tends to induce the other downstream. Therefore, in the Northern Hemisphere winter, there is a semi-permanent pattern of alternating high and low anomalies in the geopotential height field (Lynch and Cassano, 2006, Figure 8.3), known as a planetary Rossby wave (Holton, 2004, Section 7.7). The pattern is forced by three factors: orography, notably the Rocky Mountains; differences in heating over land and sea; and forcing from smaller-scale features, including midlatitude cyclones. In Figure 2.8, there is clearly a pattern of regions that tend to have high pressure anomalies alternating with preferred regions for low pressure, in this case with three cycles from high to low and back again occurring around the globe. These large-scale ridges and troughs will affect the path of the jet stream, generating regions ideal for cyclonic deepening, due to divergence aloft. Therefore, there is a mutual interaction between the global Rossby wave pattern and North Atlantic windstorms: Rossby waves can influence cyclone development, in terms of both location and depth, and the cyclones contributes to the forcing of the Rossby wave. Cyclones' dissipation can be linked to the breaking of Rossby waves; as the breaking occurs, the upper-level trough decreases in magnitude, and so the coupling throughout the atmosphere is less strong and cyclolysis begins at the surface (Thorncroft et al., 1993).

However, planetary Rossby waves are not the only kind of Rossby wave. One example of another is the way that the midlatitudes and tropics interact with each other, through the propagation of Rossby waves. The El Niño Southern Oscillation (ENSO) is one example. ENSO is a variation in the ocean circulation in the tropical Southern Pacific, which has an effect on sea surface temperatures in the eastern Pacific and therefore the local weather, for example the location of cloud formation and the strongest precipitation. The two phases are known as El Niño, with a warm anomaly, and La Niña with a cool anomaly. However, studies such as Trenberth et al. (1998) showed that there is also an effect on the weather felt further afield. In Europe, for example, El Niño events mean the winter storm track is more southerly than in a La Niña year (Brönnimann, 2007). This effect on European weather occurs due to the propagation of Rossby waves from the tropical Pacific across a large portion of the globe.

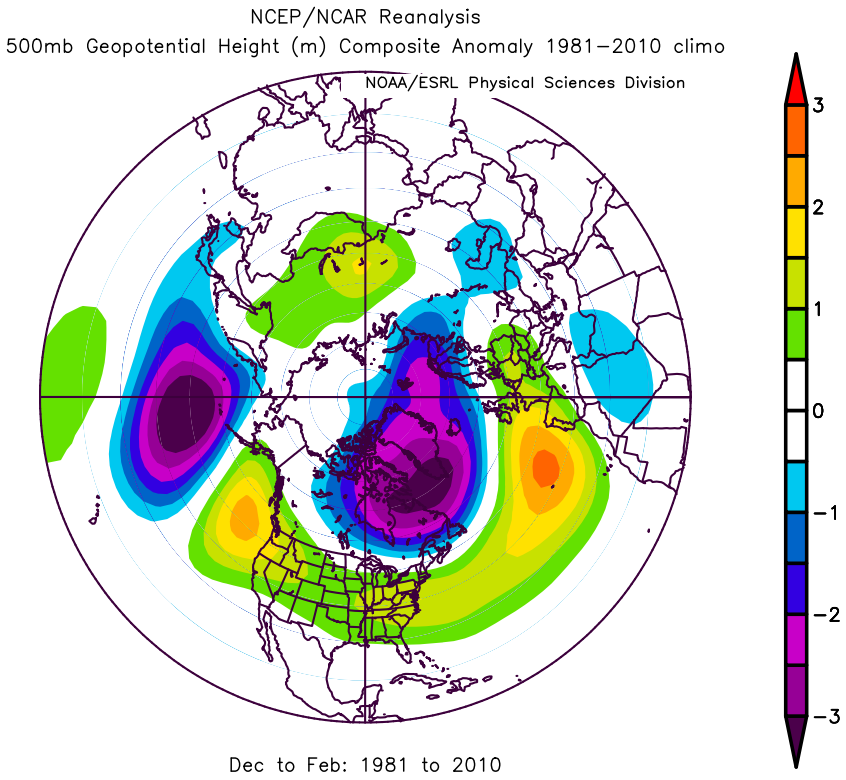


FIGURE 2.8: Northern Hemisphere 500hPa geopotential anomaly for December to February, averaged over 1991 to 2010. Made using NCEP/NCAR reanalysis data at http://www.esrl.noaa.gov/

2.2.7 Intense Cyclones

This work considers extreme cyclones, though Roebber and Schumann (2011) found that such cyclones contain similar processes to their weaker counterparts. The deepest cyclones are a 'lucky accident' when there is a large amount of available moisture to facilitate the diabatic processes, as well as strong baroclinicity and a jet stream showing strong PVA. However, intense cyclones can undergo more sudden deepening, and are called 'bomb' cyclones if deepening is greater than 24hPa in 24 hours (Sanders and Gyakum, 1980).

When considering the deepening of intense midlatitude cyclones, a large number of factors are at play. Processes that need successful modelling range from the large-scale forcing, such as the baroclinicity and jet stream, to the small-scale diabatic processes. Each of these must be modelled successfully, in order to provide a high-quality forecast for the coming days, or to predict how such systems will evolve under climate change. Explosive cyclogenesis also tends to occur near strong sea surface temperature (SST) gradients (Sanders and Gyakum, 1980), so SSTs also need

careful modelling. Sea ice concentrations needs similar consideration, because it to has an effect on midlatitude circulation.

In addition, explosive cyclogenesis is usually associated with wave activity. Lackmann et al. (1996) find that explosive cyclogenesis is associated with a planetary-scale wave. Hanley and Caballero (2012) found that before a destructive cyclone passes over Europe, there tends to be simultaneous Rossby wave breaking of both kinds; cyclonic on the polewards side and anticyclonic on the equatorward side. This induces a region of strong baroclinicity to the west of Europe, providing a potential source of energy for windstorms. Gómara et al. (2012) agreed that explosive cyclones tend to be associated with Rossby wave breaking, but find that the type of breaking is affected by the cyclone's location. Given that Rossby wave breaking can be forced by cyclones' dissipation (Section 2.2.6), then intense cyclones can be affected by the paths and direction of earlier cyclones' wave breaking.

Rivière and Joly (2006a) observed that the largest deepening rates for cyclones occur downstream of the maximum baroclinicity. They found that deformation is key to the deepening of some past intense cyclones, and identify 'barotropic critical regions' between areas of strong deformation of opposite direction, which cause strong deepening. In the second part of this work, Rivière and Joly (2006b) found that when two jets are present, their deformation fields superimpose on each other and cause a 'baroclinic critical region' in between. If a cyclone were present in this region, then the resulting baroclinic energy conversion would deepen the cyclone quickly.

Overall, a variety of factors can come together to generate an intense midlatitude cyclone, with the potential to inflict damage on Europe. Therefore, how well each of these factors is modelled will affect how well such cyclones are modelled. Next, a discussion of the level of understanding of midlatitude cyclones will ensue, by evaluating how well represented they are in several types of data.

2.3 Data Sets for Investigating Historic Cyclones

There are several sources of data, in which to investigate midlatitude cyclones. Many studies searched for trends in cyclone activity, in order to assess how such cyclones will evolve under climate change. Firstly, data from the recent past can be examined for such trends, which may indicate changes to come. These data come in two forms: direct observations and reanalysis data.

Direct observations are those made every day at thousands of weather stations, satellites, radar dishes, instrumented buoys, and so on. To investigate storms over a longer time period, ships log books, diaries, newspapers and port records have also been used (Lamb, 1991).

The modern observations are used to generate analysis charts, which provide the best estimation of the situation at a given moment given the observations, and so are used to make forecasts. The method of going from observations to analysis — known as data assimilation — has been refined over time, along with other factors such as the resolution. This inconsistency makes comparison of data of different ages difficult, as differences might be related to advances in the data assimilation. Reanalysis data are similar to analysis data, but a consistent data assimilation procedure and resolution is used, but gaps between observations by interpolating in time and space using an NWP model run for short intervals over a long period of time.

Secondly, 'weather forecasting' models that are run on short timescales using NWP. These depend strongly on the initial conditions and fast processes that occur in the model. Two different types of short-term forecast exist: deterministic and ensemble. Deterministic forecasts are run only once at a high resolution, and so are typically more accurate up to approximately four days. Ensemble forecasts differ, as the model is run several times, with slight differences in the initial conditions or governing parameters, giving a spread of results.

Lastly, climate models, both global (GCM) and regional (RCM), can be used to simulate future trends. Since they run on much longer time-scales, they are more sensitive to slow processes than NWP models. RCMs can be run at higher resolution than their global counterparts, allowing more detailed information to be gathered for impacts and adaptation studies. However, RCMs can only be run over a small region and so require a GCM to provide boundary conditions. It is also possible to run GCMs and RCMs for the recent climate (20th century), and then compare their output to surface observations or reanalysis data, to assess the accuracy of the models. Each of these data types will now be discussed in turn.

2.3.1 Surface Observations

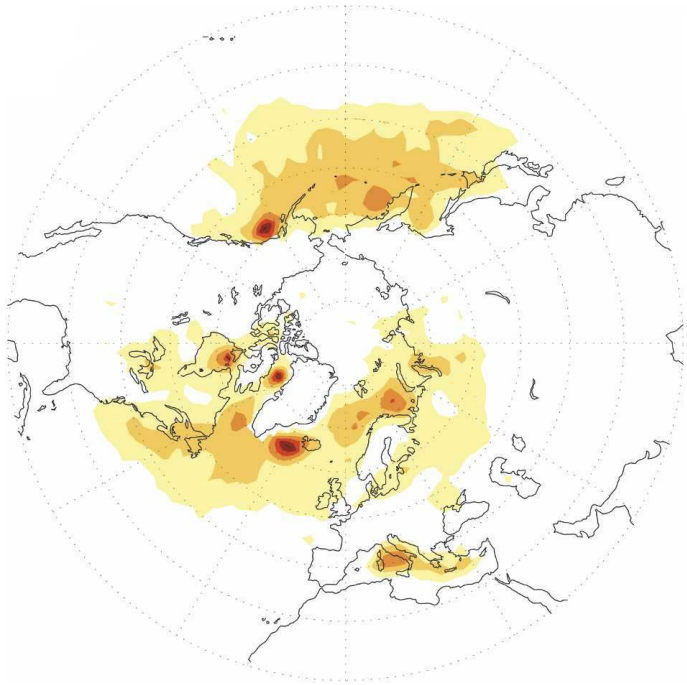
As observations in Europe are particularly dense, research has sought trends in observed cyclone patterns over Europe, but no clear conclusions can be drawn from these investigations. McCabe et al. (2001) found an increase in cyclone frequency north of $60^{o}N$, a decrease at 30 to $60^{o}N$, and an intensity increase everywhere, for the period 1959 to 1997. Harnik and Chang (2003) found a

strengthening of the storm track, particularly since the 1970s, but Alexander and Tett (2005) found no clear trends in winter storms since the 1950s. Cited reasons for this disagreement include the short observational record, relative to the timescales of internal variability; decadal variations cannot be assessed on 60 or fewer years' worth of data. Further studies have investigated changes in cyclone-related variables. However, no clear trend in the North Atlantic has emerged yet, despite investigation into a range of variables including mean sea-level pressure (MSLP) (Lambert, 1996), geostrophic wind speed (Alexandersson et al., 2000), ocean wave height (The WASA Group, 1998), temperature and precipitation (Beck et al., 2007). Longer time periods have also been investigated, for example by Phillipp et al. (2007), who grouped or clustered North Atlantic pressure patterns from 1850 to 2003. They did not find a significant trends in any particular cluster, but did find that there is a trend towards winter conditions with westerly winds, which is favourable for strong windstorms. Lamb (1991) investigated cyclones over nearly 500 years, but finds no trends related to climate change. The reasons for the absence of trends include natural variability, improvements in the instruments used in the observations, sub-grid-scale processes, variations at the synoptic scale and changes to boundary conditions. On balance, no clear trend emerges from surface observations..

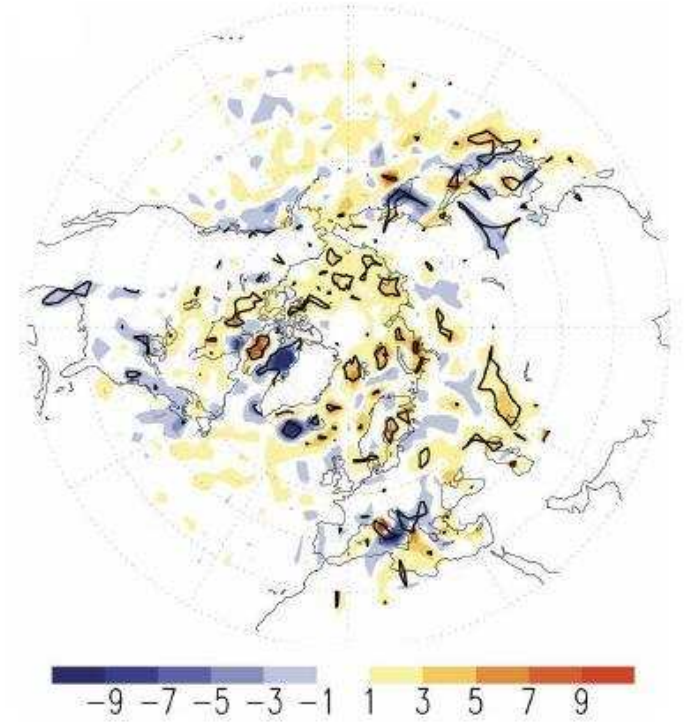
2.3.2 Reanalysis

Recent comprehensive reanalysis activities such as ERA-40 and NCEP-NCAR have facilitated trend analysis. Reanalysis tends to be run for a historical period (e.g. the ECMWF Reanalysis ERA-40, 1957-2002), but can continue from a point in history up to the present day (e.g. the ECMWF Reanalysis ERA-Interim, 1979-present). When investigating extra-tropical cyclones, the ERA-40 and NCEP-NCAR reanalyses generally agree, although ERA-40 contains more cyclones overall. According to Hodges et al. (2003), the storm tracks in the reanalyses are mainly consistent in the Northern Hemisphere, though the degree of consistency depends on the cyclone intensity. Figure 2.9 shows the discrepancies between the ERA-40 and NCEP-NCAR reanalyses, in terms of mean cyclone centre density. It shows that there are some significant differences between the two reanalyses, notably in the Baffin Sea and around Iceland because in these two areas European windstorms are often generated. However, the two reanalyses broadly agree on the location and intensity of the major storm tracks.

Raible et al. (2008) found that the cyclone trends in ERA-40 and NCEP-NCAR over the time period 1961 to 1991 also broadly agreed with each other. Generally, a northward shift of the



(a)



(b)

FIGURE 2.9: Figures from Raible et al. (2008) (their Figure 2), showing (a) total cyclone centre density in ERA-40, for storms that with a minimum lifetime of 72h for winters (DJF) 1961-1990 and (b) the difference in the same measure between ERA-40 and NCEP-NCAR reanalyses. Black contours show 5% significance level. Cyclone centre density is the number of cyclones per unit time (in this case, a season), per unit area ($1000\,km^2$)

zonal-mean storm track occurs over the time period (Trigo, 2006; Wang et al., 2006), possibly due to a northward shift in the jet streams (Hudson, 2012). Geng and Sugi (2001) found an increasing trend in 'cyclone density', which could indicate stronger cyclones, faster speeds and/or stronger deepening rates. Allen et al. (2010) examined five reanalyses that cover 1979 to 2008, and also find that they agree with each other for explosive ('bomb') cyclogenesis in the Northern Hemisphere, but no trend is found. One reanalysis, 20CR, goes back to 1871 by taking surface pressure observations and using these as input to a model, resulting in gridded data. Donat et al. (2011) and Wang et al. (2013) found an increasing trend in storminess in 20CR; however, Krueger et al. (2012) found no trend in the observations fed into 20CR. This means that the trend could be related to the limited number of observations earlier in the time period; a limitation of 20CR also noted by Compo et al. (2011). As an alternative to extending the time period of a reanalysis, it is possible to take a different approach with the analysis. Della-Marta et al. (2009) calculated the return period of European windstorms, and found an increasing frequency (i.e. shorter return periods) over the Northeast Atlantic, but noted that their statistical method has some drawbacks such as potentially oversimplifying the dynamics of a wind storm in order to sucessfully identify what constitutes an 'event'.

There are a number of factors that can affect the reliability of results from reanalysis data (Bengtsson et al., 2004a). The underlying NWP model may cause ERA-40 to have a trend of increasing temperature and atmospheric water vapour content (Bengtsson et al., 2004b). Inhomogeneities in observations are still a factor, because the number and quality of observations fed into the reanalysis increases over time. Reanalysis data is also affected by the data assimilation method used to include observations (Harnik and Chang, 2003); although in a given reanalysis data set it remains constant, it could cause differences between reanalyses. Finally, the problem of internal variability masking climate change trends remains, given that the time periods concerned are short compared to the time period of internal variability. Overall, uncertainty is high when investigating climate trends found in reanalysis data (Bengtsson et al., 2004b).

2.3.3 Forecasting on a Range of Scales

Since forecasts are made every day at dozens of centres for a range of lead times, it is possible to assess how well extra-tropical cyclones are modelled using these data. Forecasting is done in both traditional deterministic weather forecasting models, which use a single analysis, and using ensemble prediction systems (EPSs). EPSs use a set of perturbations to the analysis to generate

their initial conditions, so can give a measure of the uncertainty in the forecast (Leutbecher and Palmer, 2008). In two EPSs, made by the ECMWF and NCEP, cyclone location is predicted more reliably than intensity, but both predict cyclones to track slightly more polewards than in reality (Froude et al., 2007a). It has been shown that the speed of storms is systematically underestimated in the ECMWF EPS system (Froude et al., 2007b) and eight other EPS systems (Froude, 2010) including NCEP-NCAR, particularly over the Atlantic (Froude, 2009). However, external factors may be at play; one key factor in the performance of an EPS is the number of ensemble members (Buizza and Palmer, 1998), both in terms of spread and skill. Thus, comparing EPSs can mean those with fewer ensemble members appear to perform worse than those with more members, due to the member numbers factor alone. Forecasts are also affected by the quality of their initial conditions; uncertainty in the observations can propagate in the forecast. Another potential source of uncertainty is the data assimilation scheme, which takes the observations and processes them into the analysis or reanalysis.

In the wider meteorological community, there is an interest in the 'seamless' approach to forecasting (Hoskins, 2013; Palmer et al., 2008), whereby one computer model can be used to simulate any timescale, because the simulation would include all processes, from fast processes such as diabatic ones, to the slow processes such as the atmosphere's interaction with the cryosphere and biosphere. The next link in this chain towards a seamless approach is seasonal forecasting. However, Renggli et al. (2011)'s examination of hindcasts of windstorms on a seasonal scale in an ensemble of models found a little skill in predicting the number of winter cyclones, but no skill for the intensity. Further difficulties are anticipated on the seasonal-to-decadal time-scale; uncertainty on these time-scales was identified in the emissions scenario, in model error, and in the predictability of the climate system (Schwierz et al., 2006). Some work has already shown that running climate models for a few timesteps and comparing the result to reanalysis data indicates which 'fast' processes need to be improved (Rodwell and Palmer, 2007). This would also signpost what processes that act on seasonal or decadal timescales need improved representation or parametrisation, which could be carried forward into models that operate over longer time scales, such as climate models.

2.4 Cyclones in Climate Models

How midlatitude cyclones will react in a changing climate one of the most uncertain aspects of European weather's reaction to climate change (Beniston et al., 2007). A number of studies compare

extra-tropical storms in reanalyses and climate models, to draw conclusions about the accuracy of the climate models. However, studies take a variety of approaches in the analysis (Sections 2.5.4 and 2.5.5). In Pinto et al. (2009), when a climate model run for the recent past is compared to NCEP reanalysis, the model does successfully locate the region affected by storms (Figure 2.10). The figure shows that, although the general shape of the track is similar, it is marginally shifted south in the climate mode and that the model underpredicted the number of cyclones over the North Atlantic generally and by 50% in some regions.

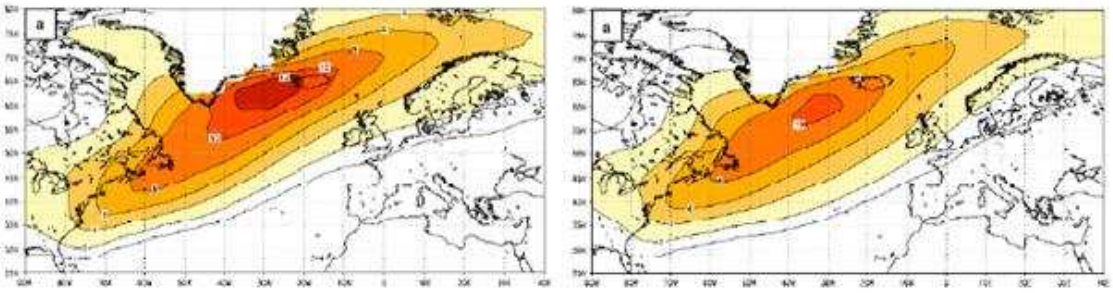


FIGURE 2.10: Track density (in cyclone days per winter) of intense cyclones for (left) NCEP reanalysis (1958-1998) and for (right) a simulation with ECHAM5/MPI-OM1 for recent climate conditions (1960-2000). The two data sets have similar horizontal resolutions. From Pinto et al. (2009), Figures 4a and 13a.

Once a comparison with past data has been made — either with reanalysis data or with observations — many studies go on to use a climate model to simulate how cyclones may differ under future climate change, despite the limitations to the climate model discussed. Furthermore, the results produced by different climate models are inconsistent; indeed, the disagreement is particularly bad over Northern Europe (Christensen et al., 2007; Woolings, 2011).

2.4.1 Extra-tropical Storms in Multi-Model Ensembles

The IPCC Fifth Assessment Report (AR5) states that 'the global number [of extra-tropical cyclones] is unlikely to decrease by more than a few percent' (Christensen et al., 2014), so any change is likely to be a small decrease in number. While the Pacific storm track is likely to shift polewards, the same cannot be said for the North Atlantic: 'it is unlikely that the response of the North Atlantic storm track is a simple poleward shift' (Christensen et al., 2014). This is a different conclusion to the Fourth Assessment Report (AR4), which states that globally, midlatitude storms

will become less in number due to a poleward shifted storm track, and that because the same conclusion is reached by several models, confidence in the result is good (Meehl et al., 2007).

Other studies have examined the data used in these reports directly (e.g. Lambert and Fyfe, 2006). However, CMIP3 (used in the Third Assessment Report and AR4) and CMIP5 (AR5) data include such a large number of models that it is difficult to critique the differences between the models. There is only weak general consensus found, though some agreement is found between some models. Furthermore, results from CMIP5 data showed a different pattern to that found in CMIP3 data; for example, the poleward shift is only found over Europe, North America and the western North Pacific. Chang et al. (2012) found that CMIP5 predicts a slight decrease in storm track activity for the Northern Hemisphere during winter, whereas CMIP3 predicts an increase. Harvey et al. (2012) compared CMIP5 and CMIP3, and generally find the CMIP5 response in terms of cyclone number and track position is less than in CMIP3 and is smaller in magnitude than interannual variations, making trends difficult to identify. It remains an open question how well extra-tropical cyclones in models compare with reality; when Zappa et al. (2013) compared storms in these model suites to those in reanalysis data, CMIP5 performs better than CMIP3. This disagreement indicates that there are deficiencies in the simulation of the North Atlantic storm track in climate models.

2.4.2 Storm Track in the Future Climate

The storm track is a region where cyclones tend to be found and move through. It is a useful measure for storminess in climate models, because it is an average over a large area and time period, so can be easily detected at GCM temporal and spatial resolution. For the North Atlantic storm track to be clearly discernible in the correct area, it is not necessary to resolve processes that act over small distances or short time periods. One method of locating the storm track is described by Hoskins and Hodges (2002); discussion of different methods can be found in Chapter 3. Hoskins and Hodges (2002) used relative vorticity (ζ) and find the storm tracks shown in Figure 2.11, which clearly shows the North Atlantic storm track as it heads over Europe.

Though many studies using climate models agree that the North Atlantic zonal-mean storm track will shift polewards in a changing climate (e.g. Bengtsson et al., 2006), this is not a universal conclusion (e.g. Pinto et al., 2007a), as can be seen in Table 2.1. This discord means that some processes that drive the storm track are not well modelled. Furthermore, Catto et al. (2011) found a northward shift of the storm track in an experiment with a doubling of carbon dioxide, but no

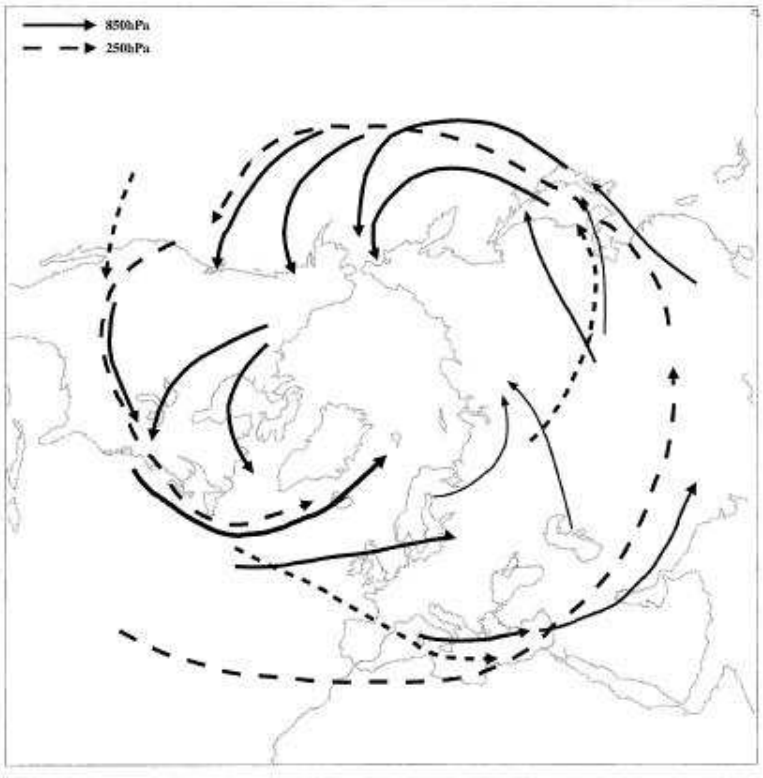


FIGURE 2.11: Main storm tracks in the Northern Hemisphere at 850hPa (solid) and 250hPa (dashed). From Hoskins and Hodges, 2002 (Fig. 14).

change with a quadrupling. Sansom et al. (2013) observed a southward shift of the storm track using a different model weighting technique, intended to minimise uncertainty; however, there is still a large amount of associated uncertainty. There are also some local differences: McDonald (2011) found a general poleward shift, but a partial southward shift causes 47% more cyclones over the UK by the end of the 21st century. In the CMIP5 data, analysis of the storm track over Europe shows an increase in storms over central Europe, but a decrease to the north and south of this area (Zappa et al., 2013). In short, there is some consensus for a poleward shift over the North Atlantic, but uncertainty in the storm track over Europe.

2.4.3 Cyclone Number and Intensity in the Future Climate

Although the storm track is a useful measure of midlatitude cyclone activity, changes to the number and intensity of cyclones also have implications. Overall, a trend towards a decrease in total cyclone numbers emerges, but there is uncertainty as to the extent and regional variation. Donat et al. (2010) compared 7 GCMs, and found local increases in the cyclone intensity of 10% ± 10%

and in the number of cyclones that will affect Europe. Over the north-east Atlantic and Britain, the recent literature points towards either little change in cyclone numbers, or more strong cyclones (Table 2.1). For example, McDonald (2011) used the Met Office's HadAM3P model to conclude that global winter cyclone numbers will fall by 3% by 2100, but found no change over the North Atlantic.

How intensity will evolve under climate change is highly uncertain, given the variety of results shown in Table 2.1. Table 2.1 shows a slight trend towards increased storm intensity or more frequent intense storms, but this conclusion is far from definite. The reasons for this discord will be discussed in Section 2.5.

2.4.4 Cyclones in Regional Climate Models

Regional climate models (RCMs) are a useful tool for examining the smaller-scale effects of climate change, but require a driving GCM to provide boundary conditions. This is because RCMs are often run over a small area, meaning some processes involved in simulating a cyclone such as the jet stream's path operate on a larger scale. Leckebusch et al. (2006) located cyclones using $\nabla^2 p$ (where p is mean seal level pressure) in four GCMs (ECHAM4/OPYC3, ECHAM5/MPI-OM1, HadAM3P and HadCM3), and four RCMs (RCAO, HadRM3P, HIRHAM4, CHRM). Overall, cyclone track density decreased over central Europe but more intense storms affected Britain. The number of extreme cyclones increased over Western Europe and the north-east Atlantic. Differences between the RCMs were found to be dependent on the driving GCM in winter, in agreement with Schwierz et al. (2010).

In summary, there is some disagreement about whether the storm track will shift, and more about how the number and intensity of North Atlantic cyclones will evolve under a changing climate. This is despite using a range of data from observations, reanalysis, forecasts, GCMs and RCMs. As McDonald (2011) noted, '*uncertainty remains high in future predictions of cyclone behaviour*'.

2.5 Reasons for Discord

The reasons for the disagreement regarding the trends in midlatitude cyclones can be divided into those that affect all aspects of such modelling, and those that only apply to such weather systems.

Source	Model	Method	Results	
			Zonal Mean Track	Cyclone Number
Lambert (1995)	CCC GCMII	Φ		Fewer cyclones More intense cyclones
Beersma et al. (1997)	ECHAM3	Φ_{500hPa}		Fewer cyclones Fewer intense cyclones
Carnell and Senior (1998)	HadCM2	p		Decreased global activity
Sinclair and Watterson (1999)	CSIR09	ζ, from $\Phi_{1000hPa}$		Fewer intense cyclones
Knippertz et al. (2000)	ECHAM4/OPY3	$\Phi_{1000hPa}$	Shift polewards and east	More intense cyclones
Geng and Sugi (2003)	JMA GSM8911	p		Fewer cyclones overall More intense cyclones
Leckebusch and Ulbrich (2004)	HadCM3	$\nabla^2 p$, p		More cyclones over W. Europe
Bengtsson et al. (2006)	ECHAM5	ζ_{850hPa}	Shift polewards	Little change
Lambert and Fyfe (2006)	IPCC suite	p		Fewer cyclones More intense cyclones
Leckebusch et al. (2006)	ECHAM4/OPYC3 ECHAM5/MPI-OM1 HadAM3P HadCM3	$\nabla^2 p$	Shift poleward Increase in wind speeds due to deeper cyclone cores and shift in storm track	Decreased activity More cyclones over UK
Pinto et al. (2006)	ECHAM4-OPYC3	$\nabla^2 p$	Small shift polewards	Increased activity
Jiang and Perrie (2007)	CCCma CGCN2	p		Little change
Pinto et al. (2007a)	ECHAM5/MPI-OM1	p	Little change	Increased activity More intense cyclones
Ulbrich et al. (2008)	IPCC suite		Shift polewards	Little change
Löptien et al. (2008)	ECHAM4/OPYC3	p	Shift polewards Slower moving cyclones	Little change
Pinto et al. (2009)	ECHAM5/MPI-OM1	p, $\nabla^2 p$		10% fewer cyclones
McDonald (2011)	HadAM3P	p	Broad shift polewards but 47% more cyclones over UK	Global 3% decrease No change over N. Atlantic
Catto et al. (2011)	HadAM3P	p	Polewards in $2 \times CO_2$ No change in $4 \times CO_2$	Decrease in intensity

TABLE 2.1: Summary table of studies considering cyclones under climate change in a GCM, having compared the model to analysis or reanalysis. North Atlantic and Europe considered, unless otherwise stated. p is mean seal level pressure, ζ is relative vorticity, and Φ is geopotential.

Using models to simulate the atmosphere has limitations, because models are imperfect and some processes must be parametrised. In general, there are three sources of uncertainty when modelling future climate change (Deser et al., 2011): forcing, model response and internal variability, discussed in Chapter 1. Yip et al. (2011) and Hawkins and Sutton (2009) both found that uncertainty in the emissions scenario dominates uncertainty after 2050. For the next few decades, Hawkins and Sutton (2009) identified that the internal variability and model uncertainty are the largest sources of error in the CMIP3 models, and state that better understanding of them would reduce uncertainty in climate projections.

Internal variability can cloud any trends that emerge in historic data, because it is difficult to identify the extent to which a trend may be due to climate change or due to long-term oscillations in the atmosphere. Model physics is a reason for disagreement in any modelling run, because processes may not be represented faithfully. Resolution is a limit in all types of modelling, because the continuous atmosphere is being simulated on a discretised grid. As well as problems inherent to all climate model simulations, differences can also be incurred when the data is analysed for midlatitude cyclones. Differences can occur in either the method for tracking, or in the way of quantifying cyclones. Finally, there are limitations in our understanding of the dynamical influences on midlatitude cyclones. These sources of error will now be discussed.

2.5.1 Model Errors

The ways in which models simulate dynamics is one reason for the variation between cyclone prediction in climate models, and there are three main dynamical cores that can be used. Eulerian models (e.g. HadCM3) consider each grid point, and how the air over that grid point changes over time. The Lagrangian approach considers many individual air parcels and how they evolve and move over time. A semi-Lagrangian model (e.g. HadGEM1) is a hybrid of the two; it begins by considering each grid point, but then considers the trajectory of the air parcel that is now over it, and how that has evolved. For the Northern Hemisphere storm tracks, results are more realistic with a semi-Lagrangian core than an Eulerian one (Greeves et al., 2007).

In addition, models have been shown to have large-scale biases; systematic shortfalls in how they simulate the atmosphere. This indicates that there are shortfalls in the modelling, because some aspects are unrealistic. Plavcovà (2011) showed that an RCM driven by ECHAM5 has a warm

bias due to overestimating the mean zonal flow, whereas with HadCM3 a cold bias is found due to overestimating the north-easterly flow.

The way models consider dynamics on a smaller scale also varies. As many processes operate at scales smaller than the length of a grid box in a GCM (typically a few degrees latitude or longitude) or of an NWP model (typically of the order of kilometres), parametrisation schemes are developed that simulate their effect on the larger scale. Parametrisation can present problems in simulating cyclones (Lambert et al., 2002), as features such as fronts or clouds can be missed by the grid boxes. Furthermore, many of these processes feed back onto the larger scale; for example, diabatic processes can play an important role in cyclone deepening (Section 2.2.4.4) and are parametrised.

Furthermore, the identification of which factors lead to disagreement could be less clear due to the balance of errors; that is, where the effect of one process being unreliably simulated in a cyclone is compensated for by a different process, making identification of the root cause more difficult. Separating sources of uncertainty would allow the dominant sources to be established, and compensation of errors to be diagnosed.

2.5.2 Internal Variability

Internal variability refers to changes in the climate that are natural and unforced. This can be a problem for two reasons. Firstly, changes due to climate change can be small compared to changes due to internal variability (Beersma et al., 1997), so any longer term trend can be difficult to discern. As discussed in Section 2.3.1, this means that trends in observations can be difficult to identify. Secondly, the understanding of internal variability patterns is low, because they operate on decadal or multi-decadal time scales and the observational record contains too few cycles to ascertain if they are regular or predictable. An important example is the Atlantic Multi-decadal Oscillation (AMO), which is a variation of North Atlantic SSTs with a time period of around 70 years. Given the observational record only goes back to the middle of the 19th century, it only includes at most two AMO cycles, and so the understanding of its dynamical drivers is low. This has an effect on the reliability of climate projections, because these patterns are not well understood so their simulation in climate models may not be realistic. Furthermore, the effect that climate will have on internal variability patterns is difficult to predict. The repercussions for cyclones in future climate are highly uncertain.

Another example of an internal variability pattern is the North Atlantic Oscillation (NAO), which particularly causes uncertainty in how cyclones will react to climate change. Across the North Atlantic, there is a pressure gradient, with high pressure over the Azores and low over Iceland. The NAO describes the strength of this north-south pressure gradient, with positive phases indicating a stronger gradient and negative phases a weaker gradient (Wanner et al., 2001). The pressure anomalies associated with two winters which were extreme examples of each are shown in Figure 2.12. Strong correlation has been found between the NAO index and the number of North Atlantic winter cyclones (Carleton, 1988; Gulev et al., 2001); positive values of the NAO index, with a stronger pressure gradient, mean more frequent and/or more intense North Atlantic cyclones (Pinto et al., 2009). In fact, when the daily NAO index is calculated and composited for the most intense storms, it is found to peak about two days before the time when the cyclone is making landfall, so could be a key factor in the development of strong cyclones (Hanley and Caballero, 2012).

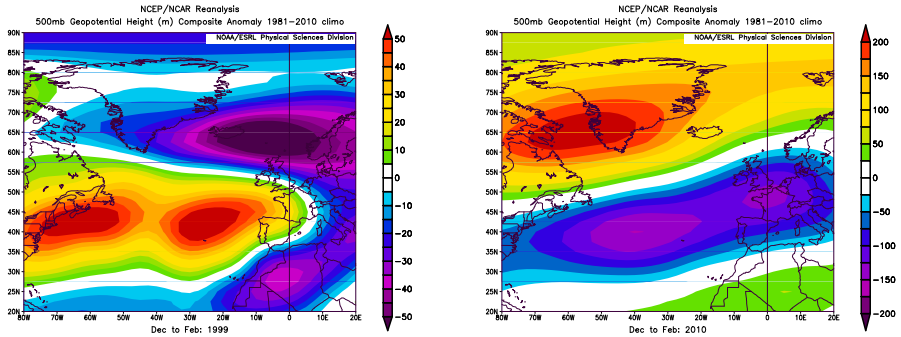


FIGURE 2.12: Northern Hemisphere 500hPa geopotential anomaly for December to February for 1999-2000 (left) and 2010-2011 (right), compared to 1981-2010 climatology, to illustrate the NAO positive and negative phases respectively. Made using NCEP/NCAR reanalysis data at http://www.esrl.noaa.gov/

There is a tendency towards more positive values of the NAO under climate change (Christensen et al., 2014, Section 14.5.1). There is a small increasing trend in the NAO index in fourteen out of fifteen climate models under climate change; however, the models generally overestimate the winter NAO index when compared to observations (Stephenson et al., 2006). When Woolings et al. (2010) investigated the UK Met Office climate model (HadCM3)'s representation of the NAO, it showed: a tendency towards negative index values in winter; a distribution spread too evenly between positive and negative values; and unrealistic time-scales of change between the two schemes. Therefore, climate models can struggle to simulate the NAO.

Overall, the evolution of internal variability patterns such as the NAO and AMO under climate

change in uncertain. Furthermore, variability further afield is known to affect European weather, such as ENSO (Section 2.2.6), but their nature or how they influence Europe could change under an evolving climate. As these factors are known to affect North Atlantic midlatitude cyclones, this adds to the uncertainty in their projections.

2.5.3 Resolution

Resolution in time and space is a limiting factor on the accuracy of all types of modelling, pertinent here as it affects the number of cyclones located. Spatial resolution comes in two forms — grid resolution and spectral resolution — referring to the point at which the truncation happens. Many climate models are not run in gridded space, but instead are transformed using a process similar to a Fourier Transform (Fourier, 1822; Riley et al., 2006, Chapter 13) into a set of smooth wave equations, which are then easier to model — this can be described as modelling wavenumber space. The truncations can therefore happen while still using the wave equations by limiting the number of wavenumbers considered, or after the transformation back to gridded space, by limiting the number of gridpoints extracted from wavenumber space.

Cyclone counts are mainly affected by altering the spectral spatial resolution (Pinto et al., 2005), with more cyclones identified at higher resolution (Stratton, 2004). According to Blender and Schubert (2000) and Tilinina et al. (2013), increasing either spectral spatial or temporal resolution leads to more midlatitude cyclones being successfully located, whereas Chang et al. (2013) found that high resolution models have weak storm tracks and low resolution models have strong storm tracks. Either way, there is clearly a relationship between a model's resolution and its ability to simulate midlatitude cyclones. It is also important to consider vertical resolution, because improving the horizontal resolution shifts the storm track polewards, whereas increased vertical resolution shifts it equatorwards (Stratton, 2004). However, it has been shown that improving either horizontal or vertical resolution in climate models leads to a more realistic distribution of jet latitude over the Atlantic in winter (Anstey et al., 2013). This could be because increased horizontal resolution improving the representation of the baroclinicity, and increased vertical resolution improving the representation of atmospheric stability.

As discussed, diabatic processes are key in the development of strong cyclones. Willison et al. (2013) simulated a case study cyclone at horizontal resolutions of 120km to 20km in the Weather and Research Forecating model (WRF). After 32 hours, the core pressure is 5hPa lower at higher

resolution (Figure 2.13). The sources of the additional deepening are unpicked, and it is found that, in the higher resolution case, there is additional latent heat release when compared to the lower resolution, and so the air flow throughout the cyclone including in the WCB is stronger, enhancing heat and moisture transport, which feeds back into further cyclogenesis. This shortfall of a lower-resolution model to simulate the core pressure of a midlatitude cyclone due to the representation of diabatic processes could indicate one area where climate models are not performing well.

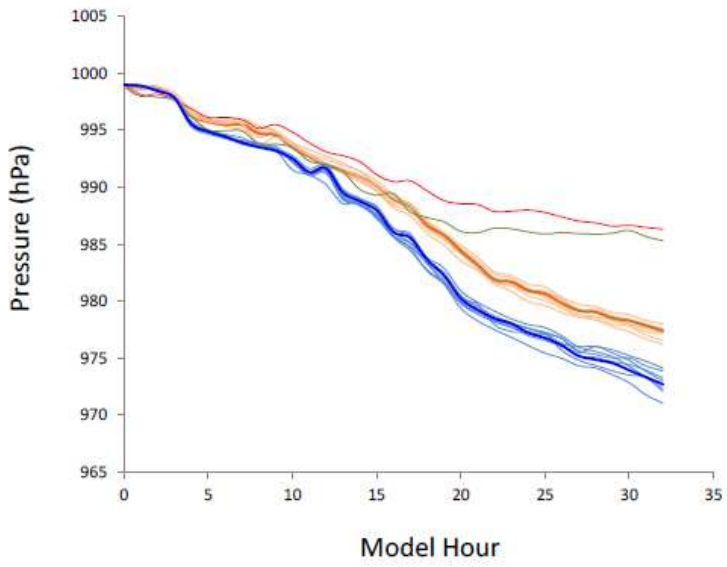


FIGURE 2.13: Willison et al. (2013), Figure 4: Minimum of MSLP for first 32 hours of WRF simluation of a case study midlatitude cyclone. Blue = 120km resolution, orange = 20km resolution, think line=control run, thin line = perturbed physics ensemble members. Two dry runs are also included, red=120km, green = 20km resolution.

2.5.4 Tracking

As well as problems that are inherent in modelling or climate modelling, there are others that are unique to the analysis of cyclones. The first of these is that cyclones need to be identified and tracked, for which there are two main methods (Hoskins and Hodges, 2002). Blackmon (1976) bandpass filtered the 500hPa geopotential height on time periods between about 2.5 and 6 days, and plot its root-mean-square deviation from climatology; in searching for the regions where the variation is largest, he finds the average midlatitude storm track. Building on this work, Hoskins and Hodges also considered filtered fields; a 2-6 day bandpass filter on a range of fields including

pressure, meridional wind, vorticity and potential vorticity to yield a general picture of the storm track.

The second tracking method is considered by Murray and Simmonds (1991) and executed by many other algorithms (e.g. Pinto et al., 2005), whereby a storm centre is located and followed. A cyclone is initially identified as a maximum in $\nabla^2 p$ (which is closely related to vorticity) and then an associated minimum in pressure is found. Once a cyclone has been identified, it is followed by looking at the next timestep, and finding the cyclonic centre there to join to the previous centre. Joining is done by considering how likely it is that the cyclone at the second time is the same entity as that at the first, to prevent the tracking method jumping between systems. When repeated over a few days, the process gives an individual cyclone's track, and when repeated over a season for several years, the storm track appears. Raible et al. (2008) considered this second type of tracking in three different ways, and concluded that they agreed in terms of location, but not in terms of track length.

The IMILAST initiative compared tracking methods (Neu et al., 2013), and finds that there are differences in the results produced by eleven different tracking methods. Although different tracking methods agree under present climate conditions, under future climate the agreement most methods produce no statistically significant signal for track density over Europe, and only one shows a statistically significant increase (Ulbrich et al., 2013). However, Ulbrich et al. (2013) speculated that reason for the lack of a signal could be due to the different treatment of weak cyclones, and find that for strong cyclones, track density increases across all methods. A consistent tracking method can be employed to avoid any possible differences, but use of different methods makes comparing different studies of cyclones difficult.

2.5.5 Quantifying Cyclones

The final source of discord is how cyclones are described numerically. As can be seen from Table 2.1, a range of metrics is used to quantify at the feature-tracking, feature-counting or intensity-measuring stages of the analysis. Different thresholds can be used at any of these stages too, for example to quantify storm intensity. The commonest parameters are mean sea level pressure (MSLP), geopotential height, and vorticity (e.g. ζ_{850}). MSLP and geopotential height tend to locate larger storms and miss the smaller or secondary cyclones, and are poor indicators of how

much damage a cyclone causes (Pinto et al., 2005). ζ_{850} tends to find the smaller cyclones, but is also a poor indicator of damage (Ulbrich et al., 2009).

The differences in these two metrics for locating cyclones is shown in Figure 2.14. There are more cyclones are found in ζ_{850}, particularly in regions prone to smaller cyclones (e.g. the Mediterranean). Even on the main Atlantic and Pacific storm tracks, twice as many cyclones are located in ζ_{850} than MSLP in some regions. Although this would mean that a larger set of cyclones would be examined in ζ_{850} than in MSLP, many of these are nearly undetectable in MSLP and so some of the processes that are described in Section 2.2 could be missing, or at least not at sufficient intensity, to produce a coherent cyclone.

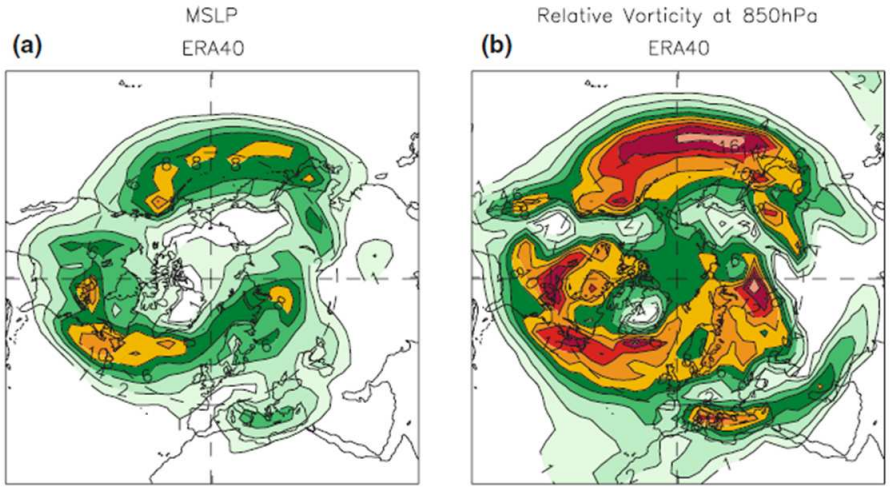


FIGURE 2.14: Track density (10^6 km^2 $month^{-1}$) for the cyclones tracked in winter (DJF) in the Northern Hemisphere. From Greeves et al. (2007), Figure 1.

Another parameter often considered is wind speed at 10m altitude (v_{10}), as it is related to the damage a cyclone causes, because the wind wreaks the damage. However, the ways of considering wind vary, using percentiles or maxima, averaged daily, hourly or not at all, and using the mean wind or the gust speed. Also, v_{10} is not a direct model output, and so is influenced by the boundary layer scheme. Furthermore, if storms are tracked in the surface winds, the result is not representative of how extra-tropical storms generally move. On the large scale, the highest wind speeds are located with the maximum pressure gradient, not with the pressure minimum or vorticity maximum.

The intensity of cyclones can be measured by different parameters, including lowest core pressure, maximum vorticity, mean pressure gradient or $\nabla^2 p$ (Ulbrich et al., 2009). The differences in trends identified in intense cyclones depend upon the definition of intensity; Ulbrich et al. (2009) consistently used the 99th percentile of present-day intensity, but still finds an increase in the number of extreme events under climate change when MSLP is considered, but a decrease when $\nabla^2 p$ is used. Other options for measuring intensity include using a different percentile as the limit (e.g. 98th, 95th), or a different variable (e.g. MSLP, wind gust speed). The differences due to various metrics could be avoided by using a metric consistently.

2.5.6 Dynamical Influences

IPCC AR5 (Christensen et al., 2014) discussed that there are dynamical reasons for the discord between future projections of midlatitude cyclones. At the large scale, there are two competing effects in a warming climate: the meridional temperature gradient decreases at the surface, but increases in the upper troposphere under climate change (Beersma et al., 1997). Both affect the hemispheric energy balance and the baroclinicity of the atmosphere. Changes to the baroclinicity under climate change over the North Atlantic is an important potential reason for the poleward shift of the storm tracks. A smaller temperature gradient between the poles and the tropics would cause less low-level baroclinicity and reduce the need for energy to be transported to the poles could affect the latitude of the storm track (Pinto et al., 2009). Catto et al. (2011) identified the reasons for changes to the storm track as:

1. Enhanced warming at low-levels over the Arctic, reducing the meridional surface temperature gradient and the baroclinicity;
2. Enhanced warming at high levels over the tropics, which moves polewards the region of strongest baroclinicity, the jet stream and therefore the storm track;
3. Changes to the Atlantic Ocean's Meridional Overturning Circulation (AMOC) — the pertinent part of the deep-water oceanic circulation — leading to changes in the SSTs and the low-level air temperatures, altering the location of regions of maximum baroclinicity.

Further to these large-scale influences, there are local effects. It is expected that North Atlantic SSTs will change under climate change, partly due to changes in AMOC; however, SSTs are another source of model bias (Keeley et al., 2013). Loss of Arctic sea ice will also occur, but the

effects are nonlinear due to feedback effects (Peings and Magnusdottir, 2014). Finally, changes to the land-sea temperature contrast are predicted. These all affect the local horizontal temperature gradients, and therefore change baroclinicity and so alter an important source of energy for midlatitude cyclones (Christensen et al., 2014, Section 14.6.2.1).

The influence of moisture under a changing climate is also discussed in Section 14.6.2.1 of Christensen et al. (2014), but the effect on midlatitude cyclones is indeterminate. The increase in global temperatures will increase the potential that the air has to hold moisture, according to the Clausius-Clapeyron Equation (Finn, 1993, Equation 9.10). This means increased potential for latent heat release, leading to stronger midlatitude cyclones (Booth et al., 2012; Dacre and Gray, 2009), due to greater potential for latent heat release by condensation and freezing. However, it also means more efficient energy transport from the equator to the poles, decreasing baroclinicity and reducing the potential for midlatitude cyclones (Schneider et al., 2010).

In summary, part of the reason for uncertainty in climate projections of midlatitude cyclones is the variety of uncertain dynamical forcings. While model errors, internal variability, different model resolutions, different tracking methods and different ways of quantifying cyclones all contribute to the uncertainty, identifying which dynamical factors are not well simulated could signpost future work.

2.6 Summary

In Europe, winter wind storms are the most damaging weather phenomenon. How such cyclones will react to climate change is uncertain; previous literature agrees that the zonal-mean storm track will generally shift polewards, but this may be the effect of low model resolution. Recent research has identified a tri-polar pattern in the change in the storm track over Europe (Zappa et al., 2013). The number of storms is generally expected to decrease, but locally numbers could increase (McDonald, 2011). Discord arises when cyclone intensity is considered, due to the use of different models using different atmospheric physics, different initialisation and different resolutions. Furthermore, the lack of an agreed measure of cyclone intensity means different outcomes are reached using various methods (Ulbrich et al., 2009).

Even when looking at trends in surface observations or reanalysis data, complications can arise. These are mainly due to internal variability on all timescales, but particularly decadal and multi-decadal, as the time period covered by both is insufficient to draw any conclusions about patterns in variation on such scales. Surface observations are susceptible to missing data and to changes in the observational network over time, such as the development of satellite data. Reanalyses are also prone to such changes, but they are also prone to model bias.

There are a variety of sources for this uncertainty, which are discussed in Section 2.5. Separating the different sources of uncertainty is key, which could be done using short model runs. The most readily available set of short model runs are those produced as weather forecasts, by dozens of NWP centres across the globe. Identifying the limitations of modelling in NWP forecasts will inform modelling on longer timescales. Furthermore, an analysis of NWP modelling will feed into the larger project within which this work fits, which will use short runs from climate models to identify sources of uncertainty. The present thesis will concentrate on the representation of historic severe European windstorms in NWP forecasts, and search for indications of where these models fall short. Improving short-term forecasts could facilitate mitigating action that could prevent human or economic losses.

Chapter 3

Data and Methods

This Chapter discusses the data and methods used to obtain the results. Following the objectives in Section 1.3, the first step is to identify the necessary data sets (Section 3.1). Analysis data are needed for the work identifying and classifying the storms, to use as a yardstick for the forecast analysis, and for the work with storm-prone situations, the large-scale atmospheric situation in the lead up to storm development. Forecast data are needed for the forecast analysis.

The next step is to identify the criteria for selecting a set of severe, historical, winter storms (Section 3.2). The storms then need to be tracked in analysis data (Section 3.3), in order to allow assessment of the entire lifetime of the cyclone. Two methods are investigated in terms of categorising the storms, based on the jet stream (Section 3.4) and the Pressure Tendency Equation (Section 3.4.2). As there are two methods of categorisation, it is important to assess whether there is any relationship between the different methods (Section 3.5.1).

After analysing the storms in analysis data only, the next steps involve the inclusion of forecast data (Section 3.6). The storms will be tracked in forecast data, and these forecast tracks matched automatically to the tracks in analysis data. Deviations between the forecast and analysis tracks will be calculated in terms of both storm location and intensity, and it is these that will be quantitatively evaluated, in order to determine forecast quality and predictability. This evaluation will be done for all of the storms together, and in the different categories of storm already discussed, to allow assessment of whether storms of one category are more predictable than storms in another category.

The final steps examine storm-prone situations (Section 3.7). A storm-prone situation (SPS) describes the large-scale atmospheric situation in the lead up to a storm's development. The aim is to ascertain whether there are features or properties common to the development of severe cyclones. The first step is to identify one or more candidate(s) to act as an SPS, and then go on to assess whether it is (they are) a useful method of identifying a situation that is likely to cause a severe midlatitude cyclone to develop. Finally, links will be sought with the categorisation methods and the forecast quality and predictability.

3.1 Data

The data need to be at a minimum of 6-hourly temporal and T63 spatial resolution, for the tracker to successfully realise a storm track (Section 3.3). Additionally, mean sea level pressure and vorticity at 850hPa are both needed by the tracker. Covering the period 1990 to 2010 would be highly desirable, as storms would be selected from that period, for reasons are discussed in Section 3.2. Therefore, the data used are mainly from the European Centre for Medium-Range Weather Forecasting (ECMWF), because the data are readily available at the required temporal resolution and for the required time period, at sufficient resolution and of high quality, in both reanalysis and forecast data sets.

Where gridded data are needed, data sets are acquired on 0.5^o resolution in both latitude and longitude, and so the native output is interpolated at the ECMWF before acquisition. For the storm tracking (Section 3.3), all of the data are acquired at native resolution, then interpolated and regridded using Climate Data Operators (CDO). Now, the reanalysis and forecast data sets used will be discussed, followed by the reasons for these choices.

3.1.1 Reanalyses

Reanalysis data are used extensively in this work. Alternatives include operational analysis or raw observational data. Reanalysis has limitations, including that it is affected by the quality and quantity of observations that are fed in, the data assimilation scheme used to process them, and intrinsic features of the model such as its resolution. However, analysis data have the same problems with observations, with the added complication that the model and data assimilation scheme are regularly upgraded. The tracks of different storms from different phases of model

development would be difficult to compare in analysis data, as differences could be due to an upgrade to the model, rather than being due to dynamical differences in the storms themselves. Raw observations are sparse, particularly over the Atlantic where the storms' tracks begin, so are not suitable either because the resolution would be insufficient. Although the observations could be interpolated, the distances between each observation are sufficiently large that this would not be accurate and would miss details. Therefore, neither analysis nor raw observations are suitable, and using reanalysis data is the best option.

The first data that are used are from the most up-to-date ECMWF Reanalysis project: ERA-Interim (Dee et al., 2011). There is an older reanalysis project, called ERA-40 (Uppala et al., 2005), but it was only run until mid-2002 and so does not cover the entire time period of interest. The native resolution also differs, with ERA-40 running at T159 (1.125^o) and ERA-Interim at T255 (0.75^o). Where possible, the storms were also tracked in ERA-40 data, in order to ascertain whether there are any systematic differences, but it quickly becomes clear that it would not be suitable in this context (not shown). Therefore, given its superior spatial resolution and coverage of all selected storms, ERA-Interim was used for the analysis discussed in Chapters 4 and 6, and as the yardstick for comparison in Chapter 5.

3.1.2 Forecasts

Both ECMWF deterministic and ensemble operational forecasts are used, and how their resolutions evolve with time is shown in Table 3.1. The Ensemble Prediction System (EPS) consists of two different types of forecast: a control run, initialised with operational analysis data but at a lower resolution than the deterministic operational forecast; and ensemble members, which are initialised with perturbed versions of the operational analysis data. These perturbations are not random, but instead are targeted to those where perturbations grow fastest (Palmer et al., 1997).

In addition to the changes described below, there was an important change to the ensemble data in October 1999, when the number of members was increased from 30 to 50. As Buizza and Palmer (1998) found that there is a strong dependence between how well an EPS performs and the number of members it contains, this work only examines the storms after October 1999 in ensemble member data.

Date	Deterministic		Ensemble	
	Horizontal	Vertical	Horizontal	Vertical
May 1985	$T106$ *95km*		N/A	
Sept 1991	$T213$ *47km*	31		
Dec 1992			$T63$ *150km*	
Dec 1996			T_L159 *72km*	31
Apr 1998	T_L319 *31km*			
Oct 1999		60		40
Nov 2000	T_L511 *19km*		T_L255 *42km*	
Feb 2006	T_L799 *13km*	91	T_L399 *25km*	62
Jan 2010	T_L1279 *8km*		T_L639 *16km*	

TABLE 3.1: Resolution of the ECMWF forecast models. The values in italics are the approximate conversion from spectral resolution to kilometres, at $60°N$.

3.1.3 Discussion

Although ECMWF data are selected for investigation due to its availability for the entire period under investigation, two alternative sources of data were considered: the National Centres for Environmental Prediction (NCEP) and the UK's Met Office (UKMO). NCEP's Climate Forecast System Reanalysis (CFSR) is an reanalysis, so could be compared with ERA-Interim; however, this is not possible in the current work. Although mean sea level pressure is readily available is CFSR, vorticity or winds aloft are not, so the tracking algorithm would need to be changed significantly in order to cope with the lack of vorticity data. NCEP also have operational models in the Global Forecasting System (GFS) and its ensemble companion (GEFS). However, both analysis and forecast data are only archived for approximately the last ten years, so are unsuitable

as they cover less than half of the required time period. Therefore, no NCEP data are used in this analysis.

UKMO data also have limited temporal coverage, as they are only archived back to 1999, so covers more than half of the selected storms and could be suitable for some limited analysis. However, although the forecast output is at sufficient temporal resolution, the analysis is only available at 12 hourly resolution for the older storms and so is not suitable for tracking. Though UKMO forecasts could be compared to ERA-Interim reanalysis, it would prove difficult to unpick the reasons why they differ. The difference could be due to the effects of the different spatial or temporal resolutions, or the different parametrisation schemes in the models. Furthermore, vorticity or winds at 850hPa are not available for the entire period since 1999, but only for the most recent portion. For these two reasons, no UKMO data are used in the present study.

Once the ECMWF data were acquired, a set of historic, intense, European windstorms needed to be selected.

3.2 Selecting Storms

In order to select only the most intense storms, it is necessary to select a measure for intensity. A meteorological index is chosen called the Storm Severity Index (SSI). The format used here is that introduced by Leckebusch et al. (2008b), who built on the work of Klawa and Ulbrich (2003). The reason for choosing a meteorological measure for storm intensity, compared to a socio-economic measure, such as insured damage caused by the storm, is that it has fewer complex factors at play. The amount of economic damage inflicted depends upon the population density it passes over, the preparedness of the population, how well insured that population is, whether there have been any other windstorms passing through recently that have primed the area for further damage, and what industries are working in the area (for example, forestry can be particularly adversely affected by windstorms).

Alternative methods of measuring storm intensity have been used in the past, including meteorological quantities, such as minimum core pressure or maximum vorticity. However, these are not linked to the damage a storm inflicts, which is important to this work not only because it is related to the societal impacts that these storms have, but also because storm damage is a particular interest of the funder of this work the AXA Research Fund, affiliated with the insurance company.

A meteorological proxy for damage is a straightforward way to evaluate storm severity and is an approach that previous studies have taken to quantify storm intensity (e.g. Haas and Pinto, 2012; Hanley and Caballero, 2012).

SSI is calculated using readily-available meteorological variables, that only produce one value for each time step; a simple measure that does not depend on whether the storm passes over high or low population densities. The relationship to the wind speed is cubic, so is proportional to the power of the wind, which is related to the potential work that the wind can do in inflicting damage. It is summed over an area (Figure 3.1), and so it is related to the damage a storm could inflict over that area. SSI then considers by how much the wind at a grid point at a given time exceeds a threshold (i.e. the 98th percentile of wind speeds at that grid point for winter), and so calculates how exceptional the wind speed is at a point. Therefore, as a measure, SSI is a hybrid of measuring severity and extremity (Beniston et al., 2007), because it considers both the damage wreaked by the storm and the high value of the wind speeds. SSI is then summed up over an area (Figure 3.1), so is also affected by area.

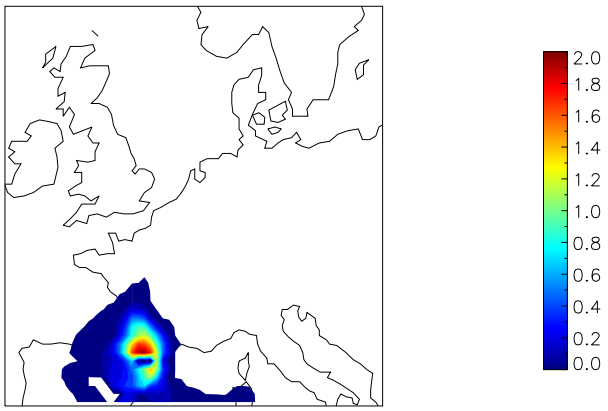


FIGURE 3.1: Map illustrating the area over which SSI was summed [40°N to 60°N, 10°W to 20°E]. Overplotted are the SSI values for 1200h on 2009-01-24, the day with the highest SSI in the time period, which was related to storm Klaus.

Alternative metrics for storm severity are available, for example that used by Lamb (1991, Chapter 2), which considers the greatest surface wind speed (cubed), the area affected by the storm, and the duration of the damaging winds. Firstly, the main strength of the SSI over Lamb's index is that SSI implicitly considers how prepared the population is for damaging winds, by comparing the wind speed to climatology. In regions that regularly experience strong winds, it is likely that the population will have adapted to stronger winds and mitigating action will have been taken against

damage; for example, European building codes state that before many buildings are constructed, the wind loading must be modelled, so that it can be allowed for in the design. However, winds that are exceptional to a region could still cause damage. Secondly, the Lamb index considers storm duration, whereas the SSI does not, but it is valid to disregard storm duration. Recent research explores the idea that the duration of a storm makes a lesser contribution to the overall damage that a storm inflicts; instead, it is speculated that the short-lived but very high wind speeds caused by smaller-scale features such as a sting jet (Browning, 2004) inflict the greater part of the damage. Furthermore, it is difficult to objectively and accurately determine storm duration. Finally, (Klawa and Ulbrich, 2003) show that SSI is related to the monetary damage a storm inflicts. For these three reasons, SSI is used in the current study.

The formula for SSI is:

$$SSI = \sum_{i=1}^{n} \left(\frac{v_i}{v_i^{98}} - 1 \right)^3 \cdot \begin{cases} 0 & \text{if } Z \leqslant 1000m \\ 1 & \text{otherwise} \end{cases} \tag{3.1}$$

where i refers to the n grid points that are summed over, v_i is the total wind speed at each grid point, v_i^{98} is the 98th percentile of the wind speed, and Z is the height of the orography.

The SSI values are calculated using ERA-Interim data, between October and March, in the years 1979 to 2011. The wind variables used are the u and v components of wind speeds at 10m (variable numbers 165 and 166), which are then added together. These wind speeds are not those where additional post-processing has been applied, to make the model more comparable to SYNOP observations by making allowances for roughness length. The wind climatology is constructed for each grid point, in order to find the 98th percentile of winter wind speed for the entire period. Then the wind climatology could be compared to the wind speed at each grid point, for every time step, meaning the SSI is calculated at each grid point before being summed over the specified area (Figure 3.1). This was only done where the altitude was less than 1000 metres, mainly because winds at ten metres are used and the values can become unrealistic in mountainous areas, so could produce very high values of SSI and skew the selected cyclones towards those that passed over mountains. A secondary reason is because these areas have low populations, meaning any incurred damage is small. Only time steps after 1st January 1990 were considered, because before then, the forecasts' resolutions were so low that storms would be difficult to track during the next stages of forecast analysis (described in Section 3.6). The dates and times are then listed by SSI

value, and the highest values taken forward to the next stage of analysis. The midlatitude cyclones in the resulting list are the selected storms.

3.3 Tracking

Once the storms are selected, the next step is to track them. A discussion of different tracking methods can be found in Section 2.5.4. The storm tracking used in this work is based on the group of methods, whereby a maximum or minimum in a meteorological parameter is identified as a cyclone centre, and connections between those at subsequent timesteps are made based on forming a coherent track (Murray and Simmonds, 1991; Pinto et al., 2005, e.g.). This is the most appropriate method to use, given that the aim of this work is to track particular storms, and this method yields the track of an individual storm most readily.

The tracking method, developed by Tomasz Trzeciak for the project in which this work is contained, uses MATLAB version 2012b. It was necessary to develop a tracking method for this project, because of the unique approach it takes; too many storms are included for tracking to occur by hand, but the selected storms must be easily identifiable in the tracker's output. The algorithm identifies minima in mean sea-level pressure and then searches for an associated maxima in vorticity at 850hPa. Therefore, only cyclones with both of these features are located, which tend to be intense systems. The vorticity is first smoothed to T63 resolution, following the work of Hodges (1994), because the vorticity is a noisy field as it is a second derivative. Hodges smoothed to T42 resolution, whereas T63 was selected for this work because it retains more of the extremes than T42. Therefore, the vorticity field still bears a strong resemblance to the pressure field, which facilitates the matching of pressure minima and vorticity maxima. Once smoothed, the vorticity maximum must be over a limiting value, and only one vorticity maximum can be identified within a set radius of the pressure minimum; this returns the largest value of vorticity in that area.

In terms of joining the points at consecutive timesteps, the first limit applied here is that the two points at consecutive timesteps must be less than a certain distance apart. The distance is measured as the shortest distance between the two points on a spherical surface, in degrees. Once two points closer than the threshold distance have been identified, then the contribution that each would have to the track's overall smoothness is considered. This is done because broadly, cyclones' tracks are a smooth curve, and only involve large changes to direction when they are dissipating, splitting, merging, or meet an obstacle (e.g. orography, high pressure). Another reason tracks can contain

sudden changes is if a secondary cyclone develops from the main centre, the tracker can pick it up and join it to the track of the primary cyclone. None of these situations are desirable, and so the smoothing criterion is included. When the storm is moving quickly, the smoothing criterion is calculated as the distance from the midpoint between the first and third points to the second point considered (Figure 3.2), illustrated by the red arrow. This distance is smaller for smoother tracks, and so it penalises sudden changes in direction or speed. The distance is divided the distance between P_1 and P_3 (Figure 3.2), and the fraction has a limit placed on it. When the storm is moving slowly, the absolute distance between the points is used, so large deviations in the later stages in the track are less likely. By allowing each point on the track to have a number of possible next points, a situation is avoided where the point that scores best for smoothness subsequently has a large diversion on its track. However, this could easily become computationally cumbersome, so there is a limit on the number of future points kept for each point, and once the tracking is complete only the smoothest overall permutation is kept.

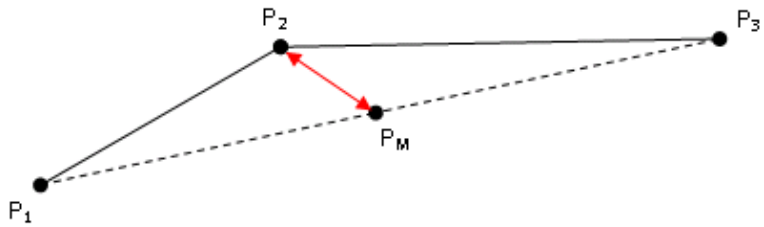


FIGURE 3.2: Diagram illustrating how the smoothing criterion is assessed. P_1, P_2 and P_3 are the first, second and third points on the track under consideration, and P_m is the midpoint between P_1 and P_3. The red arrow indicates the distance calculated as part of the smoothing calculation.

Next, the correlation is calculated on the pressure field at the two consecutive timesteps, for a specified area centred around the pressure minima, so that the same storm is more likely to be followed, as any sudden change in the magnitude or shape of the pressure field would be discouraged. In particular, this criteria is intended to suppress joining a secondary cyclone to the primary storm, as the pressure field of the secondary storm is unlikely to yield a high correlation with that of the primary cyclone. Connections between consecutive points with high correlations in the pressure field score higher.

The tracks are given a total weighted score, calculated using:

$$trackscore = W_s \sum_{i=1}^{n} S_s + W_c \sum_{i=1}^{n} S_c - W_l \times S_l \tag{3.2}$$

where W refers to weighting, S to the score, n to the total number of points along the track length, and in the subscripts, s refers to smoothness, c to correlation and l to track length. It shows that the total score for the track is related to the smoothness and correlation between each connection, summed over each point on the track, with an additional component to encourage tracks with large total length. Long tracks are encouraged because the storms of interest are intense, so tend to have long, coherent tracks, and so short tracks are unlikely to yield useful results. To this end and in line with the work of Pinto et al. (2005) and Hoskins and Hodges (2002), a minimum on the cyclone lifetime of 24 hours is set. The relative weighting of the three components in the scoring can be altered, but during this work they are kept constant.

Finally, the tracks are output. Output is either done in the order of the total score for the track, or how well they compare with a reference track. The ranking by score is particularly useful in the work with NWP data, where it is necessary to match the track in forecast data to that in analysis data, so the analysis track can be given as a reference. Track matching will be discussed more in Section 3.6.

In summary, the thresholds that are applied to the track are:

- lifetime ($24 hours$),
- vorticity value ($1 \times 10^{-5} s^{-1}$),
- vorticity radius of influence (3.75^o),
- degree of direction change between successive points (0.75^o),
- maximum distance between successive points (4.5^o latitude, 15^o longitude),
- length scale for near-stationary cyclones (2.25^o),
- number of future points kept for each point on track (2),
- area for correlation between pressure fields (3.75^o latitude, 7.5^o longitude),
- relative weighting of track length (0.3), track shape (0.5) and pressure field correlation (0.5) used to calculate track scores.

3.4 Determining and Assessing Storm Types

Once the storm events are selected and tracked, there are many possible ways to categorise the storms, in order to see if any trends emerge when forecast quality is assessed. Here, the jet stream and equivalent potential temperature (θ_e) are first examined for any possible patterns. Then a discussion of the Pressure Tendency Equation will ensue in Section 3.4.2.

3.4.1 Jet Stream and Equivalent Potential Temperature

The jet stream is key to a storm's development, both in terms of intensity and path, as discussed in Section 2.2.4.2. Where the jet accelerates, there is ageostrophic motion leading to convergence and divergence aloft (Figure 2.7). The jet is also associated with vorticity advection, which can affect the rotation of a cyclone (Section 2.2.4.2). θ_e is a measure of the temperature and moisture content of an air mass. As such, it quantifies how much latent heat there is in the air to be released, and so qualitatively represents the extent to which diabatic processes (Section 2.2.4.4) could potentially contribute to the deepening of the cyclone. Furthermore, θ_e can be used to describe different air masses, and gradients in it indicate fronts. Stronger gradients indicate a stronger contrast between the two air masses, and so more potential energy that can be released (Section 2.2.1). Therefore, θ_e qualitatively describes two sources of energy for the cyclone.

In short, both the jet and θ_e describe processes that can contribute to a cyclone's development. Therefore, these parameters are examined using a method of plotting meridional sections used in Fink et al. (2009); an example is shown in Figure 3.3. First, the storms' tracks are needed. Then meridional sections are made for each point on the storm's track; the field of interest is split at the midway point between the previous point on the track and the current point, and between the current point and the next. When continued for the length of the track, the result is a plot of the field in the meridional section where the storm is travelling. As such, it produces a concise technique for examining plots of the jet stream and θ_e fields over the duration of the storm.

These plots are made using IDL version 8.2 and the results are presented in Section 4.4.

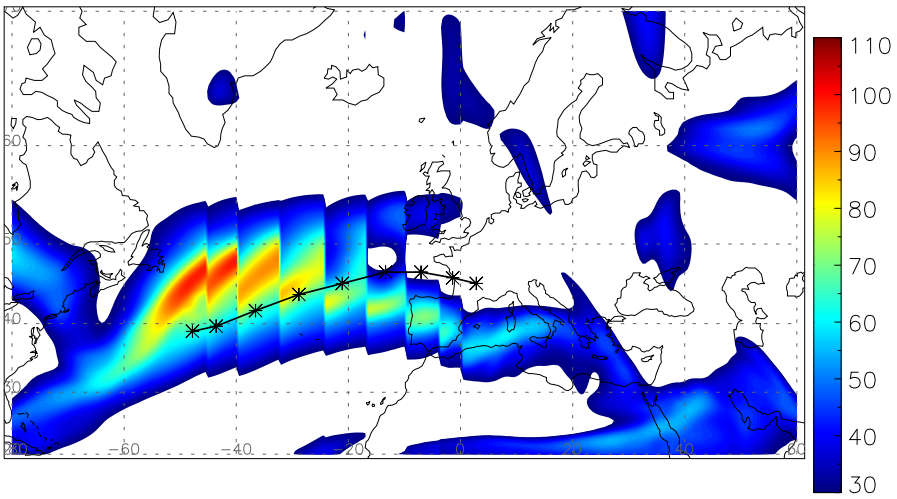


FIGURE 3.3: An example of sectional plotting for storm Klaus for total wind speed ($\sqrt{u^2 + v^2}$ in ms^{-1}) at 300hPa.

3.4.2 Pressure Tendency Equation

An alternative approach to categorising storms is desirable, to provide a comparison. Recently, Fink et al. (2012) suggested just such a method, based on the Pressure Tendency Equation (PTE), intended to provide insight into the storms' dynamics. The PTE examines which processes are contributing to a change in the core pressure of a cyclone: changes to the geopotential aloft; thermodynamic processes of horizontal temperature advection, vertical motion, and diabatic processes; and evaporation or precipitation. Fink et al. (2012) applied the PTE to five of the case studies discussed here, so this work will apply it to the remaining storms.

Other approaches have been taken to diagnose the key processes in a storm's deepening. Firstly, tendencies of other parameters have been used. Vorticity tendency has been applied to extra-tropical cyclones (Azad and Sorteberg, 2009), but contains similar terms to the pressure tendency equation so applying the vorticity tendency method in addition to PTE analysis would, in all probability, give the same results. Potential vorticity tendency could prove useful, but as the best arrangement of the equations varies with situation, which is best for extra-tropical cyclones remains an open question (Tory et al., 2012). Secondly, Black and Pezza (2013) considered which

transfers of energy in the Lorenz Energy Cycle (Figure 2.6) are dominant. While energetic analysis would provide useful insight into the deepening phase of the selected cyclones, it would not provide information regarding the contribution of diabatic processes to deepening. Since this is desirable in the current work, the PTE will be used.

The PTE considers the processes relevant to deepening in a 3^o by 3^o column that moves along the storm's track and extends from the surface to 100hPa. By assessing how the height of the 100hPa (P2) surface, the temperature advection at different heights into the column, and the moisture profile of the column evolve between each point in time on the storm's track, the causes for the cyclone's deepening or filling can be identified. The PTE is:

$$\begin{aligned}
\frac{\partial p_{SFC}}{\partial t} &= \int_{sfc}^{p2} \nabla(\rho\vec{v}) && (3.3)\\
&= \rho_{SFC}\frac{\partial \phi_{P2}}{\partial t} && (d\phi)\\
&+ \rho_{SFC} R_d \int_{sfc}^{P2} \left[-\vec{v}\cdot\nabla_p T_v + \left(\frac{R_d T_v}{c_p p} - \frac{\partial T_v}{\partial p} \right) \cdot w + \frac{T_v \dot{Q}}{c_p T} \right] dlnp && (dTemp)\\
&+ g(E-P) && (EP)\\
&+ residual
\end{aligned}$$

where p is mean sea level pressure, t is time, ρ is density, $\vec{v}$ is the wind speed vector (u, v, w), R_d is the gas constant for dry air, T_v is the virtual temperature (that is, the temperature a moist air parcel would have, if it were dried out), c_p is the specific heat of dry air, E is the mass of evaporated water, and P is the mass of precipitation. The derivation begins with the hydrostatic and continuity equations, and then integrates between the surface and an upper boundary. The local density tendency is then written as a sum between the total density tendency and an advection term (horizontal and vertical). Using the ideal gas law and first law of thermodynamics, substitutions are made that result in an equation that describes how processes contribute to the evolution of total mass in the column, and so describe the change in surface core pressure with time.

While core pressure is an important aspect of a midlatitude cyclone to consider, it does have drawbacks. For example, it does not consider the size of the storm, so it is difficult to analyse whether, for example, smaller storms have a stronger diabatic contribution than larger storms. Furthermore, when considering winds, the pressure field is only related to the geostrophic wind and

so does not reflect the ageostrophic wind or smaller-scale variations such as sting jets. Therefore, this analysis will not consider these aspects of storms, which could be examined by future work. Consideration of the storms' core pressure is sufficiently important to be worth studying in its own right. The processes that contribute towards storm deepening will shortly be discussed.

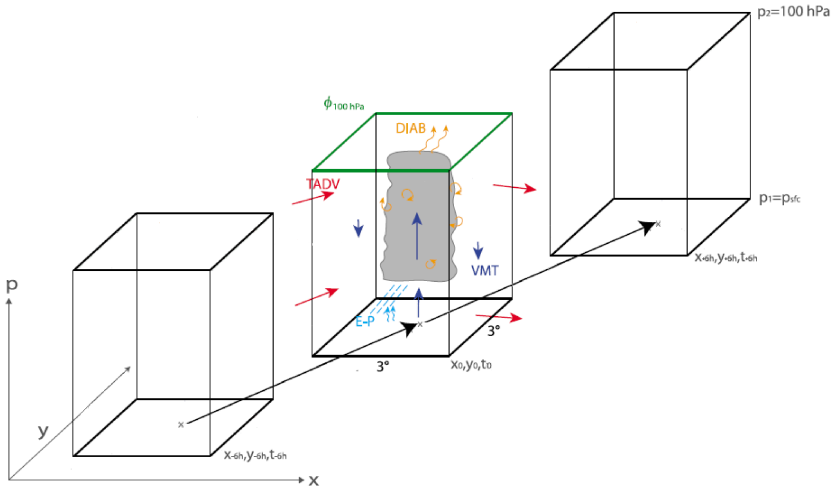


FIGURE 3.4: Schematic of how the PTE works, based on Fink et al. (2012), Figure 1. See text in Section 3.4.2 for explanation.

The first term '$d\phi$' evaluates the effect that a change in geopotential at the upper limit of the integration would have on the surface pressure. As geopotential refers to the work done against gravity to raise a unit mass from the surface to a certain height, this term essentially describes the changes in the height of the the upper limit of the integration (100hPa). This term describes any effect that the atmosphere above the limit of integration has on deepening. Ideally, the limit would act as a level of insignificant dynamics; that is, the processes occurring above this level have sufficiently low influence on cyclone development that modelling them is not worthwhile. However, the limit could well be above 100hPa (Wilbraham, 2013), but here 100hPa is used as the lid of the integration for consistency with the work of Fink et al. (2012).

The second term 'dTemp' refers to the virtual temperature tendency, and is the integral of three summed terms. The horizontal advection term ('TADV' or 'horiz') describes how the large-scale wind moves air of a different virtual temperature (T_v) into the box of interest, which could change the mass of air in the column and so change the core pressure. It is related to the baroclinicity of the atmosphere, as air of different temperatures must be advected in order to deepen the cyclone, which implies a horizontal temperature gradient and therefore baroclinicity. The vertical term

('VMT' or 'vert') describes how vertical motions can affect the vertical temperature distribution; vertical motion that changes the density distribution, so that it becomes different to the density distribution outside the column, will cause air to move in or out through the column's sides, and so alter the total column mass and the surface pressure. The diabatic term ('diab') is the term which describes the diabatic processes including latent heat release, radiative warming or cooling, diffusion and dissipation. Though some models output the parameters needed to calculate the diabatic contribution directly, in this work it is calculated as a residual from the remaining parameters, and so is termed 'diabres'.

'EP' describes how the surface pressure will change due to changes in column mass because of evaporation and precipitation. If there was 10mm of rainfall then the surface pressure would lower by 0.98hPa, but if 10mm of water evaporated, then the surface pressure would increase by 0.98hPa. This is because condensation of water vapour into rain removes mass from the air in the column and so, under hydrostatic adjustment, the pressure decreases (Spengler et al., 2011). Given that storms are often associated with precipitation, then it is anticipated that this term will mainly encourage storm deepening.

Finally, there is a residual term due to the temporal and spatial discretisations made while undertaking the calculation, and approximations made in the derivation of the equation, partly due to the finite-differencing approach needed to quantify derivatives on a grid.

It is possible to categorise the storms using the PTE method, by which processes dominate the deepening of the cyclone. Following the work of Fink et al. (2012), a qualitative approach for analysis will use bar plots of the different terms, over the course of each storm's track, made using MATLAB version 2012b. However, a more quantitative approach could be used, based on the direct output in hPa/h for each process. This is done by determining the contribution of each process as a percentage of the total deepening (in hPa/h), so storms that deepen more quickly are comparable with those that deepen more slowly. The dominant process in deepening is identified as having the largest percentage contribution. The results of the PTE categorisation are presented in Section 4.5.

3.5 Comparing Classification Methods

Once the storms have been divided into categories, it is important to assess whether there is a link between the two different methods of classification. Are storms of one jet stream type more likely to be driven by a particular process, identified from PTE analysis? After an initial, qualitative assessment of the different categories, statistical tests will be performed to assess the relationship between the two methods. Finally, a quantitative approach that compares the observed frequencies of each class of storm with the expected value, given the row and column totals, will facilitate discussion of the potential mechanisms that drive any links that become apparent. These statistical methods will now be described.

3.5.1 Relevant Statistical Tests

There are two statistical tests that are predominantly used for comparing categories of data: the χ^2 and Fisher exact tests. These are used to test whether different categories of data are related to each other. For tests like these, the data should first be presented as a contingency table. In this work, the table will show the number of storms in each category, separately by both methods, with PTE type as the columns and jet stream type as the rows. Though both the χ^2 and Fisher exact tests are useful, only the Fisher exact test will be performed here. The χ^2 test assumes the data are normally distributed. This cannot be definitely true of a small sample size, and therefore the χ^2 test is not recommended for contingency tables where some cells have fewer than five members. In situations with small data sets like this set of storms, the Fisher exact test is the best measure.

The use of the Fisher test in this work will provide information about whether the two different categorisation methods are associated; that is, whether the classification of storms by jet stream type is related to the classification by which processes dominate according to the PTE analysis. The formula in a 2x2 contingency table is relatively easy to understand, because there is only one degree of freedom (Fisher, 1973, Section 21.02). It first assesses the number of ways of arranging the frequencies in the four categories, and how likely each is to occur, for a fixed row and column total. An example is shown in Table 3.2.

It becomes apparent that some of these contingency tables are more "extreme" than others (i.e. show a stronger correlation between the two categorisation methods), and therefore are less likely when the population of the cells is determined randomly. The relevant probabilities are shown

(a)

	α	β	Σ
Φ	5	1	6
Ω	0	4	4
Σ	5	5	10

p=0.02

(b)

	α	β	Σ
Φ	4	2	6
Ω	1	3	4
Σ	5	5	10

p=0.26

(c)

	α	β	Σ
Φ	3	3	6
Ω	2	2	4
Σ	5	5	10

p=0.72

(d)

	α	β	Σ
Φ	2	4	6
Ω	3	1	4
Σ	5	5	10

p=0.26

(e)

	α	β	Σ
Φ	1	5	6
Ω	4	0	4
Σ	5	5	10

p=0.02

TABLE 3.2: Tables showing examples of contingency tables for fixed row and column totals, and the probabilities of obtaining them or a more extreme version.

underneath the permutations in Table 3.2. Given all of the possible contingency tables, the Fisher exact test calculates the probability of getting the observed contingency table or a more extreme version of it, given all of the possibilities. This means that the probability of (b) occurring contains the probability of (a) occurring, because (a) has a similar pattern to (b) but is more extreme. It can be seen that, when the p-values are lowest, the two categories are most clearly related (a and e). Therefore, a low p-value indicates that the categories are likely to be related, and a maximum cut-off of 0.05 is commonly applied.

However, the general case that will be used in this work, as the contingency table is larger than 2x2. This is significantly more complex (Mehta and Patel, 1983), because the number of permutations of numbers in each cell is much larger than in the 2x2 case. There are two approaches to determining the probability of each permutation occurring: exact and Monte Carlo. The higher number of degrees of freedom in a large contingency table can justify the random simulations of contingency tables provided by the Monte Carlo method to approximate the probabilities. Fortunately, both the exact and Monte Carlo methods are pre-programmed into statistical packages such as R (R Core Team, 2013) version 3.0.2, which will be used for this portion of the work.

3.5.2 Calculation of Expected Values and Deviations

In addition to performing the Fisher exact test on the contingency table, one further statistical approach is taken to identify links between the different categories of storms. It is possible to calculate the expected value for each cell in a contingency table, given the row and column totals.

If a contingency table is structured as shown in Table 3.3, then each of the expected values are calculated:

$$
\begin{aligned}
E_{11} &= \frac{R_1 C_1}{T} \\
E_{12} &= \frac{R_1 C_2}{T} \\
E_{21} &= \frac{R_2 C_1}{T} \\
E_{22} &= \frac{R_2 C_2}{T}
\end{aligned}
\tag{3.4}
$$

where E refers to the expected value, R to the row, C to the column, and the subscripts to the row and/or column number. These formulae describe how many storms should be in each category, given fixed row and column totals. This can then be compared to the observed number of storms in each category, to allow assessment of whether some categories favour one type or another.

	Type 1	Type2	Σ
Category 1	E_{11}	E_{12}	R_1
Category	E_{21}	E_{22}	R_1
Σ	C_1	C_2	T

TABLE 3.3: A general contingency table. C values refer to the column totals; R to the row totals; E to the expected value in each cell; and T to the grand total.

As well as these quantitative approaches, there is additional value in qualitative assessment of the contingency table. While such subjective analysis does not test the statistical significance of any relationship, there is merit is assessing the raw numbers of storms in each category based on face value. Once both qualitative and quantitative assessments of the links between the two methods of categorisation have been made, the physical reasons behind such links will be discussed.

3.6 Measuring Forecast Quality

Once the storms have been selected and classified, an assessment of the forecasts is undertaken, in terms of accuracy and spread, for both cyclone intensity and position. A quantitative approach is essential here, with statistical tests to allow assessment of the significance of any trends. It is difficult to assess such storms with a basic forecast verification approach. For example, if a large area is considered using a continuous measure (e.g. mean squared error, Joliffe and Stephenson

(2003, Section 5.4.1)) on a large-scale field such as pressure, differences away from the storm of interest could influence the overall score, which would then not be representative of the forecast quality of the storm. For this reason, a feature-based approach is taken, where the storms are objectively tracked in forecast and reanalysis data and these tracks are compared. The approach will be continuous, because if categorical measures are considered, then it is necessary to introduce thresholds to determine what would constitute a hit or a miss, which could compromise the integrity of the results.

Therefore, the differences between the forecast and the reanalysis track will be evaluated. Differences will be calculated for operational deterministic forecasts, the ensemble control, and the ensemble members (Section 3.1.2), to give an idea of both forecast quality and spread. Before taking this approach, several factors must be considered: which variable to use for measuring forecast quality; how to connect the forecast and analysis storms' tracks; at which point comparisons should be made; and how significant trends should be identified. Each of these points will now be discussed.

3.6.1 Variables

There are many potential measures to use to quantify cyclones, as discussed in Section 2.5.5. Previous work has used pressure, vorticity, wind speeds, pressure gradients, latitude, longitude, distances in degrees and distances in kilometres. For comparing forecast quality in this work, only three measures will be used: one for intensity (MSLP), and two for position (latitude and longitude). One reason for this is to simplify matters; given that the tracker uses MSLP, the core pressure of the cyclone will be the measure used for intensity. The tracker also measures position using the latitude and longitude, so these will be used for position. Froude et al. (2007b) used the distances along and across the track, but this does not allow systematic trends in north-south or east-west direction to be identified. Since not all of the storms move classically from west to east, if there were a systematic northward shift of the forecasts, this could manifest in the across track distance for one storm but in the along track distance for another. Therefore, the variables of interest in this work are the differences between forecast and analysis in pressure, latitude and longitude.

3.6.2 Track Matching

The data used for calculating differences are from the track in the forecast data that best matches the reanalysis track, subject to some quality control to ensure the same storm is being considered. Therefore, it is necessary to match the tracks in the forecasts and reanalysis, using both temporal and spatial criteria. The tracking algorithm started this work, by using a reference track (the reanalysis track) when processing forecast data. First, the tracker takes the temporal approach of rejecting any forecast that did not occur at the same time as the reference track for at least three points. Then it goes on to quantify the distance between the points on the reanalysis and forecast tracks that matched temporally. It outputs the tracks in number order, starting with the closest to the reference track.

Therefore, a quality control is needed, meaning that if the distance between any point on the forecast track and the reanalysis track (at the same point in time) exceeded 15^o, for either latitude or longitude, then the track gets rejected. The value of 15^o is a compromise between the needs to remove the irrelevant tracks and have the maximum number of tracks for comparison. The limit was determined by examining the tracks that matched, for a variety of values, for all storms. Examining a sample of track-matching plots shown in Figure 3.5, it becomes clear that 15^o is a good compromise.

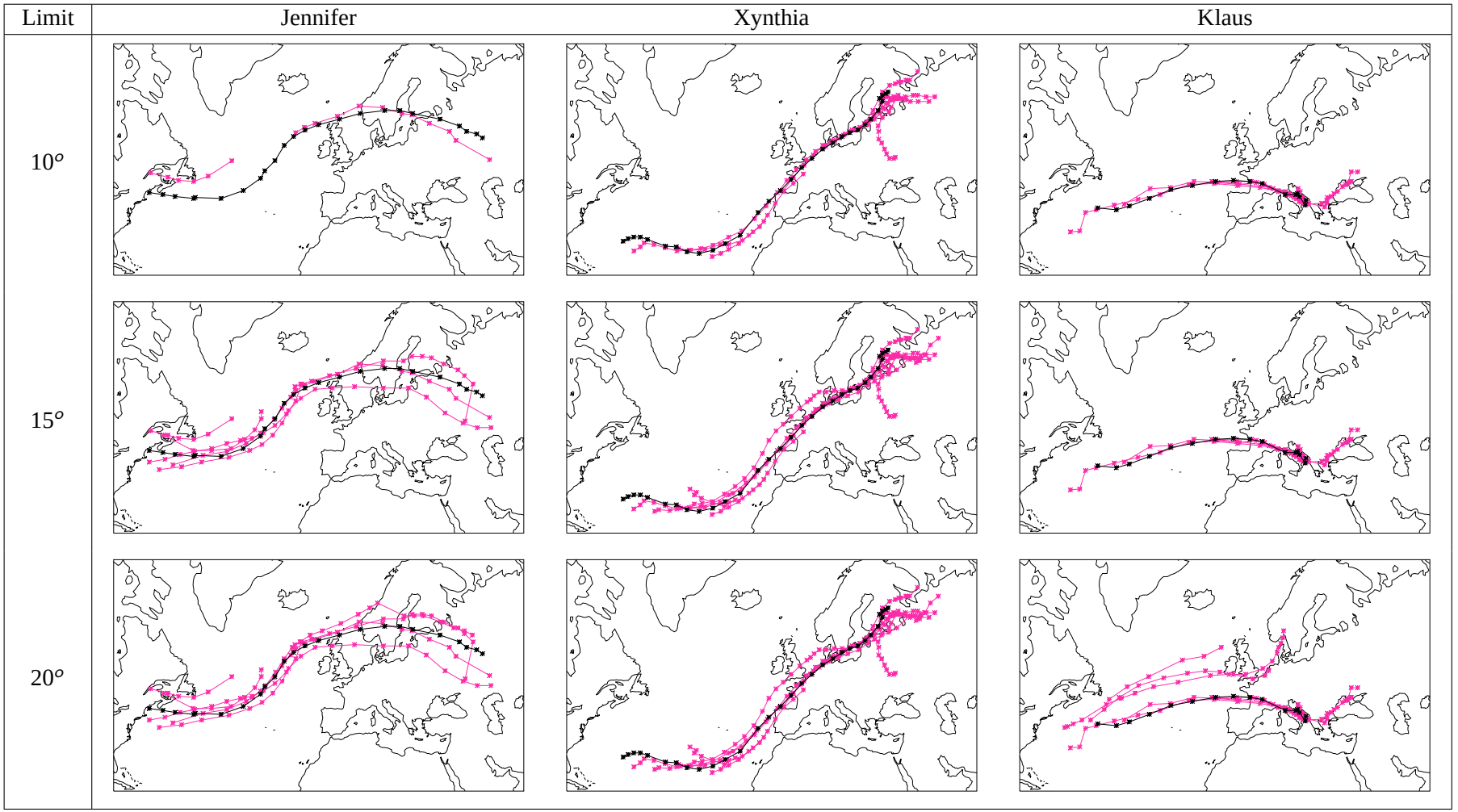


FIGURE 3.5: For storms Jennifer, Xynthia and Klaus, testing the sensitivity of different distance limits (in degrees) when matching forecast to analysis tracks. Analysis track is in black. Coloured tracks represent the best match for a variety of lead times.

3.6.3 Calculation of Differences

Once the tracks are matched, differences between the reanalysis and forecast tracks are calculated. The tracks in ERA-Interim reanalysis data are used for comparison (see discussion in Section 3.1.3). The point on the reanalysis track with the lowest pressure is selected, to be the time at which the comparison is made. The reason being that that comparing tracks along their entire length is unfair, because midlatitude cyclones are difficult to track at the beginning and end of their lifetimes. When comparing a variety of cyclone tracking methods, Neu et al. (2013) found that they tend to agree less well at the times of cyclogenesis and cyclolysis. Therefore, the comparison of forecasts will be most reliable at the time of maximum intensity. Pressure is used as the measure of intensity here, for consistency with the tracker. SSI is not used to determine which points should be compared because the portion of the track used could easily be during cyclolysis, given SSI's preference for land-based points, which is often where cyclones are filling. It could be possible to compare a number of points from around the time of maximum intensity, but this introduces a large number of complications. The lifetimes of the storms vary, so the time over which the deepening and filling differs. Selecting a single point facilitates a systematic comparison of forecast quality and spread across the variety of storms.

In summary, what results is a set of differences between forecast and reanalysis for pressure, latitude and longitude, for each storm, for only the point of minimum pressure but at a variety of lead times. The lead time used in calculations is the number of hours from the time at which the forecast is initialised, to the time of the minimum pressure. Next, it is necessary to identify any trends, and determine their statistical significance.

3.6.4 Statistical Tests

First, plots are produced of pressure, latitude and longitude against lead time, in order to determine whether there is a relationship with lead time. These will be line plots for the deterministic forecasts (made using Microsoft Excel 2007) and box-and-whisker plots for the ensemble forecasts (made using R version 3.0.2). A relationship would indicate that these variables are a useful way of measuring forecast quality, because if that is the case, then differences between forecast and analysis should decrease with time as the forecast becomes more accurate. Firstly, because the atmosphere is a chaotic system, a shorter lead time reduces the potential for random effects to cause deviations in the forecast. Secondly, the initial perturbation that grows into the storm

(Section 2.2.1) is likely to be further developed at shorter lead times, and so its development into a mature cyclone should be better represented.

Next the correlation between these variables will be quantified for the deterministic forecasts. This was done using Pearson's correlation coefficient (R), because it is the best test for continuous data. It considers the way two variables co-vary, divided by the product of the standard deviations of each variable. As such, it takes a value of zero for independent data, +1 for perfect positive correlation and -1 for perfect negative correlation (Rees, 2001, Section 14.2). A test statistic (t) tests for the significance of correlation (Rees, 2001, Equation 14.2), given by:

$$t = R\sqrt{\frac{n-2}{1-R^2}} \tag{3.5}$$

For a given number of points n used to ascertain whether there is correlation, there are $n-2$ degrees of freedom. The significance of the correlation is determined by comparing the t calculated from the data to the reference t distribution, for the number of degrees of freedom. A 'look-up table' is used, to find the values of the t-distribution for a given probability of exceedance (α, typically 0.05). Such a table can be found in many statistical textbooks, but the one used in the current work is Rees (2001, Table C.5). If t has a value greater than that in the look-up table, then the correlation is said to be significant at the α level.

An alternative test for correlation is the Spearman's rank correlation coefficient (R_S:.

$$R_S = 1 - \frac{6\sum d^2}{n(n^2-1)} \tag{3.6}$$

For the number of points in the correlation test (n), the two variables of interest are ranked ordinally from highest to lowest. Next, these rankings are compared and the difference between them (d) calculated. The reason for including Spearman's test alongside the Pearson's is that their use is slightly different. Pearson's test evaluates the linear correlation between two variables, i.e. the data points' proximity to a straight line. Spearman's test does the same for any monotonic function (each x has only one y) and so includes, for example, cubic functions. Given that it is unknown whether the data can be related linearly or monotonically, then both the Spearman's rank correlation coefficient and Pearson's correlation coefficient will be calculated where appropriate.

Statistical tests for a similar analysis but with ensemble forecasts cannot be identified. Therefore, the main tool used in ensemble analysis will be the box and whisker plots of the differences in

pressure, latitude and longitude against lead time. This analysis is performed on the entire set of storms, and then on the subsets of storms identified in Chapter 4 using the methods described in Section 3.4.

3.6.5 Simple Metric for Forecast Quality

Finally, the storms need a fair method of comparison. Summarising the quality of each storm's forecast into a single number will allow identification of correlations with severity (SSI) or connections with the categorisation of storms. The lead time must be fixed because forecast quality varies with lead time. However, this is complicated by the fact that the forecasts are initialised once a day at 12:00 but the minimum pressure for a storm can be at any of the six analysis times. Therefore, not all of the storms have the same values of lead time. For example, if the minimum pressure were at 12:00, then the forecast lead times are 24, 48, 72, etc. hours, but if the minimum pressure were at 00:00 then the lead times are 12, 36, 60, etc. hours.

Therefore, the deviations will be calculated for a particular lead time (24 hours). After performing linear regression for each storm on the differences between the forecast and the reanalysis, the formula for the line of best fit is then used to find the difference at 24 hours. For the linear regression to be reliable at a point, there should be forecasts successfully matched to the reanalysis pressure at lead times larger and smaller than the chosen time of interpolation. This is the case when 24 hours is selected. One important caveat is that each individual storm must have sufficient points for linear regression to be valid. Each point in the linear regression means the forecast track matches the analysis track at the time of minimum pressure, and therefore in a storm that is difficult to forecast, matching up tracks could happen too few times. If fewer than three points are involved in the linear regression, then the storm is rejected. This is done with the operational forecasts, so that the maximum number of storms can be included, and for each of the metrics of interest: longitude ($\Delta(lon)$), latitude ($\Delta(lat)$) and core pressure ($\Delta(p)$).

To summarise, Chapter 5 will present an assessment of the forecast quality and spread for the storms selected in Chapter 4. The correlation between forecast lead time and quality or spread will be examined. The final sections of Chapter 5 will discuss whether forecast quality or spread is affected by the different storm categories.

3.7 Storm-Prone Situations

Once the forecast quality has been assessed, the next stage is to go one step earlier and investigate the large-scale atmospheric situation before the storm developed — the 'storm-prone situation'. Once candidates have been identified, it can be determined whether storms with strong storm-prone situations are from a particular category (Section 3.4) or if they are better forecast (Section 3.6). As discussed in the literature review, there are three major sources of energy for such cyclones: baroclinicity, barotropicity, and diabatic processes. Here, the aim is to identify parameters that describe these three processes, and that correlate with the development of the selected severe cyclones.

3.7.1 Growth of Midlatitude Cyclones

As discussed in Section 2.2.4.1, there is potential energy to be released into kinetic energy by converting an atmosphere where density depends on pressure and temperature (baroclinic) into an atmosphere where density depends on pressure alone (barotropic). The Eady growth rate (Lindzen and Farrell, 1980) measures baroclinicity, and so can describe the deepening of cyclones as they tap into the baroclinicity, using two factors. Firstly, it considers stability, which measures the atmosphere's propensity towards overturning using the vertical temperature gradient. However, the concept of warmth in the atmosphere is complicated by the pressure variation with height, so potential temperature (θ) is used instead. θ measures the temperature that an air parcel would have, if it were brought to a reference pressure (1000hPa). An atmosphere with a large, positive change in θ with increasing height will tend to allow overturning. Secondly, the Eady growth rate considers vertical wind shear, as greater shear induces a faster growth rate.

This work also aims to include moisture, and so implicitly consider diabatic processes in the growth rate. Previous studies (e.g. Emanuel et al., 1987) have stated that the inclusion of moist processes will cause cyclones to grow faster. Emanuel et al. (1987) and Whitaker and Davis (1994) have endeavoured to include the effects of moisture within the atmosphere, as diabatic processes are governed by moisture and can affect the stability of the atmosphere, and therefore the growth rate. However, the formulation of the growth rate depends on whether the quasi-geostrophic or semi-geostrophic set of equations is being used to describe the atmosphere. Both the quasi-geostrophic and semi-geostrophic models are considered in Chapter 6. In addition to these factors, there is another source of kinetic energy for midlatitude cyclones: barotropic processes.

These convert energy from the large-scale kinetic energy of the flow into the kinetic energy of the cyclone. Gill (1982) provides the formula for the barotropic growth rate, which is related to the horizontal shear of the total horizontal wind but not to variations in temperature or density.

The present work aims to apply these ideas of growth rates to ECMWF ERA-Interim reanalysis data, and develop parameters that describe the potential for barotropic or baroclinic process to be the source of energy to the midlatitude cyclones. Further discussion of the method and all of the results can be found in Chapter 6.

3.8 Summary

This chapter describes the data and methods that will be used to obtain the results. Initially, the data sets that will be used throughout this project are identified: ERA-Interim reanalysis for the work where only analysis data is needed, and deterministic and ensemble ECMWF forecasts where appropriate. The first steps will be to use ERA-Interim to identify and categorise storms. Identification will be done using SSI (Equation 3.1), which is a meteorological parameter that can be used as a proxy for the damage a storm could inflict. Categorisation will be done using the configuration of the jet stream and storm track (Section 3.4), and by assessing which terms dominate the Pressure Tendency Equation (Section 3.4.2). The different methods of categorising storms will then be compared, qualitatively and quantitatively (Section 3.5). The results of this analysis will be discussed in Chapter 4.

The next portion of the work will assess forecast quality and spread. The differences between forecast and analysis storm tracks will be compared, at the time of minimum pressure, in terms of both location (measured by longitude and latitude) and intensity (core pressure). The absolute values of the differences will quantify forecast quality. The spread of the differences in ensemble forecasts will quantify the predictability. This portion of the work will then link back to the categorisation of the storms, to determine whether one type of storm is particularly well or particularly badly forecast. These results will be presented in Chapter 5.

The final results concentrate on storm-prone situations, and the identification and analysis of potential metrics. Once these parameters have been quantitatively examined on their own, they will be related to the selected storms. Finally, this work assesses which categories of storms are most strongly related to which growth rate, to establish whether one metric is strong for a particular

type of storm. Finally, discussion of storm predictability and related storm-prone situation will be discussed: are diabatically-driven storms associated with a peak in the moist growth rates more often than their dry counterparts? This analysis can be found in Chapter 6.

Finally, the summary and discussion of this work can be found in Chapter 7. This section will draw together the key themes of variation between North Atlantic midlatitude cyclones and their predictability.

Chapter 4

The Storms

In this Chapter, the results of storm selection are presented, followed by two possible methods of grouping them. The selection of storms uses the Storm Severity Index (SSI), as described in Section 3.2. Dividing the storms into groups using plots of the jet stream and storm track, and analysis using the Pressure Tendency Equation (PTE), explained in Section 3.4.2. A comparison of these categories concludes the current chapter.

4.1 Measuring Storm Severity

In this work, there is a need for a method for determining which windstorms are of interest, out of the whole range of North Atlantic cyclones. For this, a metric that is related to the damage a storm could inflict, but based on meteorological variables is needed, in order to rank days and events by their 'storminess' (Section 3.2). The SSI is chosen, and is calculated for an area over Europe (Figure 3.1), for every day since 1st January 1990, by comparing to the wind climatology from 1979 to 2011. The results for the top SSI values are shown in Table 4.1. Some high SSI days are not associated with North Atlantic extratropical cyclones and are rejected due to:

- **Duplication**, where a high SSI value is caused by the same cyclonic storm as a value ranked higher than the current value;
- **Mediterranean storms** that are dynamically different to North Atlantic storms;
- **Polar lows** that are also dynamically different;

- **High Pressure** dominating the pressure pattern over Europe, but associated with a sufficiently large pressure gradient and high wind speeds to induce a high value of SSI;
- **Orography** was locally enhancing high wind speeds and so high SSI values.

Each of these will shortly be discussed in more depth. However, this rejection process does imply that SSI could be improved from the version used here. For duplication, a simple algorithm could be included that means that high SSI value events cannot occur within a certain time period, but human intervention would be required in cases where storms are clustered very close together. For example, storms Lothar and Martin highest SSIs occur 30 hours apart, but there are other occasions when high SSI values 30 hours apart would be from the same storm, and the assessment of this would be difficult to automate. Another example is Vivian and Wiebke, where the storms occur so close together as to be causing strong winds in Europe on the same day, making them difficult to separate into two separate high-SSI events. In terms of the weather phenomena that cause high SSI values but are rejected because they are no midlatitude cyclones, the simplest way to exclude them automatically would be to make the SSI area smaller. However, this would inevitably lead to some of the weather systems of interest also being missed, so the rejection process remains subjective. While this is not ideal, the human analysis of pressure charts to identify weather systems that require rejection is quicker and better than introducing automation, which would have its own drawbacks. Furthermore, future work could include analysis of the high-SSI weather events that are not midlatitude cyclones, such as the Mediterranean storms.

TABLE 4.1: Table showing the dates and times of the top SSI values, alongside either the associated North Atlantic storm or the reason for rejection. The top 120 are analysed, but outside the top ten the duplicates are not shown.

Date	**Time (UTC)**	**SSI**	**Storm Name**	**Reason for rejection**
2009-01-24	1200	55.466	Klaus	
1990-02-26	1200	46.311	Vivian & Wiebke	
1995-03-30	1200	45.966		Mediterranean
2009-01-24	0600	45.158		Duplicate (Klaus)
1990-02-27	1200	41.420		Duplicate (Vivian & Wiebke)
2009-01-24	0000	40.488		Duplicate (Klaus)
2007-01-18	1800	39.467	Kyrill	
2007-01-19	0000	39.441		Duplicate (Kyrill)
1999-12-26	1200	37.678	Lothar	

Continued on next page

Table 4.1 – *Continued from previous page*

Date	Time	SSI (3 d.p.)	Storm Name	Reason for rejection
1999-12-27	1800	37.110	Martin	
2008-03-01	1200	34.089	Emma	
1995-03-28	0000	33.855		Mediterranean
1994-01-23	0600	33.009		High Pressure
2005-01-08	1800	29.216	Erwin	
1992-03-26	1200	28.866		Mediterranean
2004-11-14	1200	28.559		Mediterranean
2002-10-28	0000	27.507	Jeanette	
1990-01-25	1200	27.330	Daria	
1996-02-08	0000	25.510		Mediterranean
2001-03-03	1200	22.000		Mediterranean Low
2008-03-05	1200	20.986		Mediterranean Low
2006-03-12	1200	19.798		Mediterranean Low
1993-01-24	1200	19.695	Agnes	
1999-12-04	0000	18.598	Anatol	
1991-01-05	1200	17.155	Udine	
1999-02-22	1200	16.519		Polar Low
2000-11-06	1200	16.245	Oratia	
1992-12-28	1200	16.123		High Pressure
2006-03-05	1200	15.922		High Pressure
1999-02-05	1200	15.677	Lara	
2010-02-28	0000	14.939	Xynthia	
2002-01-28	1200	14.650	Jennifer	
1999-11-19	1200	14.137		Mediterranean Low
1994-01-29	1200	13.636		Mediterranean Low
2005-01-12	0000	13.530	Gero	
1999-01-30	1200	13.432		Mediterranean Low
2007-01-14	1200	13.376	Hanno	
1995-01-31	1200	13.293		High Pressure
1990-03-05	1200	13.131		High Pressure
1998-12-27	0000	12.421	Silke	
1997-02-18	0000	12.301		Orography
2005-11-23	1200	12.088		High Pressure
1997-02-19	1800	12.058		High Pressure

Continued on next page

Table 4.1 – *Continued from previous page*

Date	**Time**	**SSI (3 d.p.)**	**Storm Name**	**Reason for rejection**
2000-10-30	1200	11.837		Polar Low
2010-03-09	1200	11.477		Mediterranean Low
2002-02-22	0000	11.385		High Pressure
1998-03-10	1200	11.143		Mediterranean Low
1994-01-28	0000	11.000		Polar Low
2000-12-07	1200	10.934	Elke	
2004-11-19	1200	10.930		High Pressure
2009-10-12	1200	10.087		High Pressure
1993-12-09	0000	9.890	‘Dec 1993’	
2000-03-09	1200	9.313		High Pressure
1995-01-22	1800	9.287	Urania	
1990-02-11	1800	9.242	Nana	
2010-03-10	0000	9.053		Mediterranean Low
1997-02-13	1200	9.038		High Pressure
2009-03-21	1200	8.899		High Pressure
1998-01-20	1200	8.833		High Pressure
1998-12-20	1200	8.628		High Pressure
1999-12-17	1800	8.552		High Pressure
1993-03-27	1200	8.548		High Pressure
2009-02-10	0000	8.323	Quinten	
1993-01-14	0000	8.312	Verena	
2007-01-29	1200	8.265		High Pressure
2000-01-21	1200	8.149		High Pressure
2000-01-29	1200	7.977	Kerstin	
2002-01-02	0000	7.815	Pawel	
2005-12-16	1200	7.801	Cyrus	
2001-01-28	1200	7.765	Lukas	
1999-12-12	1200	7.699	Franz	
2002-03-07	1200	7.635	Frieda	

4.2 Reasons for Rejecting High SSI Days

4.2.1 Duplicates

Many entries in the table occur within 24 hours of a higher SSI value. These are rejected, so one storm event has only one date and time under consideration with the maximum SSI value retained. For example, storm Klaus has the highest SSI value (55.47) at 1200UTC on 2009-01-24, as well as the fourth and sixth highest values (45.16, 40.49) at 0600UTC and 0000UTC on the same day. As these are related to the same storm, the highest value is retained, and the fourth and sixth rejected as duplicates. The 48 hour period (24 hours before the storm, 24 hours after) is extended for some storms, when examining surface pressure charts (not shown) indicate that two high SSI values are caused by the same system. Overall, 49 out of the 120 analysis times are rejected for this reason.

4.2.2 Mediterranean Lows

Mediterranean lows have different dynamics from the North Atlantic cyclones of interest to this work (Trigo et al., 1999), but they cause relatively high wind speeds in the area where SSI is calculated. As these are rare events, the Mediterranean region has a lower 98th percentile of wind speed. Therefore, the wind speed does not need to be as high as in northern Europe for the SSI to be large. One example of this is 30th March 1995 at 1200UTC. High values of SSI are concentrated near Venice and Genoa, Italy (Figure 4.1c), and the associated high wind speeds occur entirely in the Mediterranean region (Figure 4.1b). Although the surface pressure chart is dominated by the high pressure over western Europe, a small low is present in the Adriatic Sea (Figure 4.1a). The high and low cause a strong pressure gradient, notably in the Gulf of Genoa and northern Italy, inducing high winds there (Figure 4.1b) and so high values of SSI. As the aim of this study is to investigate North Atlantic cyclones, the Mediterranean cyclones are eliminated, which means that thirteen of the 120 high SSI events are rejected.

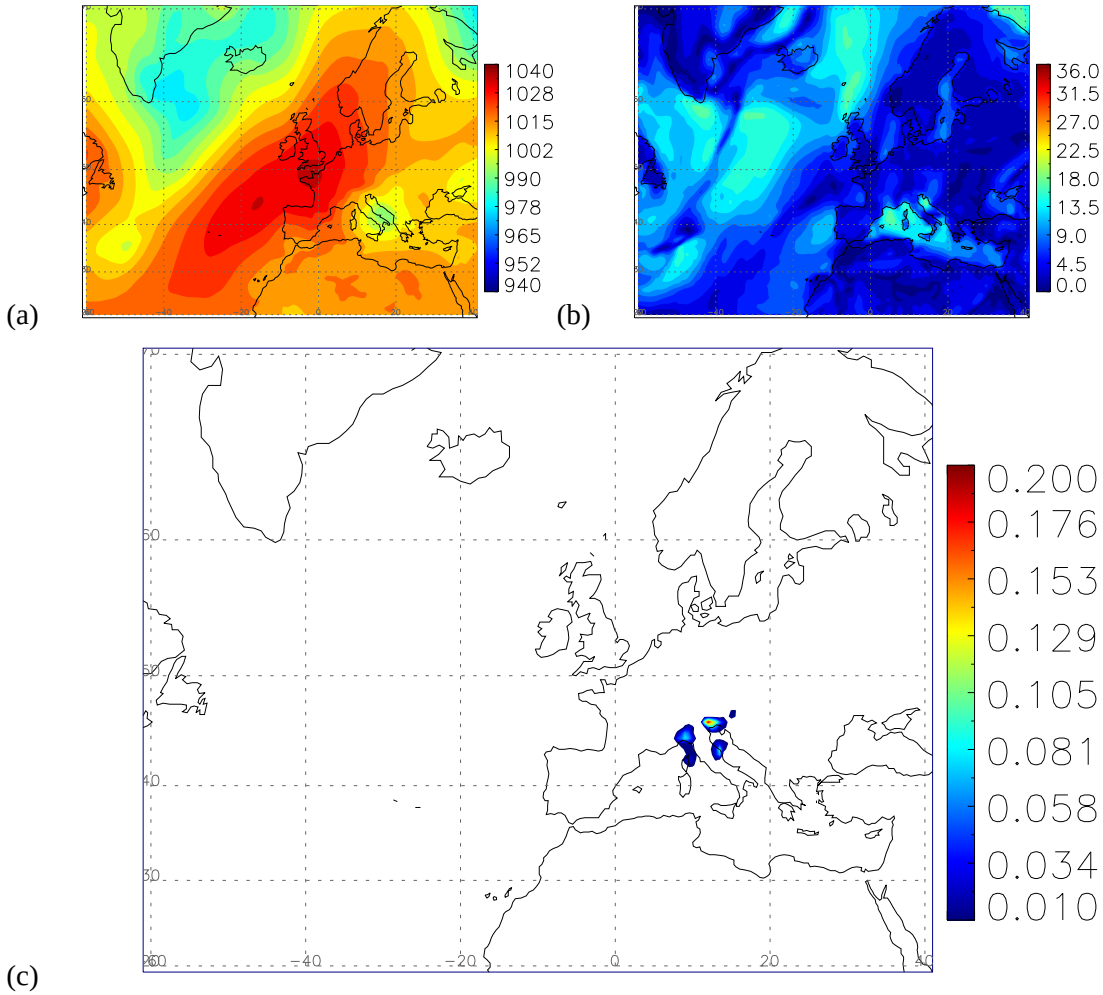


FIGURE 4.1: (a) Mean sea level pressure (hPa), (b) wind speed (ms^{-1}) and (c) SSI for 1200UTC 30th March 1995.

4.2.3 Polar Lows

There are three Polar Lows identified, which behave unlike the other Atlantic cyclones. Figure 4.2a shows storm Petra, who caused high SSI at 1200UTC on 22nd Febuary 1999. The track starts very far north and moves southeastwards into Europe. The pressure plot (Figure 4.2c) shows that Petra is filling throughout. Though this track is generally parallel to the jet stream (Figure 4.2a), the intense wind speeds are so far away from the storm that influence from the jet stream is probably small. The θ_e field (Figure 4.2b) shows that the air is unusually warm and moist for such a high latitude, indicating the potential for convective energy to be released, as is typically the case for polar lows (Rasmussen, 2003). The surface pressure charts during the days before the track of the storm is detected (not shown) indicate that the cyclone is the result of cyclogenesis between Newfoundland and Greenland. The resulting disturbance then crosses the northern North Atlantic, passing over the tip of Greenland, and moving north of Iceland. This low pressure then remains in the Norwegian Sea for three days, before forming a sufficiently coherent low for the tracker (Section 3.3) to detect it. This cyclone and the others like it are different to the systems of interest in a number of ways: no deepening phase, no clear interaction with the jet stream, and originating very far north. Therefore, they are probably polar lows, and will not be considered further.

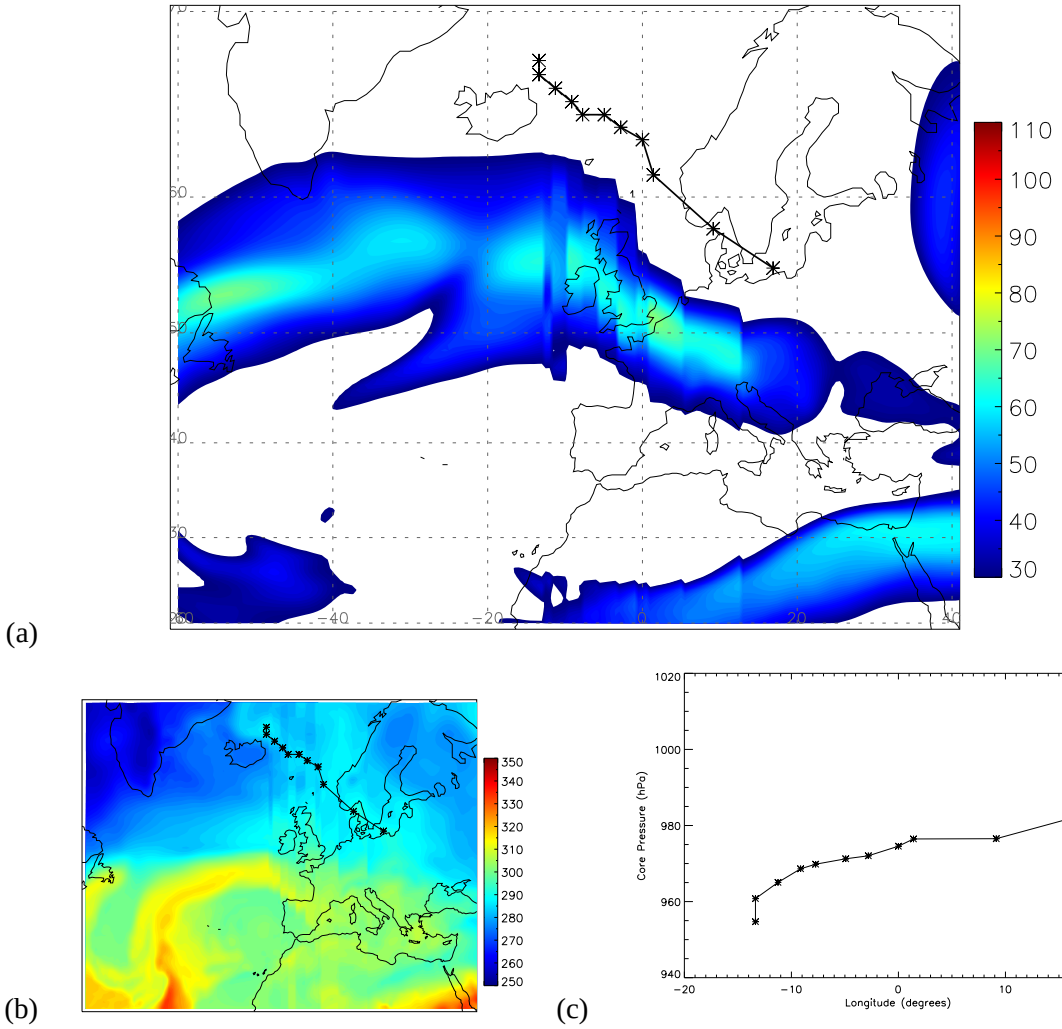


FIGURE 4.2: Polar low Petra at 1200UTC on 22nd Febuary 1999: (a) wind speed at 300hPa [ms^{-1}], (b) θ_e at 850hPa [K], (c) cyclone core pressure [hPa] against time [h]. (a) and (b) are plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4.

4.2.4 High Pressure

The fourth reason for rejecting some high-SSI days is the presence of high pressure over most of Europe, with high winds and SSI in some parts of the domain. In the case from 5th March 1990, there are strong westerlies across northern Scotland and southern Scandinavia (Figure 4.3b), induced by a pressure gradient between the unusually high pressure of over 1040hPa across Europe and the stationary low pressure in the Norwegian Sea (Figure 4.3a). There is a short-wave trough visible in the pressure field, which could increase the wind speeds enough for them to generate a high-SSI, but this kind of system is not of interest to this study for both dynamic and practical reasons: it has different factors affecting its deepening to midlatitude cyclones, and it is difficult for the tracker to detect a trough without closed pressure contours. Examining the SSI field (Figure 4.3c) indicates an additional orographic influence on the winds, as the highest SSI values occur as the westerlies reach the Norwegian mountains. There are nineteen occasions where a strong large-scale pressure gradient is the main reason for a high SSI value, and these were all checked by hand to ensure no mobile North Atlantic cyclones are present in these cases, so they are rejected.

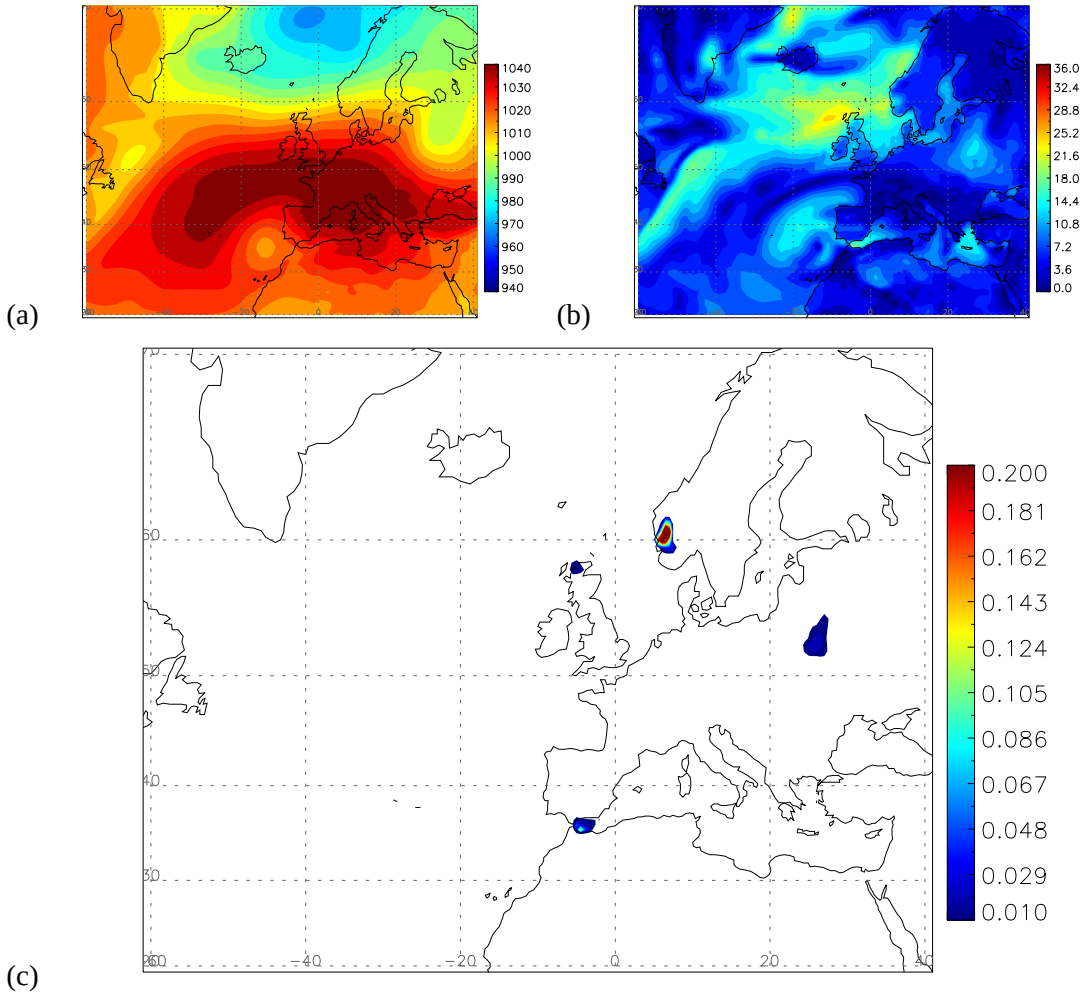


FIGURE 4.3: (a) Mean sea level pressure (hPa), (b) wind speed (ms^{-1}) and (c) SSI for 1200UTC on 5th March 1990.

4.2.5 Orography

The inclusion of the 1000m criteria on the calculation region for SSI (Section 3.2) filters out many cases with high SSI values over orography. This means that only one of the top SSI days is rejected on orographic grounds, on 18th February 1997. Figure 4.4a shows a strong pressure gradient between Iceland and the rest of Europe. This produces exceptional wind speeds off the Norwegian Coast (Figure 4.4b) and associated high values of SSI (Figure 4.4c). These exceptional wind speeds could be due to orographic blocking, where the wind is forced to speed up so it can move over the mountains or form a barrier jet parallel to the mountain range. As this is not a clear-cut case of a North Atlantic cyclone, this high-SSI day is rejected.

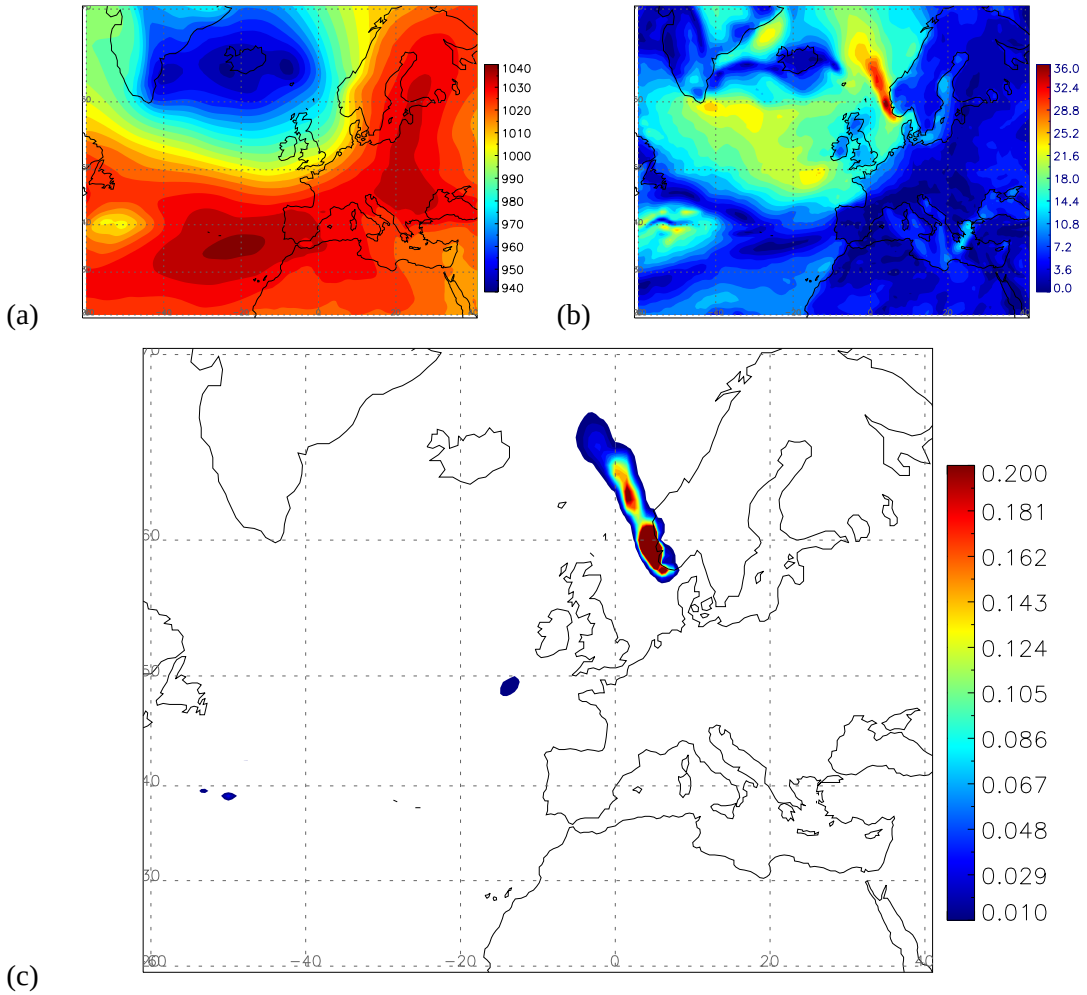


FIGURE 4.4: (a) Mean sea level pressure (hPa), (b) wind speed (ms^{-1}) and (c) SSI for 0000UTC on 18th February 1997.

4.3 Selected Storms

Once all of the rejected days are removed, the final list of storms is collated (Table 4.2). This table sheds light on several interesting aspects of storm damage. Firstly, the SSI of a storm is not clearly correlated with the damage a storm does. This is for reasons already discussed, in that the amount of damage a storm inflicts is not only based on the severity of the storm, but also on the region through which the storm passes and the socio-economic profile of that area - how densely populated is the area? How well prepared for a storm are that population? Another factor that becomes clear is that the number of people killed in a storm is not clearly connected to the economic damage. The reasons for this are: degree of preparedness; whether any weather warnings had been issued; whether those warnings were heeded; and finally luck, because for example in one case a tree could fall on the road and hurt no-one, and in another case it could fall on a moving vehicle or cause a serious accident. The complications to the damage statistics caused by socio-economic factors such as these highlight the necessity for a simple measure of the potential a storm has to inflict damage, and reiterates the reasons for selecting SSI as such a measure for this project.

One storm, Erwin, is rejected after proving too difficult to track. Despite being a destructive storm that killed 16 people with a high SSI of 29.22, the pressure contours did not close and the system remained an open wave. This type of system is difficult for the tracker to detect and follow, and so Erwin is eliminated from further investigation.

Storm Name	Dates		SSI	People Killed	Damage (Million US$)	
	Start	End			At the time	Jan 2011
Daria	1990-01-24	1990-01-28	27.33	85	6 860	13 400
Nana	1990-02-10	1990-02-12	9.24	1	190	370
Vivian	1990-02-24	1990-02-27	46.31	50	3 230	6 310
Wiebke	1990-02-26	1990-03-03	46.31	67	2 260	4 410
Udine	1991-01-02	1991-01-08	17.16	48	909	1 610
Verena	1993-01-10	1993-01-16	8.31	6	385	625
Agnes	1993-01-20	1993-01-27	19.69			
Dec 1993	1993-12-06	1993-12-11	9.89			
Urania	1995-01-21	1995-01-24	9.29			
Silke	1998-01-24	1998-01-29	12.42			
Lara	1999-02-03	1999-02-05	15.68			
Anatol	1999-12-02	1999-12-04	18.60	27	2 963	4 020
Franz	1999-12-07	1999-12-14	7.70			
Lothar	1999-12-24	1999-12-28	37.68	137	11 350	15 400
Martin	1999-12-26	1999-12-29	37.11	14	4 100	5 560
Kerstin	2000-01-27	2000-01-31	7.98			
Rebekka	2000-11-03	2000-11-07	16.25			
Elke	2000-12-03	2000-12-10	10.93			
Lukas	2001-01-27	2001-01-31	7.77			
Pawel	2001-12-31	2002-01-02	7.81			
Jennifer	2002-01-24	2002-01-30	14.65	17	150	194
Frieda	2002-03-05	2002-03-08	7.64			
Jeanette	2002-10-24	2002-10-28	27.51	39	2 531	3 240
Gero	2005-01-08	2005-01-12	13.53	7	50	60
Cyrus	2005-12-13	2005-12-18	7.80			
Hanno	2007-01-12	2007-01-15	13.38			
Kyrill	2007-01-15	2007-01-20	39.47	46	9 010	10 100
Emma	2008-02-27	2008-03-01	34.09	13	1 800	1 940
Klaus	2009-01-22	2009-01-24	55.47	28	5 100	5 310
Quinten	2009-02-07	2009-02-12	8.32			
Xynthia	2010-02-24	2010-03-03	14.98	64	6 074	6 310

TABLE 4.2: Selected storms with dates and maximum SSI. The names of the storms are obtained from Free University of Berlin charts, for all but one storm. The value of SSI quoted is the maximum reached, during the time the storm was passing over Europe. Details of fatalities and estimated damage are provided based on data from Centre for Research on the Epidemiology of Diseases (2012), where available. The values of destruction are given for both the value at the time, and the value at 1st January 2011 corrected for inflation (Office for National Statistics, 2011) to ease direct comparison.

4.4 Storm Categories

Next is an analysis of the jet stream and θ_e fields of the storms. These two variables qualitatively indicate processes that are key in the development of a cyclone, as discussed in Section 3.4. Examination of these properties shows that the storms tend to fall into groups, depending on their track relative to the jet stream, but not in the θ_e field. Therefore, the jet stream is chosen over θ_e as the quantity for this categorisation, because no clear categories were apparent in the θ_e field. While θ_e is an important for the identification of airmasses and so can be considered as a measure of the temperature contrast across fronts and the moisture content of the warm sector air, these factors refer to the potential energy of the storm and how deep it could get. The jet stream provides the divergence aloft that evacuates mass from the low pressure centre, and so is perhaps more directly related to the actual depth of the storm. Furthermore, previous studies have identified jet crossing as an important point in a midlatitude cyclones' lifetime (e.g. Liberato et al., 2011), so if and when a storm crosses the jet could be important to its development. From visual inspection, four jet categories emerge:

- **Edge:** Nana, Agnes, Lara, Franz, Kerstin, Lukas, Pawel, Frieda, Jeanette, Emma.
- **Cross late:** Wiebke, Verena, Urania, Lothar, Martin, Jennifer, Cyrus, Hanno, Klaus, Quinten.
- **Cross early:** Daria, Vivian, Udine, Silke, Anatol, Rebekka, Kyrill.
- **Split:** Dec 1993, Elke, Gero, Xynthia.

The difference between the 'cross early' and 'cross late' categories is difficult to determine subjectively, so a more objective method is utilised here. First, the point at which a storm crosses the jet is determined subjectively. The percentage of storm lifetime that the members of these two categories spent on either side of the jet stream is then calculated. If a storm crosses the jet less than 35% of the way through its track, it is 'cross early'; if it spends more than 35% of its time on the southern side of the jet, then it is 'cross late'. The main reason for selecting 35% for the limit is that the storm tracker is more effective in the latter stages of a storm's lifetime, when cyclolysis has begun, than in the beginning stages. This means that more points on the track are associated with cyclolysis than cyclogenesis. Given that this project concentrates on cyclogenesis, the limit was set accordingly. Furthermore, the limit of 35% agrees with subjective analysis of the jet stream plots such as that shown in Figure 3.3.

An example of each type will now be presented. The storms are tracked in ERA-Interim analysis data, as discussed in Section 3.1. The jet stream and θ_e plots use the same data, and the meridional slicing plotting method described in Section 3.4.

4.4.1 Edge Storms

These storms typically travel along the northern edge of a jet stream that curves anticyclonically. The storms deepen as they move north alongside the jet stream, taking advantage of the divergence aloft, as indicated in Figure 2.7b. However, as the jet stream curves back south and there is a swap to convergence aloft, the storms tend to continue eastwards and begin to fill (Figure 4.5). One example typical of the category is storm Agnes, from January 1995.

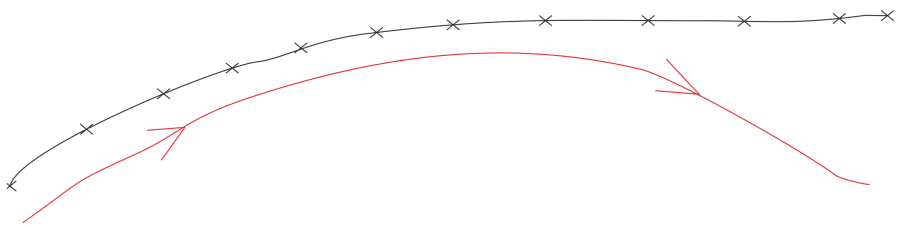
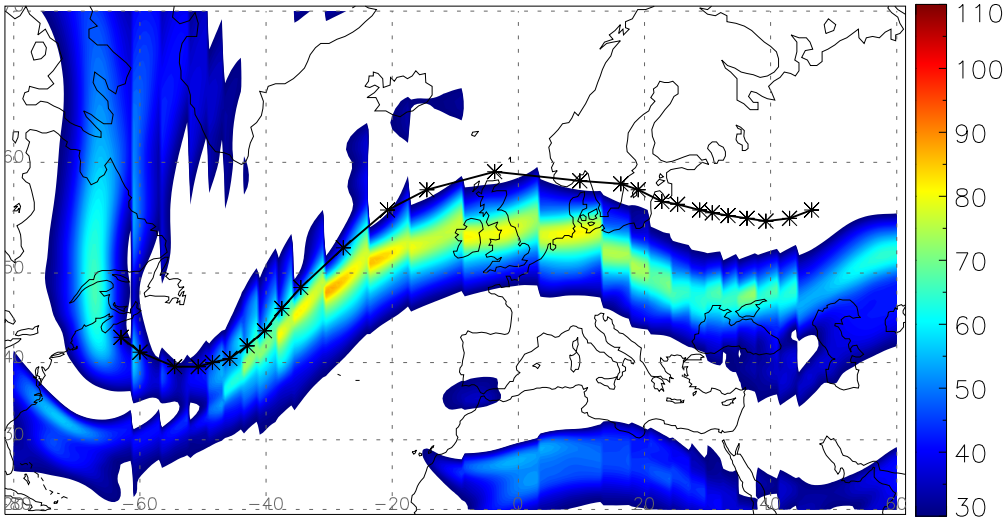
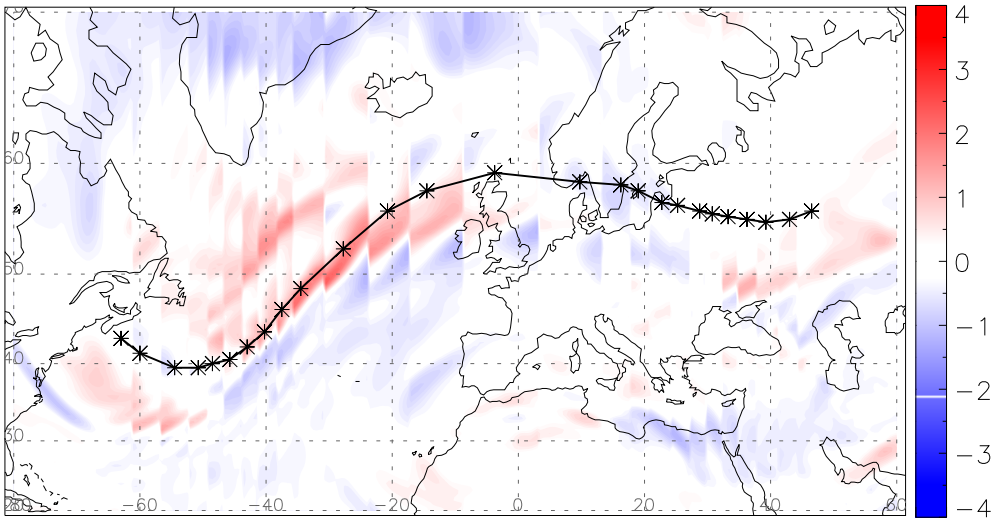


FIGURE 4.5: Schematic showing the typical track of an 'edge' storm (black) relative to the jet stream (red).

Agnes formed to the north of a curved jet stream, which is forced by a large trough over Greenland (Figure 4.6a). Agnes moved along the northern edge of the jet stream, where there is strong divergence aloft (Figure 4.6b) to aid deepening. As the jet stream moved southwards, Agnes continued in a more easterly direction towards Europe. The θ_e field (Figure 4.7a) in the region where Agnes formed shows a strong north-south gradient, and then the storm moves along a slight local maximum which is probably the warm sector. The deepening occurred before the storm reaches the most northerly point on her track, and Agnes began to fill as the jet stream shifted further southwards (Figure 4.7b). Overall, divergence aloft, due to the curvature of the jet stream, is likely to be the process that caused most of the deepening. Once the jet stream and the cyclone parted ways, then filling begins.

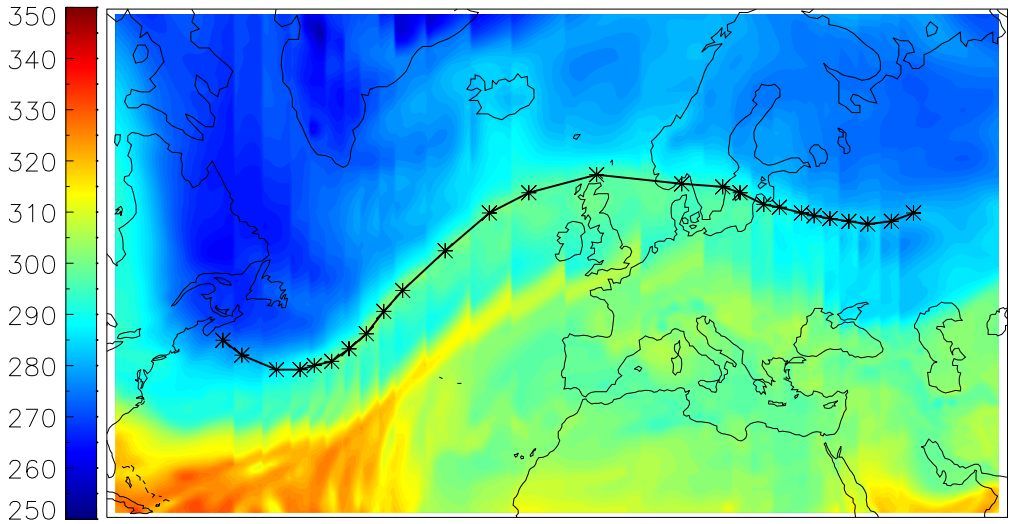


(a)

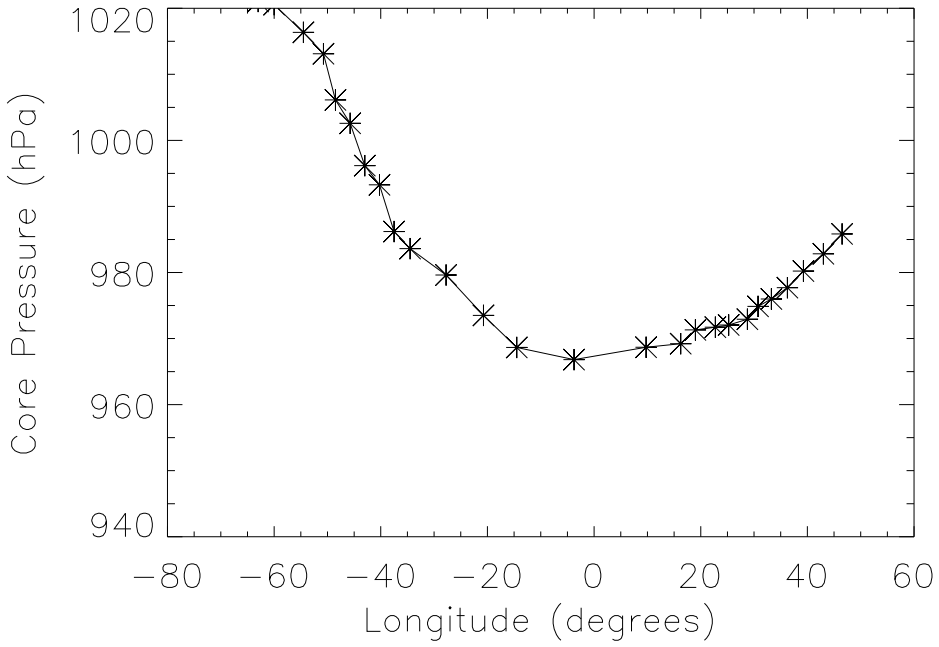


(b)

FIGURE 4.6: Storm Agnes, an example of an 'edge' storm from January 1993 (see Table 4.2 for exact dates): (a) wind speed at 300hPa [ms^{-1}], (b) divergence of the wind field at 300hPa [$10^{-}4, s^{-1}$]. Plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4.



(a)



(b)

FIGURE 4.7: Storm Agnes, an example of an 'edge' storm from January 1993: (a) θ_e at 850hPa $[K]$, (b) cyclone core pressure $[hPa]$ against time $[h]$. (a) is plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4.

4.4.2 Late Jet-Crossing Storms

Late jet-crossing storms tend to cross a straight or only slightly curved jet stream, from the right entrance to the left exit region of the jet streak (Figure 4.8), and so benefit from the divergence aloft found in both of these locations (Figure 2.7a). They tend to have straight tracks for the most part with an east-north-easterly direction, though some have a jet stream that is slightly curved (usually anticyclonically). The minimum pressure occurs just after the storm crosses the jet stream and moves into the left exit region. After the storm leaves this region, it tends to dissipate and move more slowly.

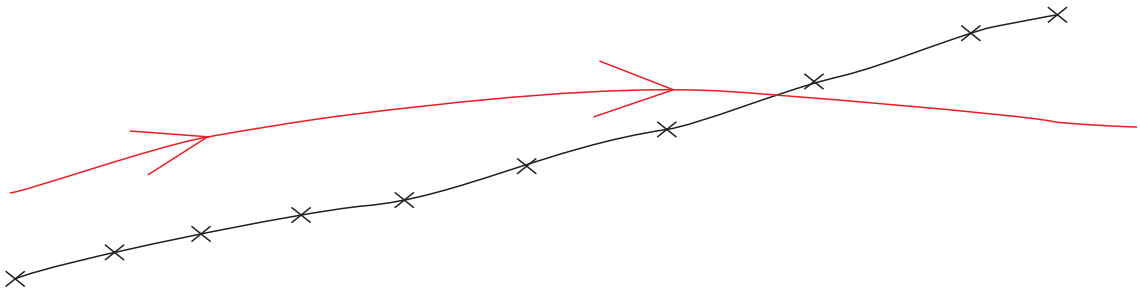
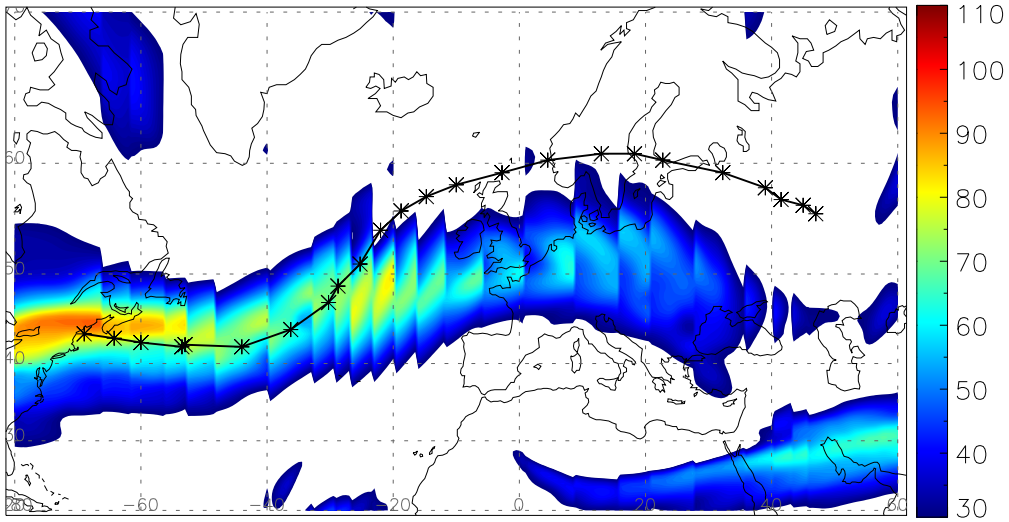
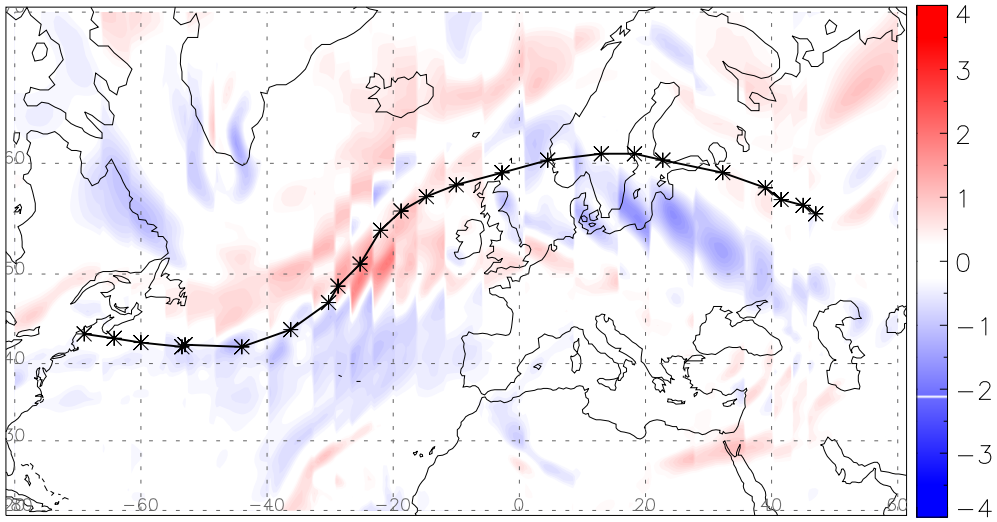


FIGURE 4.8: Schematic showing the typical track of a 'cross-late' storm (black) relative to the jet stream (red).

This behaviour can be seen in the example of storm Jennifer. The track starts off the northeastern coast of the USA, on the southern side of a jet maximum (Figure 4.9a). This is also a region of strong θ_e gradient (Figure 4.10a), indicating a baroclinic atmosphere and a source of potential energy for the cyclone. The jet stream was not perfectly zonal, but curved slightly cyclonically then anticyclonically. As storm Jennifer moved along the southern side of the jet, it deepened and the track is also curved cyclonically. The deepening was strongest in a region of strong divergence aloft, about ten degrees west of Ireland (Figures 4.9b and 4.10b), where the sense of curvature in both the jet stream and storm track changes to anticyclonic. After this sudden deepening, Jennifer's track moved further from the jet stream and filling began.

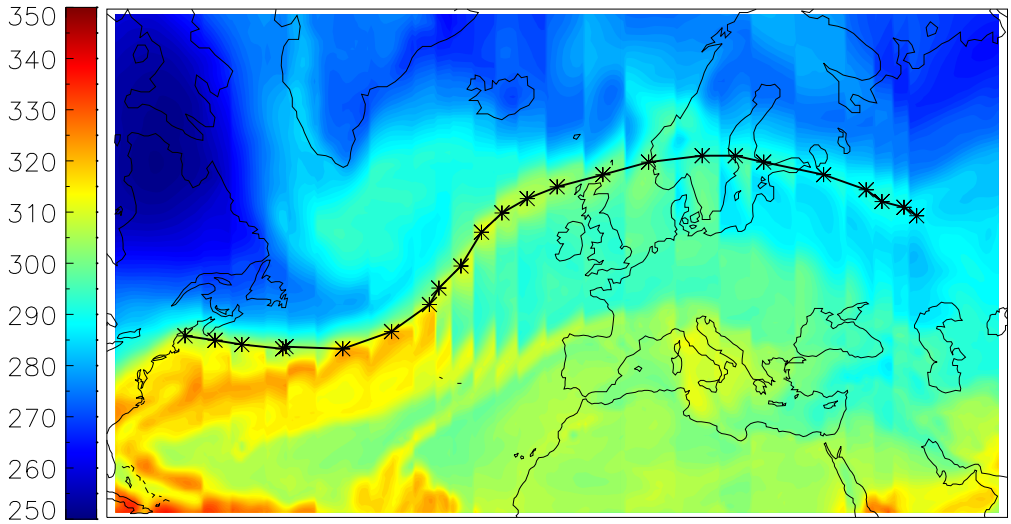


(a)

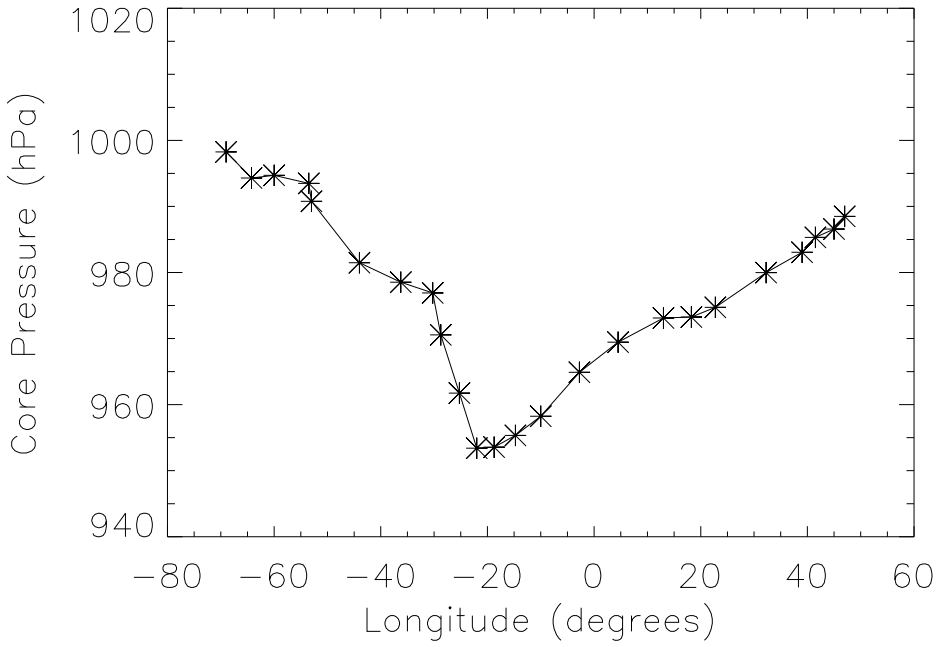


(b)

FIGURE 4.9: Storm Jennifer, an example of a 'cross late' storm from January 2002: (a) wind speed at 300hPa [ms^{-1}], (b) divergence of the wind field at 300hPa [$10^{-}4, s^{-1}$]. Plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4.



(a)



(b)

FIGURE 4.10: Storm Jennifer, an example of a 'cross late' storm from January 2002: (a) θ_e at 850hPa $[K]$, (b) cyclone core pressure $[hPa]$ against time $[h]$. (a) is plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4.

4.4.3 Early Jet-Crossing Storms

These storms cross the jet early on in the track, deepening as they pass through the right entrance and left exit regions. Then, they remain beside the northern edge of the jet stream and continue alongside the jet (Figure 4.11). Subsequently, as the cyclone and jet stream part ways, the system begins to fill. The initial location of the jet stream, as it moves from south to north, is due to the presence of a trough upstream. As an example, storm Kyrill is shown in Figures 4.12 and 4.13.

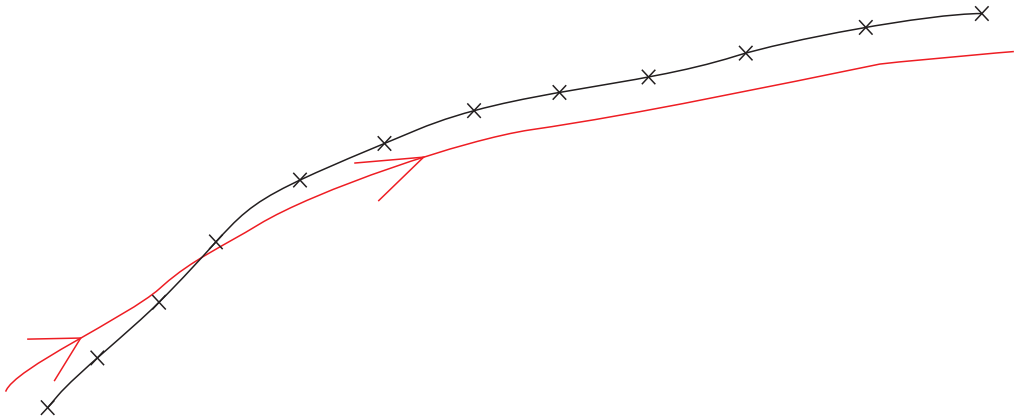
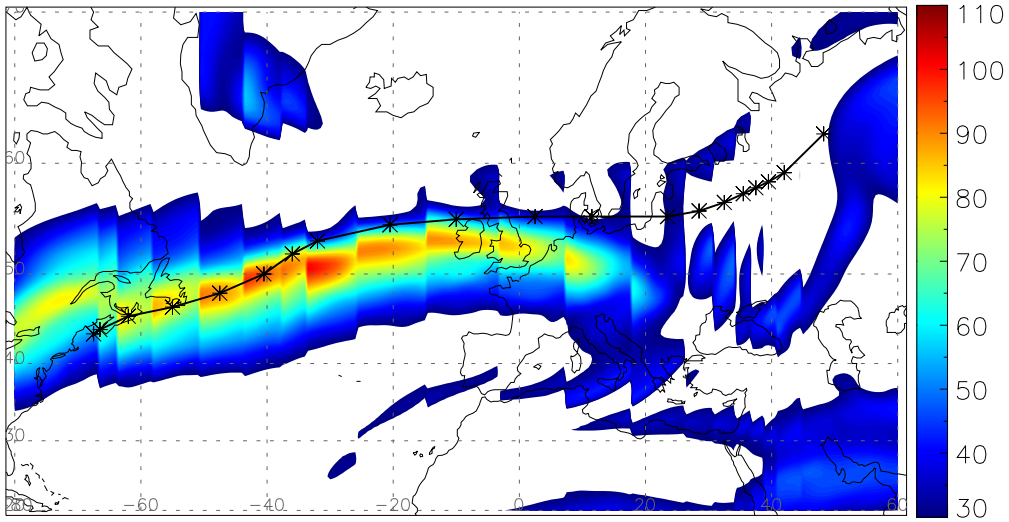
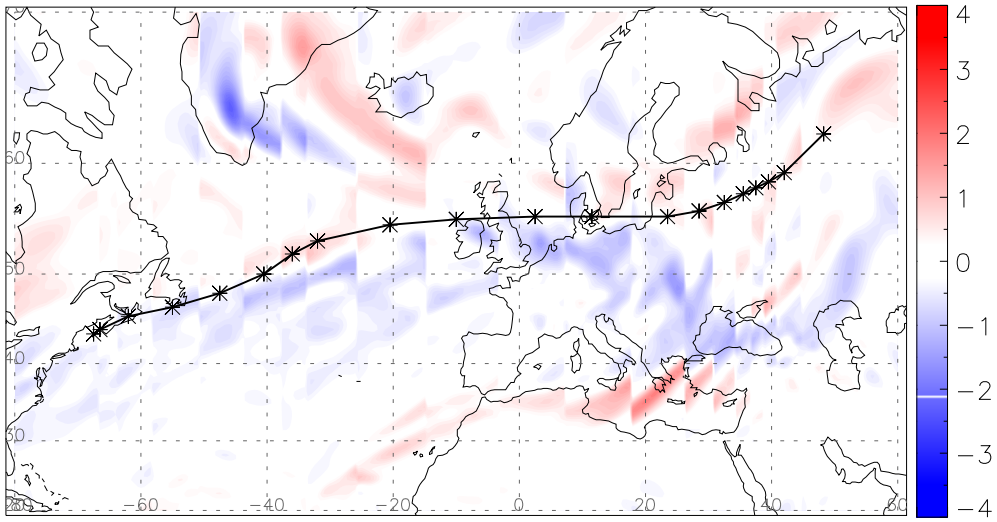


FIGURE 4.11: Schematic showing the typical track of a 'cross-early' storm (black) relative to the jet stream (red).

Storm Kyrill is used to illustrate this jet stream type. It developed over the northeastern USA coast, on the southern side of the jet stream (Figure 4.12a), in a region of neither divergence nor convergence (Figure 4.12b) but strong θ_e gradient (Figure 4.13a). In the early stages, Kyrill deepened suddenly (Figure 4.13b), dropping in 40hPa in 36 hours meaning it was a 'bomb' storm, despite passing through weak convergence (Figure 4.12b). This could indicate that the divergence is occurring at another level of the atmosphere in the early stages of storm development. On crossing 40^oW, Kyrill encountered some strong divergence aloft. The core pressure reached its minimum off the west coast of Ireland, and the track became zonal to pass north of Northern Ireland, through the Scottish Borders, across the North Sea, and over Denmark then Lithuania. During this time, Kyrill's core pressure remained low indicating that, although deepening had ceased, filling had not yet begun. Cyclolysis began as the storm passes over Belarus and Russia.

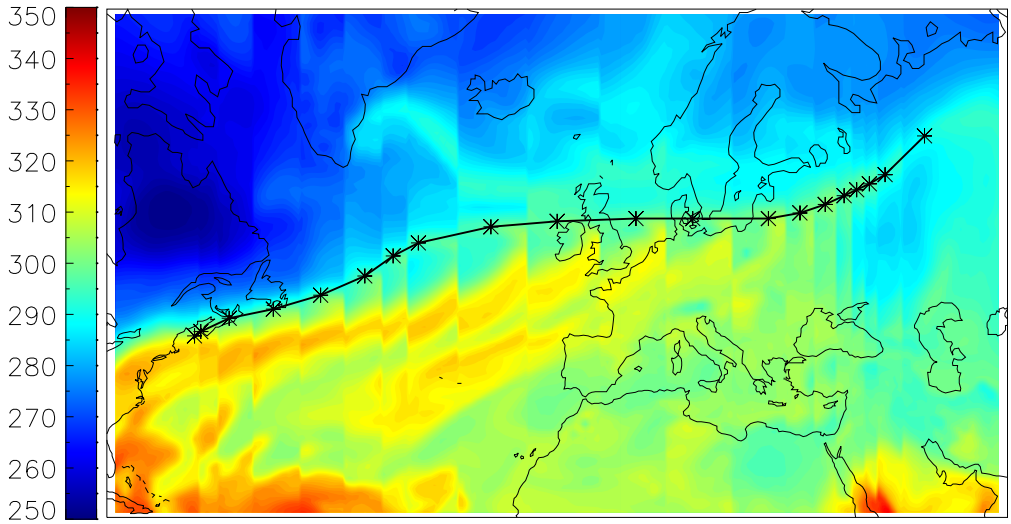


(a)

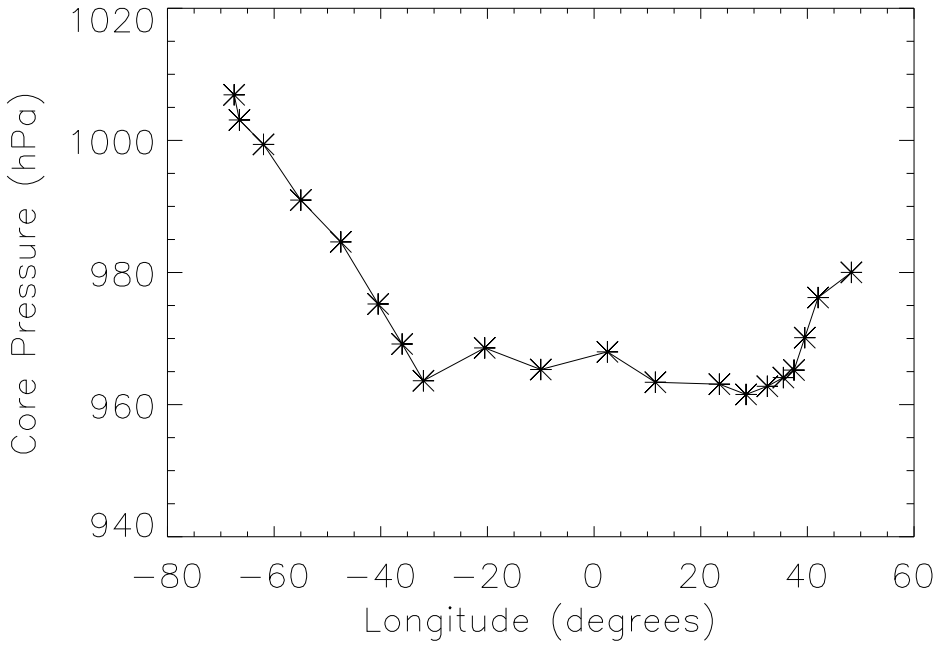


(b)

FIGURE 4.12: Storm Kyrill, an example of a 'cross early' storm from January 2007: (a) wind speed at 300hPa [ms^{-1}], (b) divergence of the wind field at 300hPa [$10^{-}4, s^{-1}$]. Plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4.



(a)



(b)

FIGURE 4.13: Storm Kyrill, an example of a 'cross early' storm from January 2007: (a) θ_e at 850hPa $[K]$, (b) cyclone core pressure $[hPa]$ against time $[h]$. (a) is plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4.

4.4.4 Storms with a Split Jet

The four storms with a split jet are unusual, as they are associated with two separate jet stream maxima. In the cases of Elke, Xynthia and a storm for which no name could be found in December 1993, the storms travel between the two jets while deepening (Figure 4.14). There are two jet streaks associated with these storms, which are typically located where the arrows have been placed on the jets (coloured red in Figure 4.14). Therefore, there are two areas of divergence aloft, to aid the deepening process. As the storms leave the region between the two jet maxima they begin to fill. This was found in snow storms over the eastern United States by Uccellini and Kocin (1987), who describe the way that two jet streaks can be arranged so that the left exit region of one and right entrance of the other coincide, inducing intense divergence and so strong deepening.

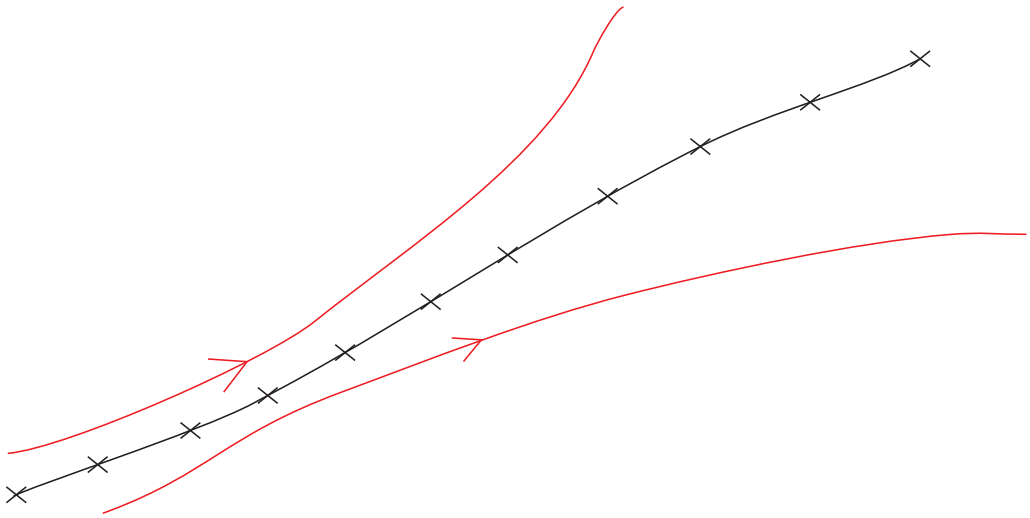
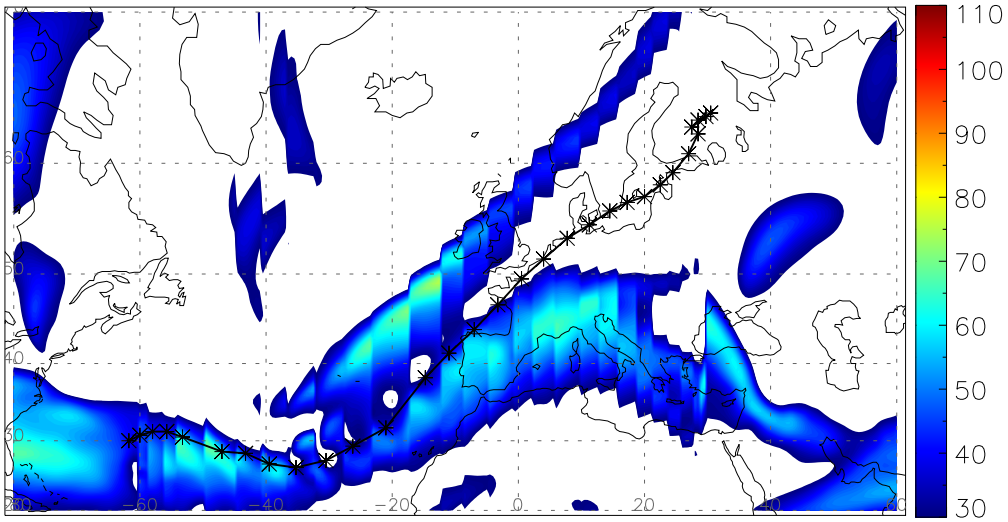


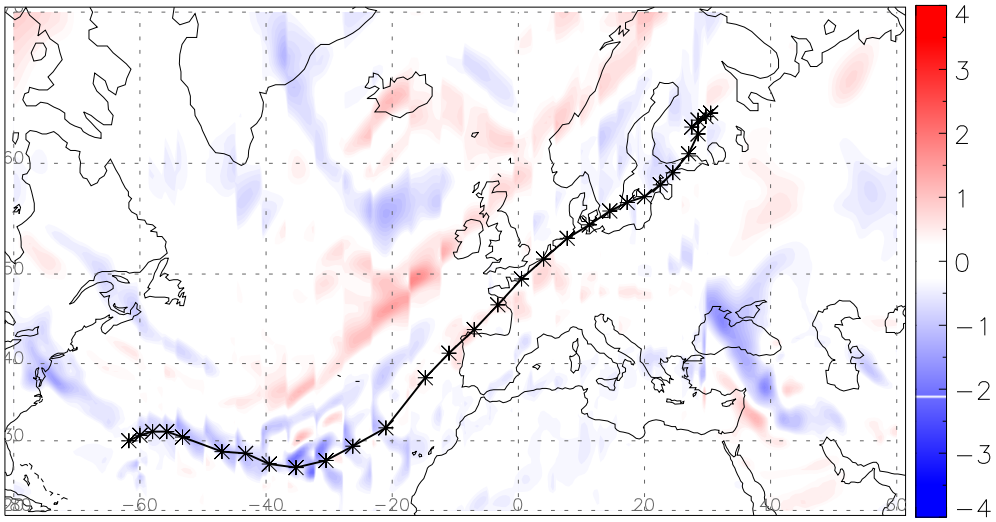
FIGURE 4.14: Schematic showing the typical track of a split-jet storm (black) relative to the jet stream (red).

An example of a split-jet system is storm Xynthia in Figures 4.15 and 4.16. As can be seen from Figure 4.15a, there are two jet streaks present during her lifetime, one centred around $47^{o}N20^{o}W$ and another at $43^{o}N10^{o}E$. Xynthia tracked along the northern edge of the southern jet stream, deepening despite convergence aloft. Xynthia reached minimum pressure off the Normandy coast (Figure 4.16b), in the region where the divergence from the left exit of the first jet was reinforced by divergence from the right entrance of the second jet (Figure 4.15b). After the minimum pressure, the two jets began to diverge and Xynthia moved between them while filling. Furthermore, very

high values of θ_e are found where Xynthia developed (Figure 4.16a), indicating large potential for diabatic processes to cause deepening.

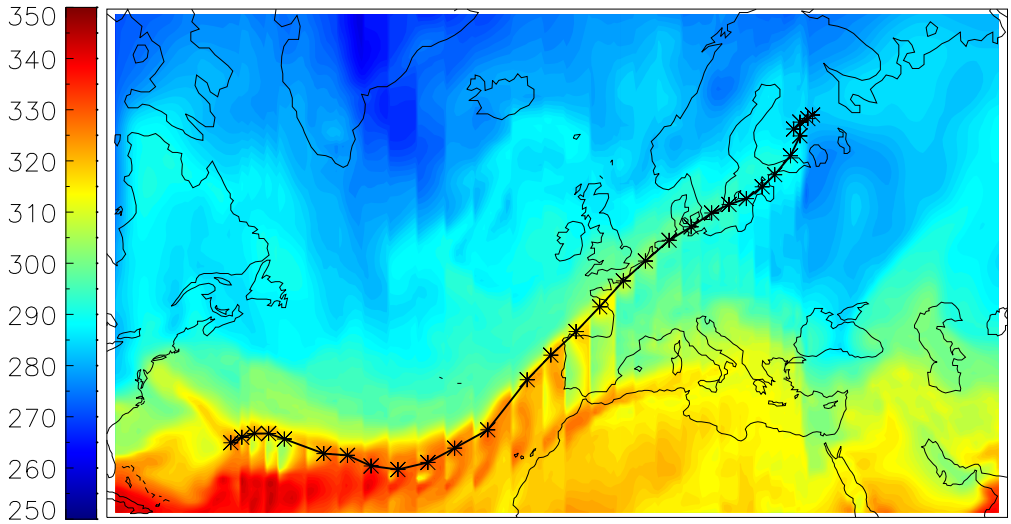


(a)

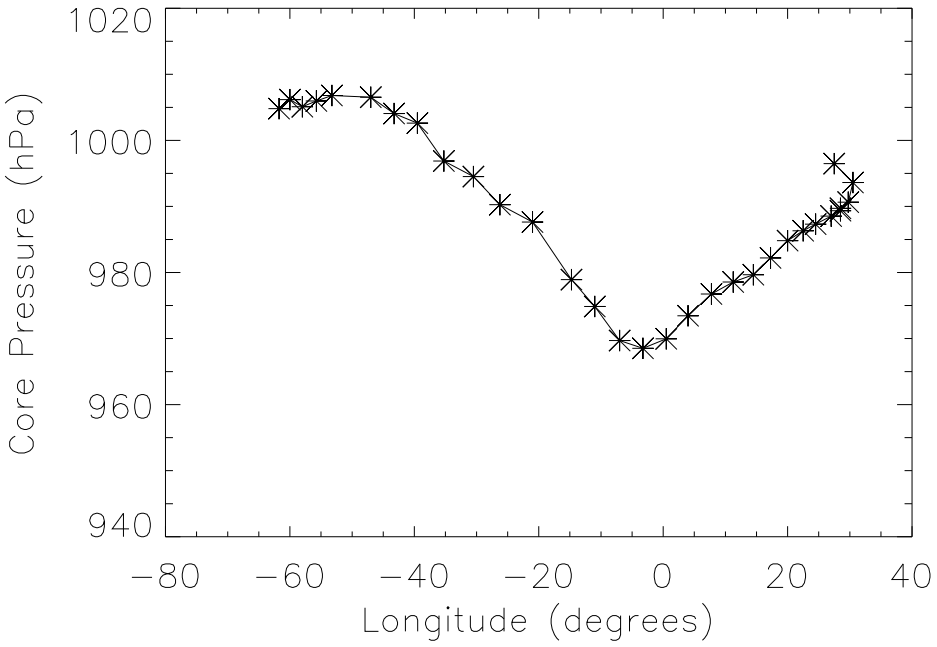


(b)

FIGURE 4.15: Storm Xynthia, an example of a 'split' jet storm from February 2010: (a) wind speed at 300hPa [ms^{-1}], (b) divergence of the wind field at 300hPa [$10^{-}4, s^{-1}$]. Plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4.



(a)



(b)

FIGURE 4.16: Storm Xynthia, an example of a 'split' jet storm from February 2010: (a) θ_e at 850hPa $[K]$, (b) cyclone core pressure $[hPa]$ against time $[h]$. (a) is plotted as meridional slices of the field, that move with the cyclone's track, as described in Section 3.4.

4.5 Pressure Tendency Equation

The pressure tendency is a method of identifying the processes that contribute to cyclogenesis and cyclolysis, and the approach is described in Section 3.4.2. While this method investigates the drivers of deepening in terms of core pressure, it should be remembered that there are smaller-scale features that can contribute to deepening or to the wind speeds (e.g. sting jets). Figure 4.17 shows example PTE bar plots, for each of the four jet stream categories. Although there is a great deal of variation between these plots, some aspects are consistent. For example, in all four cases, the evaporation-precipitation (EP) term works to deepen the cyclone, because storms are associated with precipitation, so mass is removed from the air column. The EP term is particularly large for storm Xynthia, indicating that the strongest precipitation in the area considered by the PTE calculation.

The deepening phase of all the storms is dominated by the virtual temperature tendency term. In all four cases shown in Figure 4.17, there is a point at which the dTemp term swaps from having a cyclogenetic to a cyclolytic effect. In the cases of Lothar and Daria, this occurs just after the jet crossing, which is as they move from the warm side to the cold side of the jet so likley to have an influence on the thermal PTE terms. When comparing the point at which this happens for these two storms, it becomes clear that Lothar crosses the jet stream significantly later than Daria. In the case of storm Agnes, the swap happens when the storm stops travelling alongside the jet stream and moves further from its influence. This is also the time when Agnes leaves behind the divergence aloft, but in this case the divergence is induced by the curvature of the jet stream. For storm Xynthia, the dTemp term has a strong deepening effect in the middle part of the track, which is consistent with divergence removing mass from the column. This is when Xynthia is in the region between the two jet streaks. The divergence from the left exit region of the first streak and that from the right entrance region of the second coincide to provide strong divergence aloft.

Figure 4.17 also shows that the strength of the $d\phi$ term varies between storms, but not such that categories can be discerned. This term describes the changes in geopotential that occur at the level of the lid of the integration that affect the surface pressure. In some cases, this lid term strongly supports the cyclone's deepening (e.g. Wiebke, Figure 4.18a), an idea explored previously by Colucci (2010); in others, it works the opposite way (e.g. Emma, Figure 4.18b), and storms develop despite stratospheric factors working against cyclogenesis. The drivers of this are investigated as part of a separate project (Wilbraham, 2013), which found that propagating troughs

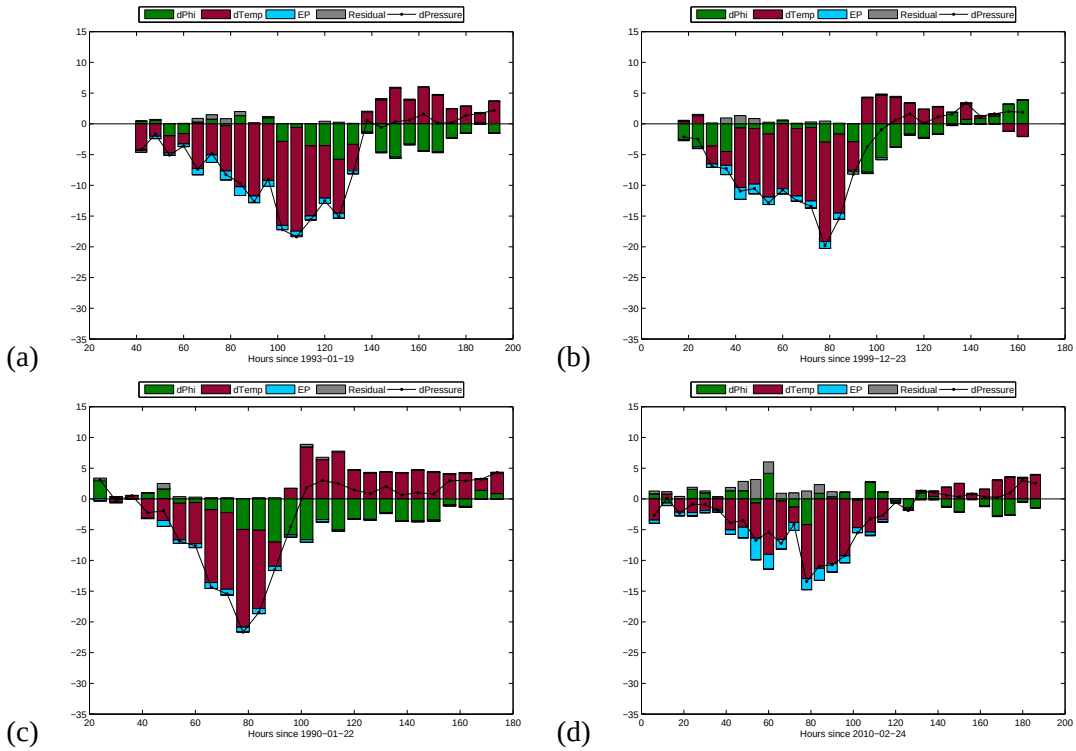


FIGURE 4.17: The components of the PTE for (a) Agnes [edge jet stream type], (b) Lothar [cross late], (c) Daria [cross early], and (d) Xynthia [split].

in the stratosphere can act to strengthen a surface cyclone. On the other hand, their absence does not necessarily suppress growth of surface cyclones, as cyclones like Emma develop despite the effects of the stratosphere. This indicates that not all intense storms develop in optimal conditions for deepening, but that what is needed for such a system to develop is a sufficient coincidence of factors – for example, very strong baroclinicty – and then the storm can overcome detrimental factors.

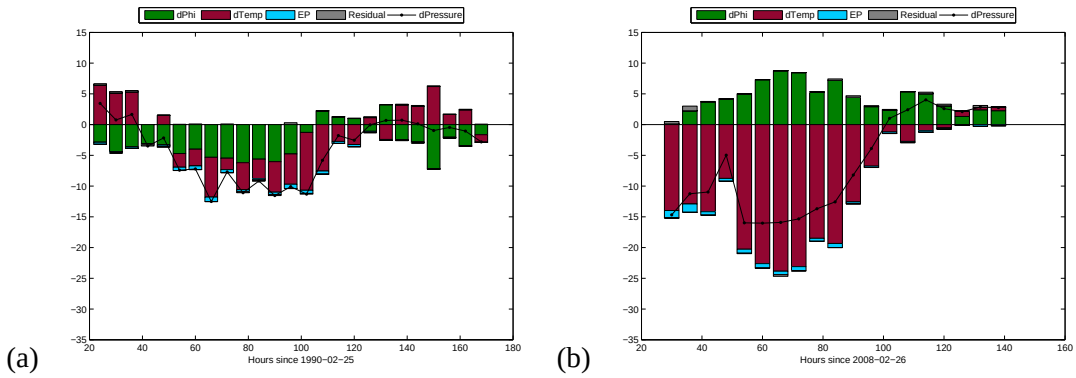


FIGURE 4.18: Components of the PTE, for storms (a) Wiebke (b) Emma.

When the temperature term is split up into its components, the expected pattern becomes clear: the horizontal and diabatic terms work to deepen the storm, whereas the vertical term works towards filling (Figure 4.19). Fink et al. (2012) suggest that the reason the vertical motion term is positive is due to ascending motion ahead of the cyclone centre. This would move warm air upwards, causing adiabatic cooling, an influx of mass into the air column, and overall working against the deepening process. Examining the four example storms in Figure 4.19, it becomes clear that the relative importance of the horizontal temperature advection and diabatic terms varies between storms. Notably, Xynthia has a much stronger contribution from diabatic processes than the other storms, which was to be expected given the high values of θ_e present in her vicinity (Figure 4.16a), indicating the presence of warm, moist air, which gives the potential for latent heat release. On the other hand, Agnes' deepening has a strong contribution from the horizontal temperature advection term, which could be due to her tracking along the cold side of the jet stream and pulling warm air unusually far north in the warm sector. However, when examining similar plots for all the storms (not shown), it becomes clear that the storms cannot easily be divided into categories. Furthermore, there is no clear relationship with the jet categories discussed in Section 4.4, because when examining the relative contribution of the horizontal and diabatic terms in these plots, there is almost as much variation within each category as between categories.

Since the Figures of the PTE terms (Figures 4.17, 4.19) do not yield any clear categories, a more quantitative approach is needed. The percentage contribution of each term to the cyclone's deepening is calculated and totalled over the course of the deepening (when the lowest pressure in the 3^o box considered is decreasing). The results of this calculation are shown in Table 4.3. The virtual temperature tendency term is divided into its separate components. Only the contribution towards deepening is considered for the percentage calculation. Therefore, if a process does not contribute towards deepening, then it appears as zero percent. One example of this is for the $d\phi$ term in the case of Emma, though it is working against deepening (Figure 4.18).

The two largest terms that contribute towards deepening are the horizontal temperature term (horiz), and the residual in the temperature term associated with diabatic processes (diabres). In some cases, there are considerable contributions from the $d\phi$ term (e.g. Lara), though contributions from the ep term are generally small. Table 4.3 shows the storms ranked by horiz. The storms are split into two categories, depending on whether horiz or diabres is the larger percentage contribution to a given cyclone's deepening. These are:

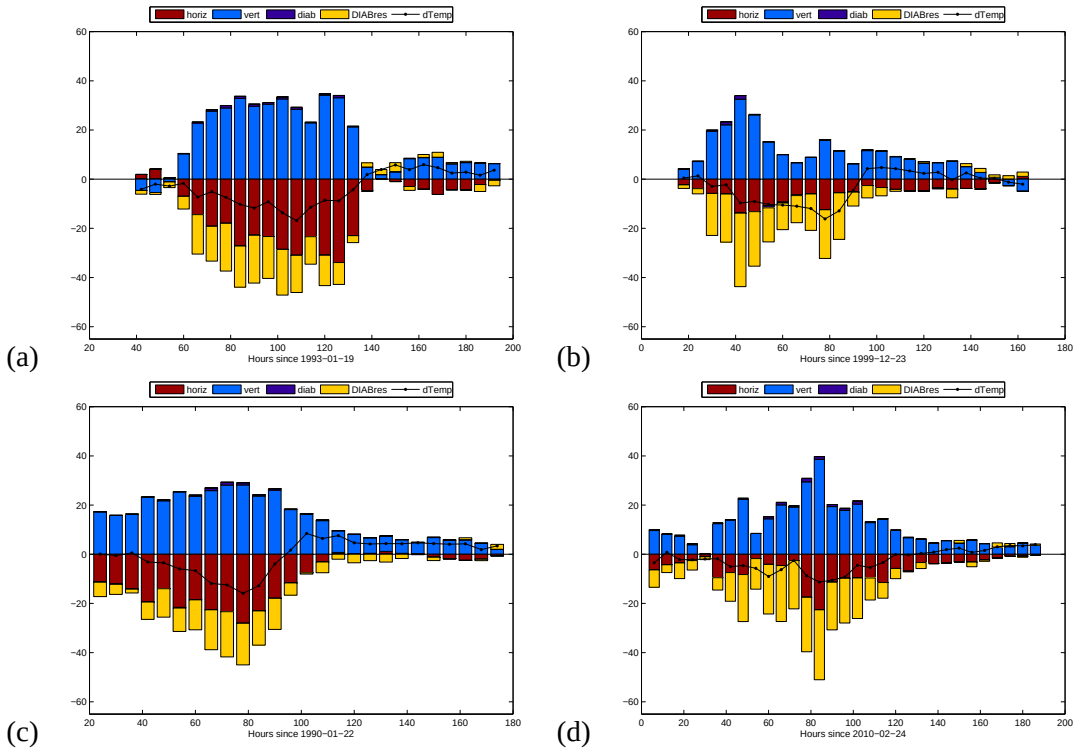


FIGURE 4.19: The components of the virtual temperature term in the PTE for (a) Agnes [edge jet stream type], (b) Lothar [cross late], (c) Daria [cross early], and (d) Xynthia [split].

- **Horiz-dominated:** Daria, Nana, Vivian, Wiebke, Udine, Verena, Agnes, Dec1993, Urania, Lara, Anatol, Franz, Martin, Kerstin, Rebekka, Elke, Lukas, Pawel, Jennifer, Frieda, Cyrus, Hanno, Kyrill, Emma.

- **Diabres-dominated:** Silke, Lothar, Jeanette, Gero, Klaus, Quinten, Xynthia.

Storm	$d\phi$	ep	horiz	diabres
Pawel	0.11	1.45	82.62	15.49
Frieda	2.81	1.70	74.28	21.11
Emma	0.00	1.41	73.80	24.72
Rebekka	7.02	1.59	73.32	17.99
Kerstin	7.21	1.20	73.09	18.41
Kyrill	2.42	1.74	66.61	29.07
Franz	1.81	2.32	66.32	29.52
Martin	0.69	1.96	63.92	33.31
Dec 93	5.15	2.06	59.01	33.76
Cyrus	8.80	1.67	58.56	30.83
Daria	7.51	1.83	56.57	33.95
Agnes	5.21	2.41	56.39	33.82
Anatol	8.79	2.11	56.02	32.88
Jennifer	5.60	2.55	55.69	35.95
Vivian	9.81	1.98	52.67	35.21
Elke	1.09	3.82	50.89	43.86
Lukas	12.16	2.94	49.22	35.64
Verena	4.29	2.62	48.12	44.84
Urania	10.09	2.82	46.39	40.38
Hanno	14.46	2.20	45.66	37.68
Silke	2.13	2.99	44.62	49.59
Lara	29.25	1.35	43.42	22.30
Wiebke	18.55	1.92	43.03	36.13
Klaus	2.46	3.01	42.83	51.29
Gero	7.68	2.46	41.89	47.60
Udine	7.90	2.52	41.26	47.60
Nana	25.77	1.93	40.85	30.93
Jeanette	8.22	3.57	39.78	45.67
Quinten	11.62	3.35	36.19	48.77
Xynthia	1.91	4.98	35.47	57.55
Lothar	11.14	3.62	29.98	54.76

TABLE 4.3: Storms with percentage contribution to deepening from selected terms of the PTE, ranked by horizontal temperature advection terms (horiz).

4.6 Comparing Categories

Two methods of categorisation are presented that assess the storms in two different ways. The first is based on the configuration of the jet stream and storm track (Section 4.4), and examines the synoptic situation at the time the storm developed and deepened. The other method uses the dominant processes that contribute towards deepening using PTE analysis (Section 4.5), which examines the processes in a volume around each storm's minimum pressure and how these affect the mass in the column and so the cyclone's core pressure. The categories of the storms are shown in Table 4.4.

First, the storms are examined for a link between storm intensity and the categories. For the jet-stream categories, there are no clear relationships with intensity (in terms of SSI). However, the six most intense storms are in one of the jet crossing categories; the edge and split-jet category storms tend to be less intense. Links between the two methods of categorisation are not immediately apparent from Table 4.4: both groups of PTE types contain members from each of the jet stream categories; all jet stream categories have members with both PTE types. From visual inspection, there is no clear relationship at all, partly because there are significantly fewer diab-type storms than horiz-type storms.

A statistical approach is now used to identify links between the categories. The Fisher exact test assesses the significance of the relationship between categories (Section 3.5.1). The null hypothesis is that there is no relationship between the categories of data. The Fisher exact test yields a p-value of 0.3116 and the alternative method involving Monte Carlo simulations converges on a similar value. This means that there is a 69% chance that the number of storms in one category affects the number of storms in another category, but this cannot be said to be certain (i.e. the null hypothesis cannot be rejected), because the p-value is larger than 0.05.

Although performing statistical tests is an important, quantitative approach to analysing data, visual inspection of Table 4.5 reveals some links:

1. **Horiz and Edge:** These storms track along the colder and drier side of the jet indicating that the potential for latent heat release is small. Therefore, baroclinicity is the predominant source of potential energy for a cyclone that generates sufficiently strong winds to cause a high value of SSI.

2. **Horiz and Cross Early:** In these storms, the deepening is strongly related to the crossing of the jet stream, and passing through the right entrance and left exit regions, so the divergence aloft removes mass from the column. This is consistent with having a large horizontal advection term, as it is calculated over the depth of the column. It is also likely that horizontal temperature contrasts are particularly strong, as the storm has advected air from the warm side of the jet stream to the cold side. Given that most of the deepening occurs on the cold side of the jet stream, a strong temperature contrast would strengthen baroclinicity and could contribute to the deepening of the cyclone.

Storm	SSI	Jet Category	PTE Category
Klaus	55.4657	Cross late	Diabres
Vivian	46.3109	Cross early	Horiz
Wiebke	46.3109	Cross late	Horiz
Kyrill	39.4672	Cross early	Horiz
Lothar	37.6781	Cross late	Diabres
Martin	37.1095	Cross late	Horiz
Emma	34.0888	Edge	Horiz
Jeanette	27.5068	Edge	Diabres
Daria	27.3295	Cross early	Horiz
Agnes	19.6948	Edge	Horiz
Anatol	18.5982	Cross early	Horiz
Udine	17.1550	Cross early	Horiz
Rebekka	16.2454	Cross early	Horiz
Lara	15.6771	Edge	Horiz
Xynthia	14.9836	Split	Diabres
Jennifer	14.6499	Cross late	Horiz
Gero	13.5304	Split	Diabres
Hanno	13.3756	Cross late	Horiz
Silke	12.4206	Cross early	Diabres
Elke	10.9337	Split	Horiz
Dec 1993	9.890	Split	Horiz
Urania	9.287	Cross late	Horiz
Nana	9.242	Edge	Horiz
Quinten	8.323	Cross late	Diabres
Verena	8.312	Cross late	Horiz
Kerstin	7.977	Edge	Horiz
Pawel	7.815	Edge	Horiz
Cyrus	7.801	Cross late	Horiz
Lukas	7.765	Edge	Horiz
Franz	7.699	Edge	Horiz
Frieda	7.635	Edge	Horiz

TABLE 4.4: The list of storms, described by the two methods of categorisation and ordered by intesity (SSI).

	Cross Early	Cross Late	Edge	Split	Total
Diabres	1	3	1	2	7
Horiz	6	7	9	2	24
Total	7	10	10	4	31

TABLE 4.5: Table showing counts of storms in each jet category (columns) and PTE type (rows).

	Cross Early	Cross Late	Edge	Split
Diabres	-0.58	0.74	-1.26	1.10
Horiz	0.58	-0.74	1.26	-1.10

TABLE 4.6: Table showing the differences between the observed and expected values for each cell in Table 4.5. Shading illustrates whether the value is positive (pink) or negative (blue).

Another way to see where the two different categorisation methods are related is to calculate the deviations from the value expected, given the row and column total (Section 3.5.2). The results of this are shown in Table 4.6. It confirms that there is a swing towards storms driven by horizontal temperature advection in both the cross early and particularly the edge categories, as discussed above. Table 4.6 also shows that there is a swing towards diabatic storms in the other two categories: cross late and split. In the case of cross late, this is consistent with a storm that spends a large proportion of its time on the southern side of the jet, where the air is warm and moist and therefore there is greater potential for specific and latent heat release. For split jet storms, the two jet speed maxima means that the dynamics of cyclogenesis are more complex. In the example of storm Xynthia, warm, moist, high θ_e air is brought up from the tropics (Figure 4.16a), which would facilitate diabatic processes such as latent heat release. However, with only four storms in this category, it is difficult to determine whether this is generally true of split-jet storms.

It remains to be seen whether these jet stream categories and PTE types are related to the degree of predictability that a storm exhibits (Chapter 5), or if they are related to any facets of the large-scale flow (Chapter 6).

4.7 Summary

First, the processes of selecting the storms is discussed. The best candidate for selection of stormy days is a parameter that can be calculated solely from meteorological variables, but still aimed at calculating the potential damage a storm could wreak. SSI was selected for this very reason, as it

measures how much greater the wind speed is than the 98^{th} percentile of the climatology at each grid point. This is then cubed, so it is related to the power of the wind and therefore is proxy for how much damage it can cause.

The SSI is calculated using ERA-Interim data, allowing values every six hours since 1st January 1990 to be compared. The next step is to choose the top 120 values and remove the duplicates. Other days are rejected for one of four reasons: because Mediterranean storms, polar lows, high pressure systems or orographic factors had driven the high values of SSI. After this, 31 North Atlantic extratropical cyclones remain.

Within these 31 storms, four different categories of jet stream configuration can be discerned: edge, cross early, cross late and split. These are determined by the relative position of the storms and the jet stream, and particularly the point when the storms' tracks crosses the jet stream. A second method of categorising the storms identifies the terms in the Pressure Tendency Equation (PTE) that dominate the deepening of each storm. However, the initial PTE analysis shows nearly as much variation in the processes that dominate deepening within each jet category as between categories. The virtual temperature term tends to dominate the deepening. When this is divided into its components, the horizontal temperature advection term and diabatic term tend to work to strengthen the cyclone, whereas the vertical motion term works to weaken it. This led naturally to two types of storms; those where the deepening is dominated by horizontal temperature advection, and those where it is dominated by diabatic processes. The percentage contribution of these two terms of the PTE is used to divide the storms in two, which allows some links became apparent between the jet stream categories and the PTE types. The strongest overlap occurs in storms where horizontal temperature advection dominates the PTE and the jet type is either edge or cross early. There is also a link between storms where diabatic processes dominate the deepening, and the jet type is either split or cross late. The next chapter will assess whether either of these methods for dividing the storms is related to their forecast quality.

Chapter 5

Storms in Numerical Weather Prediction Models

This Chapter aims to assess the predictability of the storms selected in Chapter 4, in terms of numerical weather prediction (NWP) forecast quality and spread. As discussed in Section 3.1.2, two types of forecast will be used here: deterministic (Section 5.3) and ensemble (Section 5.5). Deterministic forecasts are a single forecast, run from the operational analysis using the highest available resolution for the computing power of a single model run. Ensemble forecasts are where the forecast model is run many times, with slight perturbations to the initial conditions. Therefore, ensemble forecasts are conducted at a lower resolution, to make the most of the available computing power. Deterministic forecasts indicate the quality of the forecast, whereas ensemble forecasts indicate the spread.

5.1 Opening Remarks

Both the intensity and position of the storm will be evaluated, in terms of forecast quality and spread. The metric for intensity used is core pressure. The metrics for position are latitude and longitude. The reasons for these choices are that these measures are easy to understand, straightforward to analyse, and are direct output from the storm tracker (see Section 3.6 for further discussion). When considering all of the selected storms, it is expected that shorter lead times will have improved forecast quality and smaller forecast spread.

The measure of intensity used to select the storms is SSI (Storm Severity Index, Section 3.2). It is possible to hypothesise that the most intense storms are the least well forecast, because they are more unusual. On the other hand, intense storms are the result of strong forcing, so may be forecast better. Therefore, this work searches for correlation between storm intensity, measured by SSI, and the quality of the forecast, as measured in pressure, latitude and longitude. Also in Chapter 4, the storms were divided into different categories, in two ways (Table 4.4). This chapter will consider the forecast quality and spread of the different categories of storm, where possible. To briefly review, the first categorisation method considers four different types of jet stream:

- **Edge:** storms that move along the edge of the jet stream, and do not cross it.
- **Cross late:** storms that cross the jet late in their lifetime.
- **Cross early:** storms that cross the jet early in their lifetime.
- **Split:** storms associated with two jet maxima.

Storms accelerate and deepen strongly during interaction with the jet stream, so the quality of the forecasts, in terms of both intensity and position, is expected to increase after jet crossing has taken place.

The second method of categorisation considers the processes that contributed most strongly to the deepening phase, according to analysis by Fink et al. (2012) that used the pressure tendency equation (PTE). The storms are grouped by which of two terms dominated the deepening, namely:

- **Horiz:** storms where horizontal temperature advection is the most important process that contributes towards deepening.
- **Diab:** storms where diabatic processes are most important.

The horizontal temperature gradient is a large-scale forcing, that can be explicitly represented by the model. It is conceptually linked to baroclinicity, because a horizontal temperature gradient that is not parallel to the pressure gradient (i.e. a baroclinic atmosphere) is needed for horizontal temperature advection to occur. Diabatic processes are parametrised, and so are not explicitly represented. Furthermore, because diabatic processes are small-scale and non-linear, they are difficult to observe directly in order to further our understanding of them, so diabatic processes are a larger source of uncertainty in cyclogenesis than the horizontal temperature gradient. This

work speculates that diab-type storms will have lower quality and greater spread in their forecasts, compared to horiz-type storms.

Another factor that could affect forecast quality and spread is the native resolution of the forecast model. As discussed in Section 2.5.3, there are systematic limitations to the simulation of mid-latitude cyclones caused by insufficient model resolution. Whether there are limitations on the timescales typical of NWP models will be explored here, in terms of both forecast quality and spread. It is anticipated that lower resolution models will give a lower quality forecast, because fewer processes are considered explicitly and more need to be parametrised. Furthermore, the very strong horizontal temperature gradients and baroclinicity that are important for explosive cyclogenesis (Section 2.2.4.1) will be represented better at higher model resolution. In terms of forecast spread, higher-resolution models normally demonstrate better spread (Buizza et al., 2003). In addition, improving the resolution often coincides with other improvements to the model, such as better parametrisation schemes, which could also reduce uncertainty in the forecasts.

5.2 Review of Methods

The results presented in this Chapter mainly consider differences in core pressure, latitude and longitude against forecast lead time. The methods are discussed in Section 3.6, but are reviewed briefly here. First, the forecast and analysis tracks are matched automatically, as discussed in Section 3.6.2. This allows calculation of the differences between analysis and forecast values for each of the three variables, at the time of minimum pressure, for each storm. Therefore, positive values indicate that the value is larger in the analysis, and negative values that the value is larger in the forecast. The lead time for each of the forecasts is calculated, based on when the forecast was initialised relative to the time of each storm's minimum pressure. The results are presented for a number of different forecasts with different lead times (note that lead time is negative, so that time progresses along the x-axis from left to right). The plots for the ensemble forecast data also present differences between analysis and forecast for each lead time, but are box-and-whisker plots because there are 50 ensemble members at each lead time.

Linear regression is performed between forecast lead time and the pressure, latitude and longitude deviations, to ascertain the quality of the forecast. This will give an idea of how well predicted the storms are, and what the limitations of the forecast are. Residuals are calculated, in order to determine whether linear regression is a valid technique to perform on such data. Two tests

of correlation will be used as appropriate: Pearson's correlation coefficient and Spearman's rank correlation coefficient (Section 3.6.4). The significance of the correlation will also be determined. This will facilitate exploration of the relationship between the jet stream category or PTE type of each storm and forecast quality. Similar simple tests to explore the relationship with forecast spread cannot be identified, so this discussion will centre around the box-and-whisker plots.

The linear regression is also used to calculate a standard measure of forecast quality for all of the storms. This measure is the pressure, latitude and longitude deviations for a fixed lead time of 24 hours, and the values are referred to as $\Delta(p)$, $\Delta(lat)$ and $\Delta(lon)$, respectively. Linear regression is needed because, for a given storm, the lead times available depend on the time of minimum pressure (Section 3.6.4). The forecast is only initialised once per day, at 12:00. Since lead time is calculated from the time of minimum pressure, the available lead times depend on when the minimum pressure occurs. For example, if the minimum pressure were at 12:00, then lead times of 24h, 48h, 72h etc. would be available; but, if the minimum pressure were at 00:00, the lead times would be 12h, 36h, 60h, etc. Using linear regression on all storms to derive this simple measure of forecast quality means the measures are comparable, and that all of the storms are treated consistently. These simple measures are used to ascertain whether more intense storms are better forecast, and will be used in the work on storm-prone situations (Chapter 6).

To summarise, this Chapter presents an assessment of forecast quality and spread, for the selected storms. This will be measured by differences between the analysis and forecast tracks, at the time of minimum pressure, in pressure, latitude and longitude. These differences will be analysed to determine whether there is a relationship between forecast quality or spread, and the intensity, jet stream category or PTE type of each storm. This analysis will be performed on deterministic (Section 5.3) forecasts, in order to quantify the forecast quality. The relationship between the deviations at 24h lead time is then discussed (Section 5.4), to assess forecast quality in a standard way across all storms. The results from ensemble forecasts (Section 5.5) give insight into forecast spread, by considering the ensemble members of different storms' forecasts. Finally, these results are summarised in Section 5.6.

5.3 Deterministic Forecasts

This section will consider the results from the ECMWF deterministic operational forecasts and the ECMWF ensemble control forecast. These are made using the same input data, but are at different

resolutions. The ensemble control is always at the lower resolution so that it is comparable with the ensemble members. Therefore, when the two data sets are compared, the major source of differences will be effects of resolution. However, these data are also processed by the storm tracker (Section 3.3), and so differences should be small because the tracker interpolates the data so they have the same resolution. Another key difference between the operational and control forecasts is the time period they cover, and therefore the number of storms included in their analysis. Operational forecasts were made for the entire period of interest, and so all 31 storms are covered. However, the ensemble forecasts only began in December 1992, and so the control forecast includes 19 of the selected storms. It should also be noted that each point plotted in the figures refers to a match between the forecast and analysis tracks for a storm, at the point of minimum pressure on the analysis track, where the match meets the criteria discussed in Section 3.6.2. Therefore, some forecasts will be rejected because the forecast does not meet the criteria. Furthermore, the forecast model is only run once a day, and so each storm can only be included once in a 24 hour period. Where within the lead times a storm appears depends on the time of its minimum pressure: if it is at 1200, then the lead times that result are 24h, 48h, 72h, etc; but if it is at 0000 then the lead times are 12h, 36h, 60h, etc. For these two reasons, there are fewer than 31 points at each lead time.

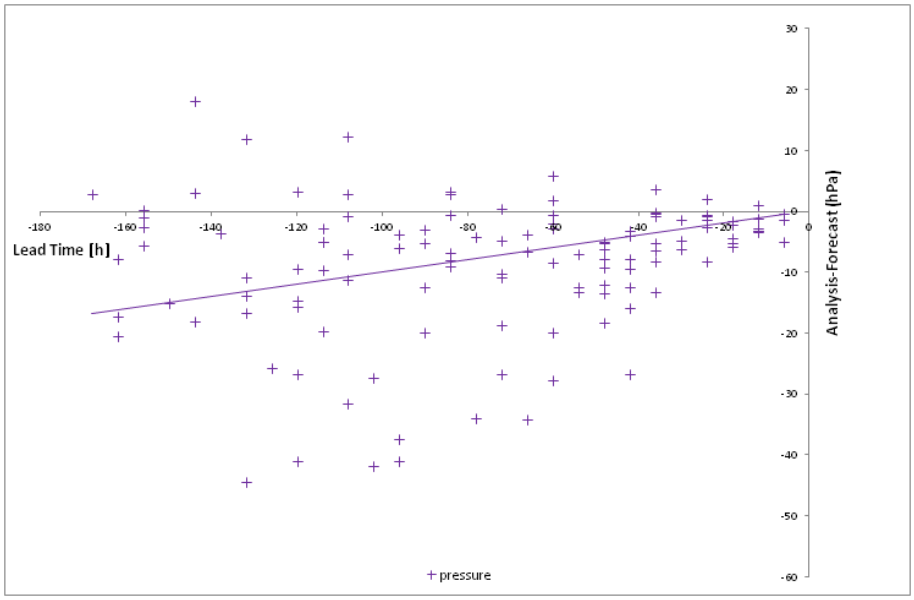
A discussion of the results when considering all storms will now ensue (Section 5.3.1), including a discussion of the validity of the regression line used (Section 5.3.2). Then the analysis will divide the storms by jet stream type (Section 5.3.3), PTE category (Section 5.3.4), and the native resolution of the model (Section 5.3.5).

5.3.1 All Storms

First, all of the storms are analysed in terms of intensity (Figure 5.1) and position (Figure 5.2). Figure 5.1 clearly illustrates that core pressure in the forecast generally has a higher value than core pressure in the analysis. More intense storms have lower pressures, and so the storms are less intense in the forecast than in the analysis. Therefore, intensity is underforecast. This is in agreement with previous work (e.g. Froude et al., 2007a). Figure 5.1 shows a relationship between the deviations and the lead time, so core pressure is a useful way of measuring forecast quality when it comes to intensity. Table 5.1 shows that there is a positive correlation between the pressure deviations and forecast lead time, because as lead time becomes more negative, so

does the deviation. This relationship is statistically significant in both the operational and control forecasts.

(a)



(b)

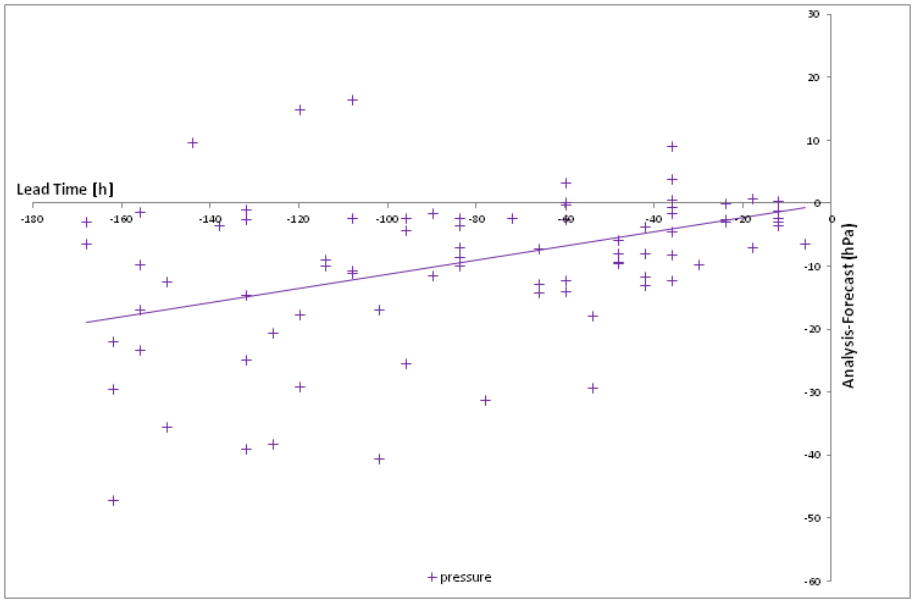


FIGURE 5.1: Pressure (analysis - forecast) against lead time for the selected storms, for (a) operational and (b) control forecasts.

The deviations for latitude and longitude are less clear-cut (Figure 5.2). For latitude, there is an indication that the storms have lower latitudes in the forecast than the analysis and so are forecast slightly too far south. However, the small values of R indicate weak correlation only, which is only significant in the control forecast. This is unexpected, because the control forecast is at a lower resolution and would therefore be expected to produce a lower quality forecast. However, lower

Variable	Operational Forecast			Control Forecast		
	R	t	Significant?	R	t	Significant?
Pressure	0.198	2.19	✓	0.380	3.68	✓
Latitude	-0.112	1.23	×	-0.230	2.11	✓
Longitude	-0.387	4.56	✓	-0.252	2.33	✓

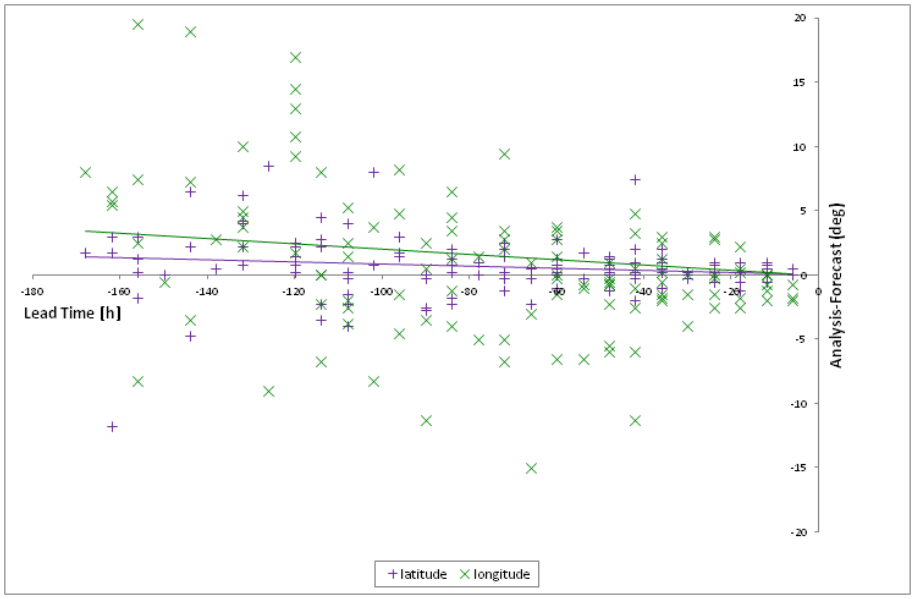
TABLE 5.1: For all storms, correlation coefficients (R) for differences against lead time, and tests for its statistical significance (t), both to 3 significant figures. As discussed in Section 3.6.4, t is compared to a reference table to test significance, based on the number of points included in the correlation (Operational = 120, Control = 82).

resolution also means some of the finer-scale features are missed, and also reduces the potential for the occurrence of large deviations because larger populations lead to larger deviations (Rees, 2001). Overall, storms tend to be forecast too far south, but whether latitude forecast quality is significantly related to lead time remains an open question.

For longitude, the correlation between the deviations and lead time is statistically significant in both the operational and control forecasts. In both the control and operational forecasts, the longitude in the analysis is larger than in the forecast, meaning that in the forecast the storm is too far west. This indicates that the storm is moving too slowly, because the difference between the analysis and forecast is positive, meaning the forecast storm is west of the analysis storm. Given that North Atlantic cyclones generally move from west to east, the speed at which the forecast storm moves from west to east is less than that for the analysis storm.

To summarise, these initial results indicate that it is true that forecast quality for these three metrics improves with shorter lead times, because the differences between analysis and forecast become smaller. Intensity is significantly underforecast, and storms are too far south and too slow in the forecasts. However, these results assume than linear regression is a valid approach for analysing these results. The next section tests this assumption.

(a)



(b)

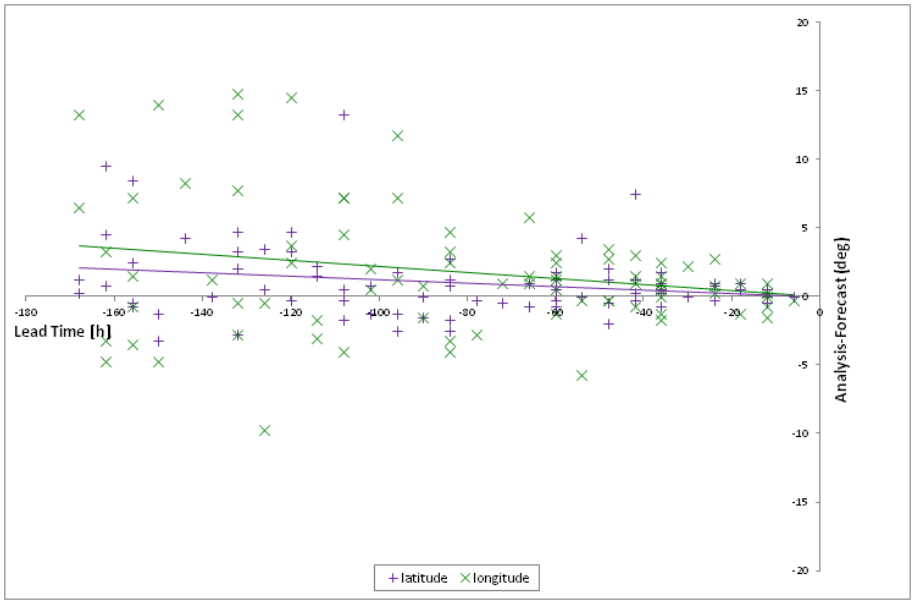


FIGURE 5.2: Latitude (purple) and longitude (green) differences (analysis - forecast) against lead time for the selected storms, for (a) operational and (b) control forecasts.

5.3.2 Testing the Linear Regression

When performing regression, it is important to analyse the residuals, to ensure that the correct type of regression is used. While linear regression is used here, it is possible to fit curves to data. Examples include a polynomial or exponential fit. However, at the outset there was no indication that curves would fit to the data any better than a straight line, and so linear regression was used. Other types of regression were tested, but did not show better results than linear regression (not shown). However, the regression was constrained to pass through the origin, because the forecast error at zero lead time should be nil. To test the linear regression further, residual analysis was performed. The residuals are the differences between each point and the line of best fit. As such, they should sum up to nearly zero, taking into account any rounding errors. This can be seen in Table 5.2.

The residuals can also be used to perform an analysis of the variance between the line of regression and the observed values. This is done using the F-ratio (Rees, 2001, Section 15.6), also known as the 'sum of squares due to a lack of fit'. This is analysed similarly to the t-tests already discussed, in that it is compared to a tabulated value (Rees, 2001, Table C.6) to determine whether the value is significant, at the 5% level. In all cases in Table 5.2, it is significant, meaning that the gradient of the line of best fit through the data points is significantly different to zero. Therefore, lead time explains a significant fraction of the variance in forecast quality of longitude, latitude and pressure.

Control Forecasts			
	Pressure	Latitude	Longitude
Sum of residuals	2.93×10^{-14}	3.59×10^{-14}	-8.88×10^{-16}
F ratio	887	21.5	22.6
Significant?	✓	✓	✓

Operational Forecasts			
	Pressure	Latitude	Longitude
Sum of residuals	4.17×10^{-14}	-1.44×10^{-14}	-1.64×10^{-14}
F ratio	211	10.8	15.0
Significant?	✓	✓	✓

TABLE 5.2: Analysis of the variance for the linear regression in the two types of deterministic forecast. Numbers are quoted to 3 significant figures.

Figure 5.3 shows the plots of the residuals, for both operational and control forecasts, and for pressure, latitude and longitude. These plots generally show a random distribution either side of the zero line, so there is no systematic error in the linear regression. Any such error would indicate

that a curve should be fitted to the data, rather than a straight line. The operational longitude (Figure 5.3e) is mostly positive for short lead times, which could indicate that a different type of line would better fit the data. However, the rest of the data shows a random distribution either side of the x-axis, so any improvement would be small. Overall, Figure 5.3 shows that linear regression is a valid method of analysing the differences between analysis and forecast pressure, latitude and longitude. Therefore, its use will be continued throughout this Chapter.

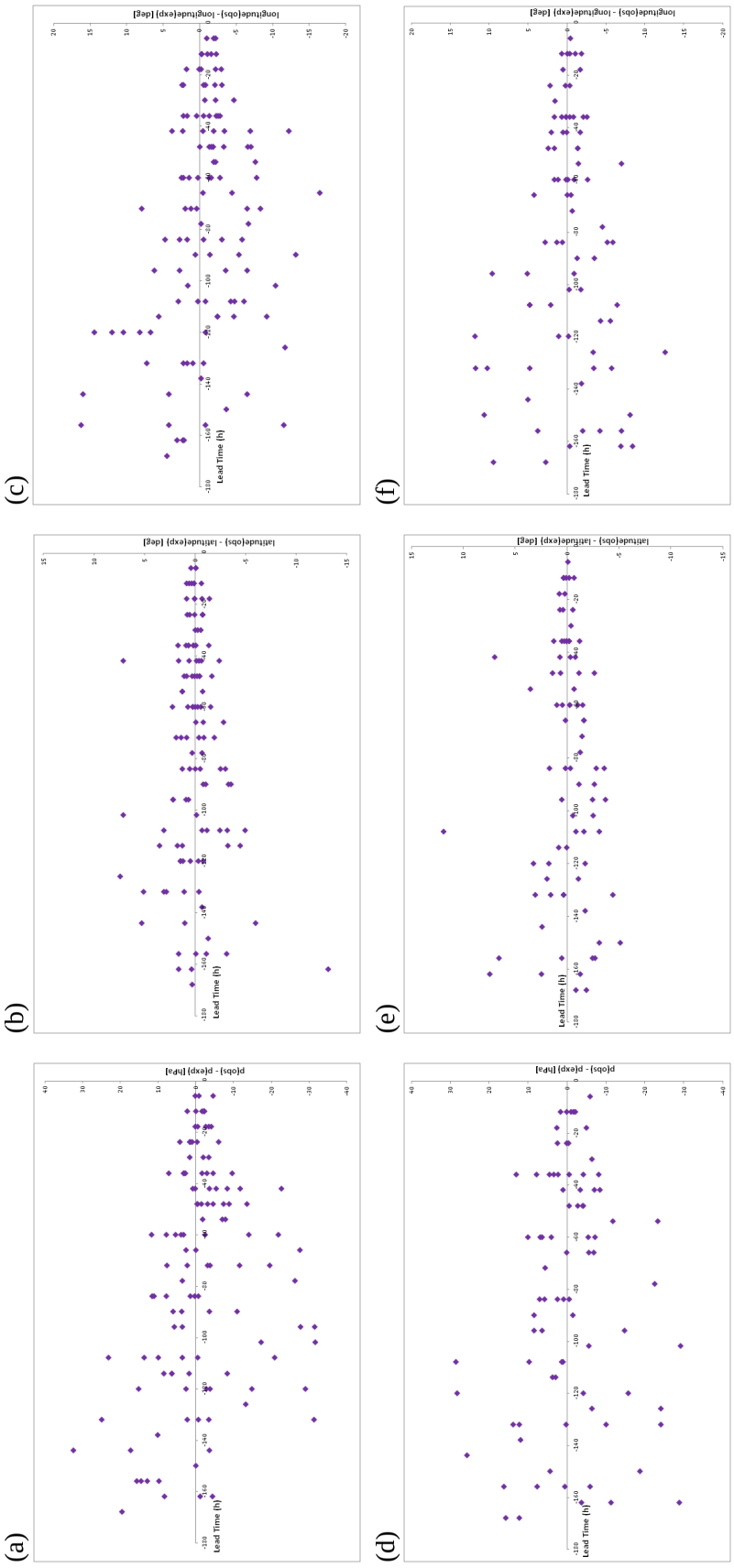
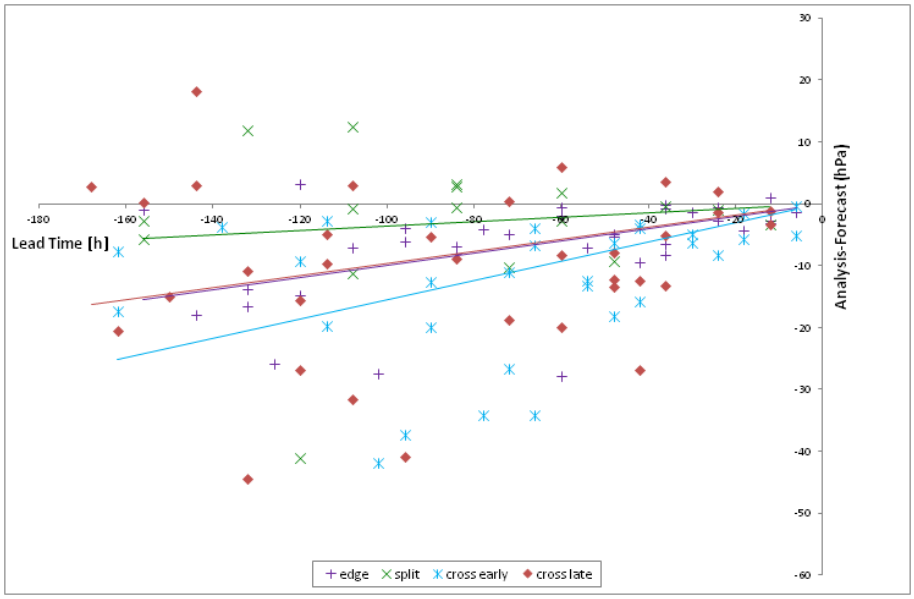


FIGURE 5.3: Residuals (observed - regression) for: (a) pressure in operational forecasts; (b) latitude in operational forecasts; (c) longitude in operational forecasts; (d) pressure in control forecasts; (e) latitude in control forecasts; (f) longitude in control forecasts. Please note the different axis scales for each variable.

5.3.3 Jet Stream Type

The next step is to ascertain whether forecast quality is affected by the configuration of the jet stream and storm track. As mentioned earlier, there are four types of this: edge, split, cross early and cross late. The forecast quality of each of these will now be discussed, in terms of pressure, latitude and longitude, in both the operational and control forecasts.

(a)



(b)

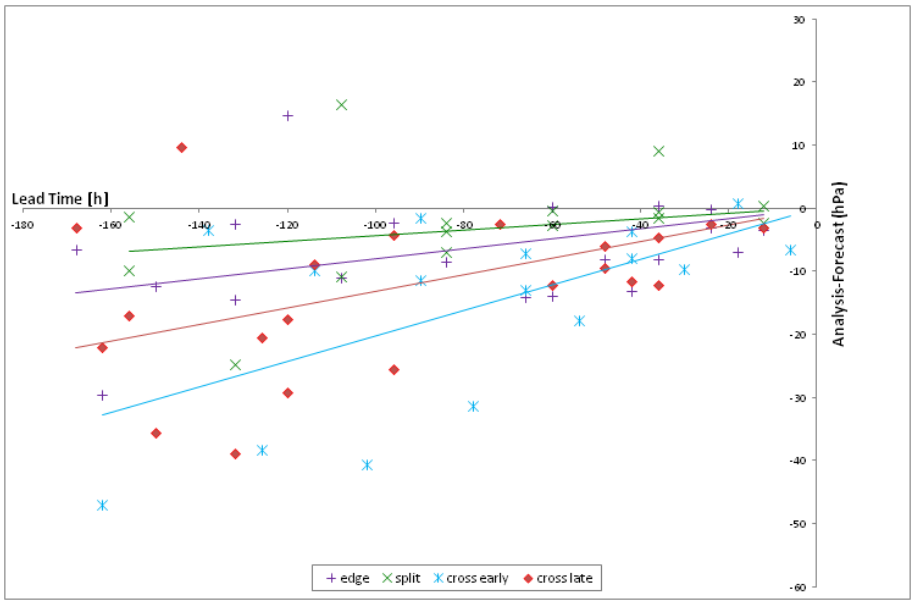


FIGURE 5.4: Pressure (analysis - forecast) against lead time for the selected storms divided by jet stream type (edge = purple +, split = green ×, cross early = blue ∗, cross late = red ♦), for (a) operational and (b) control forecasts.

Figure 5.4 shows that pressure is underforecast for all categories of storm, so all categories of storm are insufficiently intense, compared to reanalysis data. In both operational and control forecasts, the deviations are smallest for the split jet category. That the split-jet storms are relatively well forecast is unexpected, because the category of storms is small (four members). This could mean that the few that are used in this study were well forecast compared to the whole population of storms for a reason other than the jet stream, and so further analysis is required before it can be determined whether split-jet storms are generally well forecast.

The deviations in pressure shown in Figure 5.4 are largest for the cross early category, particularly at longer lead times, in both operational and control forecasts. However, when compared to the cross late category at short lead times (less than 48 hours), the range of values when comparing forecast to analysis is similar. This could mean that, once the storm has crossed the jet stream, the forecast improves compared to before jet crossing. This is reinforced by a comparison with the relatively well-forecast edge-type storms, because they do not cross the jet stream. When the cross-late storms cross the jet stream, their forecasts decrease in quality compared to the storms that do not cross the jet (edge type), and those storms that have already crossed the jet (cross early type). However, edge storms still interact with the jet stream and regions of convergence and divergence aloft, which will still affect the quality of the forecast. This is consistent with the model struggling to realistically represent the interaction of the jet stream and the storm at the time of jet crossing, because this is the point at which storms deepen most rapidly. This agrees with the hypothesis discussed earlier (Section 5.1). Further testing of this hypothesis could involve looking at case studies and their interaction with the jet stream; however, this is left to future work.

Figure 5.5 illustrates that there are no clear trends in latitude forecast quality between the different storm categories; one category does not have smaller deviations than another in both operational and control forecast. Broadly, all categories of storm are forecast too far north. In longitude (Figure 5.6), there is also no clear pattern between the different categories of storm. Although the control forecast shows that the deviations for edge-type storms are particularly large (Figure 5.6b), this is not seen in the operational forecast (Figure 5.6a). In terms of the speed at which the forecast storms move relative to the analysis storm, most types of storm are forecast more slowly than the analysis.

To summarise, the jet interaction is an important point in a storm's lifetime. The forecast quality could be related to when jet crossing occurs, particularly in terms of core pressure, but also in terms of latitude and longitude. This is manifest in the differences between the two jet-crossing

(a)

(b)

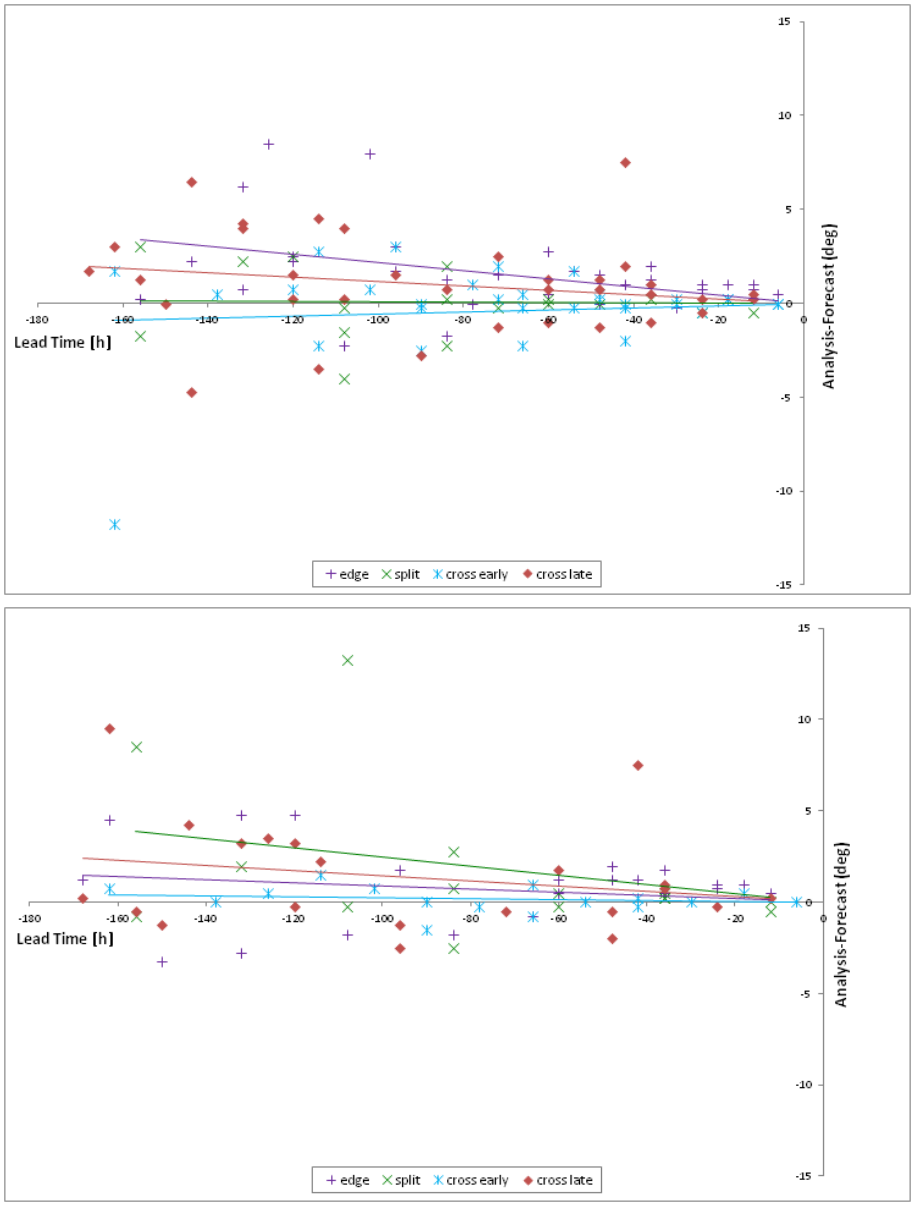
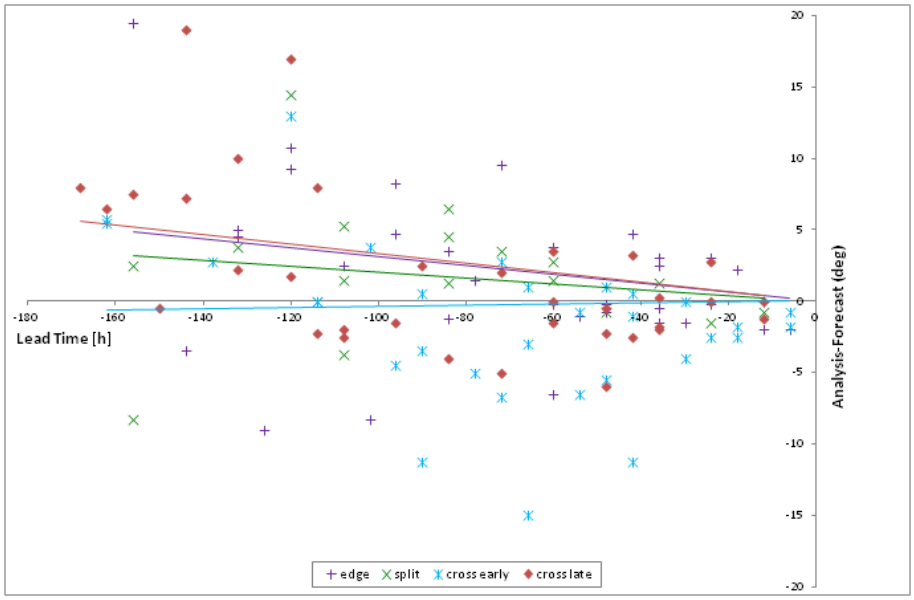


FIGURE 5.5: Latitude (analysis - forecast) against lead time for the selected storms divided by jet stream type (edge = purple +, split = green ×, cross early = blue ∗, cross late = red ♦), for (a) operational and (b) control forecasts.

categories, as one crosses the jet earlier than the other. This is because the storm is deepening most rapidly at this point, and so the uncertainty in the resulting forecast at the time of the storm's minimum pressure is high.

Forecast quality could be improved by improving the model's representation of this in one of two ways. Firstly, improving the model simulation of the interaction between the storm and the jet stream would reduce model errors. Secondly, it could be that errors in the initial conditions

(a)



(b)

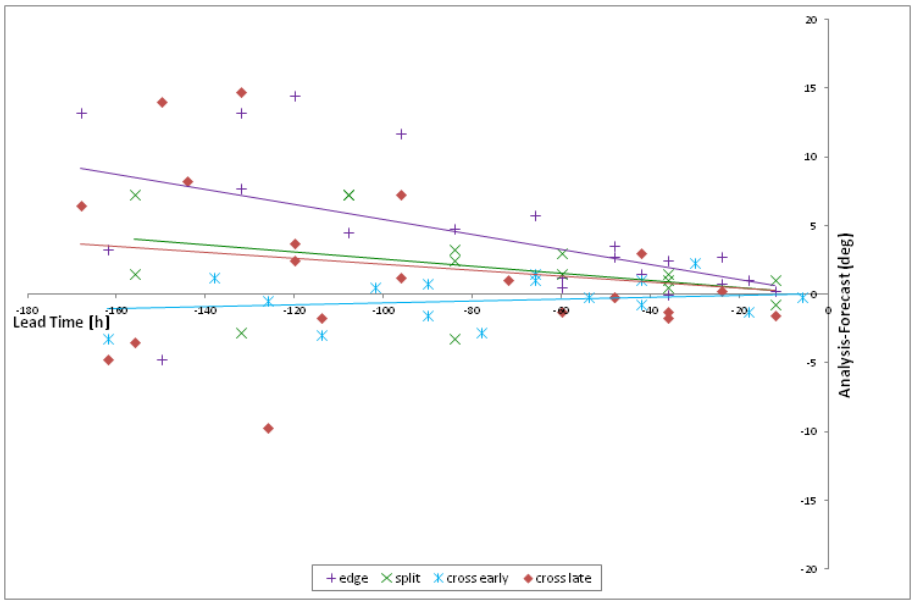


FIGURE 5.6: Longitude (analysis - forecast) against lead time for the selected storms divided by jet stream type (edge = purple +, split = green ×, cross early = blue ∗, cross late = red ♦), for (a) operational and (b) control forecasts.

propagate more quickly during the interaction because the deepening is so rapid, meaning that improving the data assimilation scheme or increasing the number of observations included would improve the forecast.

5.3.4 PTE Category

Next, the storms are analysed for their forecast quality based on which term was dominant in the PTE analysis; either 'horiz' or 'diab'. Figure 5.7 illustrates that both horiz- and diab-type storms are underforecast, as expected from the results with the jet stream categories. Furthermore, in both operational and control forecasts, horiz storms are less well forecast than diab storms, because the linear regression shows a larger difference between the analysis and the forecast for all lead times. This is unexpected, because small-scale diabatic processes are less well understood than large-scale baroclinicity and baroclinic processes, due to the latter being more easy to observe in the atmosphere than the former. For this reason, diabatic processes are the subject of current research as part of the DIAMET project.

Again, the deficiencies shown in forecasting horiz storms could be a problem of model physics or initial conditions. This could demonstrate that horiz storms are less well forecast due to a deficiency in how the model simulates the large-scale horizontal temperature gradient needed for horizontal temperature advection, or in how it simulates the conversion of this large-scale potential energy into eddy kinetic energy. Many studies have explored this relationship (e.g. Emanuel et al., 1987), so understanding of them should be high; particularly higher than for diabatic processes, which are harder to observe. Alternatively, the errors in the initial conditions could propagate more rapidly in horiz-type storms than diab, due to a misrepresentation of the temperature field over the Atlantic, where the storms develop. Although there are few observations in the Atlantic, temperature is better represented than humidity. Deficiencies in the humidity field would have more of an effect on diabatic-type storms. Therefore, this thesis speculates that the reasons are predominantly due to model error than initial condition error; however, future work is required to determine this. One caveat is that diabatic processes are parametrised in both ERA-Interim (the reanalysis used as a yardstick here) and the ECMWF forecasts with similar schemes. While raw observations are assimilated into a reanalysis, they are interpolated using an NWP model, and so a comparison of reanalysis and forecast data may yield a falsely favourable result. A comparison with observations would explore this; however, this would not be practical in this work, because it is necessary that the data are gridded to track the storms and to make the comparison.

(a)

(b)

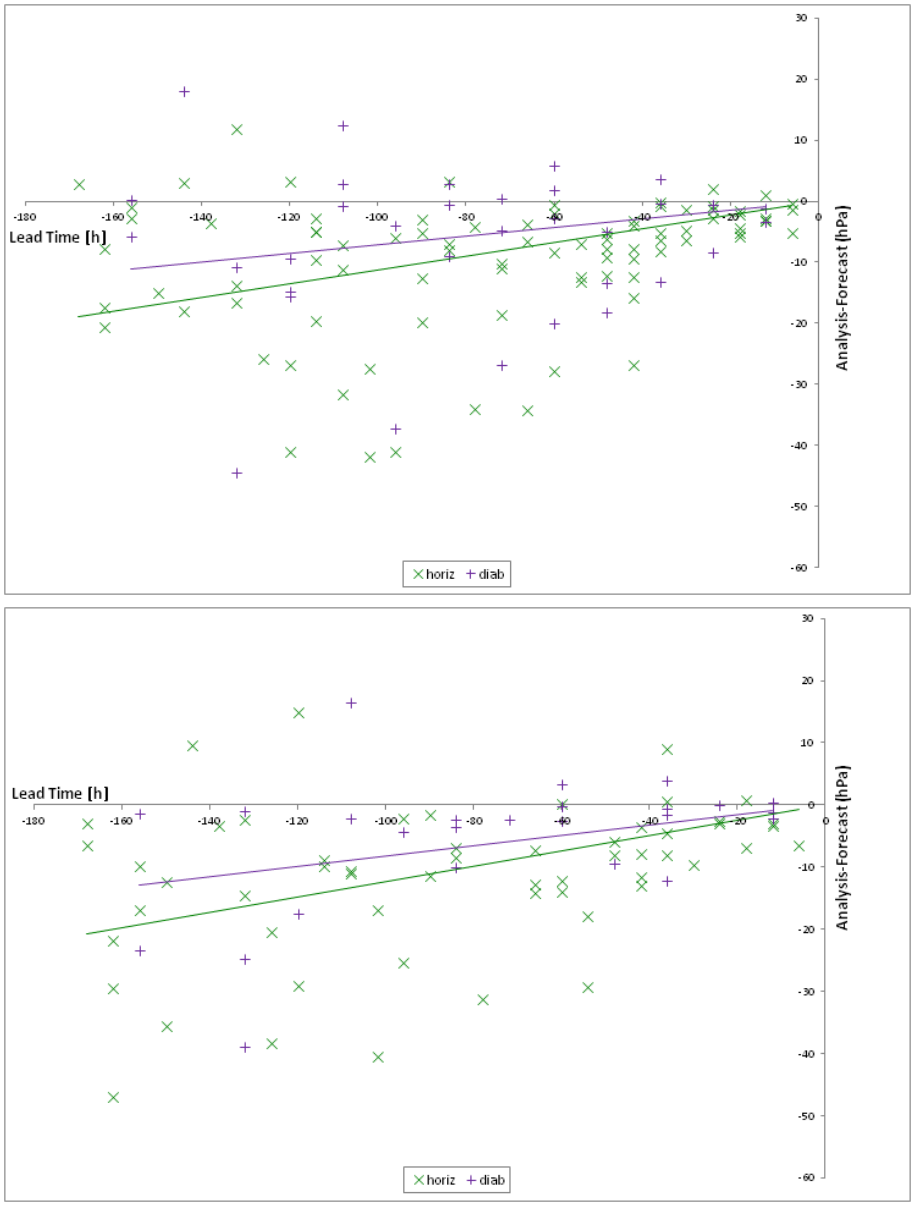
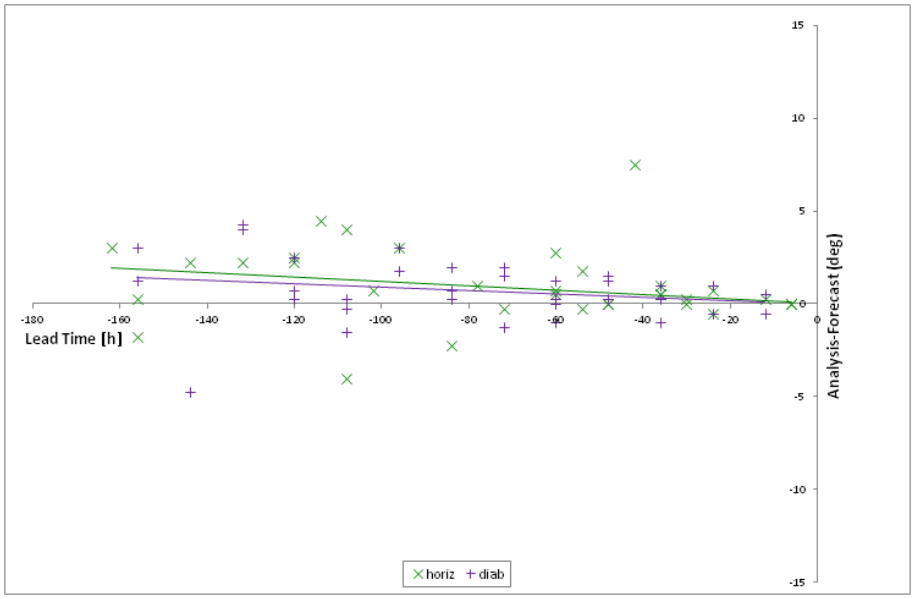


FIGURE 5.7: Pressure (analysis - forecast) against lead time for the selected storms divided by PTE category (horiz = green ×, diab = purple +), for (a) operational and (b) control forecasts.

For the intensity forecast of horiz-type storms (Figure 5.7), the line of best fit is remarkably similar in the control and operational forecasts, indicating that the resolution of the model has little effect on the forecast quality of these storms. This is consistent with the forcing coming from large-scale temperature gradients, so the large-scale structure is more important that the fine detail. At first glance, the diab-type storms appear to be better forecast in the control forecast than the operational at short lead times, because the line of best fit in the control case is very near the axis, indicating small differences between the forecast and analysis. This is counter-intuitive, because for diabatic

processes the fine detail is important, and so the increased resolution of the operational model should mean improved representation, compared to the ensemble control. However, at lower resolution, the parametrisation scheme will have a larger contribution to the overall representation of diabatic processes than at higher resolution. Therefore, this could also be due to the comparison of the reanalysis and forecast being dominated by the parametrisation scheme. Also, the gradient of the line of best fit for diab storms in the operational forecast is particularly flat, indicating that the forecast is not improving as quickly with shorter lead time than for horiz storms, which is consistent with a lower degree of understanding of diabatic processes.

(a)



(b)

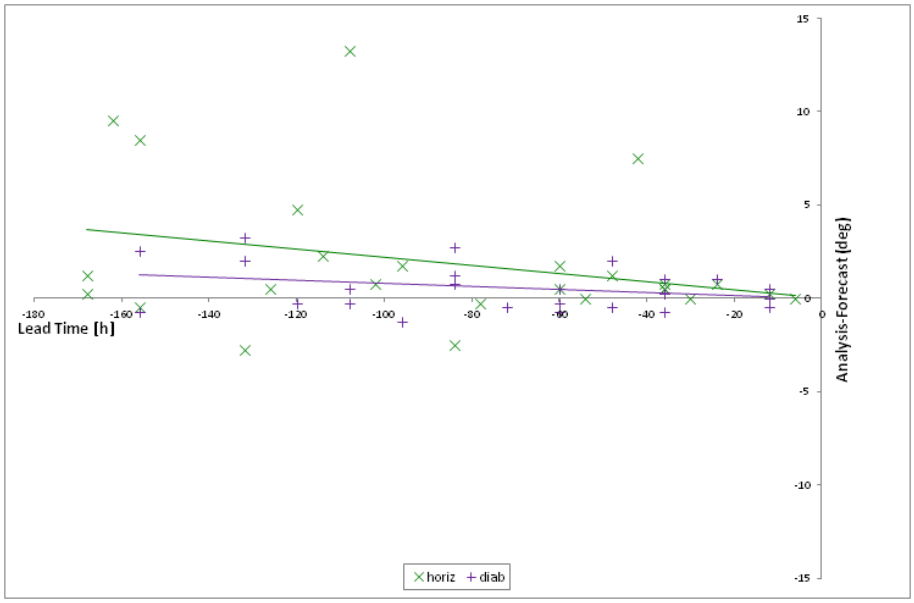


FIGURE 5.8: Latitude (analysis - forecast) against lead time for the selected storms divided by PTE category (horiz = green ×, diab = purple +), for (a) operational and (b) control forecasts.

Figure 5.8 indicates that both storm categories are forecast further north than they are in the analysis. Again, the deviations are greater for the horiz-type storms than in diab-type storms. However, the differences between the two linear regression lines are small, compared to the spread of the data. The results for longitude (Figure 5.9) are even less compelling. In the operational forecast, both PTE types of storm are forecast too far west at long lead times, and too far east at short lead times. In the control forecast, the storms are broadly too far west in the forecast, compared to the reanalysis. Importantly, in both cases the diab storms are forecast less well because the differences between the forecast and analysis are greater than those for horiz storm for most lead times. This is evidence that, in terms of position, diab storms are less well forecast than horiz storms.

To summarise, in terms of intensity, it appears that horiz-type storms are less well forecast than diab-type storms, but in terms of position, diab storms are less well forecast than horiz storms. Overall, this is consistent with deficiencies in the modelling of both the diabatic processes and the large-scale temperature structure or associated storm dynamics, although errors in the initial conditions could be another cause.

5.3.5 Model Resolution

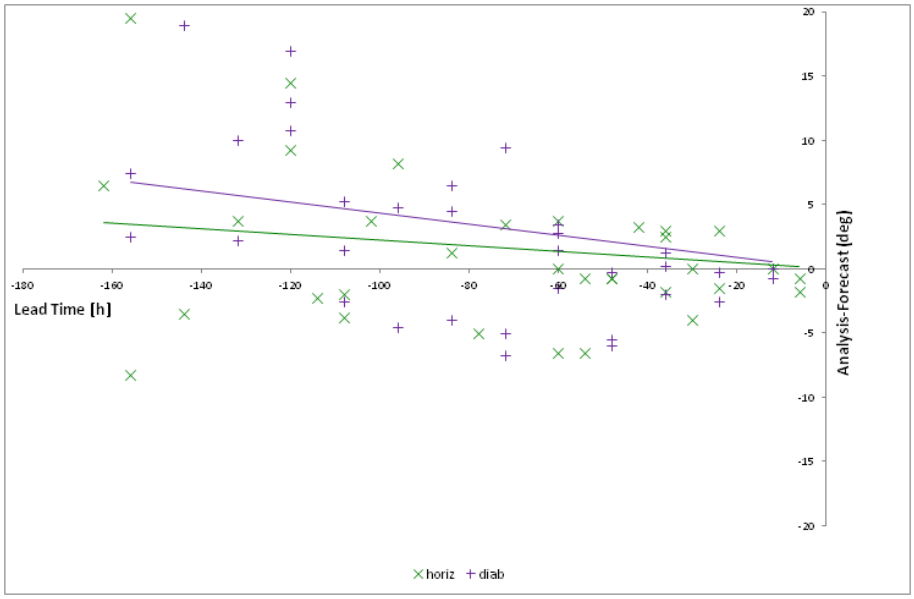
The next key factor in the forecasting of storms is that the model resolution changes with time. The native resolution is used as the measure, which refers to the resolution at which the model is run, because the data are processed ahead of tracking, so that they are all at the same resolution. This is done in wavenumber space, not in gridded space, and so native resolution is referred to in terms of T or T_L numbers. Details of the different resolutions can be found in Table 3.1.

It is anticipated that improving the native resolution of the model would improve the quality of the forecast. From Figure 5.10, there is some indication that resolution is a factor in the forecasting of pressure, because for both operational (Figure 5.10a) and control (Figure 5.10b), the lower resolutions of data generally have larger differences between the forecast and reanalysis, than the higher resolutions.

This is not the case for latitude (Figure 5.11)) or longitude (Figure 5.12), where the order of the lines of best fit in distance from the y-axis does not appear related to the native resolution of the data. This could be due to the small number of storms in each category. As shown in Table 5.3, there is only one storm at the highest resolution, and some of the other resolutions are only represented by a few storms. This increases the likelihood that one particularly well or badly

forecast storm could skew the results for the entire group of storms at the same resolution. This could be mitigated by selecting more storms initially; however, this is a problem for two reasons. Firstly, the storms would have to be analysed in less depth if there were more of them, particularly in terms of manually dividing the storms into jet stream categories. Secondly, it would mean that storms with a lower SSI were included in the analysis, and so dilute the aim to examine the most intense European windstorms. One avenue of future work is to update these results to include storms later than 2010, and have more storms at the highest resolution.

(a)



(b)

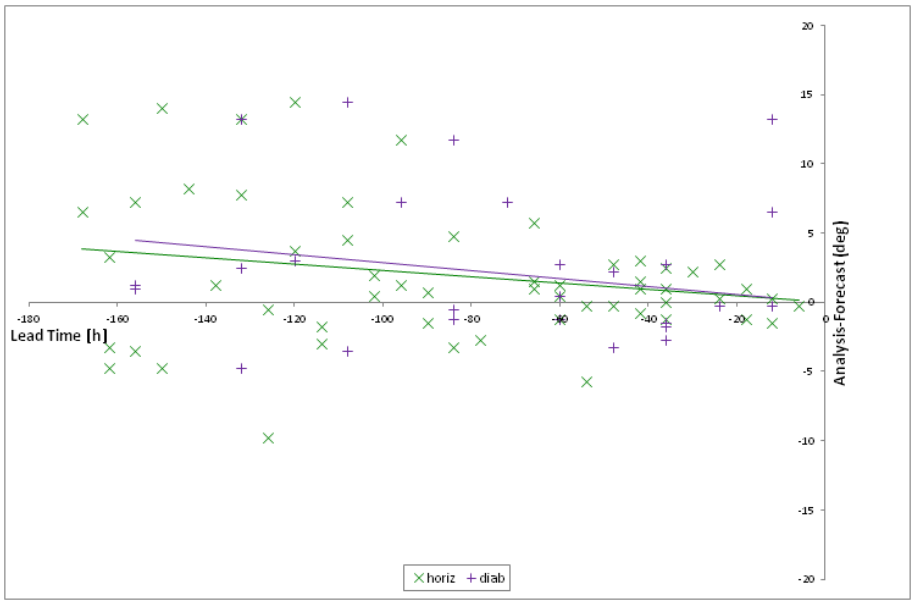
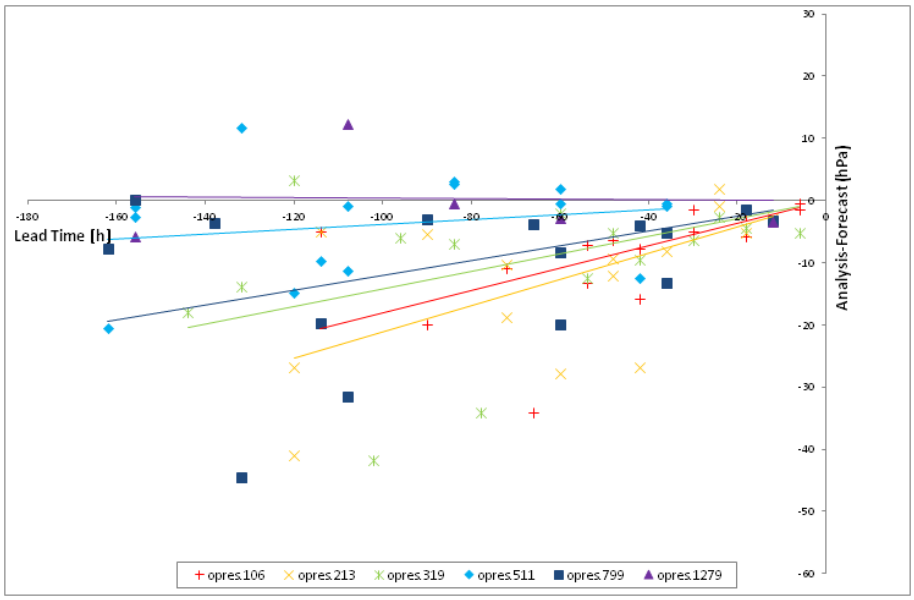


FIGURE 5.9: Longitude (analysis - forecast) against lead time for the selected storms divided by PTE category (horiz = green ×, diab = purple +), for (a) operational and (b) control forecasts.

(a)



(b)

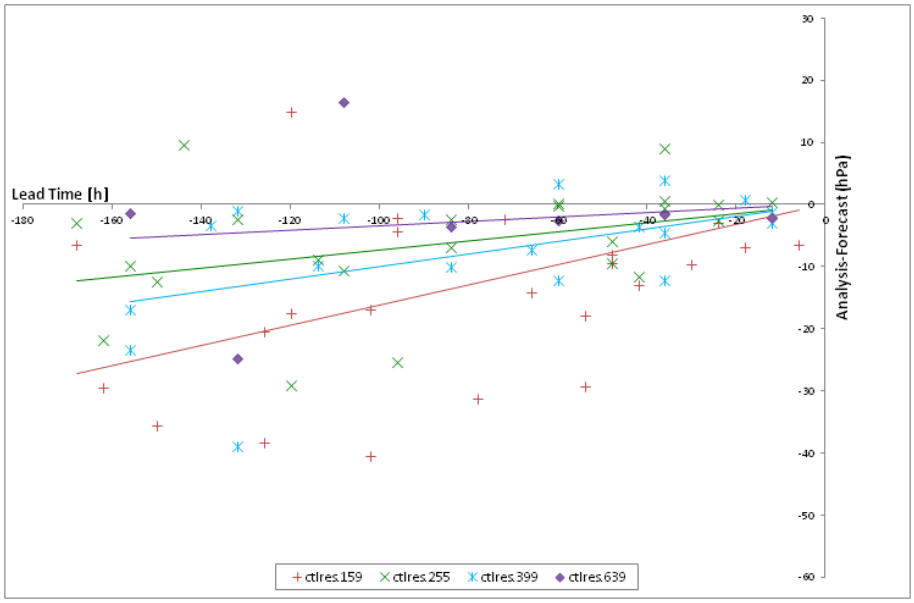


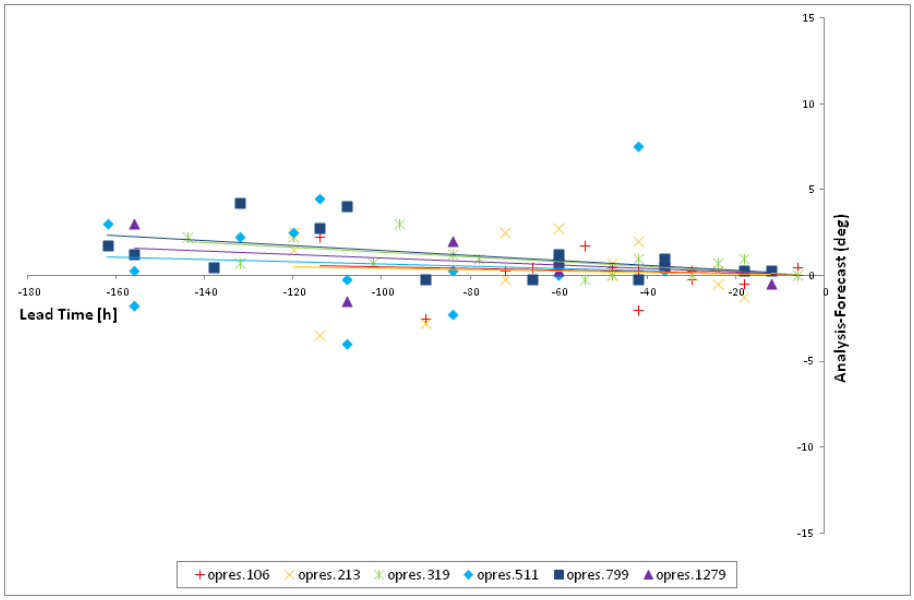
FIGURE 5.10: Storms divided by native model resolution for: (a) pressure in operational forecasts; (b) pressure in control forecasts. The numbers in the legend refer to the T or T_L native resolution of the forecast model.

Overall, the effect of resolution on the selected storms' forecast position remains unclear from these results, but intensity is slightly better forecast at higher resolution.

Operational Forecasts		Control Forecasts	
$T106$	5	T_L159	5
$T213$	4	T_L255	9
T_L319	8	T_L399	4
T_L511	8	T_L639	1
T_L799	4		
T_L1279	1		

TABLE 5.3: Number of storms at each resolution.

(a)



(b)

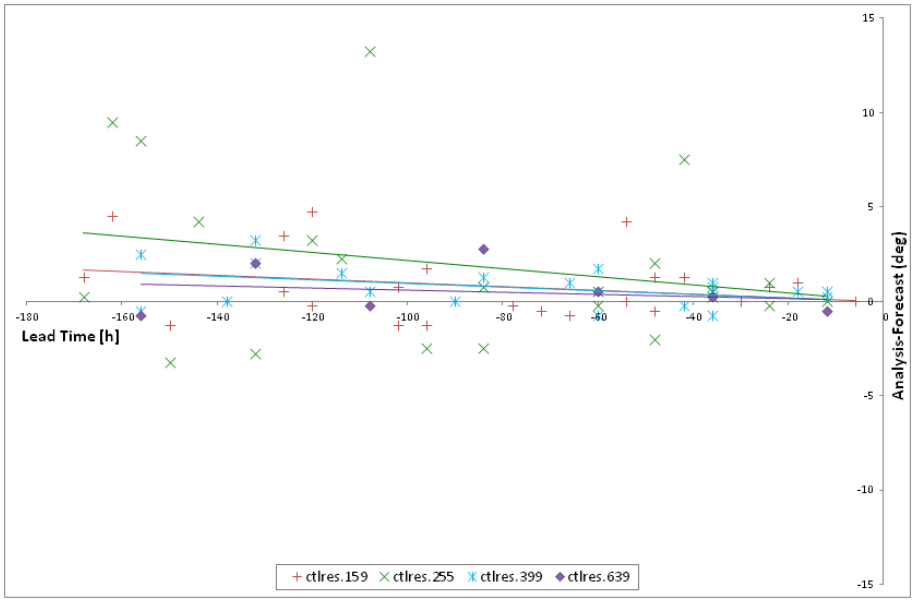
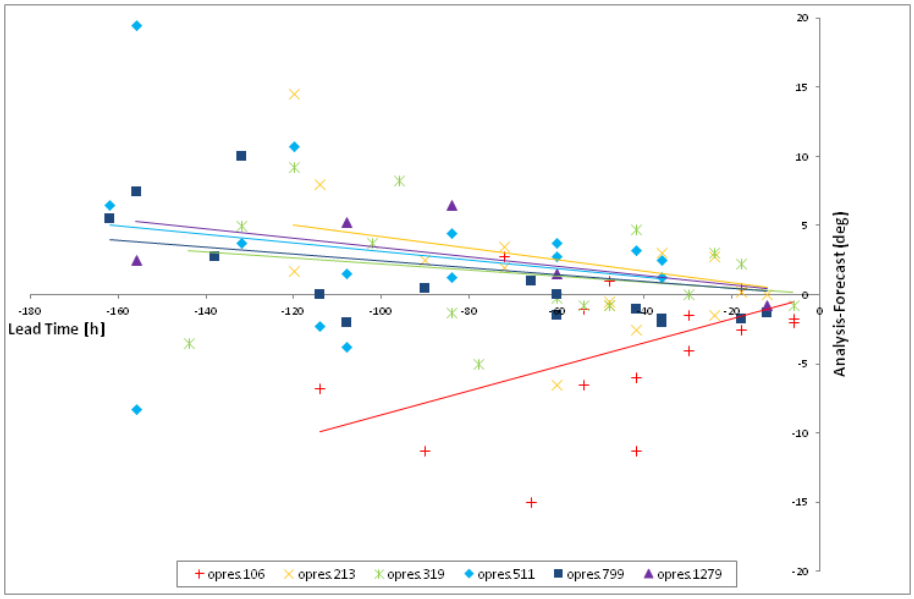


FIGURE 5.11: Storms divided by native model resolution for: (a) latitude in operational forecasts; (b) latitude in control forecasts. The numbers in the legend refer to the T or T_L native resolution of the forecast model.

(a)



(b)

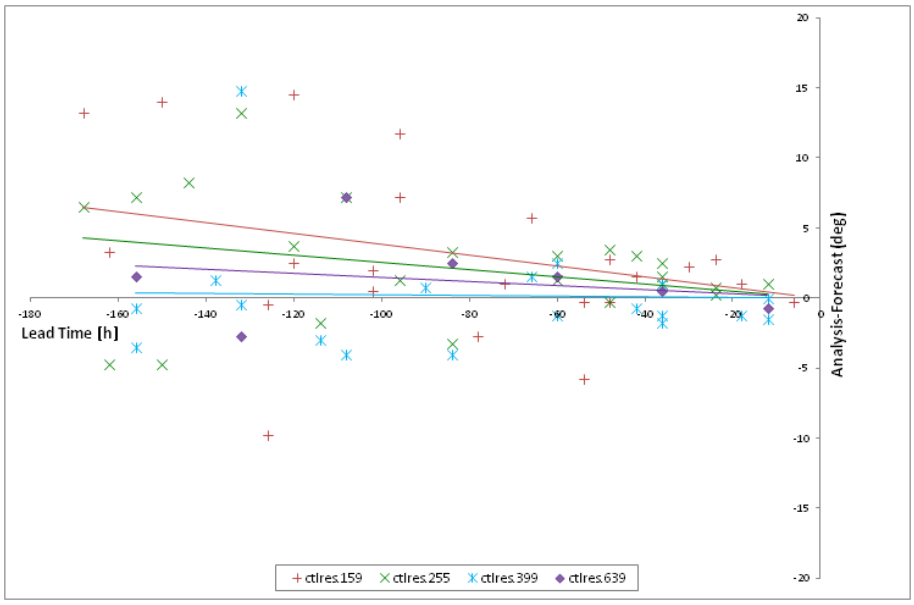


FIGURE 5.12: Storms divided by native model resolution for: (a) longitude in operational forecasts; (b) longitude in control forecasts. The numbers in the legend refer to the T or T_L native resolution of the forecast model.

5.4 A Simple Metric for Forecast Error

As discussed in Section 5.2, the longitude, latitude and pressure deviations at 24 hours are calculated using linear regression on the three metrics, and referred to as $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$, respectively. Using linear regression means that storms with fewer than three points in the regression were rejected: these storms are Dec 1993, Emma, Kerstin, Martin and Vivian. For the remaining 26 storms, the differences between forecast and analysis, calculated at 24 hours using linear regression, are shown in Table 5.4 for the storms that will be used in the analysis.

Storm	SSI	$\Delta(lon)$	$\Delta(lat)$	$\Delta(p)$
Klaus	55.47	-4.75	0.31	-8.77
Wiebke	46.31	-2.03	-0.30	-4.55
Kyrill	39.47	-1.77	-0.24	-2.89
Lothar	37.68	-15.35	2.03	-16.75
Jeanette	27.51	-0.50	1.00	-0.41
Daria	27.33	-3.49	0.45	-4.57
Agnes	19.69	0.46	0.71	-6.64
Anatol	18.60	-1.05	-0.01	-7.30
Udine	17.16	-6.25	-0.22	-11.33
Rebekka	16.25	-3.90	3.13	-1.90
Lara	15.68	-2.21	1.09	0.91
Xynthia	14.98	1.31	-0.44	-1.00
Jennifer	14.65	-2.66	-0.97	-12.82
Gero	13.53	2.00	0.31	0.94
Hanno	13.38	-1.11	0.18	-3.01
Silke	12.42	-7.65	0.05	-15.78
Elke	10.93	6.65	-4.54	-1.96
Urania	9.29	1.50	0.65	-2.06
Nana	9.24	-1.63	0.51	-2.57
Quinten	8.32	-0.83	-0.53	1.30
Verena	8.31	-1.53	0.24	-14.41
Pawel	7.82	-0.01	0.27	-4.09
Cyrus	7.80	0.95	8.10	-9.33
Lukas	7.77	11.81	-8.31	16.12
Franz	7.70	3.06	0.47	-1.86
Frieda	7.64	-0.30	0.55	-0.61

TABLE 5.4: The selected storms (ordered by SSI), with the 'standardised' forecast error at 24 hours, calculated using linear regression, for longitude, latitude and pressure.

In Table 5.5, the relationship between $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$ is explored using Pearson's correlation coefficient, and the relevant t-test for significance of the correlation (Section 3.6.4). With 24 degrees of freedom (two fewer than the number of storms included in the analysis), the t-test

result must be greater than the reference value of 1.711 (Rees, 2001, Table C.5) for the correlation to be significant.

There is a significant correlation between $\Delta(lon)$ and $\Delta(lat)$, meaning that storms with a better forecast in longitude tend to also have a better forecast in latitude. This is expected, because both are related to the position of the storm. If the storm's position is particularly well or badly forecast, then this will have an effect on both of these metrics. Significant correlation also exists between $\Delta(p)$ and $\Delta(lon)$, indicating that forecast errors in pressure are related to those in longitude. This makes physical sense because longitude is linked to storm speed, as explained in Section 5.3.1. Therefore, the correlation between $\Delta(p)$ and $\Delta(lon)$ means that storms where intensity is badly forecast are also those where speed is badly forecast.

	R	t	Sig?
$\Delta(lon)$ v $\Delta(lat)$	0.390	2.077	✓
$\Delta(p)$ v $\Delta(lon)$	0.657	4.270	✓
$\Delta(p)$ v $\Delta(lat)$	0.283	1.445	×

TABLE 5.5: The correlation (using Pearson's correlation coefficient, R) between the different 'standardised' forecast error at 24 hours; the test of that correlation using a t-test (t); and the significance (Sig?) of the correlation, obtained by comparing t to a reference value.

Overall, this metric shows promise in terms of measuring forecast quality. The next steps are to ascertain whether there is a relationship between these metrics for forecast quality and various properties of the storms and their forecasts. Two approaches will be taken. If the property could be numerical (e.g. storm intensity, resolution), then correlation is calculated directly between that and $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$. If the property is categorical (e.g. jet stream types, PTE categories), then a different approach is needed that compares the medians and inter-quartile ranges of the groups, both numerically and pictorially using box plots. Each of these four aspects will now be discussed in turn.

5.4.1 Storm Intensity

Correlation between $\Delta(lon)$, $\Delta(lat)$, $\Delta(p)$ and SSI is sought. This is done in both the Pearson's correlation coefficient (R) and the Spearman's rank correlation coefficient (R_S), and the results are shown in Table 5.6 . These two tests are described in Section 3.6.4. Testing the significance of these two coefficient is important, but the approach differs. For the Pearson's correlation coefficient, a t-test is performed and compared to the reference value (1.711) for the given number of

degrees of freedom (24). For the Spearman's rank correlation coefficient, the value of the coefficient is compared to a reference value, which for 24 degrees of freedom is 0.409 (Rees, 2001, Table C.12). In both cases, a coefficient greater than the reference values means the correlation is significant. In addition, Table 5.4 gives the raw values of $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$ ranked by SSI, so a relationship between the three forecast metrics and SSI might become apparent there.

	R	t	Sig?	R_S	Sig?
SSI v $\Delta(lon)$	0.203	1.014	×	0.028	×
SSI v $\Delta(lat)$	-0.239	1.206	×	-0.513	✓
SSI v $\Delta(p)$	0.095	0.469	×	-0.128	×

TABLE 5.6: Table showing the results of two correlation tests (Pearson, R; Spearman's Rank, R_S) of $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$ against intensity, and their significance.

However, little correlation emerges. There is no significant correlation between the $\Delta(p)$ and the intensity of the storm, measured by SSI. This indicates that intense storms are as well forecast as their weaker counterparts. This is also apparent from Table 5.4, where pressure forecast errors are clearly not ranked similarly to SSI. There is also no correlation between SSI and $\Delta(lon)$, which implies that the speed of intense storms is as well forecast as the speed of weaker storms. For $\Delta(lat)$, there is no correlation according to the Pearson's test, but there is according to the Spearman's test. That this correlation is negative means that more intense storms have smaller forecast errors in latitude. This could be due to a waveguide effect; intense storms are related to a fast jet stream, so the storm is guided by the jet stream through a particular range of latitudes. However, while the value of R_S is above the reference value indicating that it is significant, it is only slightly so, meaning the correlation might not stand up to further testing, for example if a different set of storms were used. Furthermore, it should be noted that the significance test indicates that there is a 5% chance that the correlation occurs due to random chance. Therefore, if correlation is sought twenty times, one of these will be due to random chance. Performing the significance test still leaves a 5% random chance of correlation.

Overall, there is little correlation between the storm intensity (SSI) and forecast quality in pressure, latitude or longitude. This is due to the large number of factors that affect SSI, such as the pressure gradient, the local wind speeds and the local wind speed climatology. While these factors are appropriate when considering the potential damage a storm could inflict, they may cloud the picture when a metric for storm intensity is required. Alternative metrics for intensity include

minimum core pressure, maximum vorticity, and maximum wind or gust speed. However, SSI is used in this work for consistency.

5.4.2 Jet Stream Type

The next steps are to explore the relationship between $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$ and the categories of storm identified in Chapter 4. The first of these is by the four jet stream types. Table 5.7 shows how many of the 26 storms used in this analysis are in each category. Box plots for these are shown in Figures 5.13 and 5.14. There are clearly differences in the quality and spread of these values between the different categories.

However, it is difficult to determine which are statistically significant, because some of these data fail the test for a normal distribution (Table 5.7). The test used is the Kolmogorov-Smirnov, because this is less sensitive to outliers than the alternative Shapiro-Wilk test (Gibbons, 1985, Section 7.3). The reason the data fail this test is probably because there are too few storms in each category, but as already discussed, it is not possible to increase the number of storms without decreasing the depth of the analysis. The failure of the test for normality has repercussions for hypothesis testing. As the data are not normally distributed, testing whether mean or median values for forecast quality are significantly different between the different jet categories is not possible, so this discussion will remain qualitative.

Jet Category	N	Longitude		Latitude		Pressure	
		K-S p	Normal?	K-S p	Normal?	K-S p	Normal?
cross early	6	0.389	×	0.00894	✓	0.0854	×
cross late	8	2.00×10^{-4}	✓	0.4274	×	5.09×10^{-7}	✓
edge	8	0.00901	✓	0.124	×	3.88×10^{-5}	✓
split	4	0.1876	×	0.227	×	0.00128	✓

TABLE 5.7: The number of storms of each jet stream type (n), included in the 'standardised' forecast error analysis, and the p-value of the Kolmogorov-Smirnov (K-S) test for normal distribution in each category and whether that states it is normally distributed ($p<0.05$).

Comparing the median values for $\Delta(p)$ implies that storms that cross the jet stream early are best forecast in terms of intensity; second are split-jet storms; third are edge-type storms; and the least well forecast jet stream type by this metric is the cross late type. While this is a more quantitative approach than that taken in Section 5.3.3, the results do not agree. There, the cross early storms are the worst forecast for intensity, in both operational and ensemble control forecasts. This casts

some doubt on whether the linear regression method taken here is representative of forecast quality. However, when comparing the forecast quality at 24 hours on Figure 5.4, it becomes clear that the lines of best fit are close together and so differences between the categories are small. Any differences could easily be amplified by approaching linear regression differently, as has been done for the results in Figure 5.13. The lines of best fit from Figure 5.4 take an average over the range of lead times, for each category, so it is not surprising that the two methods give different results.

There are two metrics for forecast quality here: the median and the mean. Both are shown in Table 5.8, to illustrate the difference. If the mean value is examined, then the storms that cross the jet early are still best forecast, then the edge-type, split-type, and cross late storms. The edge and split have swapped places, compared to the rankings for median, for two reasons. Firstly, there are few storms in the split-jet category, and the mean takes into account the number of storms. This implies that the mean is a more useful statistic here, because the different categories have different populations. Secondly, the mean is more sensitive to outliers than the median, because it sums over all values. This means when there are outliers, the median is more representative of the middle of the data. When considering the forecast quality of these storms, outliers are expected, because a storm could be particularly badly forecast for a number of reasons. On balance, the median is a more useful measure for forecast quality.

The spread of the results can be compared using the inter-quartile range (IQR). This is the spread of the deterministic forecasts of a number of storms in each category (Table 5.8). It is largest for the storms that cross the jet stream late, indicating that there is the widest range of forecast quality amongst the storms in this category. The IQR is smallest for the split-jet storms, indicating that the forecast quality of these storms is most similar. However, this is also the smallest category, which can mean the spread is smallest. For this reason, the standard deviation is also considered. It shows that the spread is greatest for edge-type storms, then split, cross-late, and least for cross-early type storms. However, the standard deviation is more sensitive to outliers than the IQR, so given the presence of outliers in the data, the IQR will be used for the remainder of this work.

Figure 5.13 shows that in all four jet stream cases, $\Delta(p)$ is positively skewed, given the relative position of the maximum, upper quartile, median, lower quartile and minimum. Positive skew means that there is a longer tail on the side with values lower than the median, than that for the higher-than-median values, indicating that there are a few storms with a particularly large $\Delta(p)$ and so a particularly bad forecast. Although Figure 5.13 shows that intensity is underforecast,

	Cross Early	Cross Late	Edge	Split
Median	-1.23	-4.56	-3.55	-1.88
IQR	6.07	6.99	6.84	3.97
Mean	-2.76	-6.27	-3.39	-5.38
St. Dev.	4.27	5.24	9.48	7.60

TABLE 5.8: Summary statistics for each of the four jet stream types: median, inter-quartile range (IQR), mean average and standard deviation (St. Dev).

this is not always the case, as can be seen in the examples of Lukas, Quinten, Gero and Lara (Table 5.4). It is noteworthy that these four storms have a range of intensities, come from three of the four jet stream types and both of the PTE categories, and so appear to have little in common dynamically that could make their intensity overforecast. However, the broad picture remains that intensity is underforecast.

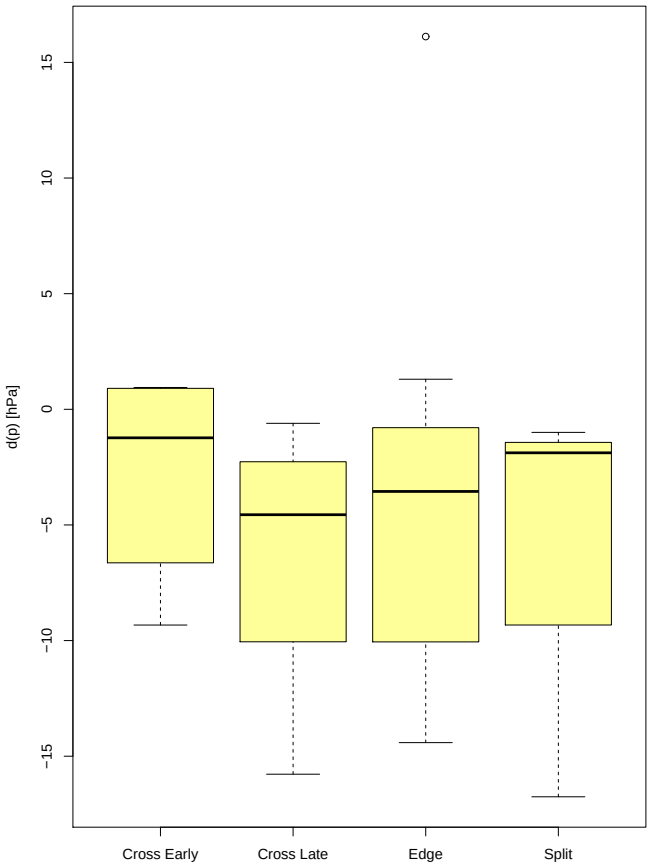


FIGURE 5.13: Box plots of $\Delta(p)$ for each jet stream type. $\Delta(p)$ is the interpolated forecast error in pressure at 24 hours lead time, based on linear regression for each storm.

	Cross Early		Cross Late		Edge		Split	
	lon	lat	lon	lat	lon	lat	lon	lat
Median	0.70	0.85	-2.76	0.18	-1.08	-0.13	-1.29	1.25
IQR	1.62	0.40	3.83	0.70	0.96	0.83	8.51	2.06

TABLE 5.9: The median and inter-quartile range (IQR) in $\Delta(lon)$ and $\Delta(lat)$ for each of the four jet stream types. $\Delta(lon)$ and $\Delta(lat)$ are the interpolated forecast error in longitude and latitude (respectively) at 24 hours, based on linear regression for each storm.

For latitude and longitude, the summary statistics are shown in Table 5.9, and the box plots in Figure 5.14. Considering forecast quality, indicated by the median values, the ranking (best to worst) of the jet stream types for $\Delta(lon)$ runs: cross early, split, edge, cross late. This broadly agrees with the same ranking for pressure (cross early, edge, split, cross late). Though edge and split have swapped places, the values for the median are very similar and so differences between the two jet stream types in terms of forecast quality are small. This reinforces the idea that storms that are badly forecast for intensity tend to be those that are badly forecast for storm speed, which is related to longitude. No similarity is found between the forecast quality in $\Delta(lon)$ and $\Delta(lat)$; for the latter, the the ranking from best to worst quality of forecast is: edge, cross late, cross early, split. This is unexpected because both are related to storm position. However, longitude is also related to storm speed, so this could be a more important factor in forecast quality for longitude, compared to storm position.

For forecast spread of the deterministic forecasts, the results are inconsistent. For longitude, the ranking from largest to smallest spread goes as follows: split, cross early, cross late, edge. For latitude, it goes: cross early, split, edge, cross late. This does indicate comparing the forecast position to the analysis position for split- and cross early-type storms generally has a larger spread than that of edge- or cross late-type storms. However, this does not agree with the spread found in pressure, so analysing forecast spread is impossible using this method, and so will only be considered during analysis of the ensemble forecasts.

Overall, $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$ are useful measures of forecast quality, and allow the differences in forecast quality between the different jet stream types to be identified. For intensity (pressure) and speed (from longitude), storms that cross the jet stream early are best forecast, and those that cross the jet stream late are worst forecast. This demonstrates that the crossing of the jet stream is a key event in determining the predictability of these midlatitude cyclones. There the similarities between the different metrics end: in $\Delta(lat)$, the edge type storms are best forecast and split-jet

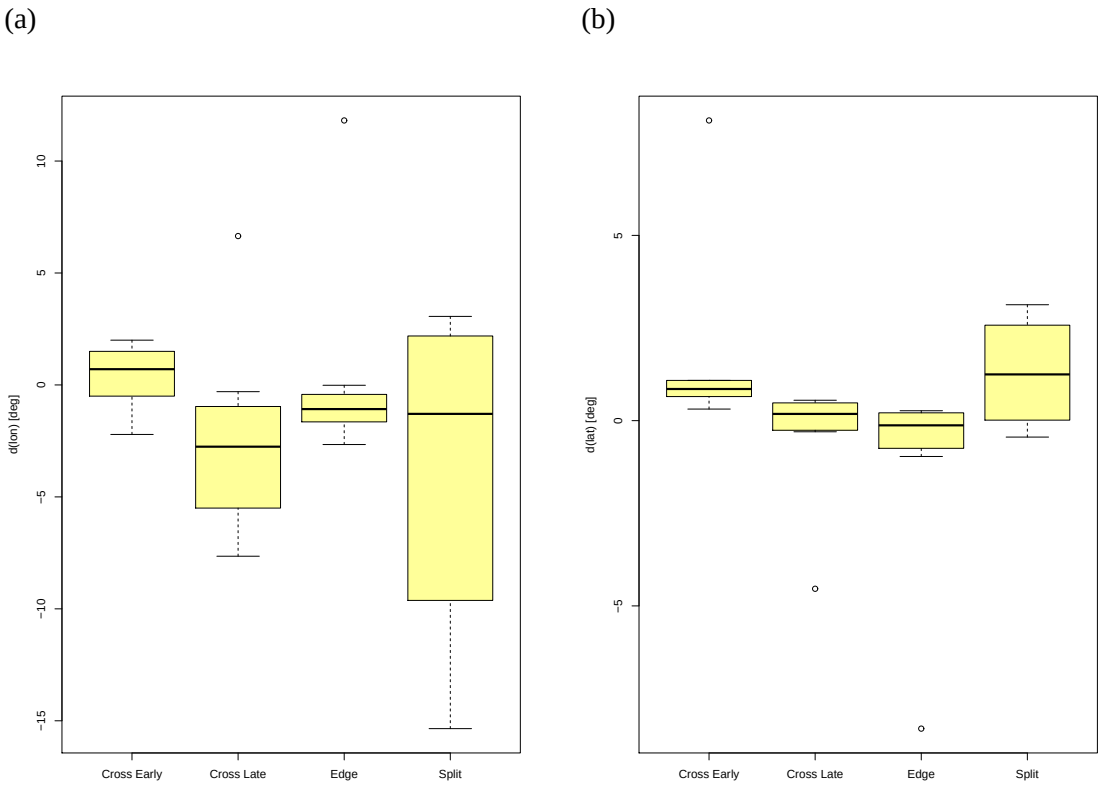


FIGURE 5.14: Box plots of (a) $\Delta(lon)$ and (b) $\Delta(lat)$ for each jet stream type.

storms are worst. This implies that the effect that the jet stream has on storm intensity and speed is different from the effect it has on the latitude band through which the storms pass. It is likely that the speed of the jet stream affects storm intensity and speed, whereas jet location affects the storm track's latitude.

5.4.3 PTE Category

PTE analysis determined that two processes are particularly important for deepening: horizontal temperature advection ('horiz') and diabatic processes ('diab'). There are more storms where horiz is the dominant process (21) than diab (5). The Kolmogorov-Smirnov test for normality (not shown) indicates that the data for 'diab' storms are not normally distributed, likely due to how few storms are in this category, so testing for a significant difference in the means or medians of the two PTE categories is not possible.

	Pressure		Longitude		Latitude	
	diab	horiz	diab	horiz	diab	horiz
Median	-1.04	-5.24	-0.69	-1.23	0.43	0.11
IQR	-1.00	-4.09	-0.83	-1.11	-0.30	0.31

TABLE 5.10: The median and inter-quartile range (IQR) for each of the PTE categories.

For pressure (Figure 5.15), both the median and the spread of the data is much greater for horiz-type storms than for diab (Table 5.10). The medians are sufficiently different indicates that the quality of the forecasts is better for diab storms than horiz. As discussed in Section 5.3.4, this could be due to model error in simulating baroclinicity or baroclinic energy conversion, uncertainty in the initial conditions; on the other hand, the diab storms could appear to be forecast unrealistically well, because the model that generate the reanalysis uses the same parametrisation for diabatic processes as the forecast model.

For location (Figure 5.16), the difference between the two categories is less clear. Table 5.10 shows that the medians and IQRs are similar for the two categories, in both latitude and longitude. No significant differences appear between the forecast quality for position of horiz and diab storms.

In summary, according to this approach, the forecasts of horiz-type storms show lower quality and greater spread than those for diab-type storms. This means that improving simulations of the large-scale horizontal temperature gradient or the baroclinic processes could significantly improve the forecast quality of midlatitude cyclones. Alternatively, it could be due to the propagation of errors in the initial conditions.

5.4.4 Model Resolution

The section analyses the relationship between $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$ and the model resolution, by searching directly for correlation. This is done in terms of Pearson's correlation coefficient, the significance of which is ascertained by performing a t-test and comparing the value to a reference value (1.711). Spearman's rank correlation coefficient could be used but, when the values of resolution are ranked, there would be a large number of tied ranks, which reduces the usefulness of the test considerably. Therefore, only Pearson's correlation coefficient will be analysed here.

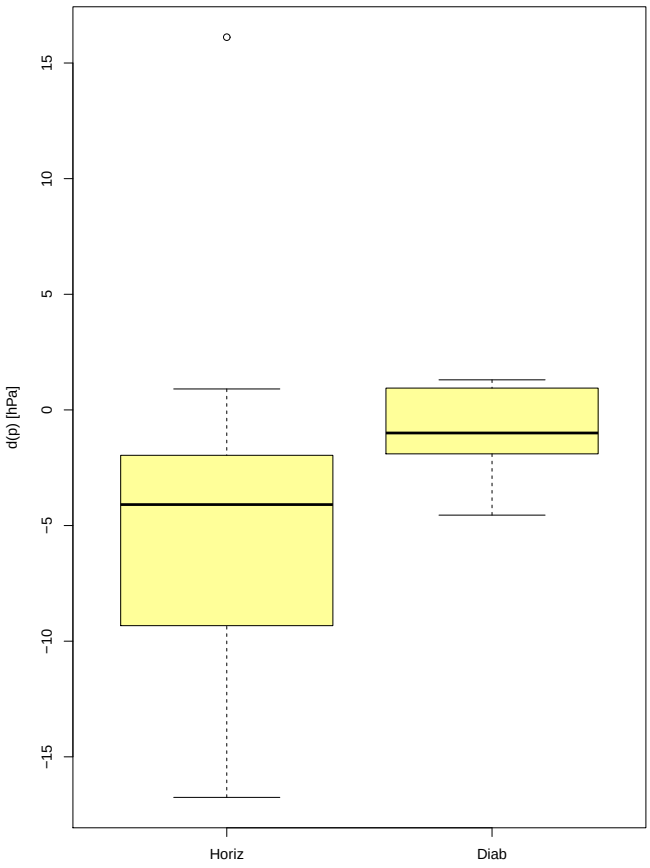


FIGURE 5.15: Box plots of $\Delta(p)$ for each PTE category. $\Delta(p)$is the interpolated forecast error in pressure at 24 hours, based on linear regression for each storm.

These results are presented in Table 5.11, and indicate that there are significant relationships to analyse. The link between resolution and the difference in intensity between forecast and reanalysis ($\Delta(p)$) shows significant, positive correlation. This implies that, as the resolution gets higher, the difference between the forecast and analysis becomes less negative. This is checked by using absolute values of $\Delta(p)$, when a negative correlation appears, meaning smaller values of $|\Delta(p)|$ at higher resolution. In short, intensity is better forecast at higher resolution. This is also the case for latitude, because $\Delta(lat)$ shows negative correlation. However, the correlation between $\Delta(lon)$ and resolution is not significant.

To summarise, calculating correlation directly reveals that increasing resolution significantly improves the quality of forecasts for intensity (pressure) and latitude, but no significant improvement is seen for longitude.

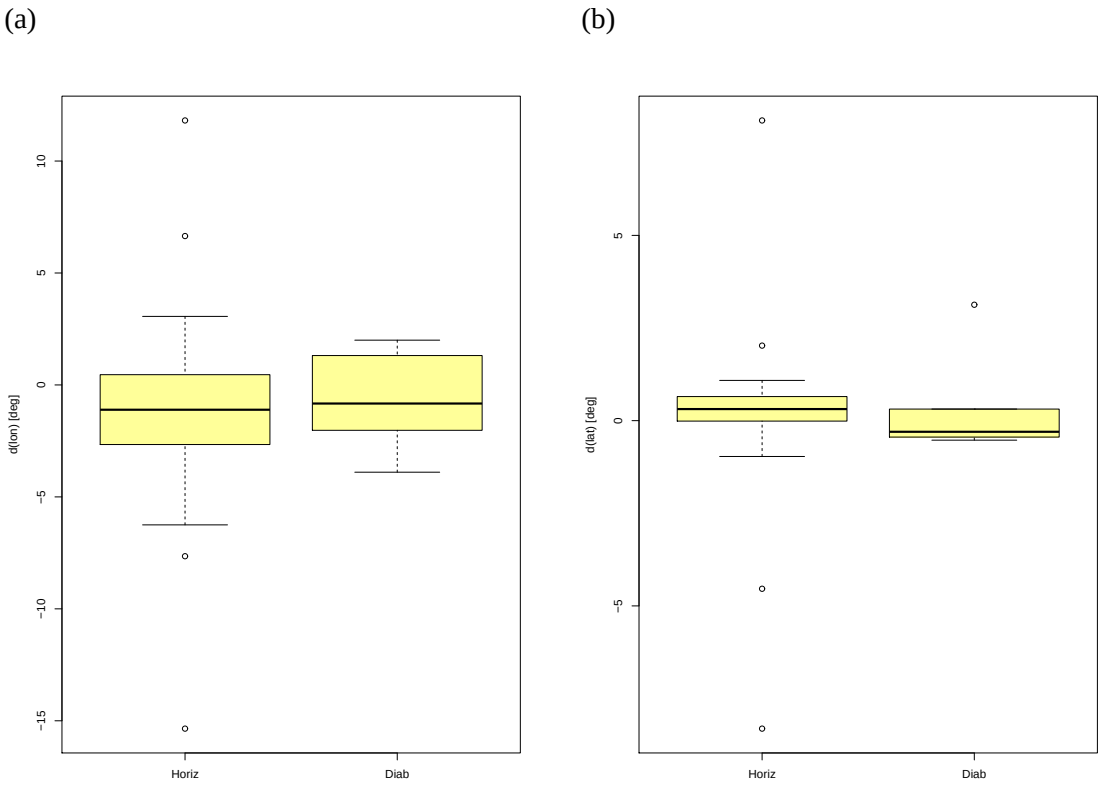


FIGURE 5.16: Box plots of (a) $\Delta(lon)$ and (b) $\Delta(lat)$ for each PTE category. $\Delta(lon)$ and $\Delta(lat)$ are the interpolated forecast error in longitude and latitude (respectively) at 24 hours, based on linear regression for each storm.

	R	t	Sig?
Resolution v $\Delta(lon)$	0.211	1.06	×
Resolution v $\Delta(lat)$	-0.603	3.70	✓
Resolution v $\Delta(p)$	0.255	1.29	✓

TABLE 5.11: Table showing the results of two correlation tests (Pearson, R; Spearman's Rank, R_S) of $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$ against resolution, and their significance.

5.5 Ensemble Forecasts

As previously discussed, ensemble forecasts are those where the initial conditions are perturbed. This provides information on the uncertainty in the forecast, by an examination of forecast spread. Forecast spread differs from forecast error, but the two are normally related because spread in an ensemble indicates an event is more difficult to forecast and so would have larger forecast error, meaning that there should be a positive correlation between error and spread in an ensemble forecast. It is possible to use the forecast median or mean to measure forecast quality. However,

as two forecasts have already been used to examine the accuracy of the forecast, this section will concentrate on forecast spread and the relationship with lead time. Note that lead time is still made negative, so that lead times get shorter further along the x-axis.

There are many ways to verify an ensemble forecast (see Joliffe and Stephenson (2003, Chapter 7) for a review). Many have been developed by institutions that run ensemble prediction systems, to provide a bulk measure of performance. These are not appropriate for use in the current context, because this work considers specific storm events and how well they are forecast. Therefore, this work concentrates on analysing the forecast tracks of the storms. Reliability of a probabilistic forecast is another measure used, determining that if an event is forecast as 40% likely to happen, then it does happen 40% of the occasions when this is the case. Quantifying reliability is not possible in the current work for two reasons. First of all, the small sample size means that the results are unlikely to prove statistically significant. Secondly, to execute this type of analysis, it would be necessary to introduce a threshold (e.g. what is the chance of core pressure falling below 960hPa?), which adds an additional level of complexity to the analysis. Given that extreme events have already been selected based on having a high value of the Storm Severity Index (SSI), but many events needed to be manually excluded due to duplication or being dynamically different from midlatitude cyclones (Chapter 4). Therefore, a threshold for these storms would be difficult to determine, so will not be used here. Overall, forecast reliability will not be examined in the current work. Another point of interest to model developers is whether an ensemble is over- or under-dispersive: that is, do the initial conditions realistically sample the uncertainties in the atmosphere? If the uncertainties of the atmosphere are insufficiently sampled to make the perturbations for the initial conditions for the ensemble model, then a forecast is said to be under-dispersive; however, if the perturbations are too large, then the model is said to be over-dispersive. Analysis of the dispersive nature of the ensemble will remain qualitative, and a more quantitative approach is left for future work.

As with the deterministic forecasts, the analysis of the ensemble forecasts concentrates on the time of minimum pressure at a variety of lead times, but now considers the spread of the forecast. All storms are included at first, but then these are divided into by jet stream type, by the results of the PTE analysis, and by resolution. This is achieved by use of box plots to illustrate the variation of forecast spread with lead time.

5.5.1 All Storms

First, all of the storms are grouped together, to ascertain whether there is a link between forecast spread and lead time. When the big picture is examined, Figure 5.17a indicates that there may be a weak relationship, because the smallest spread is found at the shortest lead times. However, though the spread (IQR) is generally larger at longer lead times, the largest spread of all is found at the relatively short lead time of 24 hours. This implies that the link between the spread of forecasts for intensity and lead time is not straightforward. However, because the forecasts are initialised once every 24 hours (at 12 noon), then the forecasts of different storms are included in the 12h forecast than in the 24h forecast. Furthermore, the forecast tracks are matched automatically with the analysis track, and any forecast tracks that do not match successfully are rejected. This means that the number of storms included in the analysis of neighbouring box plots differs. This leads to some lead times having a larger spread than would be expected based on the general trend. Unsuccessful attempts to mitigate this were made, by binning the lead times by day, which then didn't give enough boxes to identify a trend. Another approach was to average the spread of the results, but while this did damp out the anomalies, it also made any signal harder to identify.

The decreasing spread with decreasing lead times is more clear in the forecasts for latitude and longitude (Figure 5.17b,c), than in the forecasts for pressure. This unusually large spread at 24h lead time is still apparent here, so it is clear that a few storms are making the link between forecast spread and lead time less distinct. Broadly speaking, the spread is greater in longitude than in latitude. This is consistent with the location of the storms being relatively well forecast, in terms of the range of latitudes a storm will pass through, compared to the speed of the storm, which is linked to longitude.

An alternative method of considering the data is to collect the lead times together by day, so that lead times up to and including 24 hours are grouped into 'Day 1', up to 48 hours into 'Day 2', and so on. For pressure, Figure 5.18 shows that generally, forecast error and spread increase with longer lead times. However, this relationship is not consistent at all lead times for all variables, particularly at longer lead times, because some of the poorly forecast storms might just meet the track matching criteria on, say, Day 6, but be filtered out by Day 7. While this could be mitigated by having more storms in the group, this is impractical because grouping of storms involves detailed analysis of each individual storm, and including more storms would reduce the level of detail. Figure 5.18 also reiterates the point that pressure is underforecast; that is the

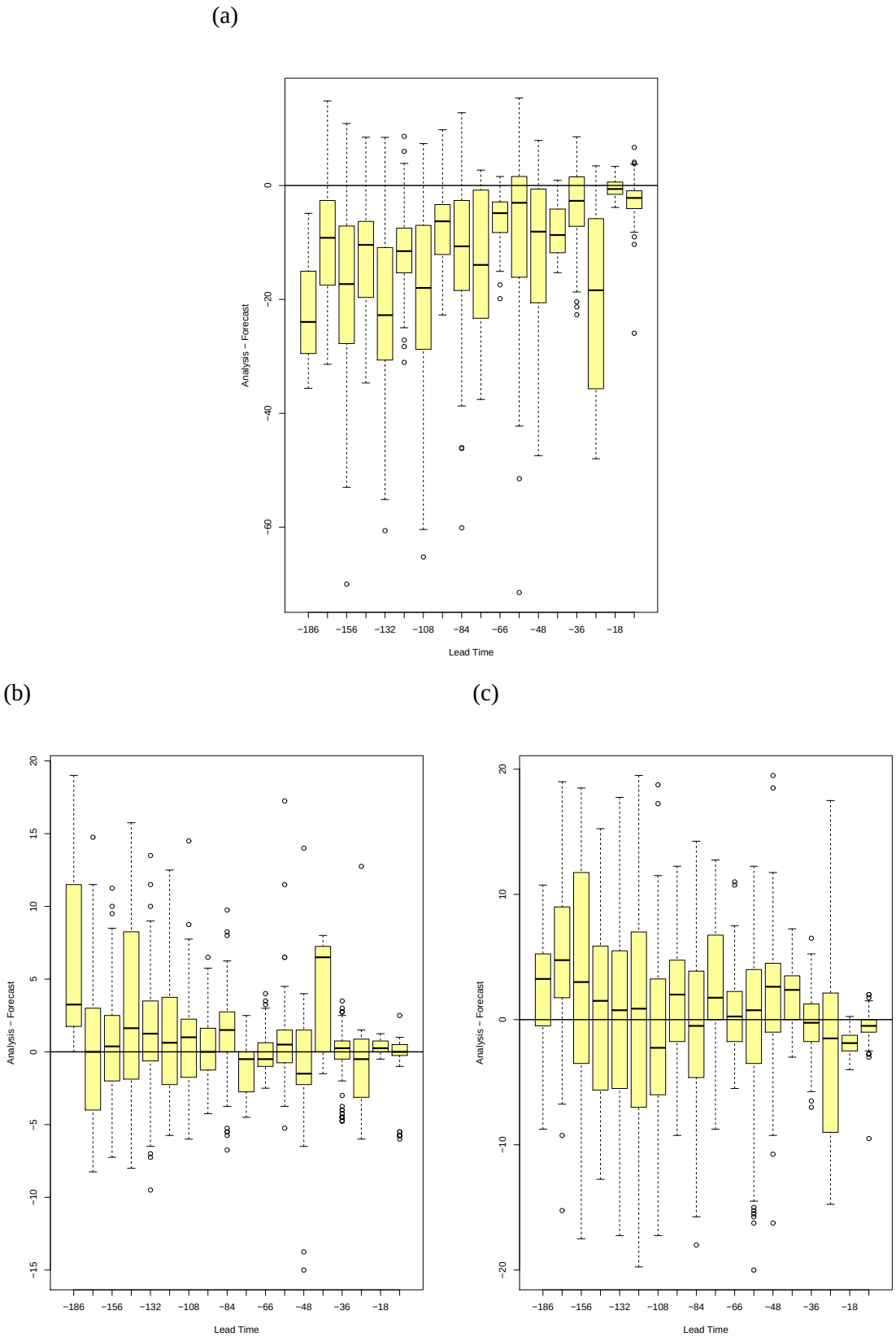


FIGURE 5.17: Differences between the ensemble forecasts and reanalysis for all storms, for (a) pressure (in hPa), (b) latitude, (c) longitudes (both in degrees)

forecast core pressure is higher than the analysis core pressure, so the forecast is less intense. Whether latitude and longitude are under- or over-forecast is less clear from these plots, because the boxes are distributed fairly evenly either side of zero.

Figure 5.19 shows box plots of the absolute values of the differences between the forecast and the analysis and the root-mean-square error; that is, the differences between the analysis and forecast pressure all squared, mean averaged, and then square rooted. For comparison with this value, the box plots are of the absolute values of the differences between analysis and forecast pressure. Figure 5.19a shows that, for pressure, forecast error and spread generally increase with lead time. Both metrics of forecast error considered — the root-mean-square error or the median value — increase with increasing lead time. Both metrics of forecast spread illustrated — the range of the values and the inter-quartile range — also increase with increasing lead time. The exception is that Day 7 has smaller root-mean-square error, median value and spread than Day 6, which is likely because these storms will be those that are generally well-forecast, as the poorly forecast storms would not have a successful match between to the analysis track at these long lead times. Figures 5.19b and 5.19c show that, for longitude and latitude, forecast error and spread also increase with increasing lead time. More so than pressure, this increase is not smooth from day to day; for example, in longitude the spread and root-mean-square error are larger at Day 5 than at Day 6. This is likely skewed by a few storms being particularly badly forecast in terms of longitude at these lead times. However, it is the general increasing trend in the metrics for error and spread that is key.

The idea that the data is skewed by a few very badly forecast storms is corroborated by the unusually large value of root-mean-square error in latitude for Day 2. At this lead time, there are a several particularly extreme outliers (shown as circles in the plot). Root-mean-square error is particularly sensitive to such outliers, because the values are squared before averaging, magnifying their extremity. Given the prevalence of outliers in the data set, root-mean-square error will not be analysed further. For this reason, further analysis will consider the raw values of the differences between forecast and analysis for pressure, latitude and longitude, by lead time.

To summarise, the key point from this analysis is that forecast spread decreases with lead time. The next stages of the analysis will ascertain whether this is the case for all categories of storm, and whether some categories have greater spread than others. In order to examine the detail of forecast quality with respect to lead time, plots similar to Figure 5.17 will be used with the lead

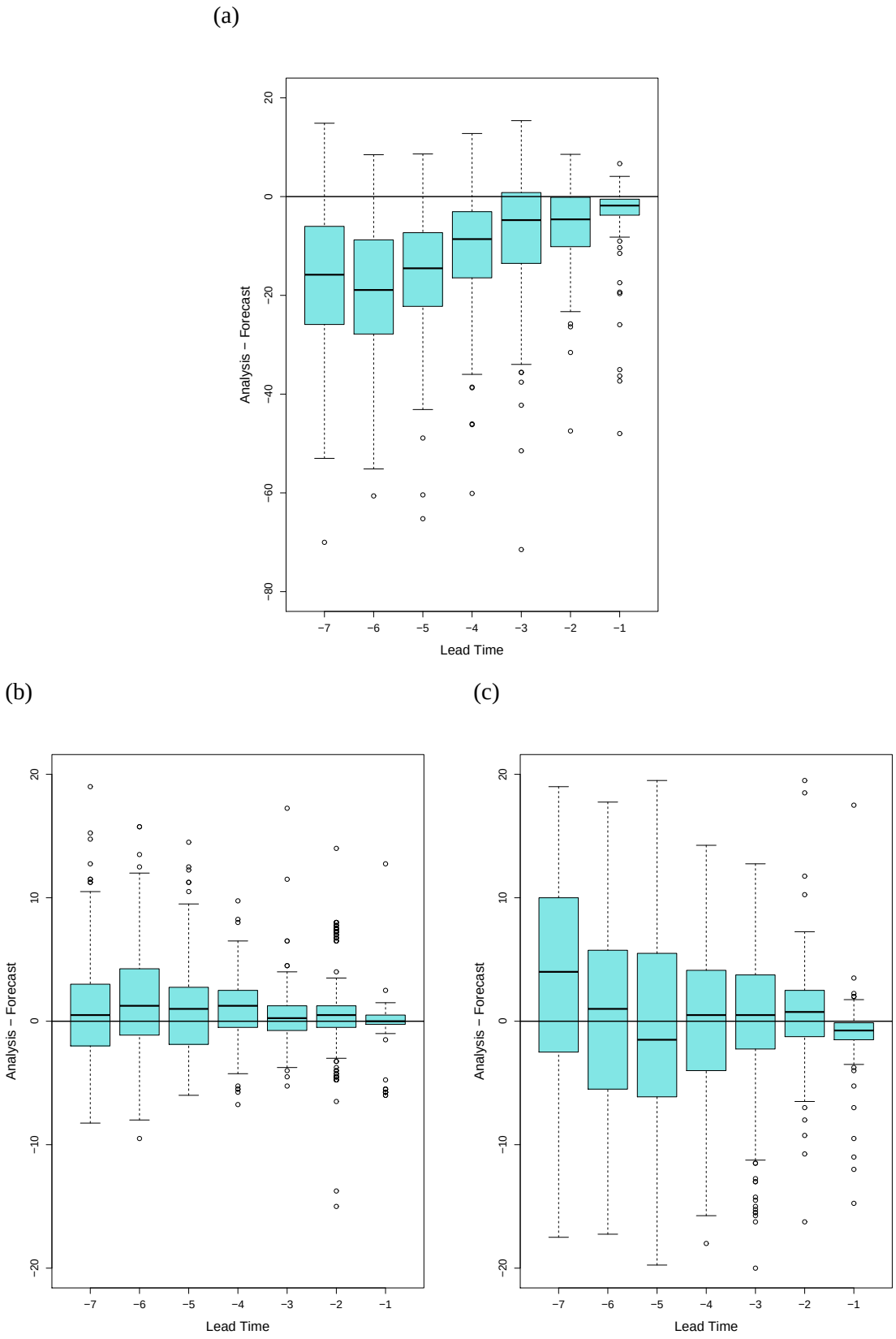
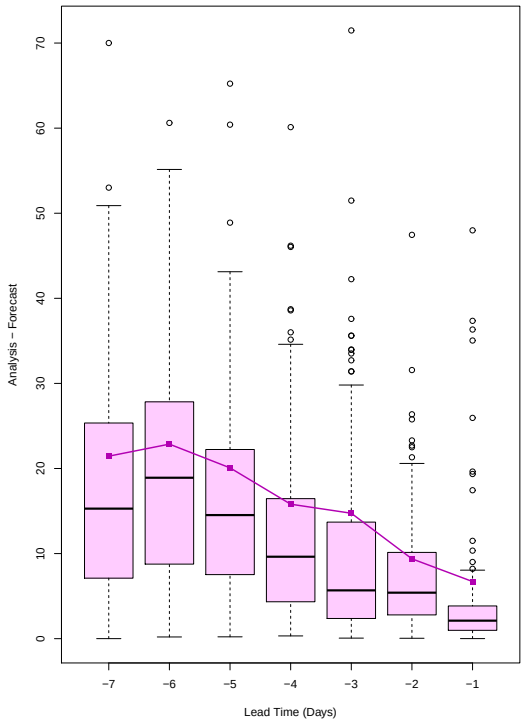
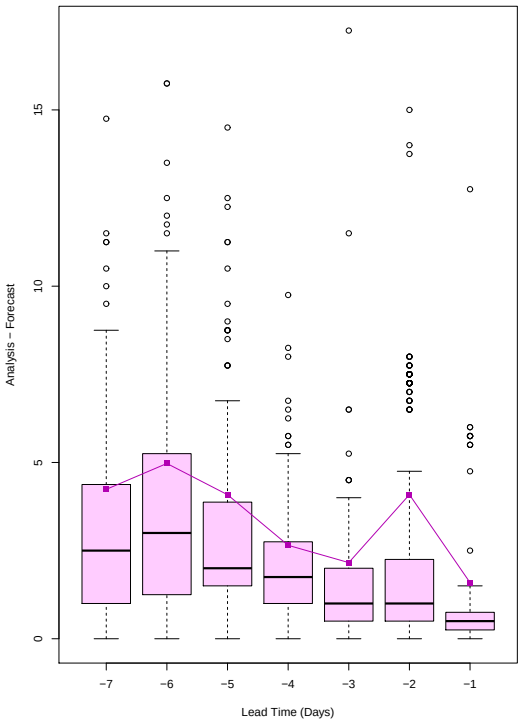


FIGURE 5.18: With lead times binned by day, box plots showing the differences between the ensemble forecast and reanalysis for (a) pressure (in hPa), (b) latitude, (c) longitude (both in degrees), for all storms.

(a)



(b)



(c)

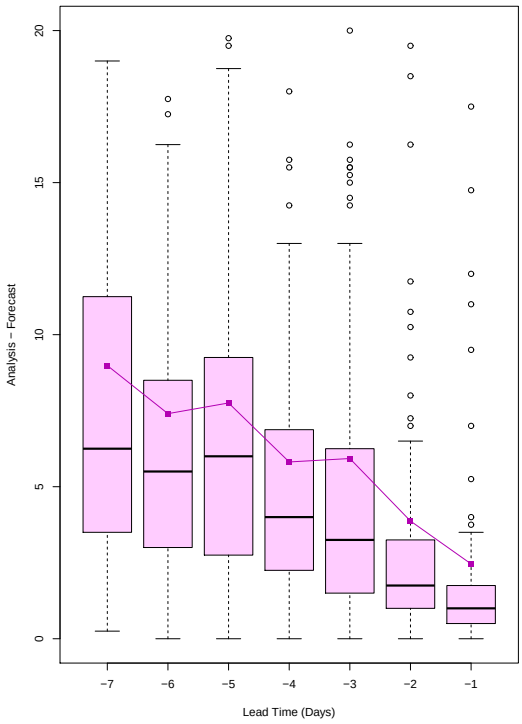


FIGURE 5.19: With lead times binned by day, box plots showing the absolute values of the differences between the ensemble forecast and reanalysis for (a) pressure (in hPa), (b) latitude, (c) longitude (both in degrees), for all storms. Line shows the root-mean-square error.

times not grouped. For plots similar to Figure 5.18 with lead times binned by day, please see Appendix A.

5.5.2 Jet Stream Type

The next steps of the analysis are to compare the forecast spread for the groupings of storm by jet stream type. This will allow determination of which categories of storm have a greatest uncertainty, and identification of which categories have the furthest to go in terms of improving their simulation to reduce forecast errors. Please note that the scales of the y-axes differ between the three variables considered, because this format is more informative than when the axis scales are consistent.

For pressure, the IQR generally decreases with lead time in all four categories, again with exceptions for reasons already discussed. The spread for split-jet storms is generally greatest, despite the fact that this category contains the smallest number of storms, which would generally mean that the variance is less well-sampled and so the IQR is smaller. However, the other three categories have similar IQRs, meaning they have similar degrees of uncertainty, the uncertainty falls off particularly irregularly for storms that cross the jet stream late, which is consistent with the crossing of the jet stream being a key factor in storm development. The uncertainty in the split and cross early cases reduces earlier than in the cross late case, meaning that the uncertainty could reduce once the jet has been crossed. However, the uncertainty remains relatively high in the storms that cross the jet stream late even at short lead times, because the storm has not yet crossed the jet and so the effect that would have on the depth of the storm is not yet determined. The spread is greatest in the case of two jet streams, which could be because the additional jet streak increases the uncertainty.

For latitude (Figure 5.21), the reduction of the IQR with decreasing lead times is again apparent. The spread is generally smaller in the edge and cross-early cases, which can be determined by comparing the IQR of 'typical' lead times (i.e. not those with an unusually large spread, e.g. 48h for the edge category). This indicates the importance of the crossing of the jet stream in reducing the uncertainty, because the uncertainty is lower in the groups that either cross the jet stream early in their development or do not cross it at all, compared to the groups where the storms cross the jet stream late or have two jet maxima to cross.

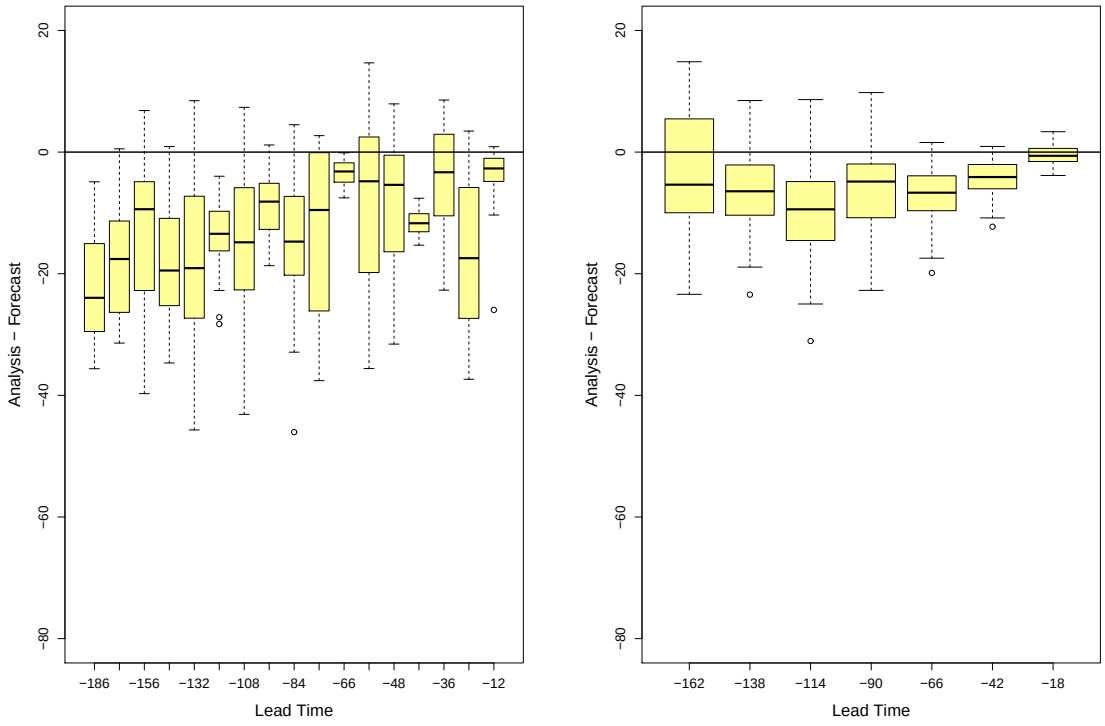
For longitude (Figure 5.22), the forecast spread generally decreases with lead time in all four categories, though some obvious exceptions remain (e.g. 24 hours in the cross late group). The

spread is comparable in all four categories, indicating that the type of jet stream has little effect on the uncertainty in the spread for longitude and therefore speed. This is consistent with the storm's speed being most closely related to the speed of the jet stream, rather than the jet stream type.

Overall, this analysis of the jet stream categories reveals that crossing the jet stream is an important factor when assessing forecast spread. As previously discussed, this could be due to limitations in the model or deficiencies in the initial conditions.

(a) Cross Late

(b) Cross Early



(c) Edge

(d) Split

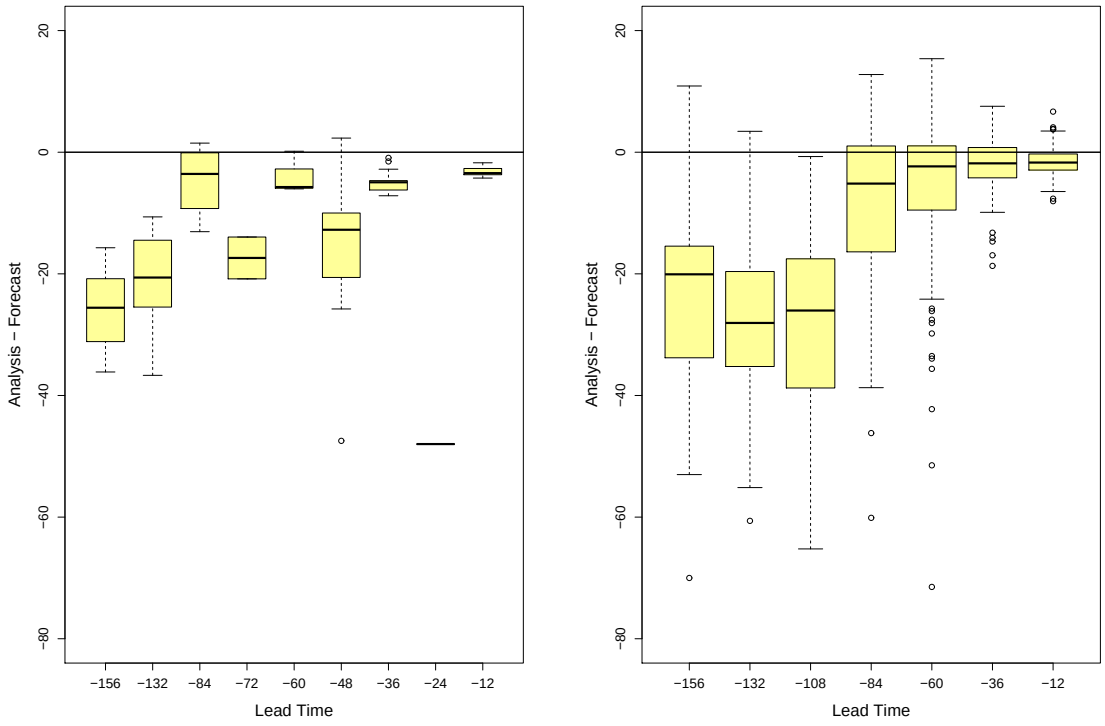
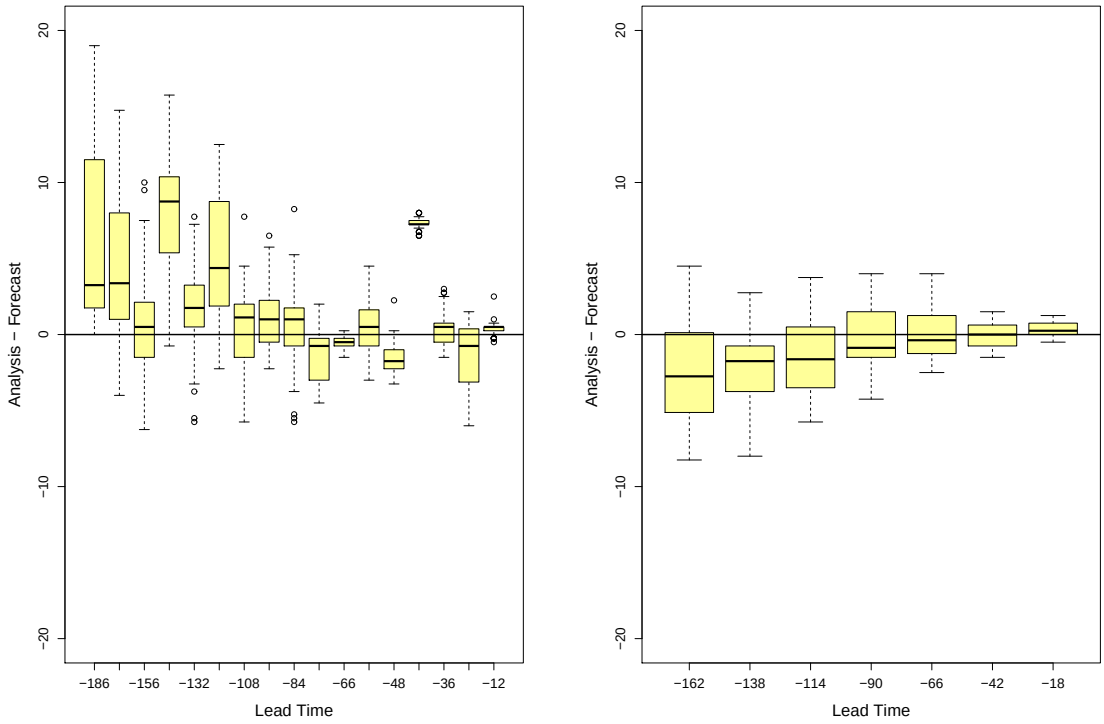


FIGURE 5.20: Differences between the ensemble forecasts and reanalysis pressure, with the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split.

(a) Cross Late

(b) Cross Early



(c) Edge

(d) Split

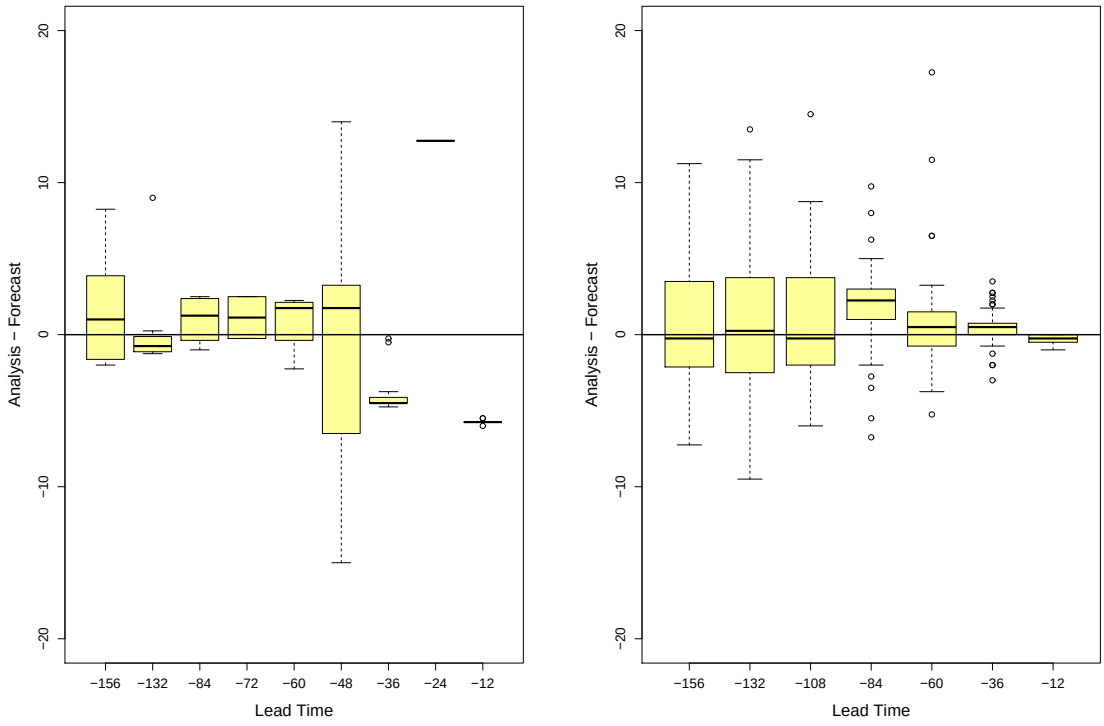
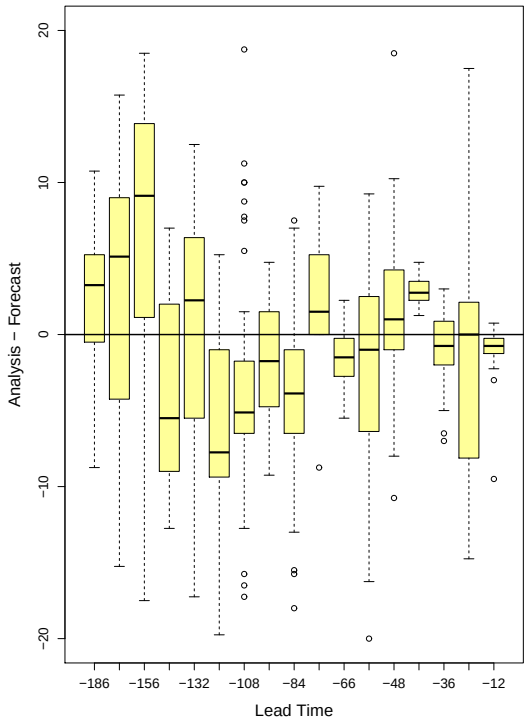
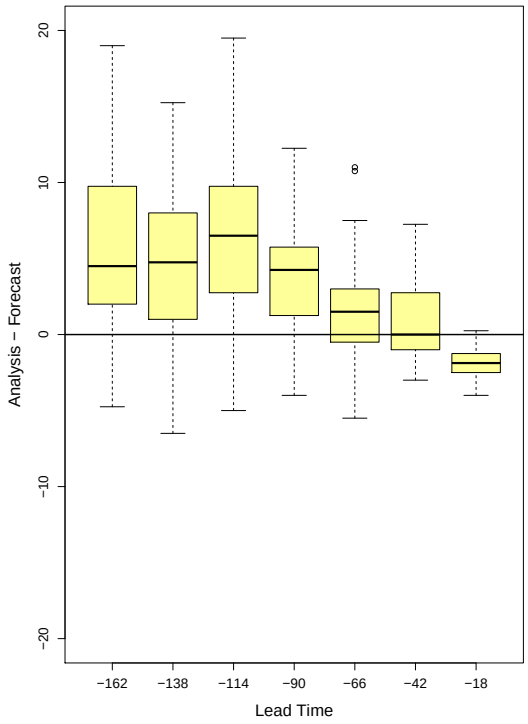


FIGURE 5.21: Differences between the ensemble forecasts and reanalysis latitude, with the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split.

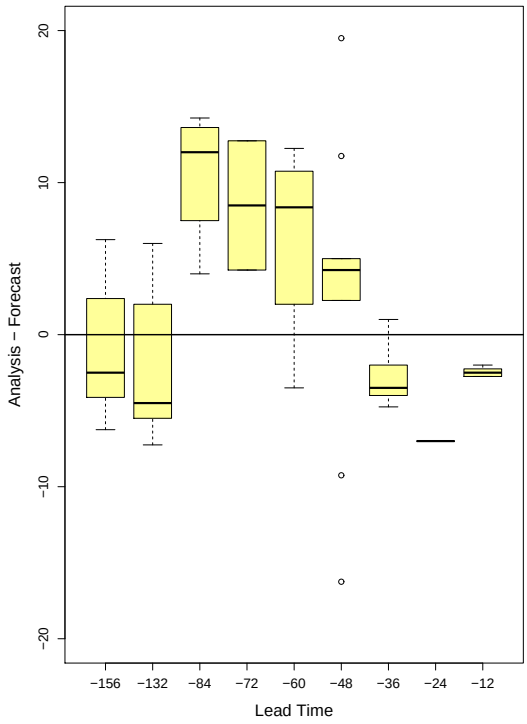
(a) Cross Late



(b) Cross Early



(c) Edge



(d) Split

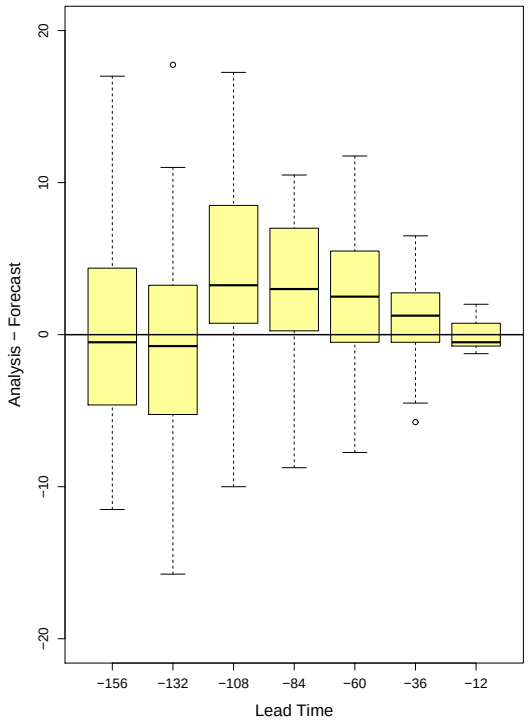


FIGURE 5.22: Differences between the ensemble forecasts and reanalysis longitude, with the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split.

5.5.3 PTE Category

The forecast spread of the two categories identified through PTE analysis (‘horiz’ and ‘diab’) will now be examined. It is noteworthy that there are more horiz-type storms compared to the number of diab-type. As with the jet stream categories, Figures 5.23, 5.24 and 5.25 show a general trend towards smaller spreads at shorter lead times. However, this is not a smooth progression, because as discussed previously, the spread can be increased when some storms’ forecasts are particularly uncertain. Therefore, when comparing the IQR of the two types of storm, this should be done at a fixed lead time, but one at which neither type of storm has an anomalously small or large spread, given the broad trend.

(a) Horiz (b) Diab

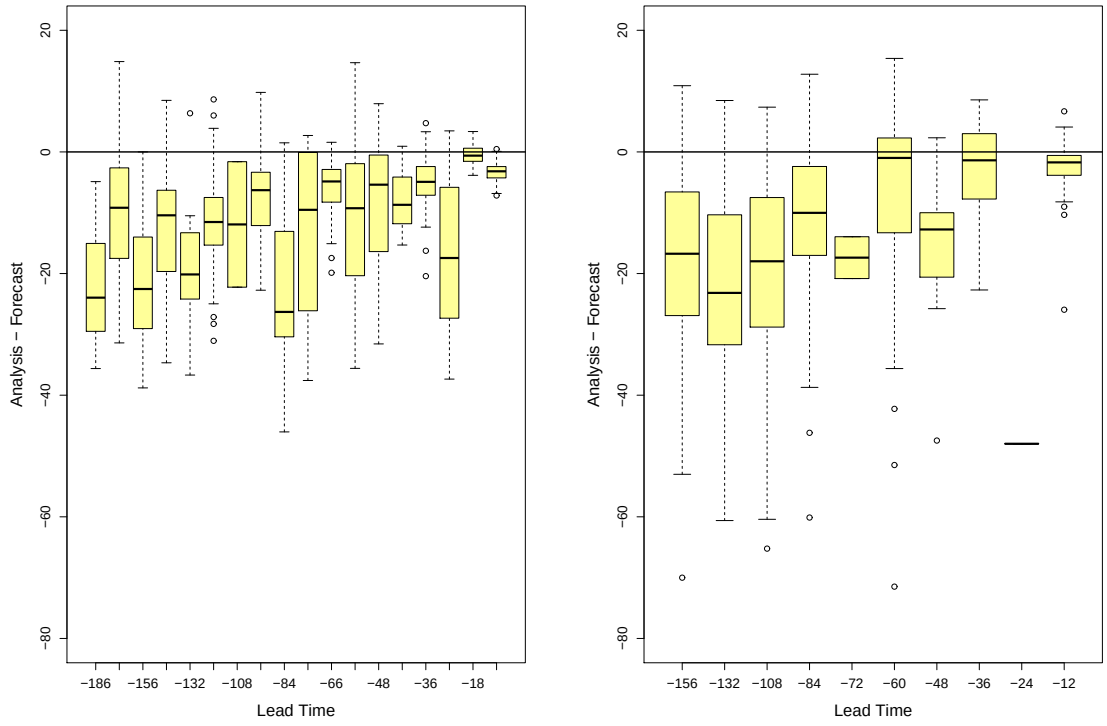


FIGURE 5.23: Differences between the ensemble forecasts and reanalysis pressure, with the storms divided by the terms that dominate the PTE: (a) ‘horiz’, (b) ‘diab’. Please note that the y-acis scales’ differ, in order to better present the data.

To concentrate on pressure (Figure 5.23), the forecast spread for intensity is greater in the diab-type storms, compared to the horiz-type storms, when the IQRs are compared. This is unexpected, because larger populations typically have a larger range and IQR. This implies that there is particularly large uncertainty for diabatially driven storms, compared to baroclinically driven storms.

Therefore, the forecast models' notable uncertainty in the forecasts of diabatically-driven storms' intensity is consistent with a limitation in the simulation of diabatic processes. However, for latitude (Figure 5.24) and longitude (Figure 5.25), the spreads are similar. Therefore, any limitations in the simulation of diabatic processes affect the intensity significantly, but this is not so for the location of the storm.

(a) Horiz (b) Diab

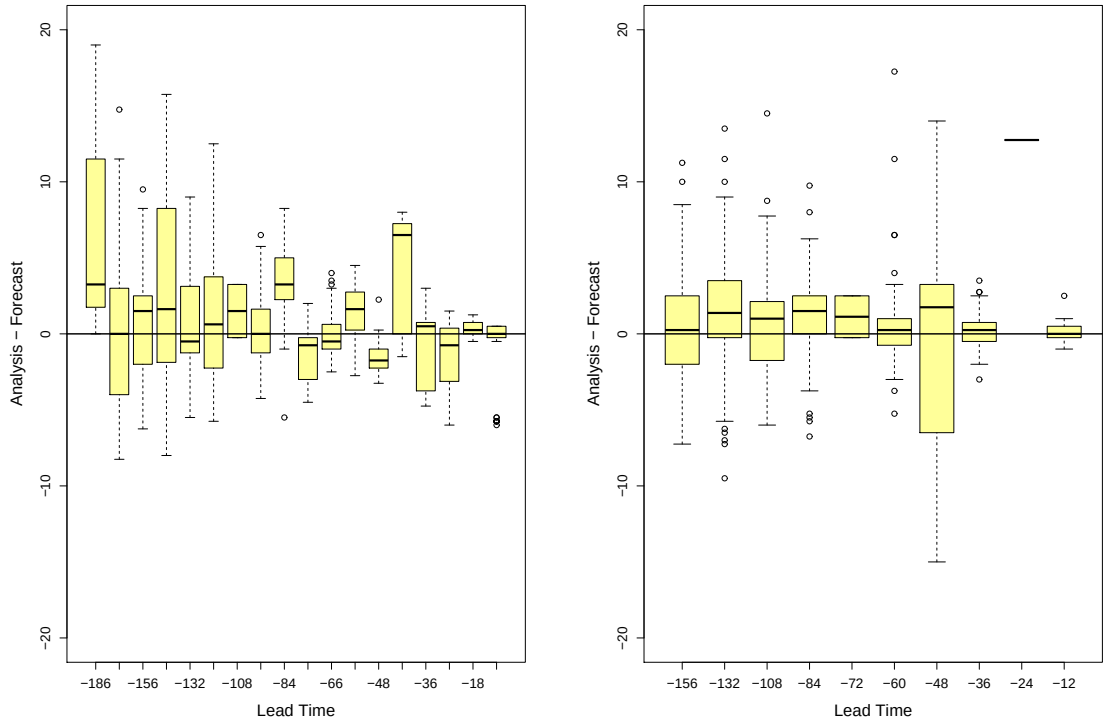


FIGURE 5.24: Differences between the ensemble forecasts and reanalysis latitude, with the storms divided by the terms that dominate the PTE: (a) 'horiz', (b) 'diab'.

5.5.4 Model Resolution

The effect of model resolution on forecast spread will now be explored. This is done for the three highest resolutions of the ensemble data available, of T_L255, T_L399, and T_L639. While ensemble data is available at lower resolutions, the temporal resolution for the ensemble members is insufficient for the automatic storm tracker and so cannot be analysed here.

In terms of pressure (Figure 5.26), increasing the resolution does reduce the spread of the forecasts, again by comparing the IQRs of similar lead times (note the difference scales on the y-axis).

(a) Horiz (b) Diab

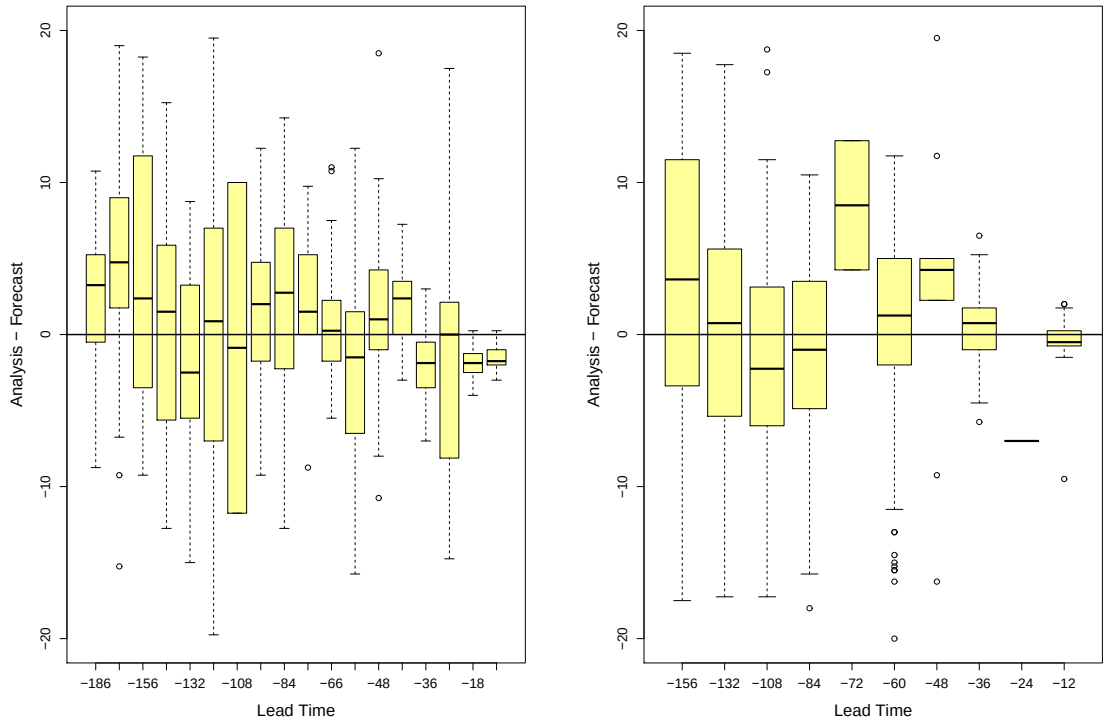


FIGURE 5.25: Differences between the ensemble forecasts and reanalysis longitude, with the storms divided by the terms that dominate the PTE: (a) 'horiz', (b) 'diab'.

However, the transition from large spread at long lead times to smaller spread at shorter lead times is not smooth, especially at the lowest resolution. This indicates that some storms were well forecast and others were not, and that these occurred at different times of day and so have different lead times, as already discussed. That it is so bad at the lowest resolution is not surprising, because the forecast model has been improved considerably since then to perform more reliably under a range of circumstances.

This irregularity is also seen in latitude (Figure 5.27) and longitude (Figure 5.28). The decrease in spread with increasing resolution is only clearly apparent in latitude, when the resolutions of T_L255 and T_L399 are compared. There is only one storm simulated at T_L639 (Table 5.3), which is too few to make any firm conclusions as to whether this trend continues. However, the overall impression is that improving the resolution does decrease the spread of the forecasts, though more clearly in storm intensity (pressure) than location (latitude and longitude).

(a) T_L255 (b) T_L399 (c) T_L639

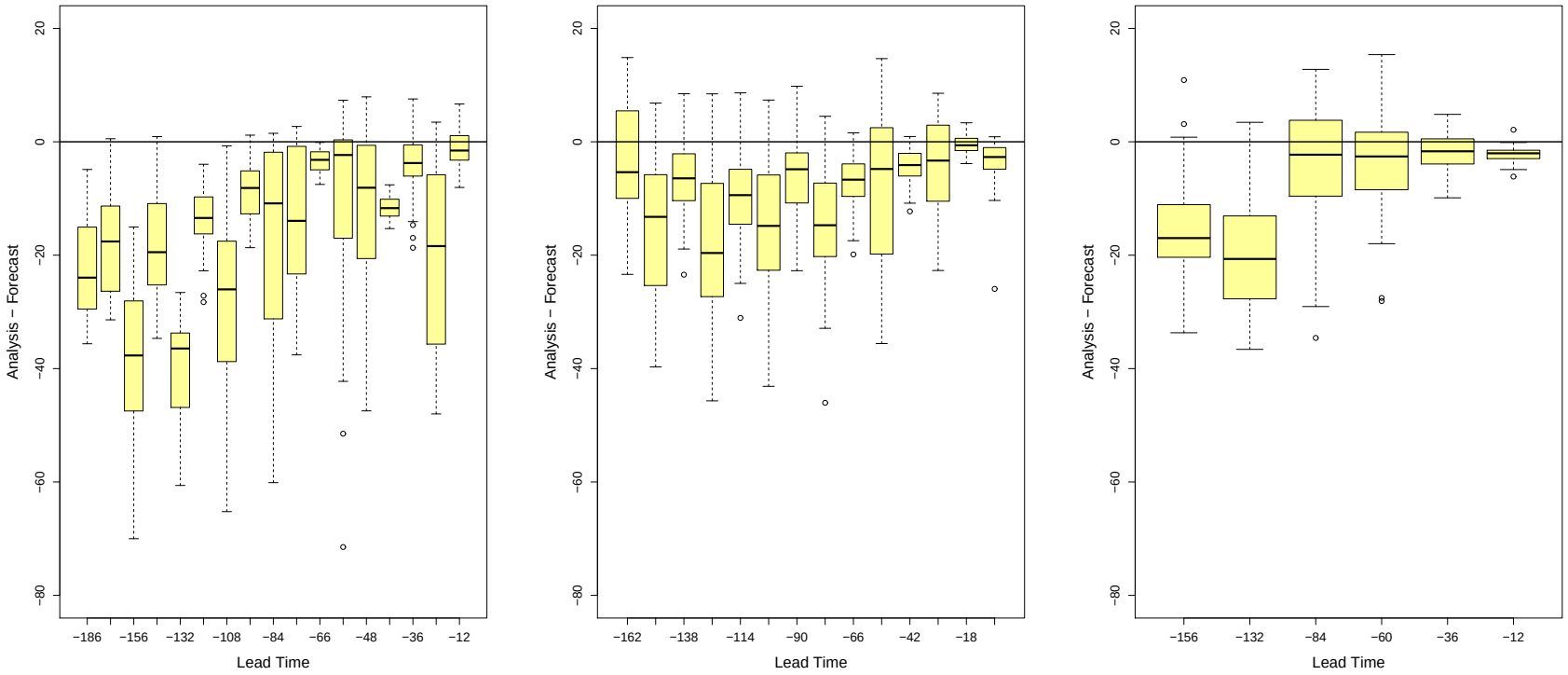


FIGURE 5.26: Differences between the ensemble forecasts and reanalysis pressure, with the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639.

(a) T_L255 (b) T_L399 (c) T_L639

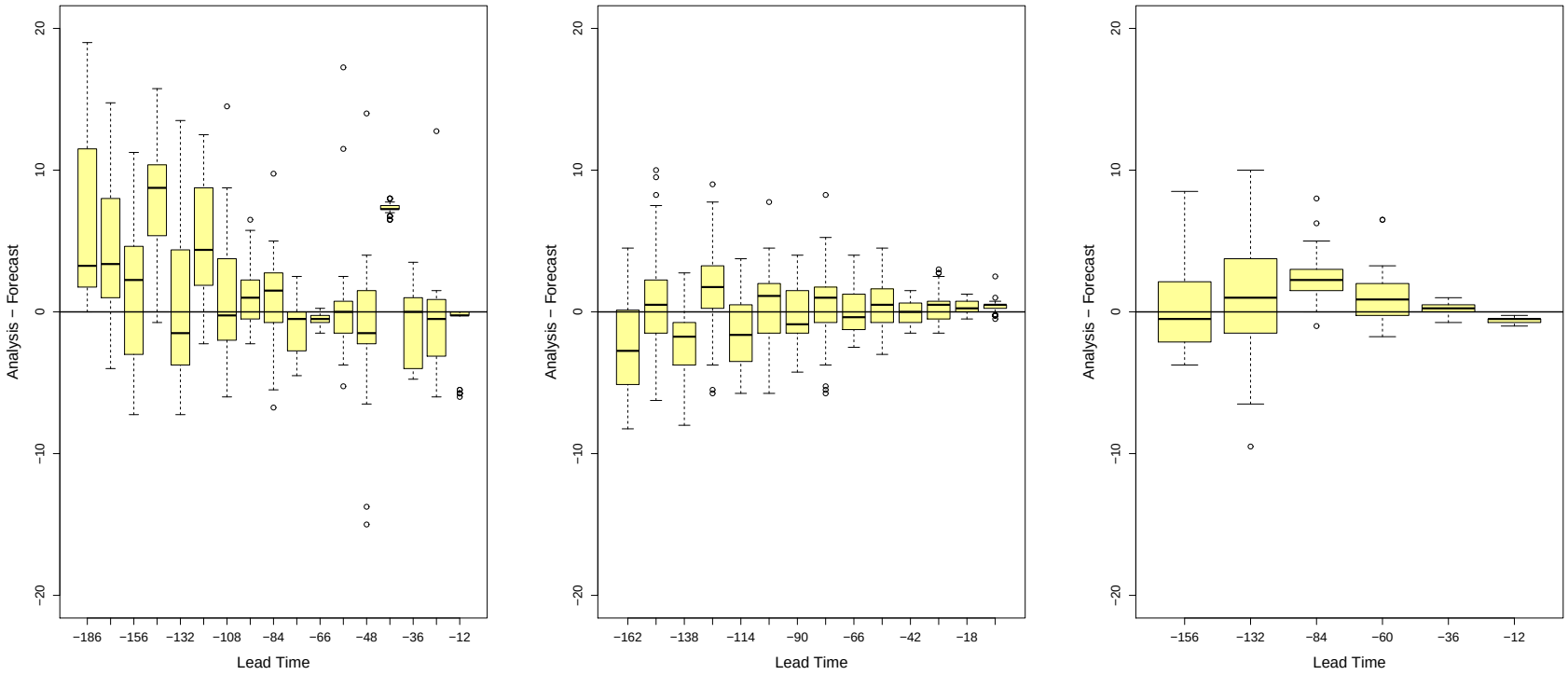


FIGURE 5.27: Differences between the ensemble forecasts and reanalysis latitude, with the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639.

(a) T_L255 (b) T_L399 (c) T_L639

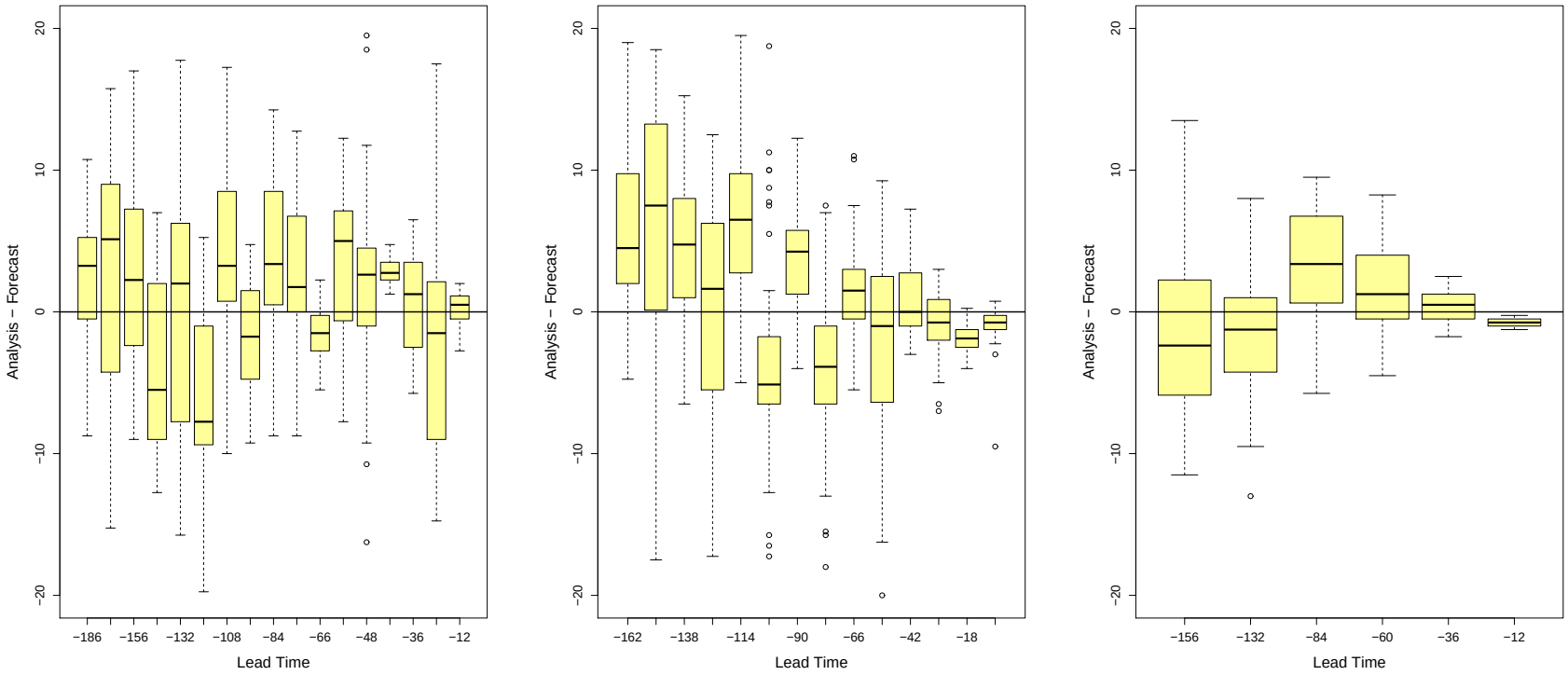


FIGURE 5.28: Differences between the ensemble forecasts and reanalysis longitude, with the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639.

5.6 Summary

These results evaluate the chosen storms' predictability in terms of forecast quality and spread. This chapter is divided into three sections; those that come from deterministic forecasts, those from ensemble forecasts, and those using the analysis of a 'simple metric' developed here. The deterministic forecasts provide insight into the quality of the forecasts, which could be considered as the forecast accuracy because they quantify how close the forecast is to the 'true' value, obtained from reanalysis. The simple metric is the difference between the forecast and analysis at 24 hours lead time, calculated using linear regression. Linear is best type of regression to perform, as discussed in Section 5.3.2. The ensemble forecasts inform about the forecast spread, which gives an idea of the uncertainty in the forecast. The results of each of these three techniques will now be discussed, in terms of how they behave when all of the storms are considered, as well as the different groupings of the storms by jet stream type, the output from PTE analysis, and the native resolution of the model used to make their forecast.

5.6.1 All Storms

When looking at all of the storms in deterministic forecasts, intensity in terms of core pressure is underforecast. The analysis for longitude indicates that the forecast storms are generally too slow. For latitude, the results show that storms are forecast too far south. All of these deficiencies decrease at shorter lead times, so these are useful measures of forecast quality. This is consistent with the work of Froude (2010), though the approach differs because fewer storms are examined in this analysis. Furthermore, this work considers winter storms (October - March), whereas Froude considered storms between February and July 2008.

In this chapter, the simple metrics are used to investigate correlation. Correlation is found between $\Delta(lon)$ and $\Delta(lat)$, which is as expected because these two metrics quantify how well a storm's position is forecast. Correlation is also found between $\Delta(p)$ and $\Delta(lon)$. $\Delta(lon)$ is related to the speed at which the storm moves, because North Atlantic cyclones tend to move from west to east. Therefore, this correlation indicates that storms where speed is poorly forecast tend to be those where intensity is also poorly forecast. This could be because both are affected by cyclogenetic factors such as the strength of the jet stream. Ensemble spread generally becomes smaller at shorter lead times in all three metrics. However, this occurs irregularly, because different storms are included at different lead times. If the forecast of one storm has a particularly large or small

spread, then this will affect the spread of those lead times in which it is included. Correlation between $\Delta(lon)$, $\Delta(lat)$ and $\Delta(p)$ and storm intensity (measured by SSI) is also investigated; however, little is found.

5.6.2 Jet Stream Type

The next portion of the work examines the data in terms of the jet stream types identified in Chapter 4. Analysis of the deterministic forecasts shows that the quality of intensity and longitude forecasts improve once the jet is crossed, but that the jet stream type has little effect on the latitude forecast. When the three simple metrics are considered, $\Delta(lon)$ and $\Delta(p)$ have similarities in which types of storm are better or worse forecast, but $\Delta(lat)$ is different from them. This is consistent with the different facets of the jet stream affecting different properties of the storm. While the storm track's latitude is linked to the position of the jet stream, the storm's intensity and speed (related to longitude) is linked to the jet stream's speed.

The ensemble forecast spread reduces in pressure and latitude earlier in the storms where the jet is crossed early, compared to those in the cross late category. This is consistent with the model having particular limitations in simulating the interaction of the jet stream and the storm when the storm is crossing the jet. This is unsurprising, because complex non-linear processes are causing the storm to deepen very rapidly at this point, meaning that the effect on the final intensity of the storm is uncertain. The rapid deepening also allows any errors in the initial condition to propagate rapidly. For latitude, jet crossing affects the quality but not the spread of the forecast. For longitude, little difference is seen in either quality or spread between the categories, which could be because jet speed is key, not jet type.

5.6.3 PTE category

The results are more difficult to analyse for the PTE categories, partly because there are many more storms where the deepening is dominated by horizontal temperature advection ('horiz'-type storms) than those where diabatic processes dominate ('diab'-type). The deterministic forecasts show that horiz-type storms are less well forecast than their diabatic counterparts, in both intensity and latitude. Little difference is seen in longitude. In the ensemble forecasts, the spread is greater in forecasts of diabatic storms than horiz-type storms, for intensity, but little difference is seen in

position. In short, for intensity, 'diab' storms have higher forecast quality but larger forecast spread than 'horiz' storms. There is an indication of this in the 'simple metric' approach too; however, any differences are small.

There are three potential reasons for the differences in forecast quality and spread between the two categories. Firstly, there could be errors in the NWP model, so improving simulations of the large-scale horizontal temperature gradient or the baroclinic processes could significantly improve the forecast quality of horiz-type midlatitude cyclones. Secondly, there could be rapidly-growing uncertainties in the initial conditions of the forecast, which improving the number or quality of observations or the data assimilation scheme would would reduce. Finally, the diabatic processes could seem to be better simulated than they are in reality, because their parametrisation in the NWP model and the model used to generate the reanalysis is the same, meaning further investigation of diabatic parametrisation schemes could be needed. Therefore, modelling of such storms as those examined here could be improved by improving any one of these three aspects of the modelling; however, deciding which would prove most fruitful is left for future work by modelling specialists.

5.6.4 Model Resolution

The final grouping for the selected storms was the native model resolution at which the forecast was made. In the deterministic forecasts, intensity is better forecast at higher resolution, but there is little effect for latitude or longitude. In the ensemble forecasts, the 'irregularity' of the spread reduces at higher resolution, which could indicate more reliable forecasts. With the simple metrics, calculating correlation directly reveals that increasing resolution significantly improves the quality of forecasts for intensity (pressure) and latitude, but no significant improvement is seen for longitude. Overall, improvements in the forecast with resolution are seen more clearly in storm intensity (pressure) than location (latitude and longitude). Therefore, future increases in model resolution are likely to improve storm intensity more than location. This result should be considered alongside the effect that model resolution has on climate model simulations (Section 2.5.3). Improving the resolution of climate models also seems to improve the location of the storm track (Stratton, 2004). This thesis shows that improving model resolution also improves storm intensity, and so continuing to increase model resolution will improve climate projections of midlatitude cyclones.

Chapter 6

Storm-Prone Situations

This portion of the work aims to identify useful measures for storm-prone situations (SPSs); that is, a method for quantifying the large-scale situation before a storm develops. This could prove useful for future work that aims to improve forecasts and model simulations, because it breaks down the potential sources of error. It could be that models are limited by their generation of SPSs, in that there are not enough or they are not sufficiently intense. On the other hand, the models could generate SPSs, but then fail to generate a storm from it. Therefore, the link between the SPS metrics and the types of storm will be explored, to see what types of storms are most strongly related to which SPS. The search for correlation between the SPS metrics and storm predictability will illuminate how well forecast the storms associated with strong SPSs are.

The growth rate is proposed as a candidate for a storm-prone situation for these severe midlatitude cyclones. The Eady growth rate, first developed by Lindzen and Farrell (1980), has been used in many studies (e.g. Greeves et al., 2007) because it quantifies the baroclinicity of the atmosphere. It has often been used as a measure of 'storminess' in climate models, whereas here it is used in a more short-term context by relating it to the selected storms. This will aid discussion of whether the growth rate is a reasonable measure of storminess.

However, the Eady growth rate does not consider the effect that moisture can have on the instability, because the effects of latent heating change the potential that the atmosphere has to release potential energy into the kinetic energy of a cyclone. Therefore, the present work assesses whether additionally considering moisture captures more storms than just using the dry case, building on

the work of Whitaker and Davis (1994) and Emanuel et al. (1987). The current work also considers the contribution from barotropic instabilty (Gill, 1982), which describes the kinetic energy of the large-scale flow. This can be converted into the kinetic energy of the cyclonic eddy through the release of barotropic instability.

First, a thorough grounding in the theory (Section 6.1) and methods (Section 6.2) used in this Chapter is presented. Within the latter, sensitivity tests will be performed for heights of the levels used in the calculation (Section 6.2.1.1), the area used for averaging (Section 6.2.1.2), and some possible ways of taking moisture into account (Section 6.2.1.3). The latter half of this Chapter discusses the results (Sections 6.3 to 6.9), before summarising the findings (Section 6.10).

6.1 Theory

Broadly, there are three major sources of energy for extratropical cyclones: baroclinic instability, barotropic instability, and diabatic processes. As discussed in Section 2.2.4.1, the first two are well-illustrated by the Lorenz Energy Cycle (Figure 2.6). Baroclinic instability describes a conversion from the large-scale potential energy of the system to the kinetic energy of the eddy. Barotropic instability describes a conversion, from the large-scale kinetic energy of the system to the kinetic energy of the eddy. It is important to note that, as discussed in Section 2.2.4, these are necessary but not sufficient conditions for a midlatitude cyclone to develop: a small perturbation must be initiated, which these processes can then encourage to grow into a mature cyclone.

Lindzen and Farrell (1980) revisited the model proposed by Eady (1949) (Section 2.2.4.3), to investigate the patterns of growth within the fastest-growing wave, and proposed the formula for its growth rate:

$$\sigma_{qgd} = 0.3125 \frac{f}{N} \frac{dU}{dz} \tag{6.1}$$

where f is the Coriolis parameter, U is the total wind speed, z is the geopotential height, and N is the Brunt–Väisälä frequency given by:

$$N^2 = \frac{g}{\theta_0} \frac{d\theta}{dz} \tag{6.2}$$

where g is the acceleration due to gravity, θ_0 is the potential temperature at the surface, θ is the potential temperature, and z is the geopotential height. N is a measure of atmospheric stability, having positive values when the atmosphere is unstable, givbing larger growth rates where the atmosphere is unstable. Equation 6.1 shows that growth rates are faster with larger vertical wind shear (dU/dz), so the air in the upper layer is moving faster, relative to the bottom layer. Physically, this means that the jet stream is stronger than the surface winds, which agrees with observations that cyclones with a strong jet stream develop more quickly, and with the ideas discussed in Section 2.2 of the jet stream's contribution towards the deepening process. The thermal wind relationship (Lynch and Cassano, 2006, Equation 6.13) means that the vertical wind shear is strongly related to the horizontal temperature gradient and therefore to baroclinicity.

6.1.1 Moisture

Including moisture is broadly expected to increase the growth rate of cyclones (Emanuel et al., 1987). The condensation of water leads to the release of latent heat and so provides a further source of energy for cyclonic deepening. Previous studies have investigated the large-scale θ_e field, as a proxy for available moisture, normally on a case-study basis; this approach is taken in Chapter 4. However, investigating the effects of moisture within the storm's environment can be improved by including its effects on the growth rate implicitly.

Whitaker and Davis (1994) developed the idea of moist atmospheric stability, because the degree of stability that moist air has differs from that of dry air. An air parcel of similar temperature to its surroundings but containing water vapour could undergo condensation if pressure and temperature allow, thereby releasing latent heat and warming the air parcel (Finn, 1993, Equation 9.10), allowing the air parcel to rise further than if it were dry. Whitaker and Davis (1994) proposed changing the dynamical equations by replacing a measure for the dry stability (N) with moist (N_m) in ascending air. A similar replacement is done here to calculate a moist Eady growth rate, σ_{qgm}:

$$\sigma_{qgm} = 0.3125 \cdot \frac{f}{N_m} \frac{dU}{dz} \tag{6.3}$$

where f is the Coriolis parameter, U is the total wind speed, z is the geopotential height, but N_m is the moist stability given by:

$$N_m^2 = \frac{g}{\theta_0} \frac{\theta \Gamma_m}{\theta_e \Gamma_d} \frac{d\theta_e}{dz} \tag{6.4}$$

where g is the acceleration due to gravity, θ_0 is the potential temperature at the surface, Γ_m is the moist adiabatic lapse rate, Γ_d is the dry adiabatic lapse rate, θ_e is the equivalent potential temperature, and z is the geopotential height (for the formulae for Γ_d, Γ_m and θ_e, see Appendix B).

The inclusion of moisture leads to a larger growth rate than in the dry case. As latent heat is released, the thermal energy of the system increases. This means there is more energy available to contribute towards the deepening of the cyclone. However, there is one difference in the replacement made here compared to that of Whitaker and Davis (1994). In their model, the moist growth rate is only used where the air is moving upwards, and the dry growth rate is used in downdraughts. In this work, the replacement does not depend on vertical velocity. The idea was developed in an idealised model, with broad regions of updraught and downdraught caused by large-scale pressure systems. If the idea of different growth rates in different vertical velocity regions were to be applied to observations, then smaller-scale features could dominate the picture and make the results less clear. The aim here is to examine scales broader than cells of updraught and downdraught, so the dependence on the direction of vertical wind is removed.

6.1.2 Semi-geostrophic Model

Whitaker and Davis (1994) also note that Emanuel et al. (1987) had developed some of the ideas to include moisture, but for a semi-geostrophic model. The semi-geostrophic equations are more complex, as they allow ageostrophic circulations, which can feed back into frontogenesis. Therefore, the model is particularly useful in regions of strong baroclinicity, such as where fronts are developing. The calculation of the growth rate in the semi-geostrophic model is significantly different to that in the quasi-geostrophic model, because the consideration of vorticity differs between the semi-geostrophic and quasi-geostrophic models. The quasi-geostrophic equations contain either vorticity or stability (Holton, 2004, Equations 6.38, 6.39, 6.40), whereas in the semi-geostophic model stability and vorticity are multiplied together to form potential vorticity (PV, Section 2.2.1). The quasi-geostrophic model's strengths lie in its ability to simulate the large-scale development of midlatitude systems, whereas the semi-geostrophic model performs better on a slightly smaller scale and, in particular, describes frontogenesis well. Given that diabatic processes occur mainly

in frontal regions, both models will be considered here. The semi-geostrophic dry growth rate is included for comparison with its moist counterpart.

The co-ordinate system of the semi-geostrophic model also differs from that in the quasi-geostrophic model; the latter is in normal Cartesian co-ordinates, but the former involves a transformation into geostrophic co-ordinates. Therefore, there must be a transformation back into Cartesian co-ordinates during the calculation of growth rate in the semi-geostrophic system, which mean that the coordinate system moves along with the geostrophic flow. The geostrophic co-ordinates are therefore described by:

$$X = x + \frac{v_g}{f}; Y = y - \frac{u_g}{f}; Z = z; T = t \tag{6.5}$$

where the capital letters refer to distance, height and time in geostrophic co-ordinates; small letters to the same quantities in Cartesian co-ordinates; u_g and v_g to the geostrophic wind speed (zonal and meridional, respectively), and f is the Coriolis parameter.

Emanuel et al. (1987) developed a growth rate (Equation 6.10) in an idealised semi-geostrophic model. They included the effect of vorticity (η_g), given by:

$$\eta_g = f + \frac{dv}{dx} - \frac{du}{dy} \tag{6.6}$$

where f is the Coriolis parameter; and u and v are the zonal and meridional wind speed. This is then multiplied by the vertical temperature gradient to give quantities similar to PV. There are two versions of this: dry (q_g):

$$q_g = \frac{g\eta_g}{f}\frac{d(\ln\theta)}{dz} \tag{6.7}$$

and moist (q_{ge}):

$$q_{ge} = \frac{g\eta_g}{f}\frac{d(\ln\theta_e)}{dz} \tag{6.8}$$

The essential difference between the two is similar to the difference between N and N_m (Equations 6.2 and 6.4), in that where potential temperature (θ) is used in dry q_g, this is replaced by equivalent

potential temperature (θ_e) in moist q_{ge}. This difference leads to two versions of the growth rate, with a dry version σ_{sgd}:

$$\sigma_{sgd} = 0.586 \cdot \frac{f}{\sqrt{q_g}} \frac{dU}{dz} \tag{6.9}$$

and a moist version σ_{sgm}:

$$\sigma_{sgm} = 1.484 \cdot \frac{f}{\sqrt{q_{ge}}} \frac{dU}{dz} \tag{6.10}$$

The derivation of the constant in Equation 6.10 is based on the ratio, r:

$$r = \frac{\Gamma_m}{\Gamma_d} \frac{q_{ge}}{q_g} = \frac{\Gamma_m}{\Gamma_d} \frac{\frac{dln\theta_e}{dz}}{\frac{dln\theta}{dz}} \tag{6.11}$$

Emanuel et al. (1987) use this ratio as a measure of the relative strengths of the dry and moist temperature gradients (Equation 6.11). Its derivation is based on solving the equations stated in Emanuel et al. (1987, Appendix C) numerically, separately in the updraught and downdraught regions, because r is non-linearly dependent on the direction of the vertical wind. Assumptions need to be made in order to derive this, including that there are only two regions: one of updraught, one of downdraught. In the atmosphere, it is not clear that this assumption is realistic.

Figure 6.1 shows that the maximum growth rate (σ) depends non-linearly on r. As $r \rightarrow 1$, the vertical temperature gradients in θ_e equal those in θ, and so the atmosphere is becoming dry. At the point $r = 1$, the maximum growth rate is $0.586 \times 10^{-5} s^{-1}$, providing the constant in σ_{sgd} (Equation 6.9). As $r \rightarrow 0$, the temperature gradient in θ_e is much lower than that in θ, and so the atmosphere is becoming increasingly moist. At $r = 0$, the maximum growth rate is 1.484, which is then used in the moist version of the growth rate (Equation 6.10). For a discussion of the wavelength of the maximum growth rate (λ) depends on r, see Emanuel et al. (1987).

One subtlety of the approach of Whitaker and Davis (1994) was that they calculated σ_{sgm} in the regions where there are updraughts and σ_{sgd} in regions of subsidence, as condensation and latent heat release will only happen in rising air. However, this calculation is not made in this work. One reason for this is the same as for the σ_{qgm} case; these growth rates are developed in idealised models and not all of the same ideas can be applied to the atmosphere in reality. Furthermore, the

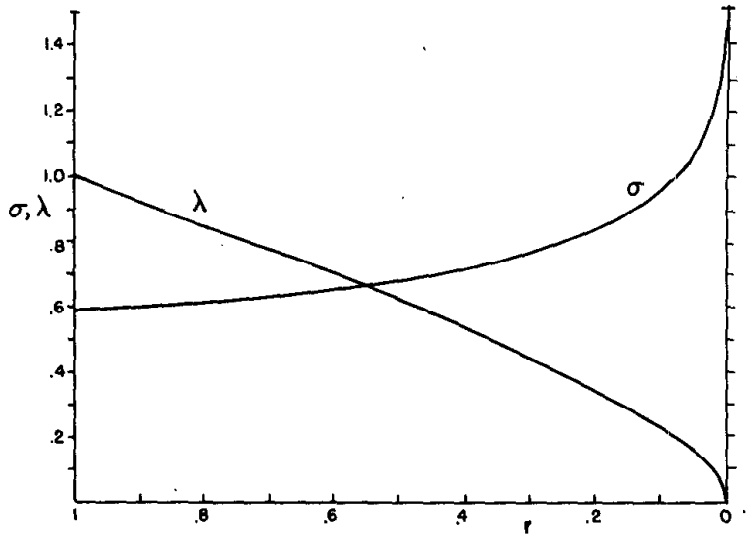


FIGURE 6.1: (Emanuel et al., 1987, Figure 3a), showing how maximum growth rate (σ) and the wavelength of that growth rate (λ) vary with the ratio r (Equation 6.11). Note that $r = 1$ is at the left and $r = 0$ is at the right.

values of σ_{sgm} are much larger than those of σ_{sgd}, in agreement with Emanuel et al. (1987), who find that inclusion of moist processes in this increases the growth rate by a factor of about 2.5. Therefore, σ_{sgd} and σ_{sgm} are not easily comparable because σ_{sgm} is considerably larger, and so σ_{sgm} will be calculated everywhere, not just in the regions of updraught.

In short, there are four baroclinic versions of the growth rate. These are derived from two different models of the atmosphere, quasi-geostrophic and semi-geostrophic, which differ in their treatment of the vorticity. For each model, there is a dry and moist version of the baroclinic growth rate. However, baroclinicity is not the only source of kinetic energy for a midlatitude cyclone.

6.1.3 Barotropic Growth Rate

Barotropic processes convert energy from large-scale kinetic energy to kinetic energy of the eddy. This is in contrast to the baroclinic case, where the energy is converted from large-scale potential energy to eddy kinetic energy. There are conflicting theories as to how barotropic and baroclinic processes interact in midlatitude cyclones. On the one hand, Kucharski and Thorpe (2000) suggested that barotropic energy conversions can encourage a cyclone to develop. Although previous work concentrates on case studies, Kucharski and Thorpe (2001) showed that barotropic growth primes the atmosphere for the growth of baroclinic waves such as midlatitude cyclones. On the

other hand, James and Gray (1986) proposed a mechanism whereby the conversion of baroclinicity into the energy of the cyclone is limited by barotropic processes, known as the 'barotropic governor'. The current work aims to further explore the relationship between barotropic growth rate and midlatitude cyclones.

The barotropic growth rate is derived by considering an atmosphere with uniform meridional shear (i.e. shear in the y direction), which generates vorticity. There is no need to consider temperature, because temperature variation is not applicable to barotropic processes, only to baroclinic ones. Furthermore, because horizontal temperature gradients and vertical wind shear are related by the thermal wind equation (Lynch and Cassano, 2006, Equation 6.13), there is also no need to consider the vertical variation of wind in the barotropic case. The maximum barotropic growth rate is given by:

$$\sigma_{BT} = 0.2012 \frac{dU}{dy} \tag{6.12}$$

where U is the total horizontal wind speed, and dy refers to the distance between to latitudes on which this is calculated (Gill, 1982, Equation 13.6.12). It is important to note that this is very sensitive to the domain chosen, particularly the location of the northern and southern boundaries, because this formula calculates the difference in wind shear across these two boundaries. For consistency, the domain determined using the baroclinic growth rate (Section 6.2.1) will be used for the barotropic growth rate. However, this could prove to be a limitation of using σ_{BT}.

The derivation of σ_{BT} includes the assumption that the Coriolis parameter is constant, and does not vary with latitude as it does in reality (Gill, 1982, Section 13.6). Although the Coriolis parameter quantifies the apparent turning force felt by an air parcel due to the rotation of the Earth, it also quantifies path differences due to the Earth's sphericity. This means that if two similar air parcels are moving at the same velocity but at different latitudes, then in a fixed time, the one at higher latitude will cover a greater percentage of the Earth's circumference. When considering the box over the North Atlantic over which the averaging process is made (Figure 6.2), the Coriolis parameter varies between 0.834×10^{-4} and 1.318×10^{-4}. If we consider two air parcels, moving at the same zonal velocity but one at the upper boundary of the box and one at the lower, then in the time that the one at the upper boundary has covered the entire width of the box, then the one at the lower will have covered only 63% of the lower boundary. Disregarding this difference could prove to be another limitation.

Despite this, σ_{BT} still describes a source of energy that mid-latitude cyclones can tap into, which is not described by its baroclinic counterparts. The conversion of energy from large-scale kinetic to eddy kinetic could prove key in some cases. Therefore, despite the f-plane approach, it will be examined alongside the baroclinic versions.

6.2 Method

Five versions of the growth rate will be used: dry quasi-geostrophic; moist quasi-geostrophic; dry semi-geostrophic; moist semi-geostrophic; and barotropic. For convenience, the abbreviations used are shown in Table 6.1. As discussed in Section 6.1, these five versions encapsulate the different processes that govern the deepening of midlatitude cyclones. If they are to be a successful candidate for SPSs, the metrics' interaction with the selected storms should be consistent, without producing too many null cases (i.e. when there is a peak but no selected storm). Ideally, they peak before the storm develops, when the energy (potential in the baroclinic cases, large-scale kinetic for barotropic) in the atmosphere builds up, and then drop suddenly as the storm passes and converts the large-scale potential or kinetic energy into eddy kinetic energy.

Dry quasi-geostrophic growth rate	Lindzen and Farrell (1980)	Eqn. 6.1	QGD
Moist quasi-geostrophic growth rate	Whitaker and Davis (1994)	Eqn. 6.3	QGM
Dry semi-geostrophic growth rate	Emanuel et al. (1987)	Eqn. 6.9	SGD
Moist semi-geostrophic growth rate	Emanuel et al. (1987)	Eqn. 6.10	SGM
Barotropic growth rate	Gill (1982)	Eqn. 6.12	BT

TABLE 6.1: List of abbreviations used throughout this chapter, for the five different versions of growth rate.

These five parameters are calculated over a large area covering Europe and the North Atlantic, for each winter (October-March) in the ERA-Interim data set (1979/1980 - 2011/2012). The values are then averaged over a box ($35 : 65^{o}N$, $-60 : 20^{o}E$), through which the selected cyclones travel, to provide a time series of growth rates. The sensitivity of the value to the box used is studied (Section 6.2.1.2), but the box used in most calculations is shown in Figure 6.2.

6.2.1 Sensitivity Testing

Once these five metrics were selected as candidates for storm-prone situations, sensitivity testing is performed to determine the optimal way of calculating them. Three factors are important in

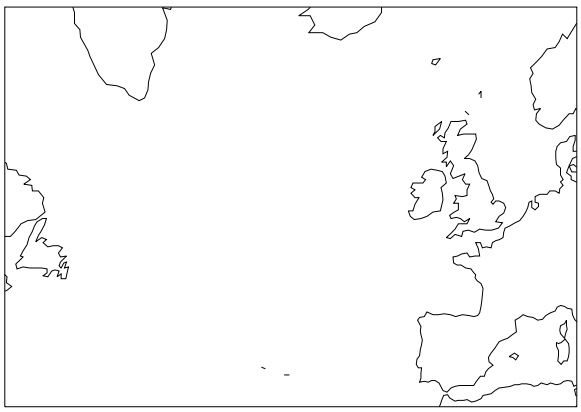


FIGURE 6.2: Figure showing the area over which the four different versions of growth rate are averaged.

these calculations:

- the pressure levels used in the calculation, because QGD, QGM, SGD and SGM contain height derivatives;
- the area box over which the average is taken;
- the different versions of moist temperature available, because the two papers used to derive the moist versions (Emanuel et al., 1987; Whitaker and Davis, 1994) refer to different measures of moist potential temperature.

Each of these will now be discussed. Sensitivity testing is only performed for the baroclinic versions, as once the two levels used in that calculation are selected, then the barotropic growth rate is simply calculated on the level in the middle of the two. In the cases of the height and area testing, the sensitivity is only shown for the QGD versions because QGM, SGD and SGM produce similar results.

6.2.1.1 Height Sensitivity Testing

The four baroclinic versions of the growth rate contain derivatives with respect to height, and the levels used to calculate the derivative will affect the final value of growth rate. This is illustrated in Figure 6.3 for the QGD growth rate. Three levels are used: the top and bottom for vertical

derivatives, and the middle layer for horizontal derivatives. The figure shows that the five different combinations of levels have similar shapes, and at first appear to be shifted vertically. In particular, the 200/500/800hPa combination is significantly lower than the others, possibly due to the lower stratosphere being included in the calculation. This would make the air column appear more stable, than if the upper troposphere were used as the upper limit in the calculation, because the stratosphere is warmer than the upper troposphere. On closer inspection of Figure 6.3, some peaks appear more prominent in one version than another; for example, the peak around 2011-03-05 is the most prominent in the 400/600/800hPa combination, but is less prominent in the 200/500/800hPa version. The 300/600/900hPa combination was used in the remaining calculations, because it describes a large depth of the atmosphere, and therefore the energy stored within it, with minimal interference from the stratosphere.

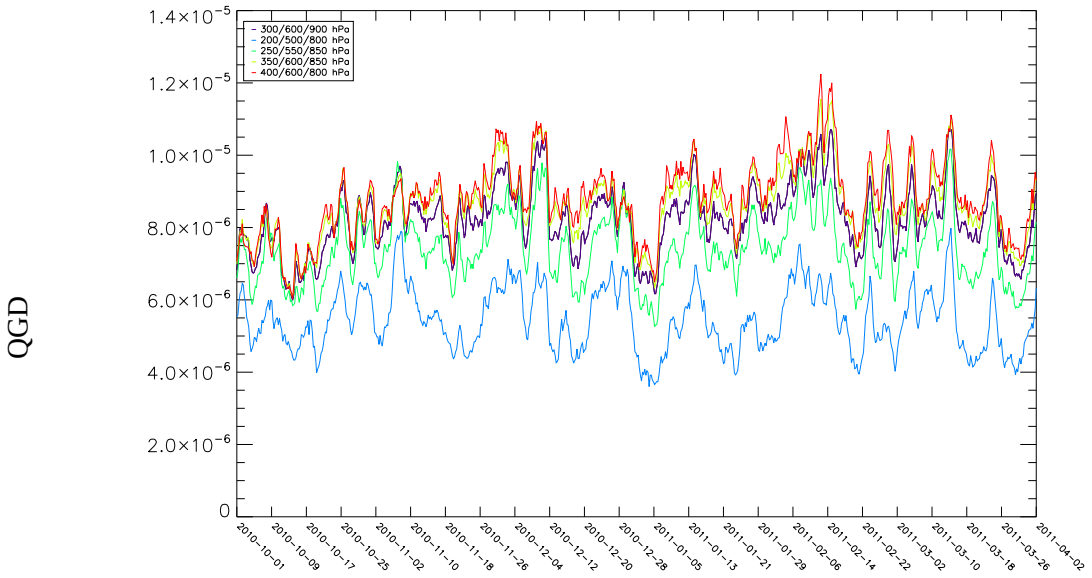


FIGURE 6.3: Time series illustrating the sensitivity of the QGD growth rate [s^{-1}] to different combinations of pressure levels used in the calculation.

6.2.1.2 Area Sensitivity Testing

Sensitivity testing of candidate areas, over which the growth rates are averaged, does not provide such a clear a distinction (Figure 6.4). All of the boxes in Figure 6.4 encompass the box used in the SSI calculation (Figure 3.1), for consistency with the approach used to select these storms. It is clear that all of the areas considered affect the relative prominence of different peaks. This is important for the results, because it is the measure that will be used to determine objectively whether a peak is there or not, and then whether the storms are associated with any such peak. Overall, deciding which area to use from Figure 6.4 is impossible. Instead, two criteria were selected for the box: that it should be as small as possible, in order to avoid averaging out the fine-scale structure (e.g. the jet core); and that it should cover a portion of Europe and the North Atlantic that all of the storms pass over for at least some of their lifetime, particularly the regions where cyclones often develop before moving towards Europe. The tracks of the cyclones in reanalysis data were compared to the box candidates, and it was ensured that they all passed through it for a portion of their lifetime. The area from 35 to 65°N and −60 to 20°E was selected, as it met these criteria.

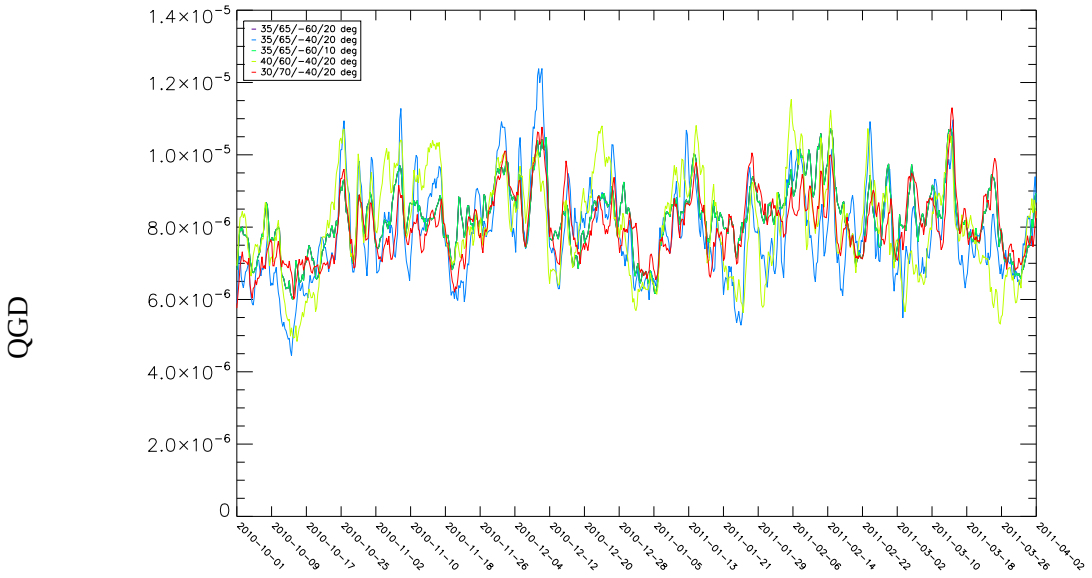


FIGURE 6.4: Time series illustrating the sensitivity of the QGD growth rate [s^{-1}] to different areas, over which the average is performed.

6.2.1.3 Different versions of moist temperature

In the literature, it is not clear which version of moist temperature should be used to calculate vertical temperature gradients in the two moist cases. Whitaker and Davis (1994) specify that saturated equivalent potential temperature (θ_{es}) should be used in the calculation of such gradients, whereas Emanuel et al. (1987) use equivalent potential temperature (θ_e). For consistency, the same measure should be used in the two versions. Therefore, this work studies the sensitivity of using different combinations of θ_e and θ_{es} in calculating SGM. The results of this are shown in Figure 6.5. The different use of θ_e and θ_{es} has an effect, although the $\theta_e - \theta_e$ version is covered by the $\theta_{es} - \theta_e$ version. A shift is the main difference between these two and the $\theta_{es} - \theta_{es}$ version, with the two visible time series having similar shapes. However, the relative prominence of different peaks is also affected slightly. In conclusion, the $\theta_{es} - \theta_{es}$ version in both QGM and SGM versions was chosen, because it was specified for this use in the more recent of the two papers specified (Whitaker and Davis, 1994), and it may be that Emanuel et al. (1987) was saying θ_e but referring to θ_{es}; at the time Emanuel et al. (1987) was published, 'equivalent potential temperature' sometimes referred to θ_{es}.

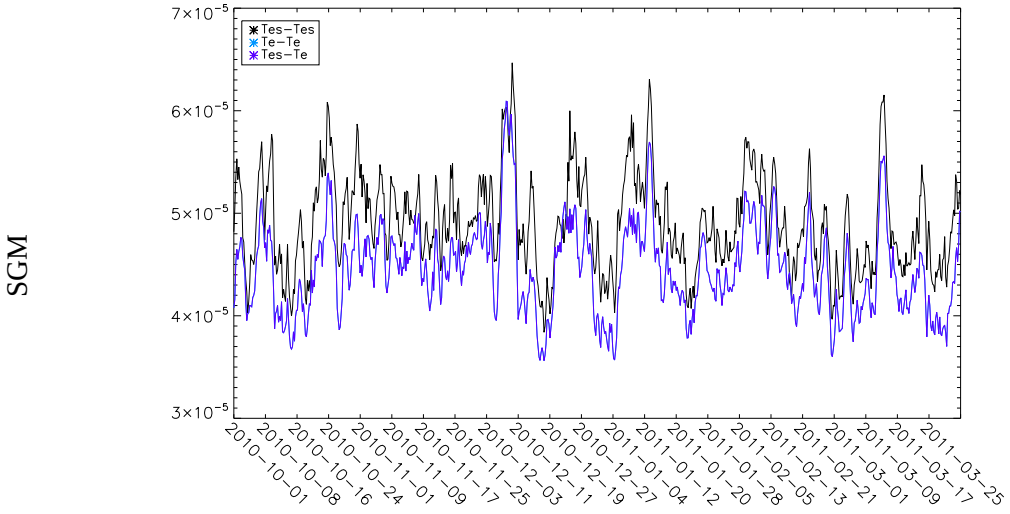


FIGURE 6.5: Time series illustrating the sensitivity of the SGM growth rate [s^{-1}] to different combinations of θ_e and θ_{es} on the vertical levels selected in Section 6.2.1.1. For example, 'Tes - Te' refers to the difference between θ_{es} at 300hPa and θ_e at 900hPa.

6.3 Comparison of Growth Rates

First, this work will investigate the similarities and differences between the five different versions of growth rate. Frequency density plots will allow detection of correlation between the variables, which will then be tested. The aim of this is to ascertain the degree of similarity between the different versions of growth rate. Then the temporal and spatial similarities between the metrics will be considered.

6.3.1 Correlation between the Growth Rates

Correlation tests are performed, and the results are shown in Table 6.2. Correlation is calculated using Pearson's Correlation coefficient, because the data are nearly normally distributed. As applied in Chapter 5, the correlation is tested using the t-test (Equation 3.5), which quantifies whether the correlation is significant given the number of data points (Rees, 2001, Section 14.3). This value is compared to that in a look-up table (Rees, 2001, Table C.5), and so whether it is significant or not is also included in Table 6.2. These values will be examined alongside the frequency density plots (Figures 6.6, 6.7 and 6.8). The frequency density plots illustrate the correlations between the different metrics, shown in Figures 6.6, 6.7 and 6.8. These are similar to histograms but two-dimensional, so the approach taken to analyse them is similar to that for scatter plots. For a high degree of correlation, there should be a high frequency of points along a diagonal line. For positive correlation, this diagonal should have positive gradient; for negative correlation, it should have negative gradient. The correlation between two variables will be weaker if there is spread either side of the diagonal line, or a significant number of outliers. Please note that the colour scales in Figures 6.6, 6.7 and 6.8 differs between the different combinations, so that the fine detail of the spread can be seen in each variable, and that the axis scales differ depending on which variable is under consideration.

Variables		Correlation coefficient	t-test	Significant?
QGD	QGM	0.958	510.3	✓
QGD	SGD	0.916	348.6	✓
QGM	SGM	0.992	1230.3	✓
SGD	SGM	0.853	250.1	✓
QGD	BT	-0.252	-39.8	✓

TABLE 6.2: Test of correlation between variables

The plots in Figures 6.6 and 6.7 clearly show that the four baroclinic versions of growth rate (QGD, QGM, SGD, SGM) all bear a striking correlation to each other. They are all strongly, positively correlated, and when tested it is found that this is significant (Table 6.2). This is expected, because the physical processes driving these four versions are very similar. The differences between the four are based on their treatment of moisture and vorticity, and the former means that their treatment of stability is subtly different. Therefore, it must be the other factor - vertical wind shear - that dominates the picture. It is important to note that the aim of this work is to identify a set of metrics for SPSs that ideally describe different SPSs. Therefore, highly correlated metrics are not desirable, because they could describe the same SPSs. Metrics that are little correlated with each other are likely to describe different SPSs. However, Figures 6.6 and 6.7 show that the degree of correlation between the four baroclinic versions does vary, so none of these metrics will be rejected until their relationship with the selected storms has been explored (Section 6.4).

The strongest correlation is found between the two moist versions (SGM and QGM), indicating that the difference between the quasi- and semi-geostrophic models is less important when equivalent potential temperature (θ_e) is considered, compared to when its dry counterpart (θ) is used. This could be because the difference in treatment of the vorticity is a smaller component when the larger, moist growth rates are considered. However, given the strong correlation also shown by the dry versions in Figure 6.7a, this is a small effect. In addition, the correlation between the dry and moist quasi-geostrophic growth rates is stronger than that between the dry and moist semi-geostrophic growth rates. This is explained by the differences between dry and moist potential vorticity being greater than the difference between dry and moist stability. When the formulae are compared, a clear difference appears of a factor of θ different between the dry stability and potential vorticity, or θ_e between their moist counterparts. In both the quasi- and semi-geostrophic cases, there is a large degree of scatter when comparing the dry and moist cases, indicating that moisture plays an important role in modifying the growth rate. Therefore, considering both the dry and moist versions is key, because they describe different SPSs.

It is also found that there is a negative and statistically significant correlation between the baroclinic versions and the barotropic version, though only the correlation between QGD and BT is shown here. This could imply an inverse relationship between horizontal wind shear and either vertical wind shear or stability. However, when the frequency density plot is examined (Figure 6.8), any correlation between QGD and BT seems to be too weak to indicate any physical relationship between the variables, though it is tested to be significant. This supports the idea of the

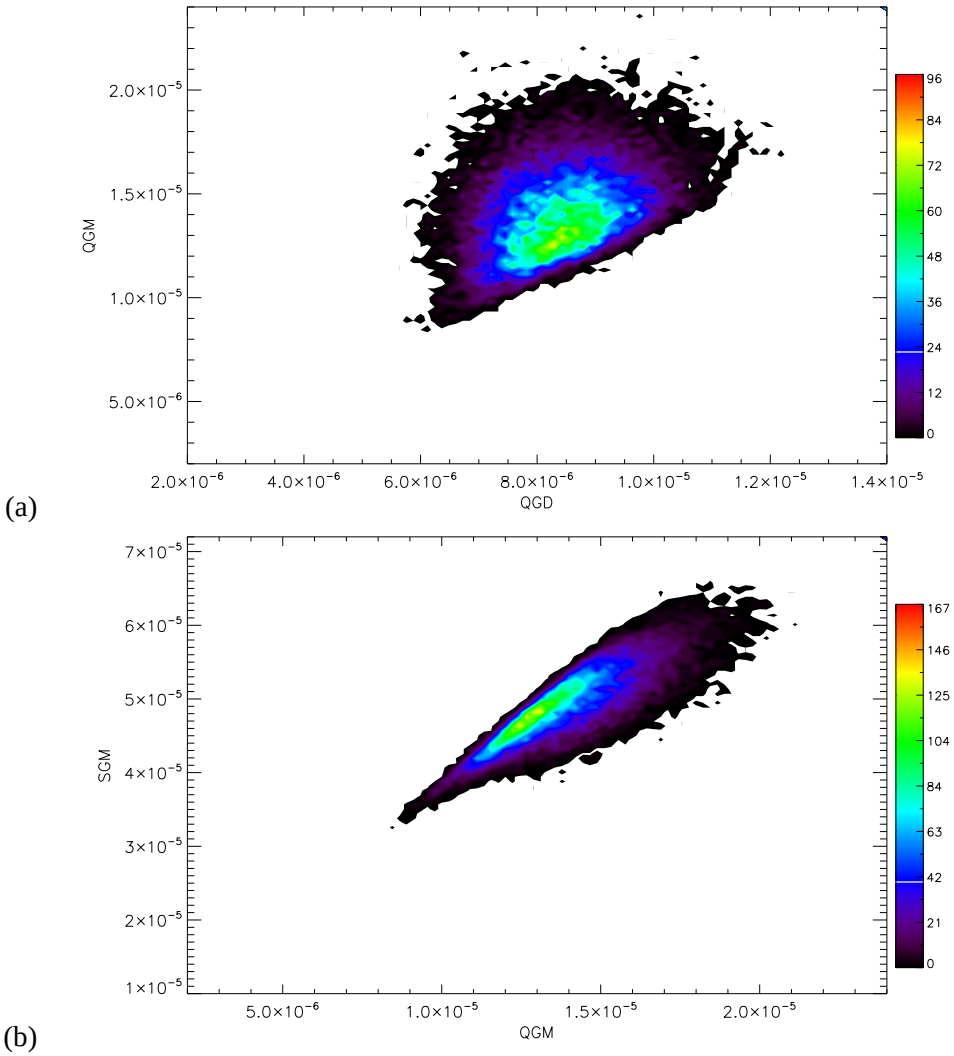


FIGURE 6.6: Frequency density plots for the different versions of baroclinic growth rate $[s^{-1}]$ for the period 1970/1980-2010/2011: (a) QGD against QGM, (b) SGM against QGM.

barotropic governor proposed by James and Gray (1986), because larger barotropic growth means smaller baroclinic growth, as the latter is inhibited by the horizontal wind shear. However, this requires further analysis before it can be concluded definitely.

In summary, there is a strong correlation between the four different baroclinic growth rates. While the degree of correlation depends on which variables are being compared, the correlation is high and positive, and tested to be significant. This indicates that atmospheric stability is the dominant factor, while the treatment of vorticity and moisture are secondary. However, comparing the different growth rates will illuminate the subtleties related to these secondary factors. There is a weak negative correlation between the baroclinic and barotropic growth rates. Most importantly,

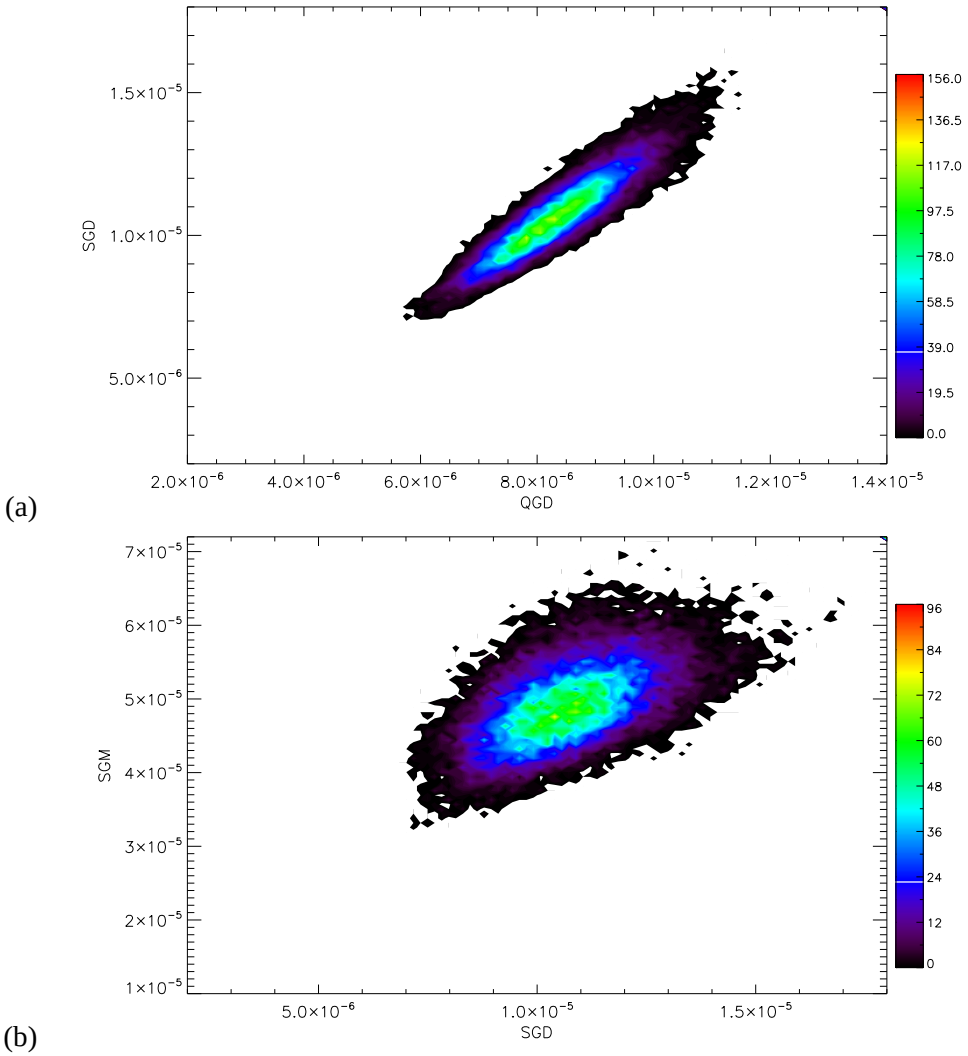


FIGURE 6.7: Frequency density plots for the different versions of baroclinic growth rate [s^{-1}] for the period 1970/1980-2010/2011: (a) SGD against QGD, (b) SGM against SGD.

the lower correlation between the dry and moist versions of the growth rate implies that including moist effects in this way will be of some value, as different SPSs will be represented.

6.3.2 Temporal Variation in the Growth Rates

Time series illustrate how different the five versions of growth rate are, in terms of their evolution over a season. An example time series is shown in Figure 6.9. The four different versions of the baroclinic growth rate have a similar shape, but are shifted and have different amplitudes (Figure 6.9a). This implies that the stability is the main term when considering the shape of the time series, rather than the vorticity or moisture. If the consideration of vorticity was dominant

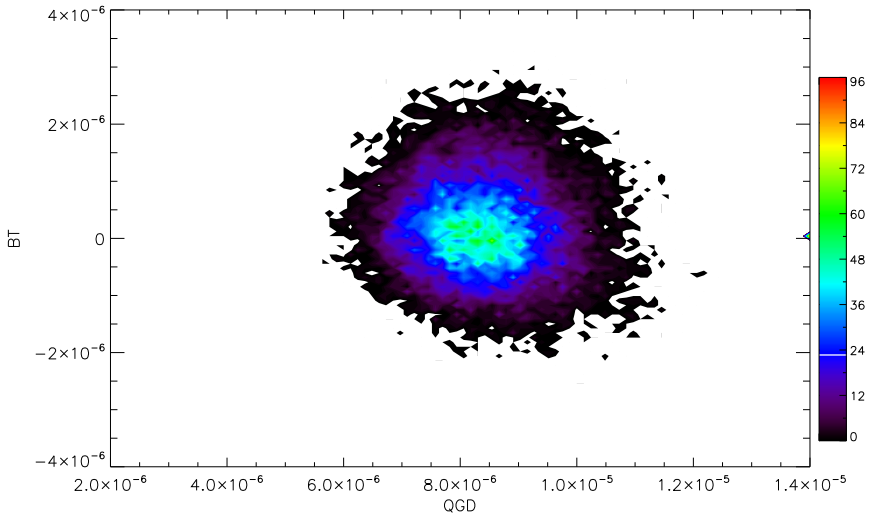
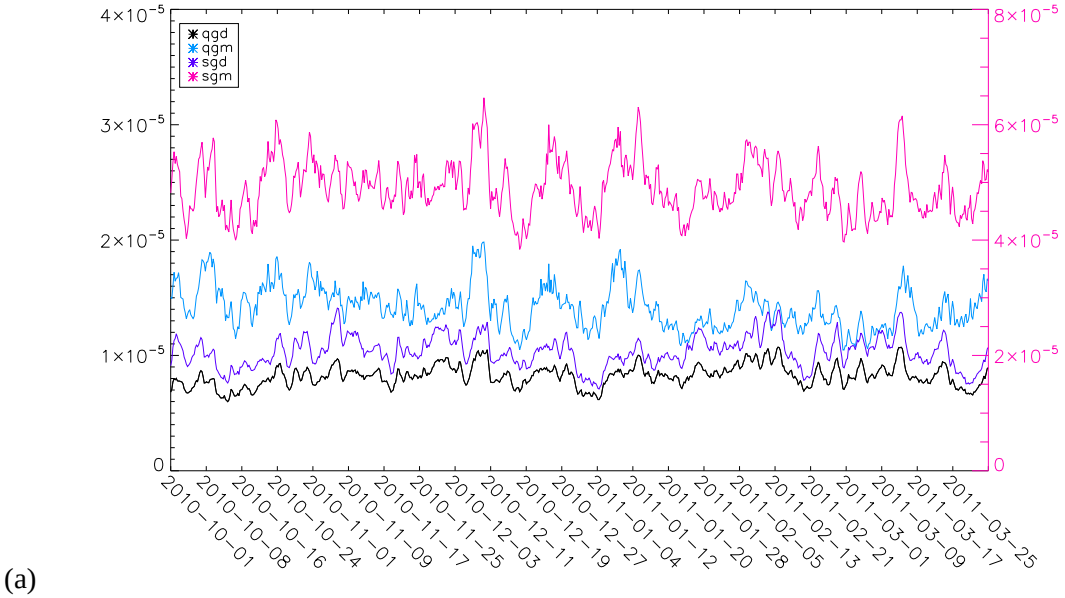


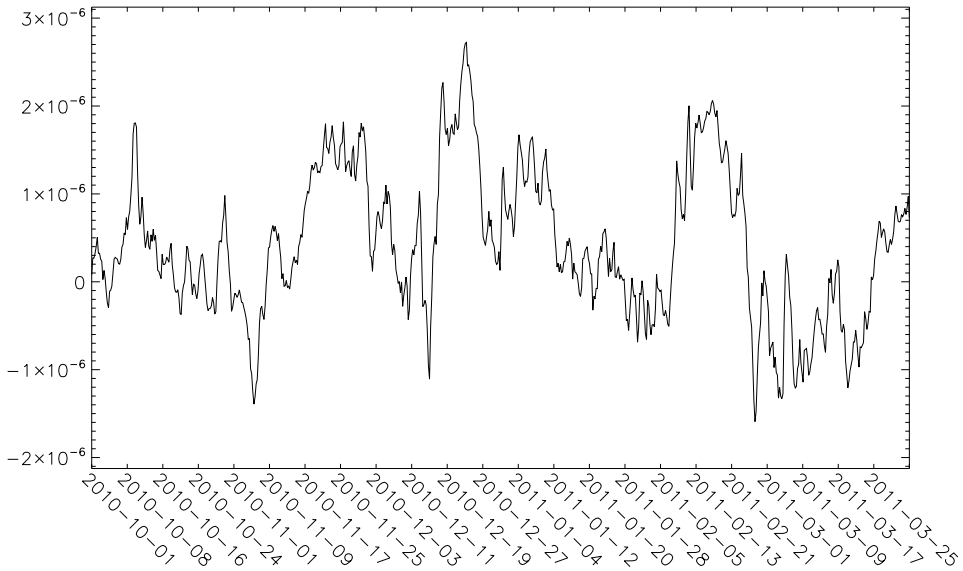
FIGURE 6.8: Frequency density plot of QGD and BT.

then there would be more of a difference between the quasi-geostrophic versions (where only planetary vorticity is considered) and the semi-geostrophic versions (where both planetary and relative vorticity are considered). There are some small differences in the shape of each of the four versions, notable when assessing the prominence of the different peaks. This can be illustrated by the two highest peaks in SGM and SGD, one in November and one in March. The November peak is the most prominent peak in the SGD timeseries, whereas it is the peak in March that is the most prominent SGM peak. Given that the storms should be associated with peaks in these metrics, there small differences are important. Further discussion of this will ensue when considering the different versions at the time when a storm is developing (Section 6.4.1).

The order from smallest to largest illustrated by Figure 6.9a can be explained physically: QGD is the smallest, as it only considers the dry stability and planetary vorticity. The addition of moisture (QGM) increases the values of growth rate, as this means more of the available energy in the atmosphere is included. Alternatively, the additional use of relative vorticity in the semi-geostrophic dry case (SGD) also increases the growth rate, as local vorticity maxima encourage cyclogenesis. The inclusion of both moisture and vorticity in the SGM case means it is the largest. Furthermore, a key part of the derivation of SGM is the constant r (Figure 6.1 and Equation 6.11). When r equals one, the maximum growth rate is the $1.484 \times 10^{-5} s^{-1}$ value used in the calculation of SGM (Equation 6.10). When $r = 1$, then Equation 6.11 implies that the ratio of the temperature gradients is at the idealised limit at all points across the domain. This seems unlikely, as temperature



(a)



(b)

FIGURE 6.9: Time series for 2010-10-01 to 2011-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate [all in s^{-1}]. For (a), QGD, QGM and SGD are all plotted using the scale on the left of the plot, but SGM uses the scale on the right.

and humidity vary for many reasons. Therefore, SGM should be thought of as a theoretical limit, if $r = 1$ were true at all points over the domain.

The barotropic growth rate (Figure 6.9b) is clearly of a very different shape to its baroclinic counterparts. The horizontal wind shear is, therefore, a source of energy largely independent from the baroclinicity of the atmosphere. Although some of the peaks occur at the same time as peaks in the baroclinic growth rates, there are also troughs in one at the same time as peaks in the other. Also,

the barotropic growth rate can be positive or negative, depending on the direction of the meridional wind shear.

6.3.3 Spatial Variation of Growth Rates

Figures 6.10, 6.11 and 6.12 show example maps for each of the five versions of growth rate, for 24th January 2009. The four baroclinic versions of growth rate again have similar structures to each other (Figures 6.10 and 6.11), but there is also some variation; the areas with the highest values in one version are not necessarily particularly high in another version. For example, all four have an area of high values over the Sinai Peninsula, which is of a similar relative magnitude and shape in all for versions. However, there is also a distinctive pattern of three maxima over western Europe: one near Brittany, a second near Barcelona, and a third between Sardinia and Sicily. Though all three areas are present in all four versions of the growth rate, the relative strengths vary widely. Figure 6.10a shows that, for QGD, none of the maxima are particularly large. Figure 6.10b shows QGM has larger maxima near Barcelona and Sardinia. Figure 6.11a shows SGD has a strong maximum over Brittany. Finally, Figure 6.11b shows SGM, like QGM, has strong maxima near Barcelona and Sardinia, but these are weaker than for QGM. The variations between the growth rates could be used to unpick the different drivers of the maxima, given which factors feed into each of the different versions of growth rate. For example, moisture is clearly key for the two maxima over the Mediterranean, which is unsurprising given the warm underlying waters.

When the barotropic growth rate (Figure 6.12) is compared to its baroclinic counterparts, there is a large difference. This is to be expected, as barotropic growth considers only the horizontal wind shear, not a factor considered by the baroclinic growth rates. There is one strong maximum near the northern Spanish coast, and a second off the Newfoundland coast. These are in the same region as maxima in the baroclinic growth rates, but not in exactly the same location. If this continues in regions where the selected storms are developing, then this would indicate that it is a combination of factors that generate a storm-prone situation.

6.4 Storms in the Growth Rates

As a storm approaches Europe, the growth rate should increase before dropping suddenly as the storm passes, to indicate the energy building up in the atmosphere, before being removed and

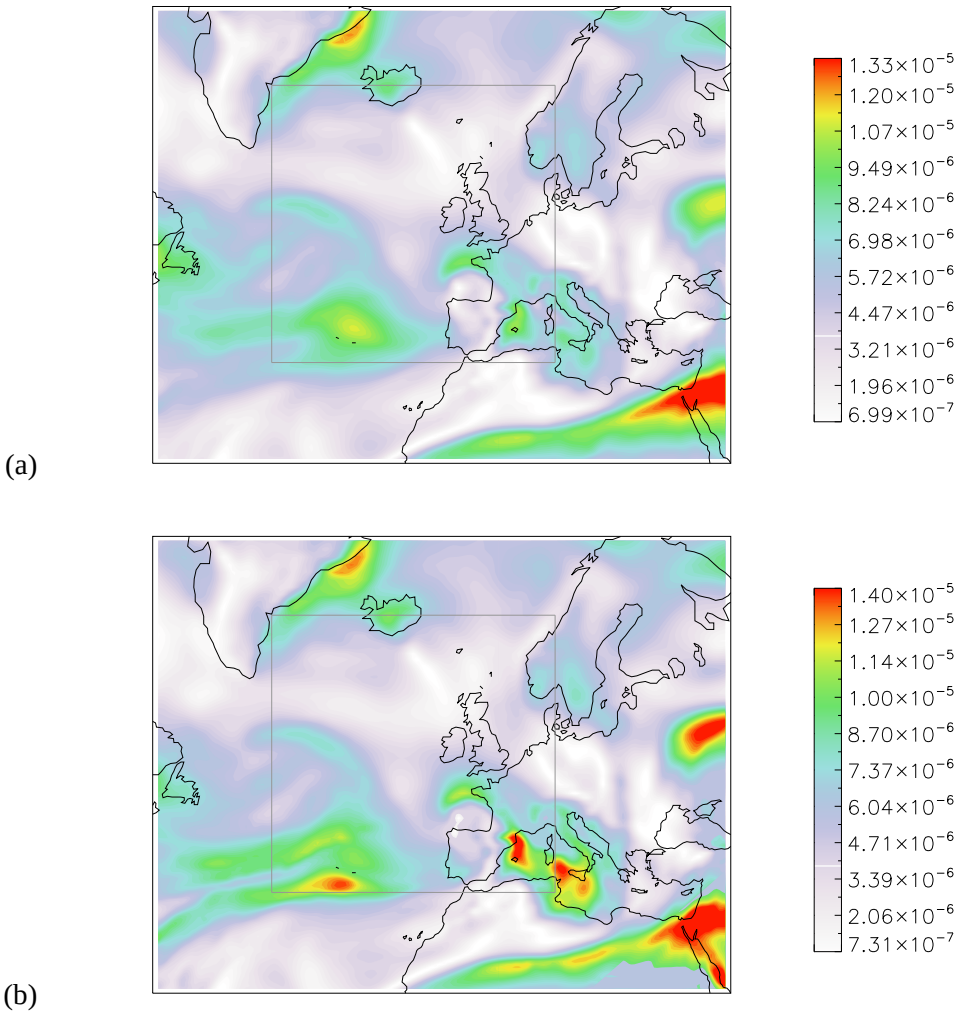


FIGURE 6.10: Maps of two of the four baroclinic growth rates [s^{-1}] for 0000 on 2009-01-24: (a) QGD, (b) QGM. The box marks the area over which the values are averaged. Please note that the colour scales are different in these plots, because the raw values are different (see Figure 6.9). The scales are determined using a percentiles approach.

converted into kinetic energy by the storm. The peak and sudden drop should be discernible in a time series plot of the growth rate; however, as the drop is related to the storm passing, a potential storm-prone situation would be a peak in growth rate in the days before a storm hits Europe. Therefore, as a first step, peaks in these five versions of the growth rate are identified automatically. All values above the 98th percentile are located, by calculating the parameters for October to March from 1979/1980 to 2011/2012, and those within a 24 hour period of each other are removed. This avoids double-counting high values, and so that a single peak is identified for each potential storm event, but does have limitations when two storms are very close; for example, Vivian and Wiebke.

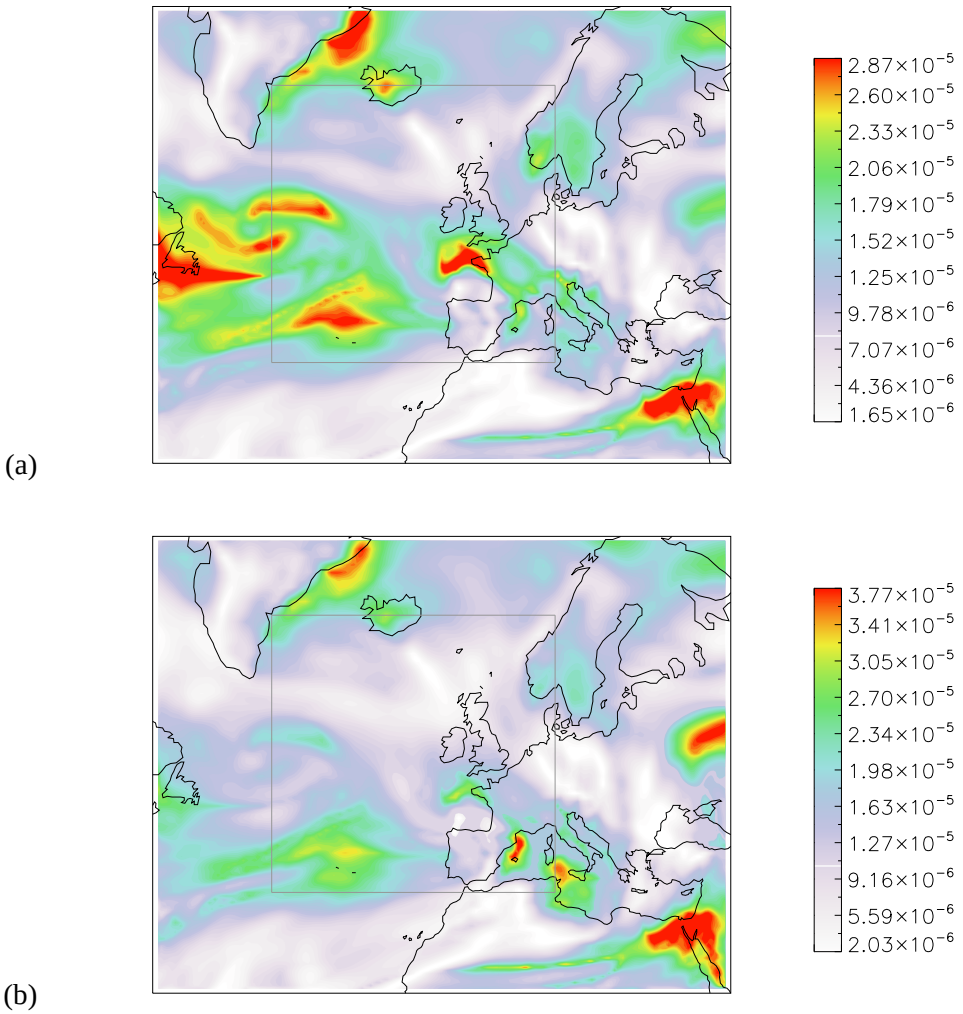


FIGURE 6.11: Maps of two of the four baroclinic growth rates [s^{-1}] for 0000 on 2009-01-24: (a) SGD, and (b) SGM. The box marks the area over which the values are averaged. Please note that the colour scales are different in these four plots, because the raw values are different (see Figure 6.9). The scales are determined using a percentiles approach.

Next, a method of automatically identifying which storms are associated with a peak is used: the growth rate has to peak between 96 hours before and 24 hours after the date of maximum SSI. The date of maximum SSI is used for consistency with the storm identification portion of this work (Section 3.2). Furthermore, this work endeavours to draw together two ideas about midlatitude cyclones that are not necessarily related: storms that are strongly influenced by baroclinicity are not necessarily the same storms that exhibit strong potential for destruction. Although dynamically, baroclinicity affects a storm's core pressure, vorticity and therefore pressure gradients and wind speeds, there are many other factors that affect these, as discussed in Section 2.2.4. In addition, SSI – the measure of potential damage – is also affected by the local wind climatology (Equation

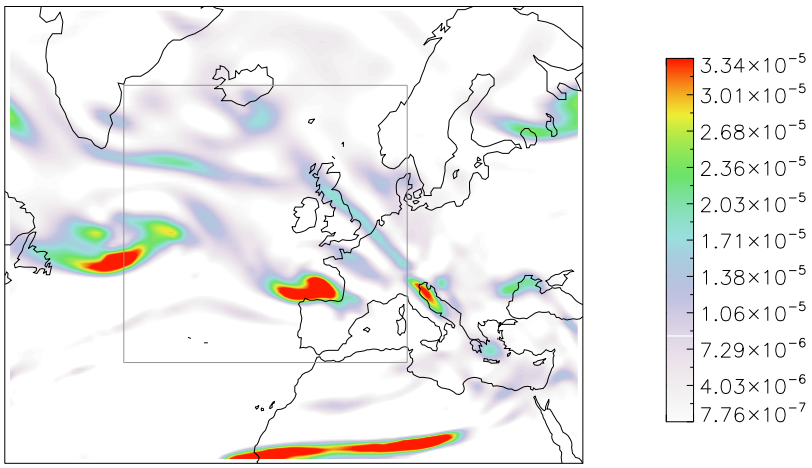


FIGURE 6.12: Map of the barotropic growth rate [s^{-1}] for 0000 on 2009-01-24.

3.1). This work examines the link between large-scale drivers of storm deepening and the potential such storms have to inflict damage. The time period of up to 96 hours before was selected because storms are most likely to deepen over the Atlantic. They then track over Europe and potentially cause damage in the box used to calculate SSI, and this movement takes 3-4 days at average speeds of storm motion, according to observations. However, the box used for averaging growth rate and that used to calculate SSI overlap, so a storm's peak growth rate could occur at the same point or even shortly after the maximum SSI. This was the reason why peaks in growth rate up to 24 hours after could also be associated with a storm. The use of this automatic, objective method means that the same criteria can be applied consistently across all storms, which may not be possible with a subjective approach.

Once the different versions have been associated with the identified peaks, the analysis of the results can begin. Section 6.4.1 contains an analysis of the temporal variation of the parameters, and how this related to the selected storms. The usefulness of a contingency table is explored in Section 6.4.3. The relationship between SPSs and the selected storms is examined using examples (Section 6.5) and 'null cases' (Section 6.6). The final stages of these results ascertain whether there is a relationship between storm-prone situations and storm intensity (Section 6.7), jet stream category (Section 6.8.1), PTE type (Section 6.8.2), or forecast quality (Section 6.9).

6.4.1 Time Series

These objective methods of associating storms and SPSs is applied to determine whether the storms are associated with a high value (greater than the 98th percentile) of the SPS metrics. Table 6.3 states which storms are successfully associated with a peak in the different versions of growth rate. It shows that that for 29 out of 31 cases at least one of the versions of growth rate has a peak associated with the storm. Clearly, these measures of a storm-prone situation are of some value, but further analysis is needed to determine the extent of their usefulness.

Table 6.3 illustrates that the combination of SPS metrics gives a 94% success rate; that is, 94% of the storms (29 out of 31) are associated with a peak in at least one of the metrics at or above the 98th percentile. It is also clear that on an individual basis, the QGD metric performs best, with SGM and SGD performing nearly as well, but with BT performing poorly. However, two storms (Urania and Franz) that are only associated with a peak in BT, so it adds value to the total percentage of storms successfully identified. It is important to note that it is in combination that these metrics perform exceptionally well. There are only two storms that are not associated with a peak in any of the metrics: Daria and Gero.

It is also clear from Table 6.3 that, even if a storm is not associated with a peak (greater than the 98th percentile) in a given baroclinic metric (QGD, QGM, SGD, SGM), it is often associated with a high value of it. This is the case with both Daria and Gero. Also, in the barotropic growth rate, there are only nine storms associated with BT growth rates greater than the 98th percentile, there are a further nine associated with BT between the 90th and 98th percentile. However, ten storms are associated with unexceptional values of BT. Therefore, the baroclinic growth rates perform better than the barotropic growth rate, with QGD performing best individually.

6.4.2 Combining Metrics

It is possible that some metrics duplicate the work of others, so could be eliminated. However, it is not possible to assess every combination of the five variables alone, in pairs, threes, fours and all together, because there are 126 combinations. Therefore, an examination of Table 6.3 shows combinations likely to prove fruitful. The combinations explored further are shown in Table 6.4, along with the original number of storms associated with the single metrics and the

Storm	QGD	QGM	SGD	SGM	BT	Overall success
Daria	97.7	92.7	91.7	97.5	89.6	~ (BC)
Nana	98.4	97.8	99.0	98.4	96.6	✓
Vivian	98.6	99.0	99.1	98.9	27.0	✓
Wiebke	98.6	97.9	99.1	92.9	27.0	✓
Udine	99.1	96.5	98.8	98.6	95.8	✓
Verena	99.9	98.2	99.2	99.1	88.3	✓
Agnes	98.6	85.5	98.7	93.5	36.3	✓
Dec 1993	98.1	94.1	97.2	96.5	45.4	✓
Urania	92.0	93.4	32.7	88.7	98.0	✓
Silke	99.3	95.0	99.7	96.7	92.3	✓
Lara	99.8	99.9	99.7	99.5	33.7	✓
Anatol	99.0	98.2	99.3	98.4	14.4	✓
Franz	93.9	87.2	93.9	90.3	98.0	✓
Lothar	98.7	96.6	97.8	96.9	94.3	✓
Martin	98.7	97.9	97.8	98.3	98.5	✓
Kerstin	99.1	97.2	99.1	98.7	32.5	✓
Rebekka	92.6	99.9	89.4	99.2	95.6	✓
Elke	98.6	98.8	97.6	99.2	99.1	✓
Lukas	93.5	94.5	94.9	97.7	98.9	✓
Pawel	98.8	98.1	98.1	98.2	98.3	✓
Jennifer	99.3	97.4	99.1	98.9	98.9	✓
Frieda	99.2	95.4	99.4	95.4	45.8	✓
Jeanette	94.6	99.8	85.2	98.9	98.7	✓
Gero	95.7	88.3	93.7	90.6	81.5	~ (BC)
Cyrus	97.9	94.6	98.2	98.2	10.0	✓
Hanno	99.6	98.7	99.3	99.3	29.9	✓
Kyrill	99.3	99.4	99.9	99.5	96.1	✓
Emma	98.3	96.7	98.4	98.3	93.7	✓
Klaus	99.6	99.7	99.4	99.9	99.4	✓
Quinten	98.6	98.7	98.4	99.6	98.9	✓
Xynthia	97.9	98.3	90.6	98.8	99.2	✓
Number >98th perc.	22	15	19	20	9	29
Percentage >98th perc.	70%	48 %	61%	65%	29%	94%
Number >90th perc.	31	28	28	30	18	31
Percentage >90th perc.	100%	90%	90%	98%	58%	100%
Number >80th perc.	31	31	30	31	22	31
Percentage >80th perc.	100%	100%	98%	100%	70%	100%

TABLE 6.3: The list of storms (ordered by date), and the percentile of the growth rate with which they are associated, in a window that stretches 96 hours ahead at 24 hours after maximum SSI. The degree of shading depends on the value; those values larger than the 98th percentile are shaded most intensely; 90th to 98th percentile shaded medium intensity; 80th to 90th low intensity; less than 80th percentile not shaded. The overall success column specifies whether the 98th percentile is exceeded with a tick, or if it is nearby with a ~. In the latter case, it is stipulated whether the baroclinic (BC) or barotropic (BT) is nearly at the 98th percentile. The overall success of each metric is shown at the bottom, for the different percentiles (perc.).

number associated with all five metrics combined. Only the 98th percentile is used here, because for the vast majority of storms, one or more metric has a value greater than this threshold.

Metrics used	Number Identified
QGD	22
QGM	15
SGD	19
SGM	20
BT	9
QGD, QGM	25
QGD, SGD	23
QGD, SGM	26
QGD, BT	27
QGD, SGM, BT	29
QGM, SGD	25
QGM, SGD, BT	29
QGD, QGM, SGD, SGM, BT	29

TABLE 6.4: Assessing the number of storms identified by different combinations of the SPS metrics.

This shows that the combining QGD, SGM and BT or QGM, SGD and BT is as successful as combining all five measures. This can be explained physically, because these three metrics together describe the variety of processes that contribute towards the deepening of midlatitude cyclones: baroclinic, barotropic and moist processes. This also illustrates that the different treatment of vorticity in the quasi- and semi-geostrophic models is important, because more storms are identified when QGD and SGM are used in combination, than when QGD and QGM are used together.

6.4.3 Contingency Table

Another approach to analysing whether these are successful metrics for storm-prone situations is to use a table similar to a contingency table, which are used extensively in forecast verification (Joliffe and Stephenson, 2003). Table 6.5 shows the number of times both an SPS and a storm occur, one or the other occurs, or neither. The values in the absence of a storm were reached by calculating the number of analysis timesteps (4 daily) occurring in the winter seasons (October - March) of the years 1989/1990 to 2009/2010, which is the time period under consideration here. This gives 15308. Though the SPS events occur 2% of the time (306), an SPS can include one or more or the five metrics. Therefore, comparing the dates of peaks in each metric are compared and duplicates removed, giving the number of unique events as 1172. Of these, 29 are associated

with a storm, so the final number of 'null cases' is 1143. The remainder of analysis times are not associated with an SPS (14,136), but there are two that are associated with a storm.

	SPS ✓	SPS ×
Storm ✓	29	2
Storm ×	1143	14,134

TABLE 6.5: Comparing the instances of the presence (✓) or absence (×) of an SPS and the presence or absence of a selected storm.

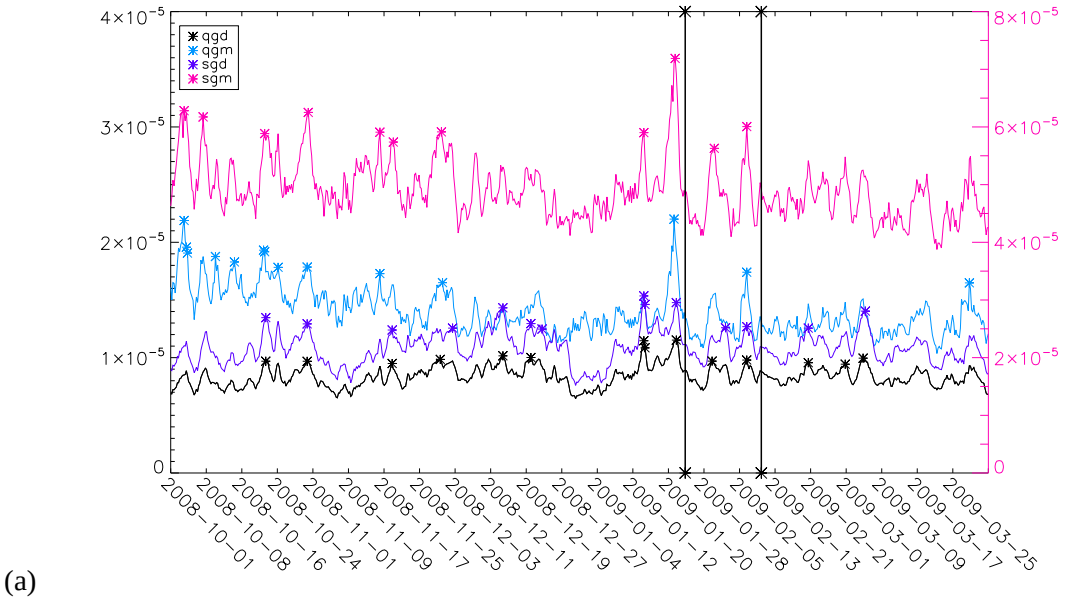
This shows that, if a forecast measure were to be used to quantify the relationship between SPSs and storms, such a metric would indicate a very strong relationship. There is a low 'null case rate', where an SPS is not associated with one of the selected storms. The miss rate, where there is a storm but no SPS, is also small. However, forecast measures will not be used to quantify this relationship, because they will be skewed by the large number of cases where there is no SPS and no storm. This number should be high, because this work deals with rare, severe events, but this does impede the effectiveness of many forecast measures.

6.5 Examples of Storms and their Storm-Prone Situation

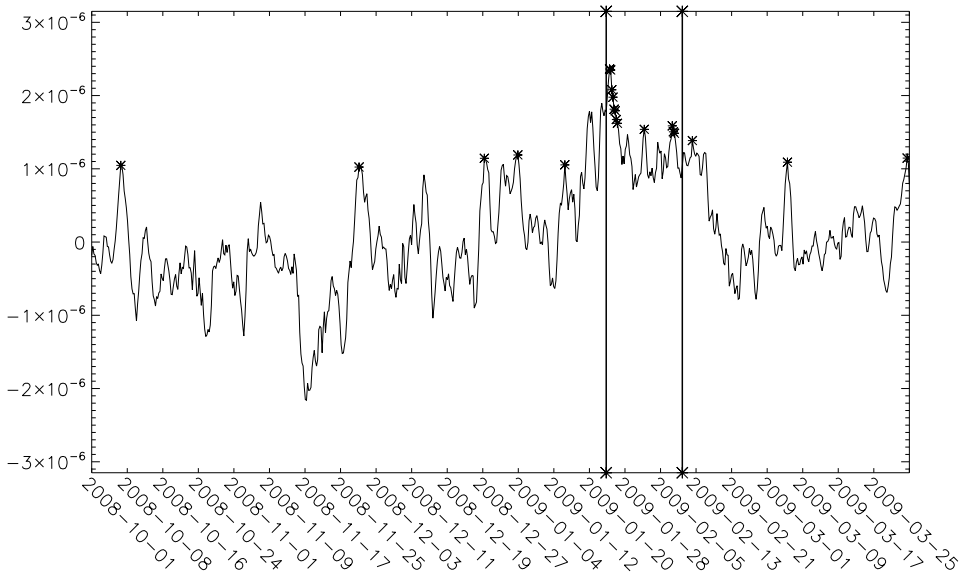
Having examined the broad picture, this section will consider the SPSs of individual storms. It will investigate the relationship between each storm and its SPS in the combination of metrics already discussed, and endeavours to probe the relationship between SPSs and storm types by examining several examples. This latter idea will be explored in greater depth in Section 6.8.

6.5.1 Klaus and Quinten

Figure 6.13 shows the baroclinic and barotropic growth rates in the winter season that included storms Klaus and Quinten. The objectively-identified peaks are indicated by the stars, and the two storms within this season (Klaus on 2009-01-24; Quinten on 2009-02-10) are indicated using the black vertical lines. Both are preceded by a strong peak in the four baroclinic growth rates, indicating a strong build up of large-scale potential energy as baroclinicity before they develop. Both are also related to peaks in the barotropic growth rate, indicating a build up of the large-scale kinetic energy. This large-scale energy is then converted to the kinetic energy of the storm, aiding deepening and strengthening winds.



(a)



(b)

FIGURE 6.13: Time series for 2008-10-01 to 2009-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate. The lines indicate when storms passed through Europe: Klaus on 2009-01-24; Quinten on 2009-02-10.

Klaus is the most intense storm in the set selected, when SSI is used as the metric (Section 3.2). The peak in the barotropic growth rate happens on the same day as the storm has its maximum intensity, but occurs slightly after the maximum SSI. While this might indicate a link between the peak in the barotropic growth rate and the decay of the cyclone, this seems unlikely, as there would be no physical explanation for the drop in BT after Klaus has passed. It seems more likely that Klaus causes the BT growth rate to drop after the high peak, but that the barotropic energy conversion is more important than its baroclinic counterpart in the later stages of Klaus's development, in agreement with Kucharski and Thorpe (2000). However, in the PTE categories (Table 4.4), Klaus is a diabatic-type storm, indicating that diabatic processes dominate deepening. To summarise, it would seem that baroclinic energy conversion dominates the early stages of Klaus's development, barotropic energy conversion dominates the later stages, and diabatic processes are important throughout deepening. Furthermore, the baroclinic and barotropic peaks are the highest or nearly the highest observed during the winter 2008-2009. Based on plots similar to the jet stream and θ_e shown in Chapter 4 (not shown), the jet stream has unusually high wind speeds, and unusually warm moist air has been drawn up from the equator. This would imply that, for the atmosphere to produce a storm as intense as Klaus, then a number of factors combine to generate a 'perfect storm'.

This idea of a perfect storm is reinforced by Quinten (Figure 6.13). Again, not only are there peaks in all four baroclinic growth rates, but also there is a peak in BT. Although Quinten is much less intense than Klaus in terms of SSI, the PTE analysis shows that Quinten's deepening is also dominated by diabatic processes. Furthermore, Quinten, like Klaus, crosses the jet stream later on in his lifetime. This leads to the notion that if all of the cyclogenetic factors discussed – baroclinic and barotropic instability, and moisture – work together, they can generate a cyclone, even if they are not unusually strong. If they all exhibit extremely high values then a storm as damaging as Klaus can develop; if the values are moderately high, a storm like Quinten is more likely.

While these storms are both related to peaks in all five versions of the growth rate, the relative prominence of the peaks varies between the versions. In SGM, Klaus is associated with what is clearly the highest peak, but in QGM the peak is only just the highest. For QGD and SGD, the peak associated with Klaus is the second highest. This lends weight to the argument that Klaus was driven by exceptionally strong storm-prone characteristics, working together to generate an exceptionally strong midlatitude cyclone. However, a similar difference is found for storm Quinten. Although Quinten's peaks are not amongst the highest in the season, they are also more

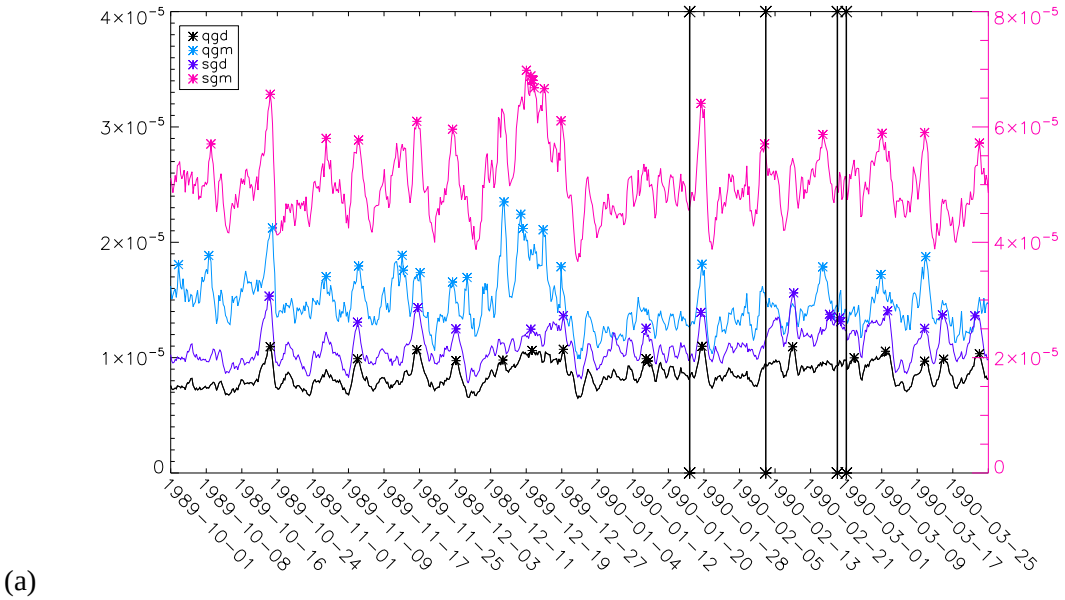
prominent in the moist growth rates than in the dry. These differences in peak prominence can be used to unpick the drivers of storm deepening. Since the peaks in the moist growth rates are more pronounced than their dry counterparts, this signals that moisture was particularly important in the development of Klaus and Quinten. This is consistent with the PTE analysis, where diabatic processes dominate the deepening in both cases.

6.5.2 Daria and Gero

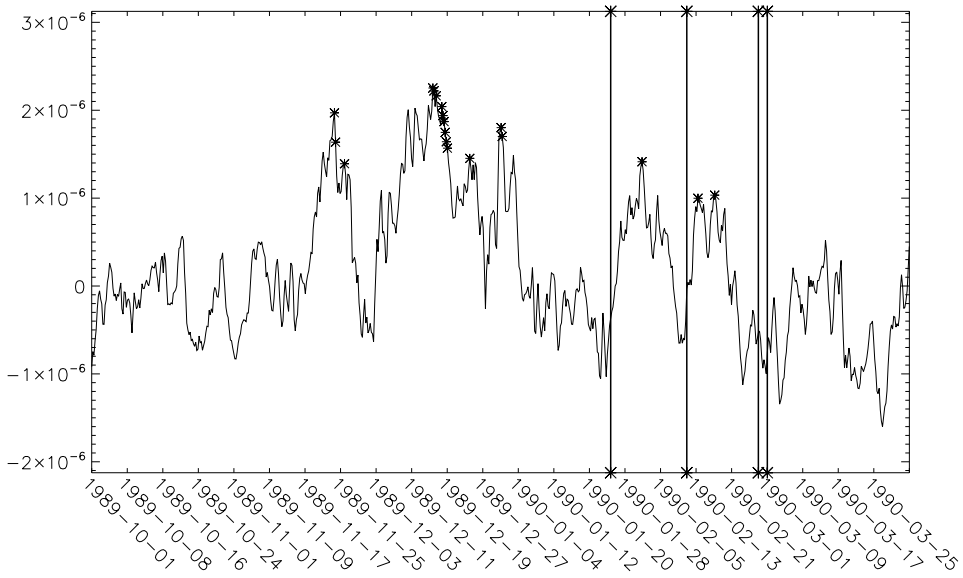
These are the two cases that are not associated with a peak that exceeds the 98th percentile in any of the metrics. However, in both cases, the most of the values exceed the 90th percentile.

Figure 6.14 shows that Daria occurs just before a very large peak in all five versions of the growth rate. Daria crosses the jet stream early, and is in the 'horiz' category from the PTE analysis. Daria is the ninth most intense storm in the set of 31, in terms of SSI. Franz is also in the 'horiz' PTE category, but is an edge storm and is second least intense. Franz also occurs just before a peak in all five versions of the growth rate. In both cases, the peak after the storm is higher, indicated by the stars. Therefore, the reason why these storms are not associated with a high (greater than the 98th percentile) value is that these peaks are too far after the storm to be within the 24-hour, post-maximum SSI window. This could mean that the percentile values shown for these storms in Table 6.3 are on the rising arm of the later peak.

This would mean that the window during which a storm and a peak can be associated (96 hours before the storm, 24 hours after) should be widened after the storm. However, this is not viable for two reasons. First, this is not practical during times of storm seriality (e.g. January/February 1990), because the peak just after one storm could be related to a different storm. Secondly, a damaging storm could not be explained by a storm-prone situation significantly after the event; it would compromise the concept of a storm-prone situation. Therefore, the window of interest will be left as it is.

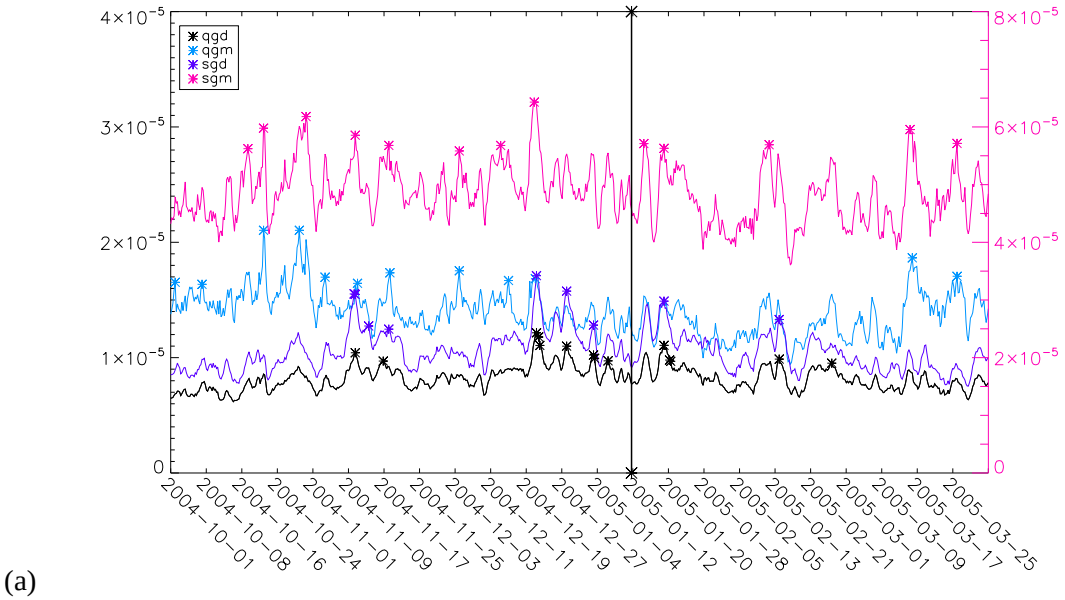


(a)

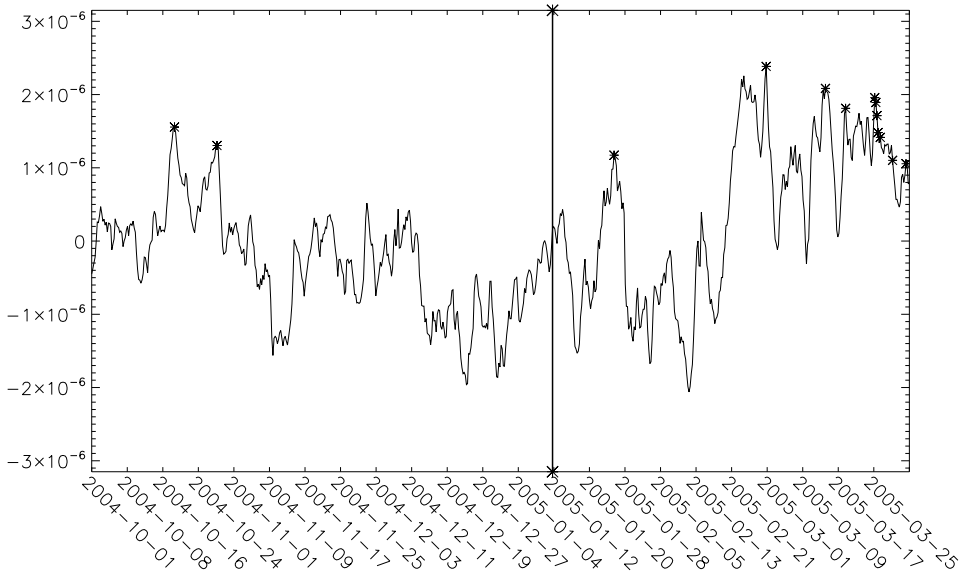


(b)

FIGURE 6.14: Time series for 1989-10-01 to 1990-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate. The lines indicate when storms passed through Europe: Daria on 1990-01-25; Nana on 1990-02-11; Vivian on 1990-02-27; Wiebke on 1990-03-01.



(a)



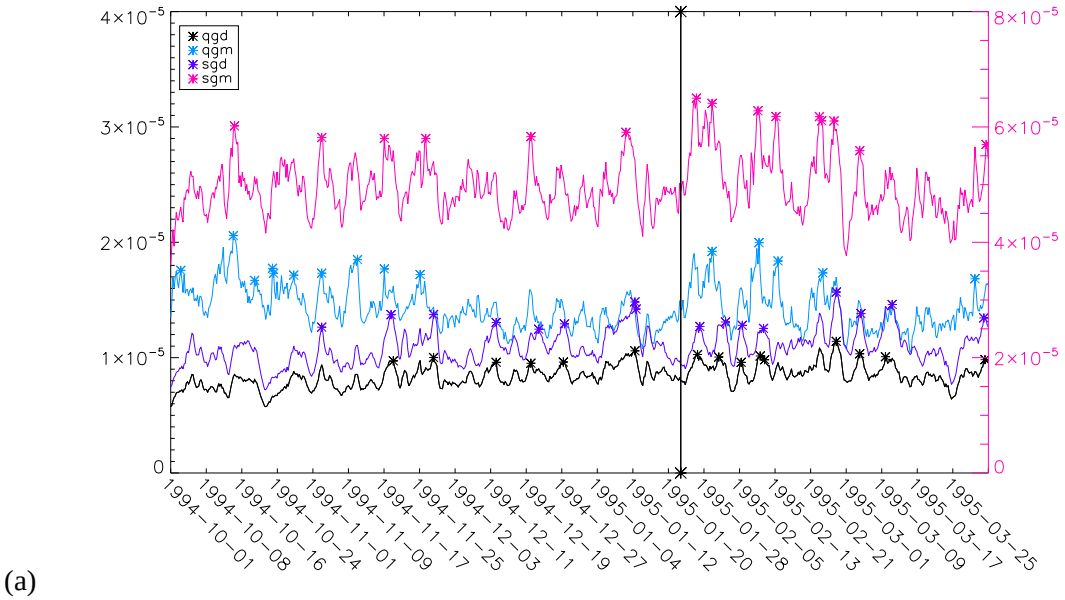
(b)

FIGURE 6.15: Time series for 2004-10-01 to 2005-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate. The line indicates when storm Gero passed through Europe, on 2005-01-12.

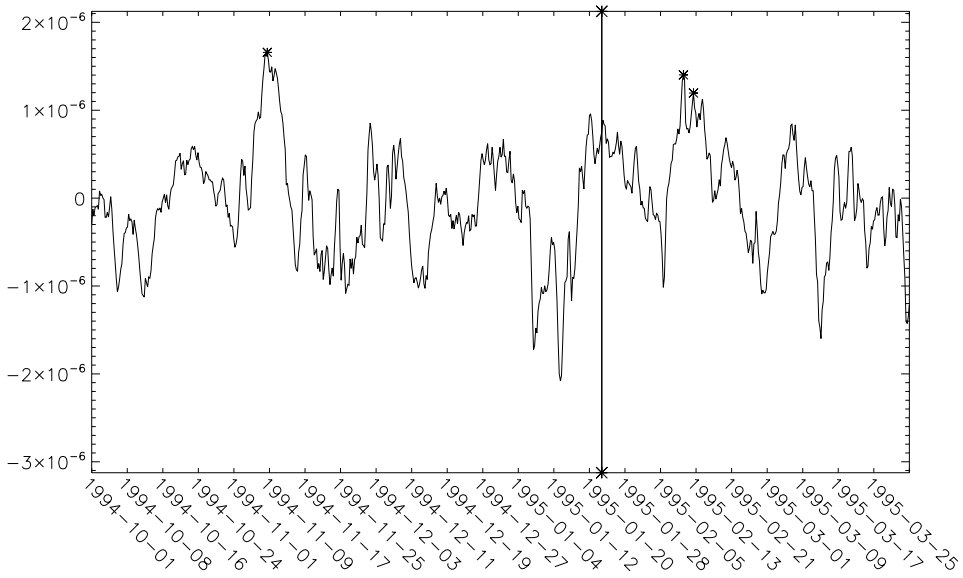
6.5.3 Urania and Franz

These two storms are only identified in the barotropic growth rate. Figure 6.16a shows that Urania is not associated with a high value of the baroclinic growth rates, but does have a peak shortly after the storm has passed through Europe. Figure 6.16b shows that there is a high value of the barotropic growth rate in the area when Urania passes through Europe. Figure 6.17 shows a similar story for storm Franz: the barotropic growth rate is high, but the baroclinic growth rate peaks afterwards. This could indicate that the best way to identify storms in the baroclinic growth rates is to look for sudden drops in the time series, rather than the peaks. This will be discussed further in Section 6.5.4.

Both of these storms are associated with baroclinicity that does not meet the 98th percentile criteria. Furthermore, neither Urania nor Franz was associated with an exceptionally high value of the Storm Severity Index (SSI, Section 4.1). The next steps are to search for baroclinicity associated with these storms, and assess its strength. Franz is presented as an example, because the weather charts are not available for Urania. The Berlin weather charts are shown in Figures 6.18 (at the surface) and Figures 6.19 (at 500hPa). The day before the highest SSI is shown, in order to capture the development of the storm. When Franz's location is identified on the surface chart, translating this to the 500hPa chart shows that there is some baroclinicity in the area, because the thickness contours are not parallel to the geopotential height contours. However, the baroclinicity is not strong, because the thickness gradient is weak. This is consistent with the ideas of James and Gray (1986), whereby a strongly barotropic atmosphere prevents the development of strong baroclinicity through the 'barotropic governor' mechanism. The implications of this are that the barotropic growth rate does not necessarily contribute towards the deepening of storms, but rather is an indicator for storms that could cause damage given their SSI, but are not the most intense storms in the set selected. It is possible that storms like Franz and Urania initially gain energy through barotropic energy conversions and then from the baroclinicity, as proposed by Kucharski and Thorpe (2000). Even if the wind shear was working to reduce the baroclinicity through the barotropic governor effect, the two types of energy conversion working together could be sufficient for an intense storm to develop. This is consistent with the discussion of storm Klaus, which is associated with high values of all growth rates and the result is the most intense storm in the set.



(a)



(b)

FIGURE 6.16: Time series for 1994-10-01 to 1995-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate. The line indicates when storm Urania passed through Europe on 1995-01-23.

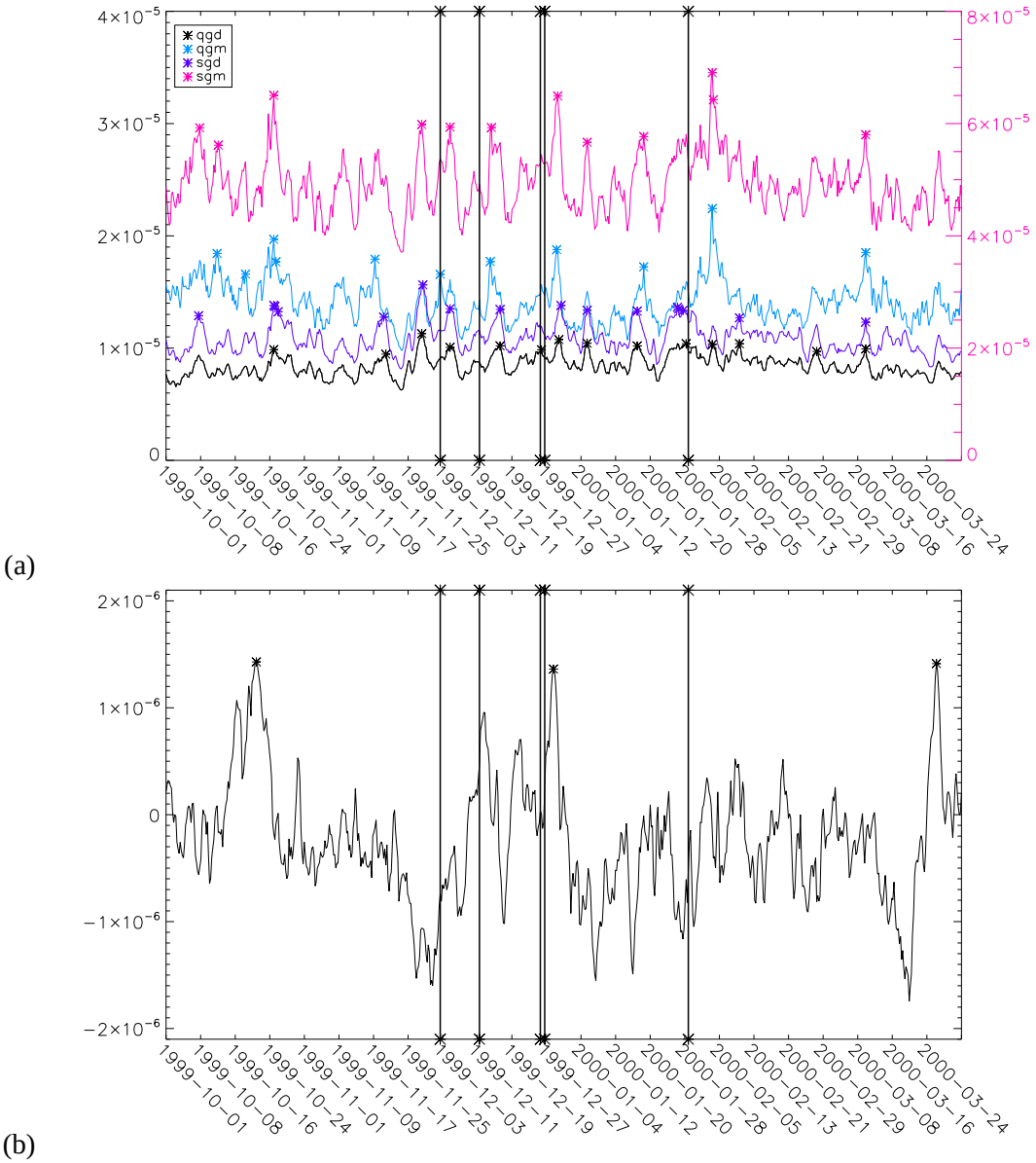


FIGURE 6.17: Time series for 1999-10-01 to 2000-03-31 of (a) the four versions of baroclinic growth rate and (b) barotropic growth rate. The lines indicate when storms passed through Europe: Anatol on 1999-12-03; Franz on 1999-12-12; Lothar on 1999-12-26; Martin on 1999-12-27; Kerstin on 2000-01-29.

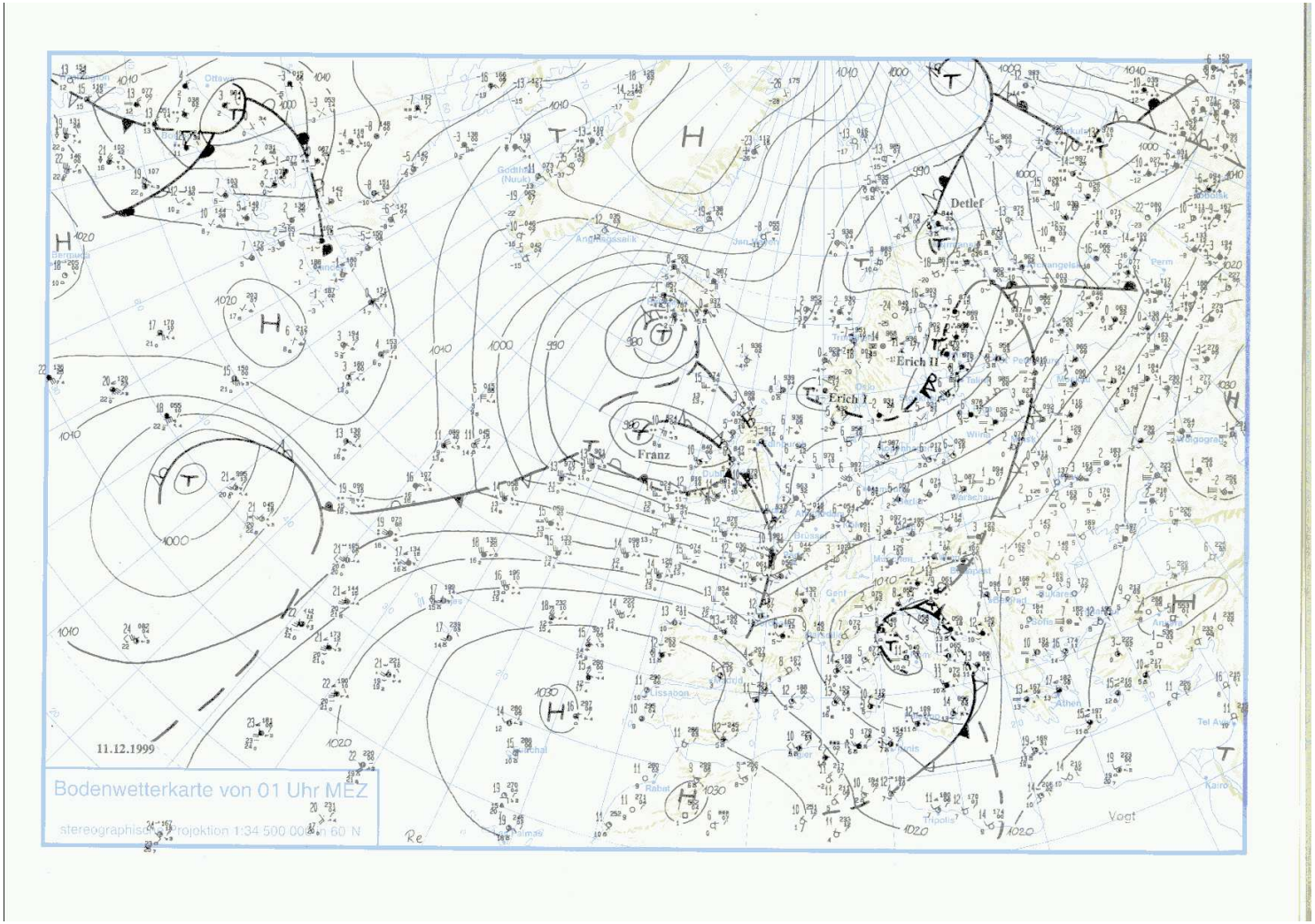


FIGURE 6.18: Surface weather chart for 1999-12-11, illustrating storm Franz. Contours of mean sea level pressure, diagnosis of fronts and station plots also shown.

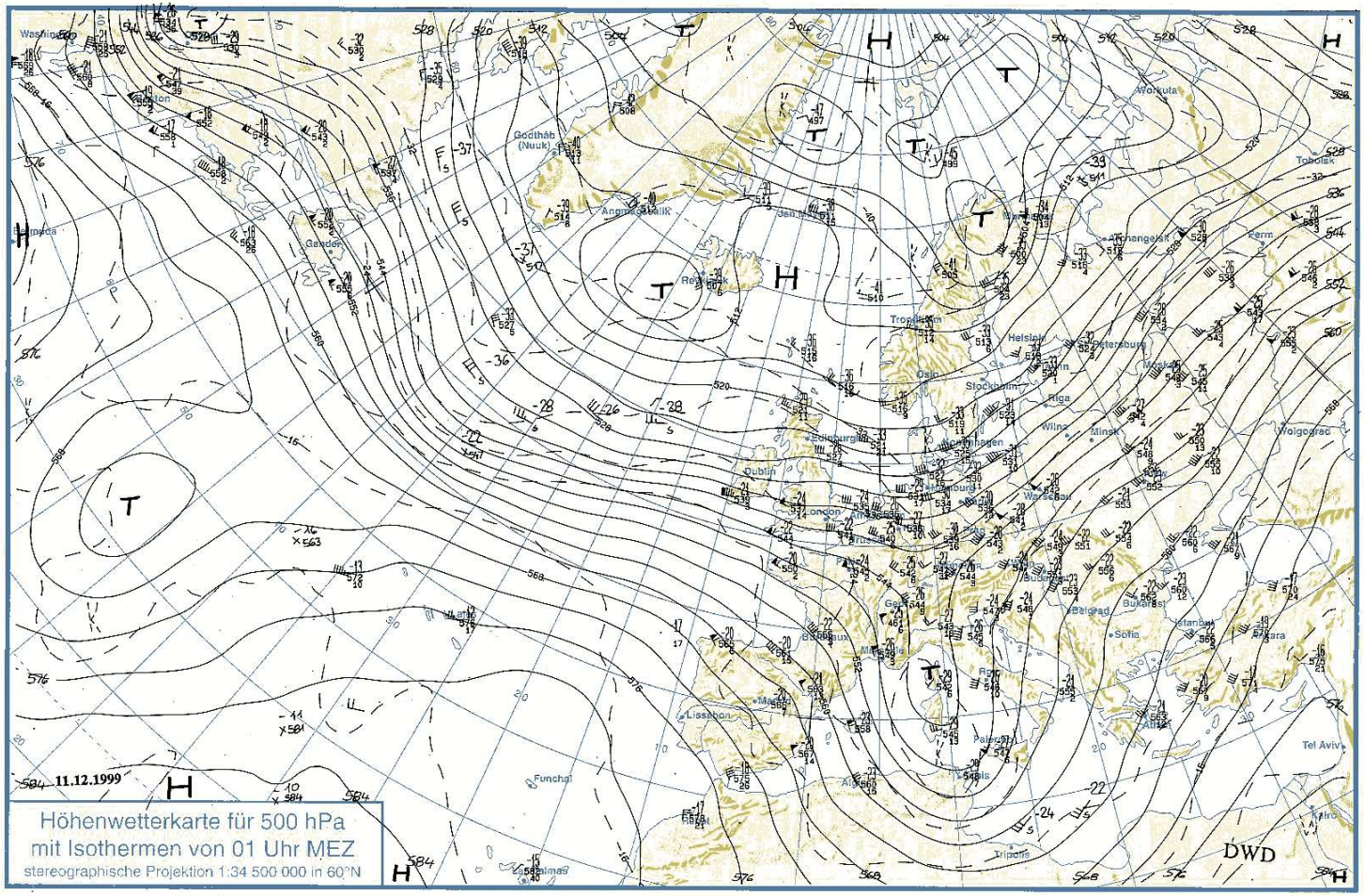


FIGURE 6.19: Weather chart on the 500hPa isobaric surface for 1999-12-11. Solid contours of geopotential height of the 500hPa surface, dashed contours of 1000 to 500hPa thickness. Data from radiosonde launches also shown.

6.5.4 Anatol

As can be seen from Figure 6.17, Anatol is related to an increasing value of the barotropic growth rate. Therefore, part of the reason that he is not related to an exceptionally high value is because the value at the start of the increase is very low. This could indicate that the best way of identifying storms in the barotropic growth rate could be to search for sharply increasing values. As mentioned earlier, this could also be the case for the baroclinic growth rate. However, when this was tested (not shown), it did not improve the hit rate for associating either the baroclinic or barotropic growth rate with the selected storms. Furthermore, this would not make physical sense as a metric for a storm-prone situation. Energetically, the peak represents the build up of energy available for storm development, whereas the drop is associated with the storm removing that energy and converting it to eddy kinetic energy. Therefore, the identification of peaks will remain as the method used.

6.5.5 Summary of Examples

Broadly speaking, the storms are generally associated with peaks in at least one of the metrics. The analysis of Klaus and Quinten indicates that there are many factors that come together to cause cyclogenesis. Storm Klaus in particular, being the most intense storm selected, illustrates the idea of a 'perfect storm', where these factors work together to generate an exceptionally intense mid-latitude cyclone. This may also be the case with Urania and Franz, which are identified as peaks in the barotropic growth rate but not the baroclinic growth rates. However, when upper air charts are examined, baroclinicity is present, which could mean that the barotropic and baroclinic processes both contribute towards deepening but that the baroclinicity is not above the 98th percentile of climatology.

Those storms that are not are associated with a peak are generally associated with a high value of the metric, as shown in Table 6.3 by the percentiles of the SPS metrics with which the storms are associated. In the cases of Daria and Gero, there are peaks shortly after they occur. While this could indicate that the window used to distinguish whether a storm is associated with a peak should be widened, this was rejected for practical and physical reasons. The example of Anatol allows discussion of whether peak or drop identification would be the optimal approach. The outcome is that peak identification is better, because it represents the build up of potential energy ahead of the storm, whereas the drop represents the storm removing the energy.

Despite the mostly successful link between the selected storms and the SPS metrics, there are many peaks in the growth rates that are not associated with one of the selected storms. The next section will analyse these null cases.

6.6 Null Cases

There are 1143 instances where there is a peak in the growth rate, but the peak is not associated with any of the selected storms. From these, a subset is discussed, in order to examine a variety of cases in sufficient depth as to ascertain why they are null cases. Baroclinic peaks are discussed, because more storms are identified in baroclinic rather than barotropic growth rates. For simplicity, only peaks in QGD are included, due to the similarity between the four versions of baroclinic growth rate compared to surface charts.

The subset of null cases – where there is a peak but not a selected storm – are selected based on the following criteria:

- One random peak per year, so that a particular weather pattern is not sampled multiple times,
- After June 1999, because this is the earliest date when surface pressure charts are archived at the BADC,
- Not within a week of any selected storm, to ensure that a selected storm is not the reason for the peak.

This results in eight dates for further investigation. On examination of their charts, it becomes clear that these null cases generally fall into three categories:

1. High and Low
2. Occluded Low
3. Mature Low

Representative members of this set are analysed further, with the surface pressure charts and maps of QGD shown in Figures 6.20, 6.21 and 6.22. Each of these will now be discussed in turn, along with the implications for each of these categories.

6.6.1 High and Low

Two typical examples of this are shown in Figure 6.20.The first example of this, (1999-11-21, Figure 6.20a) shows only a weak trough starting north of the Faroe Islands, which does not produce a sufficiently strong pressure gradient to drive the strong winds that could inflict damage. The reason that this trough did not become intense, despite there being baroclinicity locally (Figure 6.20b), could be because no other deepening processes were there to assist; for example, the jet stream could have been too weak. Alternatively, Figure 6.20b implies that the baroclinic wave is already breaking (Section 2.2.6), which as it is associated with cyclolysis means the storm is likely filling.

On 2006-03-24, the midlatitude cyclone is blocked (Figure 6.20c). While the high pressure centred at $20^{o}W$, $50^{o}N$ is not particularly strong, over the preceding days (not shown) it has impeded the progress of the cyclone sufficiently that by the time it reaches Europe, the cyclone's intensity is diminished and it no longer has the potential to inflict damage. The high QGD is in the area of the storm, indicating that the storm is still developing (Figure 6.20d). However, the blocking is sufficiently strong and the storm is sufficiently weak that it does not have the potential to inflict damage on Europe.

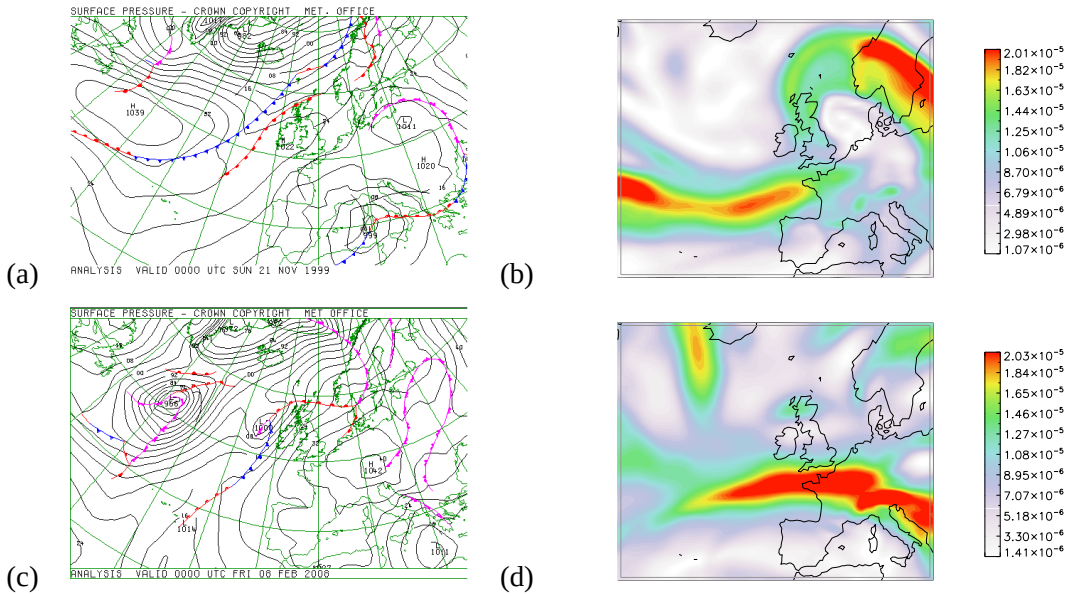


FIGURE 6.20: Two of the random null case dates, showing how high pressure can block midlatitude cyclones: (a) the Met Office surface pressure chart for 1999-11-21, (b) QGD for 1999-11-21, (c) the Met Office surface pressure chart for 2008-02-08, (d) QGD for 2008-02-08

In short, these examples are too weak, in the wrong place, or moving in the wrong direction to inflict damage on Europe, due to the effect of a blocking high pressure. That this blocking situation is associated with high baroclinicity is a notable caveat.

6.6.2 Occluded Low

There are two of these randomly-selected cases where a peak in QGD occurs on the same day that Europe is being affected by an occluded low: 2002-11-24 and 2010-02-06. The surface pressure charts and QGD maps for these are shown in Figure 6.21.

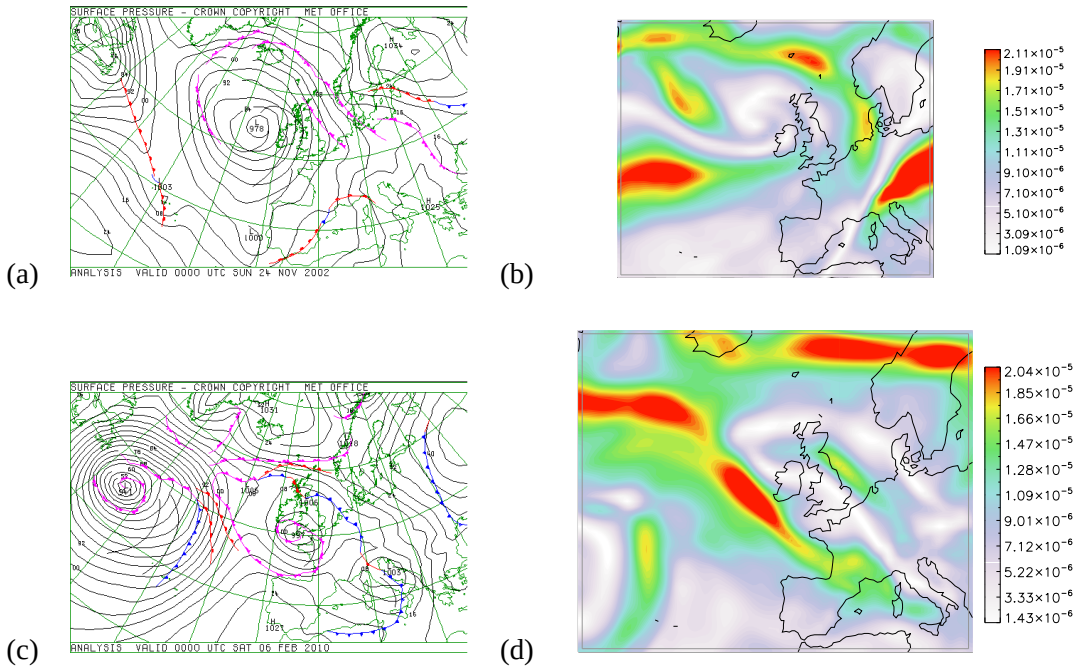


FIGURE 6.21: Two of the random null case dates, both showing an occluded midlatitude cyclones: (a) the Met Office surface pressure chart for 2002-11-24, (b) QGD for 2002-11-24, (c) the Met Office surface pressure chart for 2010-02-06, (d) QGD for 2010-02-06.

Figure 6.21a shows that, on 2002-11-24, there was a low pressure to the west of Ireland, associated with some occluded fronts. There is some evidence of cyclogenesis on the southern flank of the main low, off the west coast of Portugal and over the Azores. However, neither of these has an unusually low core pressure nor a strong pressure gradient, so their potential to inflict damage if they track over Europe is low. The values of QGD are not high in a single, coherent area near either storm (Figure 6.21b). The high baroclinicity is most likely related to the low over Canada, centred in the Gulf of St Lawrence. However, this cyclone is in a mature stage, and subsequent

charts (not shown) show that by the next day, this low is filling, and so does not reach Europe but stays west of 10^oW. Therefore, there is an argument that the western boundary of the box used for averaging the growth rates should be brought further east, to avoid capturing cyclogenesis in this area that is not relevant to Europe. However, cyclogenesis in this area can lead to storms that could inflict damage on Europe, such as storm Agnes (Figure 4.6), and so the boundary will remain at its current location.

The situation is similar in the case of 2010-02-06; while an occluded cyclone hangs over the UK, in the west Atlantic there is a very low (941hPa) cyclone. However, this cyclone is also in the occluded stages of development, and so it is filling. Subsequent charts (not shown) reveal that this cyclone remains in nearly the same location throughout its filling, developing into a complex range of fronts and secondary low-pressure centres over the course of a few days. Despite the presence of quite high growth rates in the area of interest (Figure 6.21d), these storms are not strengthening. In short, this storm did not affect Europe.

Overall, these two examples reveal that there are cyclones that develop in the western Atlantic that do not reach Europe. However, they are often slow moving and at a mature stage before reaching the area used for averaging the SSI. This illustrates an important concern in a forecasting situation, in terms of the mobility of the cyclones on the western edge of the domain. One improvement could include a criterion regarding the speed at which the cyclone is travelling.

6.6.3 Mature Low

Of the eight random dates, two appear to be associated with coherent, mature cyclones, both with high values of QGD over Europe. The first example of this, 2001-11-30 (Figure 6.22a), shows a low pressure system at 20^oW, off the Irish coast. The high values of QGD are in the vicinity of the storm (Figure 6.22b), so it is not surprising that during this time the storm is deepening. However, subsequent charts (not shown) reveal that this storm's track has a strong northerly component, and it subsequently moves harmlessly between the UK and Iceland.

The second example from 2006-11-30 is undergoing explosive deepening; in the 24 hour period centred on the chart in Figure 6.22b, the cyclone there shown at 20^oW off the west of Ireland undergoes 31hPa of deepening. QGD is also high near to the storm, indicating baroclinicity and the potential for the storm to develop further, However, the subsequent charts (not shown) illustrate that this storm's track also has a strong northerly component, and it stalls over Iceland. It is also

important that there is a high pressure covering BeNeLux and western Germany, which blocks the storm from encroaching on Europe. The only region afflicted by a strong pressure gradient and, by inference, strong winds is the Western Isles of Scotland. Since storms were selected based on SSI (how exceptional the wind speeds are for each grid point, summed up over an area covering Europe), a storm that affects a small fraction of the area will not have an exceptionally high value of SSI. This is the reason why this storm is a null case, rather than a selected storm.

6.6.4 Discussion

Broadly speaking, these null cases are easily explained. They are mostly associated with a midlatitude cyclone, but one that is not moving at the right speed or in the right direction to potentially inflict damage on Europe. Therefore, there are caveats to using QGD as an SPS metric: that the storm must be moving in the right direction and at the right speed to reach Europe while still a strong storm.

An improvement to this as a way of identifying potentially damaging storms before they reach Europe could be to include feature-tracking elements. This could determine whether the storm is moving in the right way to reach Europe while still an intense storm. However, this would involve

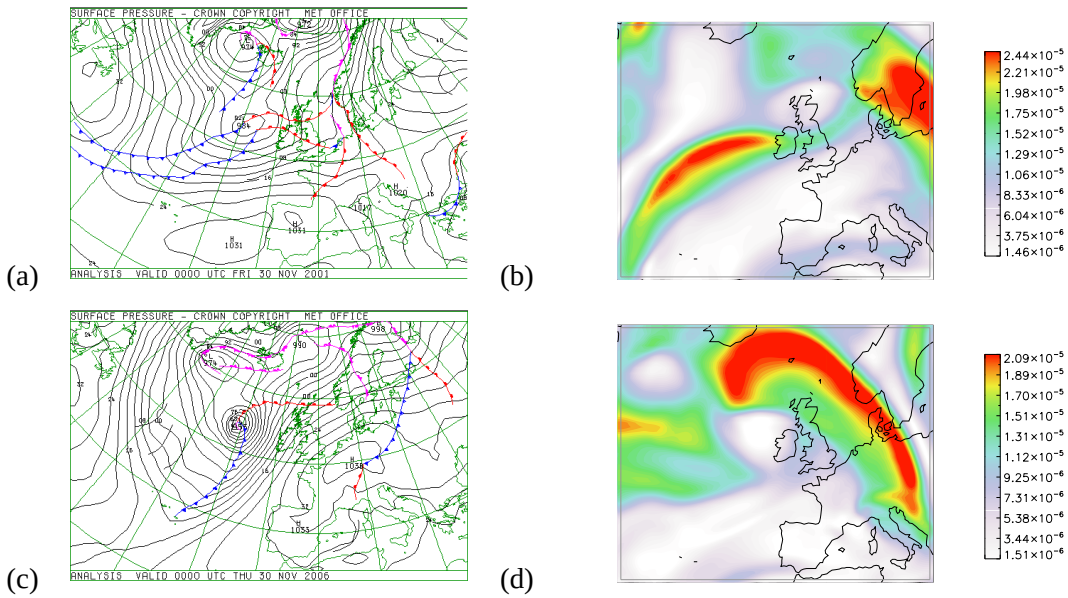


FIGURE 6.22: Two of the random null case dates, both showing Europe being affected by a mature cyclone: (a) the Met Office surface pressure chart for 2001-11-30, (b) QGD for 2001-11-30, (c) the Met Office surface pressure chart for 2006-11-30, (d) QGD for 2006-11-30.

linking the algorithm that detects peaks in the baroclinic growth rate with the algorithm that tracks the storms. This goes against the aim of having a simple metric for a storm-prone situation, and begins to filter the storms that are associated with peaks in the SPS metric. Such post-processing is left for future work.

One further caveat that is not observed within this set of null cases – or during those studied but not included here – is that baroclinicity is a necessary but not sufficient condition for cyclogenesis. The atmosphere being arranged so that density is a function of pressure and temperature (not just pressure) does not cause a cyclone to form; however, if that arrangement is perturbed slightly, then the perturbation is ripe to grow into a midlatitude cyclone. However, this analysis implies that the development of a storm with an imperfect track for potentially inflicting damage on Europe is more likely to be the cause of the null case, than the absence of the initial perturbation.

6.7 Storm-Prone Situations and Intensity

Table 6.6 shows the storms and the SPS metrics they are associated with, but sorted by their intensity according to the storm severity index (SSI, Section 3.2). On examination, it becomes apparent that there is no clear link between intensity and whether or not a storm is identified by the SPS metric, or between intensity and which metrics are successful. Although the most intense storm (Klaus) is located in all five versions of the growth rate, this is also the case for storm Quinten, which is in the bottom third of this list. This lack of a relationship could be for a number of reasons. Firstly, SSI may not be the best measure of storm intensity to use in this context, because it measures how exceptional the wind speed is at each point. It might be more appropriate to use a measure for eddy kinetic energy, as this is what the baroclinic and barotropic energy become. Alternatively, minimum core pressure or maximum vorticity could be used, as in previous studies, but this would be inconsistent with the rest of this work. On balance, SSI is used here for consistency.

Secondly, the absence of a clear link between intensity and the SPS metric could signpost that there are aspects of storm development that are not captured by this method of measuring a storm-prone situation. This is likely, because some aspects of cyclogenesis cannot be accounted for explicitly on this scale. The notable example is of diabatic processes, which are not accounted for analytically, only implicitly by use of the moisture field. These themes will be explored in Section 6.8.2. This approach does also not explicitly consider the position or strength of the jet stream,

storm	SSI	Jet Category	PTE type	QGD	QGM	SGD	SGM	BT
Klaus	55.47	Cross late	Diabres	99.6	99.7	99.4	99.9	99.4
Vivian	46.31	Cross early	Horiz	98.6	99.0	99.1	98.9	27.0
Wiebke	46.31	Cross late	Horiz	98.6	97.9	99.1	92.9	27.0
Kyrill	39.47	Cross early	Horiz	99.3	99.4	99.9	99.5	96.1
Lothar	37.68	Cross late	Diabres	98.7	96.6	97.8	96.9	94.3
Martin	37.11	Cross late	Horiz	98.7	97.9	97.8	98.3	98.5
Emma	34.09	Edge	Horiz	98.3	96.7	98.4	98.3	93.7
Jeanette	27.51	Edge	Diabres	94.6	99.8	85.2	98.9	98.7
Daria	27.33	Cross early	Horiz	97.7	92.7	91.7	97.5	89.6
Agnes	19.69	Edge	Horiz	98.6	85.5	98.7	93.5	36.3
Anatol	18.60	Cross early	Horiz	99.0	98.2	99.3	98.4	14.4
Udine	17.16	Cross early	Horiz	99.1	96.5	98.8	98.6	95.8
Rebekka	16.25	Cross early	Horiz	92.6	99.9	89.4	99.2	95.6
Lara	15.68	Edge	Horiz	99.8	99.9	99.7	99.5	33.7
Xynthia	14.98	Split	Diabres	97.9	98.3	90.6	98.8	99.2
Jennifer	14.65	Cross late	Horiz	99.3	97.4	99.1	98.9	98.9
Gero	13.53	Split	Diabres	95.7	88.3	93.7	90.6	81.5
Hanno	13.38	Cross late	Horiz	99.6	98.7	99.3	99.3	29.9
Silke	12.42	Cross early	Diabres	99.3	95.0	99.7	96.7	92.3
Elke	10.93	Split	Horiz	98.6	98.8	97.6	99.2	99.1
Dec 1993	9.89	Split	Horiz	98.1	94.1	97.2	96.5	45.4
Urania	9.29	Cross late	Horiz	92.0	93.4	32.7	88.7	98.0
Nana	9.24	Edge	Horiz	98.4	97.8	99.0	98.4	96.6
Quinten	8.32	Cross late	Diabres	98.6	98.7	98.4	99.6	98.9
Verena	8.31	Cross late	Horiz	99.9	98.2	99.2	99.1	88.3
Kerstin	7.98	Edge	Horiz	99.1	97.2	99.1	98.7	32.5
Pawel	7.82	Edge	Horiz	98.8	98.1	98.1	98.2	98.3
Cyrus	7.80	Cross early	Horiz	97.9	94.6	98.2	98.2	10.0
Lukas	7.77	Edge	Horiz	93.5	94.5	94.9	97.7	98.9
Franz	7.70	Edge	Horiz	93.9	87.2	93.9	90.3	98.0
Frieda	7.64	Edge	Horiz	99.2	95.4	99.4	95.4	45.8

TABLE 6.6: The list of storms, described by the two methods of categorisation and ordered by intensity (SSI), along with the percentile values of each growth rate with which each storm is associated, shaded as per Table 6.3.

though it does consider the vertical wind shear so strength is considered implicitly. How the storm track and jet stream are configured facilitated the division of storms into jet categories, and their relationship with the SPS metrics will be discussed in Section 6.8.1.

Table 6.6 hints that there might be a weak relationship between intensity and the SPS metrics. Tests for correlation (not shown) show very weak positive correlation between SSI and the percentile values, which is not significant under the relevant t-test (Equation 3.5). One way to assess whether there is a weak relationship is to compare the four most intense and four least intense storms. When

this is done in QGD all four high-intensity storms exceed the 98th percentile criterion, but only one out of the four low-intensity storms do so. Similar comparisons can be made for QGM, SGD and SGM, indicating a weak connection between intensity as measured by SSI and all four baroclinic metrics. No such relationship appears for BT. In short, a slight connection appears between the baroclinic SPS metrics and the intensity (SSI), but it should be noted that any relationship is weak, because it does not stand up to statistical testing.

6.8 Storm-Prone Situations and Storm Type

This work now concentrates on the potential for a relationship between the categories a storm is in, and the SPS metrics associated with each storm. Two methods of categorisation were used in Chapter 4: one based on the configuration of the jet stream and the storm track, and the second based on the terms that dominate the deepening according to the Pressure Tendency Equation (PTE). These will now be discussed in turn.

6.8.1 Jet Stream Categories

The selected storms were divided into categories, based on the configuration of their tracks and the jet stream. The method used for this is discussed in Section 3.4, and a discussion of each jet stream type can be found in Section 4.4. Four types were identified:

- storms that cross the jet early in their lifetime (‘cross early’),
- storms that cross the jet late in their lifetime (‘cross late’),
- storms that do not cross the jet, but travel along its perimeter (‘edge’),
- storms that are influence by more than one jet stream (‘split’).

The analysis of the association between the SPS metrics and these categories will follow a similar pattern to that described in Section 3.5. The raw numbers and percentages for each category that each SPS metric identifies at the three thresholds are shown in Tables 6.7.

In terms of the percentage of each category identified at or above the 98th percentile threshold (Table 6.8), QGD is the best metric for three jet-stream categories: cross early, cross late, and

QGD

	98th	90th	80th	Fail
CE	5	3	0	0
CL	8	1	0	0
ED	7	3	0	0
SP	2	2	0	0

QGM

	98th	90th	80th	Fail
CE	4	4	0	0
CL	4	5	0	0
ED	3	5	2	0
SP	2	1	1	0

SGD

	98th	90th	80th	Fail
CE	6	1	1	0
CL	6	2	0	1
ED	7	2	1	0
SP	0	4	0	0

SGM

	98th	90th	80th	Fail
CE	6	2	0	0
CL	6	2	1	0
ED	6	4	0	0
SP	2	2	0	0

TABLE 6.7: Tables showing the number of storms associated with each of the four baroclinic growth rates at the different percentile thresholds, for each jet stream category. 98th refers to the number of storms associated with a value at or above the 98th percentile; 90th refers to those between the 90th and 98th percentiles; 80th to those between the 80th and 90th percentiles; Fail to the number of storms associated with a value less than the 80th percentile.

QGD

	98th	90th	80th	Fail
CE	62.5%	37.5%		
CL	88.9%	11.1%		
ED	70.0%	30.0%		
SP	50.0%	50.0%		

QGM

	98th	90th	80th	Fail
CE	50.0%	50.0%		
CL	44.4%	55.6%		
ED	30.0%	50.0%	20.0%	
SP	50.0%	25.0%	25.0%	

SGD

	98th	90th	80th	Fail
CE	75.0%	12.5%	12.5%	
CL	66.7%	22.2%		11.1%
ED	70.0%	20.0%	10.0%	
SP		100.0%		

SGM

	98th	90th	80th	Fail
cross early	75.0%	25.0%		
cross late	66.7%	22.2%	11.1%	
edge	60.0%	40.0%		
split	50.0%	50.0%		

TABLE 6.8: Tables showing the percentage of storms associated with each of the four baroclinic growth rates at the different percentile thresholds, using the raw numbers from Table 6.7. The jet stream types are abbreviated: CE = cross early, CL = cross late, ED = edge, SP = split.

edge. It is expected that it would perform best for the storms that spend most or all of their track on the north side of the jet, because the cold, dry air found there would suppress the potential influence of moisture. However, that it performs best for storms that cross the jet stream late is unexpected. Since storms that linger on the south side of the jet are likely to spend longer in the warm, moist air found there, the moist SPS measures should be more strongly associated with these storms. For the split-jet storms, half of them are located in QGD at the 98th percentile threshold, but this is also the case with QGM and SGM. This is consistent with moisture playing a more important role for these storms than for the other jet stream categories. However, there are

only four storms in this category, and so firm conclusions cannot be drawn. Broadly speaking, that QGD is the metric associated most strongly with most categories of storm indicates the importance of dry dynamics and baroclinicity in cyclogenesis.

Comparing the different thresholds, QGD remains the most successful metric for all storm categories; by the 90th percentile, all of the storms in all categories have been identified. SGD is the least successful, with two storms that are not associated with a high value of the metric, and one that is associated with a value above the 80th percentile threshold. Broadly, the 98th percentile appears to be the best threshold of the three used here, because it has a sufficiently high success rate for associating the SPS metrics with the selected storms that statistical tests can be performed and the different categories can be compared. Using a threshold above which all of the different categories of storms are identified would impede the assessment of a relationship between the SPS metrics and the different storm categories.

(a)

	98th	90th	80th	Fail
CE	0	4	1	3
CL	5	1	1	2
ED	4	2	0	4
SP	2	0	1	1

(b)

	98th	90th	80th	Fail
CE	0%	50.0%	12.5%	37.5%
CL	55.6%	11.1%	11.1%	22.2%
ED	40.0%	20.0%	0.0%	40.0%
SP	50.0%	0%	25.0%	25.0%

TABLE 6.9: For the barotropic growth rate, (a) the raw numbers and (b) the percentages of the storms in each jet stream category identified in each percentile threshold, as per Tables 6.7 and 6.8. The jet stream types are abbreviated: CE = cross early, CL = cross late, ED = edge, SP = split.

The BT metric (Table 6.9) is successfully associated with more than half of the storms that cross the jet stream late, indicating that horizontal wind shear is important for the development of these storms. The BT metric is also associated with half of the split-jet storms. It is hardly surprising that storms with more than one jet maximum involved in their development are encouraged to develop by horizontal wind shear, although an important caveat is that there are only four storms in this category.

The final piece of analysis uses a similar approach to that used in Section 4.6, where the expected number of storms in each category are compared to the observed values. The expected value is calculated on the row and column totals for the number of storms in each category, identified in each metric, as described in Section 3.5.2. This analysis is only performed for the 98th percentile, for reasons discussed earlier in this section. The differences between the number of storms observed and expected is shown in Table 6.10. Positive values indicate that more storms in a given

	QGD	QGM	SGD	SGM	BT
Cross Early	-0.44	0.79	1.31	1.06	-2.72
Cross Late	0.49	-0.44	-0.48	-0.82	1.25
Edge	0.01	-1.13	0.96	-0.35	0.51
Split	-0.07	0.78	-1.79	0.12	0.96

TABLE 6.10: Difference between the number of storms observed and expected to be associated with each SPS metric at the 98th percentile threshold, for each jet stream category. Shading illustrates whether the value is positive (pink) or negative (blue).

category are observed to be related to a high value of a given metric than expected; negative values indicate fewer storms than expected are similarly related, to illustrate the relative performance of each metric for each category.

Table 6.10 shows the differences between the observed and expected values of successful identification in each metric, for each jet category. indicates that the QGD metric performs relatively well for the storms that cross the jet stream late, but relatively badly for those that cross early. The SGD metric is more likely than expected to be associated with cross early and edge type storms, but less likely for cross late and split-jet storms. The differences between the QGD and SGD metrics illustrates that vorticity is a key factor in the development of cross-early and edge storms. QGM and SGM both perform well for the cross early and split-jet storms, which points towards moisture being important for these categories of storm. Comparing QGD and QGM also demonstrates this, because more storms than expected are associated with a high value of the moist version, whereas fewer than expected are related to such a value in the dry counterpart. This is unexpected for cross early storms, because they spend their time on the northern, drier side of the jet stream. The barotropic measure of SPSs behaves differently from its baroclinic counterparts, identifying fewer than expected storms that cross the jet stream early, but more of the other three categories. However, these results may not be significant, given the small number of storms. Few of the storms are associated with a value of BT greater than the 98th percentile, and when these are divided by jet stream category it means there are insufficient storms to draw firm conclusions.

Overall, QGD is the best or joint-best metric for all four jet stream categories of storm. While some of the subtleties of interactions between the jet stream and storms are captured, the different SPS metrics do not fairly represent the range of such interactions. However, this is not the only method of categorising the storms under consideration; the next section will perform similar analysis for the different PTE types.

6.8.2 Pressure Tendency Equation Type

The second method of dividing the storms was by which terms dominated their deepening, according to PTE analysis. Section 3.4.2 describes the method used to identify these processes, and the results are discussed in Section 4.5. While several processes are key to development, the deepening of the storms was predominantly due to two of them:

- horizontal temperature advection ('horiz'),
- diabatic processes ('diab').

This leads naturally to two types of storms, depending on which of these processes dominates the deepening. The relationship between these types and which of the SPS metrics the storms are associated will now be explored.

QGD

	98th	90th	80th	Fail
Horiz	18	6	0	0
Diab	4	3	0	0

QGM

	98th	90th	80th	Fail
Horiz	9	13	2	0
Diab	4	2	1	0

SGD

	98th	90th	80th	Fail
Horiz	15	7	1	1
Diab	4	2	1	0

SGM

	98th	90th	80th	Fail
Horiz	16	7	1	0
Diab	4	3	0	0

TABLE 6.11: Tables showing the number of storms associated with each of the four baroclinic growth rates at the different percentile thresholds, for each PTE type. 98th refers to the number of storms associated with a value at or above the 98th percentile; 90th refers to those between the 90th and 98th percentiles; 80th to those between the 80th and 90th percentiles; Fail to the number of storms associated with a value less than the 80th percentile.

In terms of the raw numbers of storms related to high values of each baroclinic SPS metric (Table 6.11), horiz-type storms appear to do well in most metrics, and diab less so. However, the picture is blurred somewhat by the relative populations of the two PTE categories: horiz has 24 members, but diab only 7. Therefore, percentages are a better approach for assessing the relationships.

These percentages are shown in Table 6.12. For horiz-type storms, QGD is the metric that most often has values over the 98th percentile during storm development. This is as expected, because it reinforces the link between large-scale, dry baroclinicity and storms where horizontal temperature advection is the most important deepening process. For such advection to occur, the atmosphere should be baroclinic. This links together the large-scale SPS approach and the small-scale PTE

QGD

	98th	90th	80th	Fail
Horiz	75.0%	25.0%		
Diab	57.1%	42.9%		

QGM

	98th	90th	80th	Fail
Horiz	37.5%	54.2%	8.3%	
Diab	57.1%	28.6%	14.3%	

SGD

	98th	90th	80th	Fail
Horiz	62.5%	29.2%	4.2%	4.2%
Diab	57.1%	28.6%	14.3%	

SGM

	98th	90th	80th	Fail
Horiz	66.7%	29.2%	4.2%	
Diab	57.1%	42.9%		

TABLE 6.12: Tables showing the percentage of storms associated with each of the four baroclinic growth rates at the different percentile thresholds, using the raw numbers from Table 6.11.

approach. The percentage of storms associated with values over the 98th percentile is significantly lower in QGM than QGD, indicating that moisture is not particularly key to their development. The percentages for SGD and SGM are similar, but both lower than QGD, indicating that considering only planetary vorticity not relative, local vorticity is sufficient for horiz-type storms.

For diabatic storms, the picture is less clear. Exactly the same fraction of storms are located above the 98th percentile in all four baroclinic metrics, so the effects of including moisture in the SPS are not apparent here. This analysis could be improved by including more diabatic storms; however, this would lessen the potential depth of the analysis considerably, and so is not done as part of the current work. That the same fraction of diab storms are identified in each metric means that, compared to horiz-type storms, moisture and local vorticity play a more important role. However, their relative importance is difficult to unpick from Table 6.12.

(a)

	98th	90th	80th	Fail
Horiz	7	5	2	10
Diab	4	2	1	0

(b)

	98th	90th	80th	Fail
Horiz	29.2%	20.8%	8.3%	41.7%
Diab	57.1%	28.6%	14.3%	

TABLE 6.13: For the barotropic growth rate, (a) the raw numbers and (b) the percentages of the storms in each PTE type identified in each percentile threshold, as per Tables 6.11 and 6.12.

For the barotropic growth rate, one interesting feature does appear: all of the storms that are 'fails' (i.e. where the value of BT is less than the 80th percentile) are all horiz-type. This implies that the barotropic growth rate is more important for diab-type storms than horiz-type storms. Again, this is indicative of the idea of a perfect storm in these diab cases. Since the local baroclinicity is not exceptional enough for horizontal temperature advection to dominate the deepening, then a coming together of other factors leads to the generation of a severe storm. While baroclinicity is still present for diab-type storms, it is weaker than for the horiz-type storms. This means that diabatic

	QGD	QGM	SGD	SGM	BT
Horiz	1.18	-0.94	0.47	0.71	-1.41
Diab	-1.18	0.94	-0.47	-0.71	1.41

TABLE 6.14: Difference between the number of storms observed and expected to be associated with each SPS metric at the 98th percentile threshold, for each PTE type. Shading illustrates whether the value is positive (pink) or negative (blue).

processes and horizontal wind shear work together, to deepen the storm, when the baroclinicity and associated horizontal temperature advection are too weak.

This idea is supported by the data shown in Table 6.14, which considers the relative performance of the different metrics. This is similar to Table 6.10, comparing the observed and expected values for each PTE type. More diab storms are associated with a high (above 98th percentile) value of BT than expected, illustrating that the two processes do work together to generate some severe midlatitude cyclones. As already discussed, QGD performs best with horiz-type storms, here shown by the greater number of horiz storms that are associated with a high value of QGD than expected. SGD also performs slightly better for horiz-type storms than diab-type, for similar reasons. QGM is more closely associated with diab-type storms, indicating that moisture is more important for diab storms' development, as expected. However, SGM performs better with horiz-type storms that diab-type, so the connection between the large-scale moisture fields that feed into the QGM and SGM metrics and the occurrence of small-scale diabatic processes according to PTE analysis is not clear.

However, there are some nuances to the relationship between the SPS metrics and the storms. Despite the strong link between QGD and the horiz-type storms, there are examples of storms where this is not the case. Urania is one such horiz-type storm not associated with a peak in any of the baroclinic metrics here (Figure 6.23a). This could perhaps indicate that there is not a perfect relationship between the extent to which horizontal temperature advection dominates in the PTE analysis, and the strength of the baroclinicity across the large part of the North Atlantic and Europe used to calculated the SPS metrics. This is to be expected, because the two are very different approaches. The way baroclinicity is measured here looks at a large area and averages over it, whereas the PTE analysis examines a small box in the region around the storm. This is an important caveat to remember as the analysis continues. However, there are many examples where the opposite is true. Storm Anatol is also a horiz-type storm, and is associated with peaks in

the four baroclinic growth rates, as can be seen from Figure 6.23b. This, alongside the statistical approach already taken, indicates that QGD performs particularly well with horiz-type storms.

There are also some nuances to the relationship between these metrics and the diab-type storms. It is anticipated that a storm with strong diabatic influences would be more likely to be associated with peaks in the moist growth rates. From the PTE analysis, the main driver of Xynthia's deepening are diabatic processes, indicating that moisture and the associated release of latent heat is a key factor in her development. This is reflected in the high values of QGM and SGM before Xynthia develops (Figure 6.23c), because these consider the humidity of the air in their calculation of stability. Diabatic processes have greater potential to occur and/or to be more powerful in air of higher humidity, because there is more water vapour present that can be converted to liquid water and facilitate latent heat release. However, Lothar is associated with a peak in the two versions of the dry growth rate (QGD, SGD), illustrated in Figure 6.23b. This is not what would be expected; as Lothar's deepening is also driven by diabatic processes, according to the PTE analysis, moisture should be key to his development, but this is not associated with peaks in the two moist growth rates (QGM, SGM). This illustrates that the relationship between the moist growth rates and the diabatic PTE-type storms is imperfect. This is for similar reasons to the imperfect relationship between the baroclinic growth rates and horiz PTE-type storms. The two different analyses take different approaches, particularly in that they consider very different scales.

Overall, there is a link between the PTE categorisation of the storms and the SPS metrics associated with them; however, it is not a strong relationship. It is expected that storms driven by horizontal temperature advection, according to PTE analysis, will be more strongly linked to the baroclinic rather than the barotropic growth rates. This is because, for horizontal temperature advection to occur, then the atmosphere should have a horizontal temperature gradient, and so be baroclinic. While this is the case for storm Anatol, it is not the case for storm Urania. It is also expected that storms that are diabatically driven will be more strongly related to the moist metrics than the dry, because of the necessity of the presence of moisture for diabatic processes such as latent heat release to occur. While this is the case with storm Xynthia, it is not for storm Lothar. This is probably due to the very different approaches taken by the PTE and SPS analysis; the former examines the drivers of a storm on a $3^o \times 3^o$ box, whereas the latter considers the average of a metric over much of the North Atlantic and Europe. In summary, there is a link between the QGD metric horiz-type storms, where horizontal temperature advection dominates the deepening

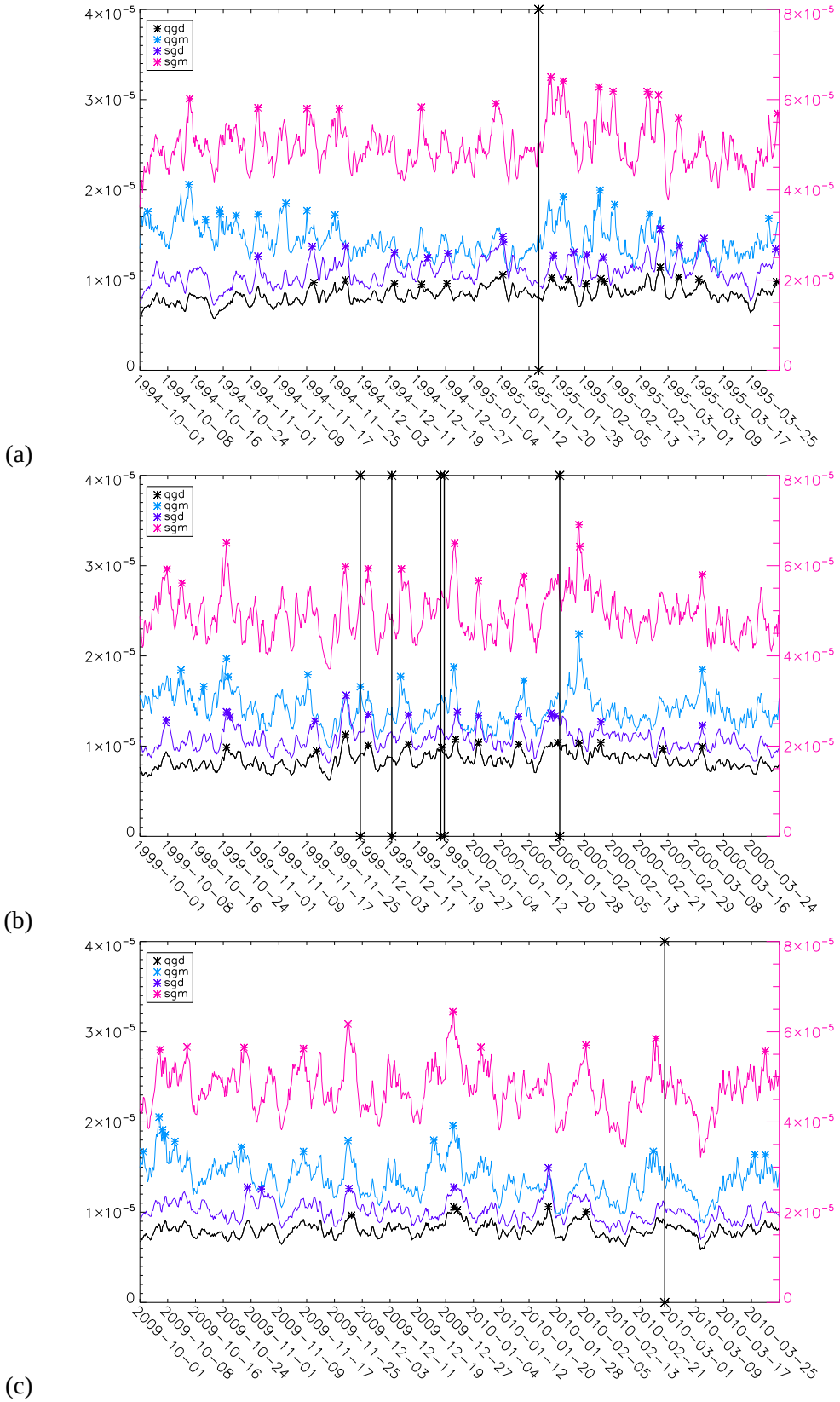


FIGURE 6.23: Time series of the four versions of baroclinic growth rate for (a) for 1994-10-01 to 1995-03-31 (the black vertical line indicates when storm Urania passed through Europe on 1995-01-23), (b) 1999-10-01 to 2000-03-31 (Anatol on 1999-12-03; Franz on 1999-12-12; Lothar on 1999-12-26; Martin on 1999-12-27; Kerstin on 2000-01-29), and (c) 2009-10-01 to 2010-03-31 (Xynthia on 2010-02-28).

according to PTE analysis. The expected connection between the moist SPS metrics and diab-type storms is only weakly apparent. However, there is a link between diab-type storms and the barotropic growth rate. This could demonstrate that in the cases where there is a deficiency of baroclinicity, the horizontal wind shear (barotropicity) and diabatic processes work together to generate a storm, sufficiently severe to be one of those selected here.

6.9 Storm-Prone Situations and Storm Predictability

The next steps examine the links between this work on SPSs, and the work presented in Chapter 5 regarding the predictability of the storms. The analysis used the 26 storms where linear regression is valid to derive the 'simple metric for forecast quality' (Section 5.4), using their operational forecasts. Here, the values for differences between the forecast and the analysis (forecast - analysis) at 24 hours calculated from linear regression are compared to the percentiles of the SPS metric values associated with each storm. This is done in terms of calculating Pearson's Correlation Coefficient (R), and using Spearman's Rank Correlation Coefficient (R_S), and the results are shown in Tables 6.15, 6.16 and 6.17. In terms of R, the significance of the correlation is tested using a t-test, and comparing the value to a reference value for the number of degrees of freedom. Here, the reference value is 1.711 (Rees, 2001, Table C.5). For R_S, the ranks of the different storms in each variable are compared and the coefficient calculated. The value is compared to a reference coefficient, to determine whether the correlation is significant, which is 0.392 (Rees, 2001, Table C.12).

Test	R	t	sig?	R_S	sig?
QGD	-0.094	-0.461	×	-0.009	×
QGM	0.004	0.021	×	-0.054	×
SGD	0.101	0.497	×	-0.010	×
SGM	0.075	0.368	×	0.022	×
BT	0.342	1.783	✓	0.229	×

TABLE 6.15: Testing correlation between the SPS metrics and predictability metric for longitude, using Pearson's correlation coefficient (R), its test for significance (t), Spearman's rank correlation coefficient (R_S), and whether the two tests are significant or not.

Table 6.15 broadly shows that there is no significant correlation between forecast quality for longitude of the selected storms and the strength of their storm-prone situations. There is one exception: positive correlation between longitude and BT, meaning storms associated with a strong value of BT are better forecast in terms of longitude. This is likely because BT represents horizontal (zonal)

wind shear, and an area with a large amount of such shear is likely to generate a cyclone. That this process is well understood means the forecast quality is better for location. The reason it is particularly in terms of longitude is because the strong horizontal wind shear could guide the storm through a particular zonal band, making its speed easier to forecast. However, it is important to note that, if correlation is sought in a lot of cases, then some correlation will be found by random chance, and so there may be no physical reason for this correlation.

Test	R	t	sig?	R_S	sig?
QGD	-0.323	-1.671	×	-0.464	✓
QGM	-0.019	-0.091	×	-0.038	×
SGD	0.026	0.129	×	-0.471	✓
SGM	0.150	0.743	×	-0.041	×
BT	-0.097	-0.479	×	0.208	×

TABLE 6.16: Testing correlation between the SPS metrics and predictability metric for latitude, using Pearson's correlation coefficient (R), its test for significance (t), Spearman's rank correlation coefficient (R_S), and whether the two tests are significant or not.

Table 6.16 shows that there is little correlation between the SPSs metrics and the forecast quality of latitude. There is some negative correlation between the two dry SPS metrics (QGD and SGD) and the latitude forecast error. This means that storms with a higher value of the SPS metrics tend to be worse forecast. This is consistent with baroclinicity being a key factor in determining the region through which the storm moves, due to the storm strengthening most in the regions of strongest north-south temperature gradient. Stronger temperature gradients are harder to model on a grid, because a high resolution is required to capture the fine detail of the location of the temperature gradient. Therefore, the location of the strongest baroclinicity is not well captured by a forecast model, and so these storms tend to be worse forecast in terms of location.

Test	R	t	sig?	R_S	sig?
QGD	0.222	1.114	×	0.286	×
QGM	-0.029	-0.141	×	-0.311	×
SGD	0.235	1.185	×	0.276	×
SGM	0.177	0.879	×	-0.076	×
BT	0.084	0.411	×	-0.112	×

TABLE 6.17: Testing correlation between the SPS metrics and predictability metric for pressure, using Pearson's correlation coefficient (R), its test for significance (t), Spearman's rank correlation coefficient (R_S), and whether the two tests are significant or not.

Table 6.17 shows that there is no correlation, ranked or otherwise, between forecast quality of storm intensity (pressure) and the SPS metrics. This is unexpected because a relationship between predictability and the SPS metrics makes physical sense. On one hand, a strong SPS is not observed as often as a weak SPS, and so in theory storms associated with a strong SPS should be worse forecast. This would be expected particularly with intensity, given that is was the variable with the strongest correlation with forecast lead time, as seen in Chapter 5. On the other hand, stronger forcing is likely to give a stronger storm, so more intense storms would be related to a stronger SPS. Either way, little correlation between the strength of the SPSs is observed, by any metric, and forecast quality, in terms of location or intensity. A couple of combinations exhibit weak but significant correlation. However, on the whole, the relationship between storm SPSs and predictability is feeble.

6.10 Summary

This work has identified five metrics for storm-prone situations (SPSs):

- Dry quasi-geostrophic growth rate (QGD)
- Moist quasi-geostrophic growth rate (QGM)
- Dry semi-geostrophic growth rate (SGD)
- Moist semi-geostrophic growth rate (SGM)
- Barotropic growth rate (BT)

These consist of four baroclinic growth rates (QGD, QGM, SGD, SGM) and one barotropic (BT). There are two different models used in the derivation of the baroclinic versions: the quasi-geostrophic and semi-geostrophic, which differ in their treatment of vorticity. The quasi-geostrophic versions consider only planetary vorticity, whereas the semi-geostrophic versions consider the absolute (planetary plus relative) vorticity. The treatment of moisture also differs, with the dry versions using potential temperature (θ) and the moist using saturated equivalent potential temperature (θ_{es}). The barotropic growth rate simply calculates the horizontal wind shear, and so it is not appropriate to consider moisture here. These different metrics are calculated and averaged over a box covering much of the North Atlantic and Europe (Figure 6.2). Sensitivity testing is performed to

optimise the levels in the atmosphere that are used in the calculation, the area used for averaging, and the best measure of moist temperature (Section 6.2.1). Comparing the growth rates (Section 6.3) shows that the baroclinic growth rates are strongly correlated with each other, but not with the barotropic growth rate. This is consistent with atmospheric stability being the determining factor for the baroclinic growth rates, with treatment of moisture and vorticity being secondary factors, and the baroclinicity not being strongly related to horizontal wind shear. Examining the relationships between these metrics and a set of severe midlatitude cyclones is novel.

The success of finding identifying the selected storms and a peak in at least one metric is high (Section 6.4.3), with 29 out of 31 storms associated with a value at or above the 98th percentile in one or more of the growth rates. Therefore, the link between the area-averaged growth rates and midlatitude cyclones is strong. On an individual basis, a high value (over the 98th percentile) of the QGD growth rate is most often associated with a storm, so from an absolute perspective, QGD is the best performing metric (Table 6.3). This underlines the importance of dry baroclinicity in the development of these severe midlatitude cyclones. On the same score, BT performs poorly, with peaks being associated with the fewest storms. This indicates the relative importance of moisture and vorticity over the barotropicity of the atmosphere.

The examples of storms and their storm-prone situations (Section 6.5) illustrate that the relationship between the storms and the value of the SPS metrics varies considerably, due to the differing relative importance of the different factors used to calculate the different growth rates (e.g. moisture, vorticity). This analysis allows discussion of the idea of a 'perfect storm' of processes, whereby baroclinicity, barotropicity and moisture combine to produce extreme deepening in the most intense case (Klaus), but also combine to produce a less intense storm (Quinten). In the rest of the case studies, these processes are seen alone or in pairs to generate other members of the set of severe cyclones. Two storms, not associated with a value of any of the SPS metrics over the 98th percentile, are associated with fairly high values of those metrics (Table 6.3), and in both cases there is a peak just outside the window (96 hours before, 24 hours after). While this could be an argument for widening the window, this is not done in this work because it does not make dynamical sense; intense midlatitude cyclones rarely take longer than five days to cross the Atlantic. Furthermore, in times of 'storm seriality' (Stephenson, 2006), when a large number of storms occur in a short period of time, then the storm-prone situations would overlap more if the window were wider. Therefore, it is left as 96 hours before maximum SSI, and 24 hours after.

Once the SPSs that are associated with a selected storms had been examined, it is important to examine the SPSs not associated with a selected storm – here called a 'null case' (Section 6.6). These are mainly associated with a storm that is moving in the wrong direction or at the wrong speed to pass over Europe. This illustrates that the SPS metrics alone could never be used to forecast a potentially damaging storm over Europe, because the direction and speed of the storm track are also important. However, they will not be included here in the development of an SPS metric, because the aim of this work is to concentrate on developing a pertinent metric for the large-scale configuration of the atmosphere at the time of cyclogensis.

Overall, the strength of the relationship between the SPS metrics and severe midlatitude cyclones is strong. The 'null cases' are mainly associated with storms that do not pass over Europe. This indicates that, if a model were trying to simulate these severe midlatitude cyclones, it does need to simulate the strong SPSs well. This has implications for climate modelling, which are known to have deficiencies in their modelling of severe midlatitude cyclones. On one hand, if they are also deficient in modelling SPSs, then there is a problem with the generation of strong, large-scale baroclinicity. On the other hand, if climate models simulate SPSs realistically, this implies there is a limitation in modelling the conversion of the large-scale energy to the kinetic energy of a cyclone. Determining which is causing climate models to not realistically represent severe midlatitude cyclones would allow the identification of model limitations and direct future model development. However, determining this is left to future work.

The next steps of this work explore whether the link between the storms and the SPSs metrics varies, depending on some properties of the storm under consideration. It examines the how the SPS metrics perform relative to each other when considering storm intensity and the two different methods of categorisation described in Chapter 4. There is only a weak connection between storm intensity and the extremity of the SPSs metrics. In most cases, there is only weak association between the different SPS metrics and the categorisation of the storms in terms of their jet stream (Section 6.8.1). When a similar analysis is performed for the PTE categorisation, two more coherent links emerge. Firstly, the horiz-type storms are more strongly related to the two dry baroclinic metrics than the moist metrics, which is consistent with the need for a baroclinic atmosphere (i.e. one with a horizontal temperature gradient) to facilitate horizontal temperature advection. This shows that the two approaches are related: the PTE analysis which examines the processes that govern a storm's deepening on the small scale, and the large-scale SPS approach taken in this Chapter. Secondly, the diab-type storms are weakly related to the two moist baroclinic measures

and strongly related to the barotropic SPS measure, compared to their relationship with the dry baroclinic metrics. This indicates that, when the baroclinicity is insufficiently strong for a storm to be associated with strong horizontal temperature advection, the diabatic processes and background horizontal wind shear work together to generate a strong midlatitude cyclone.

When predictability is considered, there is only a weak relationship with the different SPS metrics. While most combinations of the five SPS metrics and the three measures of forecast quality (longitude, latitude and core pressure) show no significant correlation, there are a couple of exceptions. The storms associated with strongly barotropic initial conditions tend to have better forecast longitudes, indicating that barotropicity affects the speed of the storm. There is also correlation between the two dry, baroclinic measures (QGD, SGD) and the latitude of the storm, so baroclinicity influences the zonal region through which the storm moves. However, this correlation is negative, meaning that storms with stronger baroclinicity are worse forecast in terms of latitude, probably because the forecast models cannot fully represent the strong temperature gradients. An unexpected result is the lack of correlation between the forecast quality for pressure and the strength of the SPS, but this may simply mean that intensity is most strongly affected by the chaotic nature of forecasting.

Overall, these proposed SPS metrics are related to the selected storms, and do illuminate some of the subtleties in storm properties. Notably, there is clearly a relationship between the large-scale approach that the SPS analysis takes, and the small-scale approach of PTE analysis. This is illustrated by the link between the storms where the baroclinicity is high as shown by the SPS analysis, and those where horizontal temperature advection dominates the deepening according to the PTE analysis. There is a similar link between storms associated with high values of the barotropic growth rate and those where diabatic processes dominate the deepening, implying that these processes can work together to generate a severe storm. However, only weak links are found between the strength of the SPS and the predictability of the storms. In summary, this chapter has achieved what was set out in Research Question 3: to develop SPS metrics, and explore their relationship with storm type and predictability. It has found that, in combination, the SPS metrics are strongly linked to the selected storms; an idea that has important implications for climate model development.

Chapter 7

Discussion and Conclusions

Midlatitude cyclones regularly inflict damage, both human and economic, on Europe. This work has investigated a set of such historical storms, which had the potential to inflict damage due to the high winds they produced. Analysis concentrated on their NWP forecasts, before exploring the relationships between forecast quality and spread, and the storms' dynamics. The dynamical analysis included an assessment of the large-scale configuration of the atmosphere around the time the storm was generated (a storm-prone situation, or SPS).

The results of this work could help to identify sources of uncertainty in forecasts of severe midlatitude cyclones. Identifying the processes that cause the uncertainty means that future work can aim to minimise them through model development, and so improve forecast quality. Improving the understanding of these processes will facilitate improvements in model simulations of these cyclones, on time scales from those typical of NWP forecasts to those of climate models. Climate model simulations disagree about the evolution of midlatitude cyclones under anthropogenic climate change (Table 2.1), indicating that current understanding of midlatitude cyclones is wanting. While some studies have investigated the jet stream in climate models, there is a complex interplay between the jet and the cyclone making sources of uncertainty difficult to identify. Taking the 'seamless' approach of using shorter simulations to identify sources of uncertainty provides insight into the limitations of climate models (e.g. Martin et al., 2010).

The examination of a set of storms of this size (31 storm events) is a novel approach that allows each storm to be investigated in greater depth than would be possible if all of the storms in a season were considered, but gives a broader picture than investigating case studies. Linking storms

and SPSs directly is also a new approach. Furthermore, exploring the relationships between the types of storm and their predictability has not been attempted in previous work, to the best of the author's knowledge. This chapter is divided into sections based on the objectives stated in Section 1.3. First, the results from the selection (Section 7.1) and categorisation (Section 7.2) of storms are reviewed, followed by the findings from the analysis of NWP forecasts (Section 7.3). The final sections discuss storm-prone situations (SPSs), first in terms of their identification (Section 7.4) and then in terms of their relationship with storm categorisation (Section 7.5).

7.1 Storms Selection

The first step was to select a set of severe historical European windstorms, explained in Chapter 4. The current work focusses on wind damage. While precipitation can also cause damage through flooding, the flood risk is affected by a large number of topographical factors (e.g. proximity to river and floodplain), so this work focusses on wind risk. The metric used to select the storms in this work was the Storm Severity Index (SSI), first proposed by Leckebusch et al. (2008b). In essence, SSI relates meteorological variables to the damage a storm could inflict over an area of interest; in this case, over a portion of Europe. Another criterion for the selected storms was that they should be recent, for reasons of data availability, and so those before 1st January 1990 were excluded.

Calculating SSI led to a list of dates with exceptional wind speeds, but some of these were rejected for a number of reasons (Section 4.2). Many dates were associated with the same high-wind event, and so were rejected on grounds of duplication. Some of the remaining dates were excluded because they were associated with meteorological phenomena that are not the topic of the current investigation: Mediterranean cyclones, high pressure, polar lows, or in which the winds were influenced by orography. Finally, a set of 31 severe North Atlantic cyclones were chosen, with the potential to inflict damage, including many of the most well-known storms from the period since 1990 (e.g. Daria, Lothar, Kyrill). These were then tracked automatically (Section 3.3), to facilitate the next stages of analysis.

7.2 Storm Categorisation

Chapter 4 also presented results that categorised the storms. Many different options are available for categorising cyclones; for example whether a storm is more similar to the Norwegian or Shapiro-Keyser model (Sections 2.2.1 and 2.2.2), or whether it exhibits cyclonic or anticyclonic wave breaking (Thorncroft et al., 1993) (Section 2.2.6). However, given this work's aim to investigate SPSs, the categorisation should consider factors that affect the deepening of the cyclone. Two approaches were identified. First, the jet streams of the 31 cyclones were examined relative to the storms' tracks, and four categories were determined:

- Storms that cross the jet stream **early** in their lifetime (7 storms),
- Storms that cross the jet stream **late** in their lifetime (10 storms),
- Storms that do not cross the jet, but moved along the **edge** of it (10 storms),
- Storms that are associated with a **split** jet (4 storms).

The second approach used a method from Fink et al. (2012): the pressure tendency equation (PTE). This quantifies the contribution of various processes to the deepening of the storm, using a $3^o \times 3^o$ cuboid around the centre of the storm, extending from the surface to 100hPa. It assesses the changes in properties of the column between successive points on the storm's track; the upper lid, the temperature and the moisture profile. The temperature term is divided into horizontal, vertical and diabatic effects. The percentage contributions of each process to the deepening were calculated, and two dominated the picture resulting in two categories:

- horizontal temperature advection ('horiz') (24 storms),
- diabatic processes ('diab') (7 storms).

While some storms have undergone PTE analysis by Fink et al. (2012), this is the first time that a set of this size has been classified. In addition, using the percentage contributions to divide the storms into these categories is a new, more objective method for classifying storms. In the future, PTE analysis could be used to identify NWP model limitations, by comparing the terms for a forecast storm to those in the analysis and identifying discrepancies.

The two methods of categorisation are linked. The strongest association between these two categories was between the 'horiz'-type storms and those that move along the edge of the jet stream, or cross the jet stream early in their lifetimes. The storms where horizontal temperature advection dominates the deepening spend much or all of their lifetime on the northern side of the jet, where there is little potential for moist processes to occur because the air is colder and drier. Therefore, there is little potential for diabatic processes to occur, so these storms are more likely to be driven by horizontal temperature advection. There is a weak link between 'diab' storms and those that cross the jet stream late or have a split jet stream, because on the warm, moist side of the jet there is greater potential for diabatic processes to occur, meaning these storms' deepening is more likely to have a greater contribution from diabatic processes. These results show that the small-scale terms of the PTE are linked to the large-scale influences on the storm, such as the jet stream and background temperature and humidity fields.

7.3 Storm Predictability

The next portion of the work investigated the predictability of these storms, in terms of forecast quality and spread (Chapter 5). The strategy followed in this section deviates from most previous studies, which have taken either a case study or statistical approach. In agreement with these previous studies (e.g. Froude, 2010), in the forecasts the storms are not as intense, are slower moving, and are further north, than in the analysis. In the current work, no link was found between the forecast quality and the intensity of the storms, as measured by SSI of the storms. This is because SSI is sensitive to a number of factors, such as the wind climatology of the area over which a storm passes, that are not influenced by forecast quality.

Next, this thesis investigated the predictability of the different categories of the storms. Analysis of the jet stream types has shown that the forecast models could have limitations simulating the interaction with the jet stream, because the forecast quality and spread improve at longer lead times for storms that cross the jet stream early, compared to their later crossing counterparts. On the one hand, this could be because the interaction of the jet stream and storm centre is not well simulated by the model, because the uncertainty in the forecast reduces once this point has passed. This could be improved by studying jet interaction in more depth, through idealised or case study-based modelling. This approach would allow exploration of the sensitivity of different factors, along the lines of the approach taken by (Willison et al., 2013) to explore diabatic processes. On

the other hand, this uncertainty could be because crossing the jet is typically a storm's period of strongest deepening, so small errors can propagate quickly; these errors could be reduced by improving the initial conditions that feed the forecast, either in terms of observations (number or quality) or the data assimilation scheme. However, further analysis is required before concluding definitively that the jet crossing is poorly simulated.

Similar analysis of the PTE categories reveals that 'horiz' storms are slightly less well forecast than their 'diab' counterparts, but the spread is greater in 'diab' type storms. This could mean that there are deficiencies in the initial conditions or model simulations of horizontal temperature advection, but considerable uncertainty in the diabatic processes remains. However, diabatic processes are parametrised similarly in the NWP forecast model, and the NWP model used to create the reanalysis. Reanalyses have been shown to be influenced by the NWP model used (Bengtsson et al., 2004b), so the validity of a comparison between reanalysis and NWP data in terms of the diabatic processes is open to question.

The current work concludes that there are limitations to the simulation of severe midlatitude cyclones in NWP models. Forecast quality and spread could be enhanced by improving the simulation of 'classical' cyclogenetic processes, as well as by improving parametrisation schemes for the diabatic processes. Classical processes include the jet stream's influence by providing a source of divergence aloft and vorticity advection, and the representation of large-scale temperature gradient that is the source of baroclinicity and a source of potential energy for a midlatitude cyclone. Alternatively, improving the initial conditions and data assimilation schemes, for example through more reliable inclusion of a satellite-derived humidity field, would mean fewer errors develop at the start, leading to decreased potential for such errors to propagate as the forecast progresses. Improving short-term forecasts would facilitate mitigating action when severe cyclones occur in the future, which could lessen the adverse human and economic impacts seen in past storms.

7.4 Identifying Storm-Prone Situations

The final results chapter (Chapter 6) concentrated on SPSs. This work identified a metric for SPSs in the Eady growth rate (Lindzen and Farrell, 1980), and drew in the ideas of Emanuel et al. (1987) and Whitaker and Davis (1994) to include moisture. These papers use two different approaches to their development of the growth rate: either using the semi-geostrophic equations (Emanuel et al., 1987), or the quasi-geostrophic equations (Lindzen and Farrell, 1980; Whitaker and Davis, 1994).

The two sets of equations treat vorticity differently (Section 6.1), in that the quasi-geostrophic equations only contain planetary vorticity, whereas the semi-geostrophic equations also consider relative vorticity. However, both models include stability and vertical wind shear in the growth rates. Moisture is an important factor in the development of midlatitude cyclones, because it implicitly describes the potential for the occurrence of diabatic processes (e.g. latent heat release through condensation of water vapour) and their effect on atmospheric stability. Therefore, the inclusion of moisture affects the growth rate: it significantly increases. In addition to these measures of baroclinicity, the barotropic growth rate was also calculated (Equation 6.12). In short, there are five candidate SPS metrics:

- Dry quasi-geostrophic growth rate (QGD)
- Moist quasi-geostrophic growth rate (QGM)
- Dry semi-geostrophic growth rate (SGD)
- Moist semi-geostrophic growth rate (SGM)
- Barotropic growth rate (BT)

These measures are averaged over a large part of the North Atlantic and Europe, for every winter in the ERA-Interim data set (1979/1980 - 2011/2012). In theory, these measures have a high value ahead of storm development, representing the energy building up before being converted into kinetic energy of the storm. The next step involved quantifying whether high values are related to the selected storms. This was calculated by associating each storm with a time window that covered four days before and one day after the storm reached maximum SSI. This was chosen to include the typical time period between when a storm undergoes cyclogenesis to when it affects Europe. The maximum of each of the five SPS metrics within this window was identified, and the percentile in which this fell was calculated. If it was over the 98th percentile, the storm was said to be associated with a high value of the SPS metric.

In combination, these metrics are a powerful indicator of storm potential. 94% of the storms are associated with a high value in one or more of the five metrics. The next portion of the work aimed to determine which combinations of the metrics were most often connected with one of the selected storms. Two combinations also had this very high success rate of 94%: QGD, SGM and BT or QGM, SGD and BT. This is because these combinations of metrics represent processes that

encourage deepening: the classic, dry baroclinicity; moist processes; and barotropic processes. It should be noted that, theoretically these processes are only a necessary but not a sufficient condition to generate a midlatitude cyclone. To release energy through these processes, there must be a perturbation, which can then use this energy to deepen and gain kinetic energy. The high success rate is promising, in terms of improving the forecast of severe midlatitude cyclones. If a severe storm occurs, in most cases it is associated with a strong SPS. Therefore, if a model does not generate strong enough SPSs, then it is unlikely to generate such cyclones. Making improvements to the modelling of SPSs would improve the modelling of intense North Atlantic cyclones.

These metrics were then examined through examples. Klaus is the storm with the highest SSI, from the storms selected. Klaus is an example of a 'perfect storm', where the initial conditions (a strong jet stream, a strong temperature gradient and a strong humidity gradient) and cyclogenetic processes (baroclinic and barotropic energy conversion, and diabatic processes) work together to generate an exceptionally intense storm. While these processes also work together to generate Quinten, a less intense storm, the initial conditions are not as strong meaning there is less potential energy for the storm to use. This agrees with previous work (Roebber and Schumann, 2011), which found that the same deepening processes occur in very intense storms, compared with those that act in an average midlatitude cyclone. Case studies that were only identified in the barotropic growth rate were also examined, because previous work discussing barotropic influences on midlatitude cyclones has provided evidence that barotropic growth contributes towards deepening (Kucharski and Thorpe, 2000) but can limit the baroclinic growth of a storm (James and Gray, 1986). The case studies show that both of these could be true; neither storm was associated with strong baroclinicity, but both had a sufficiently high SSI to be included in the set of storms. Further work could include deeper investigation of these case studies, to unpick the contribution towards their deepening of barotropic and baroclinic processes. Model simulations of these storms could also allow further unpicking of development; for example, if the barotropic contribution were lower, whether the baroclinicity would increase.

There are two ways in which the SPS metrics can fall short: cases where there is a storm but no SPS, and cases where there is an SPS but no storm. The former only occurs twice in the set of 31 storms. For the latter, there are 1,143 occasions when there is a peak in one or more SPS metric but no selected storm. Many of these peaks were associated with midlatitude cyclones that were not moving at the right speed or in the right direction, to cause high values of SSI over Europe. This means that they are unlikely to inflict any damage on Europe, so are not of interest to this study.

Overall, the link between the selected storms and the SPS metric is strong, with QGD performing best.

7.5 Storm-Prone Situations and Storm Categories

The final objective of this work was to investigate links between the SPS metrics derived in Chapter 6, and the different types of storm discussed in Chapter 4 or the predictability discussed in Chapter 5. These results are discussed in Sections 6.7 to 6.9, where the relative performance of each metric is considered. There is little association between the SPS metrics and storm intensity (SSI). This makes physical sense, because not only are other factors contributing to the deepening that are not included explicitly in the SPS (such as diabatic processes, which are only included implicitly), but also there are many factors that affect the value of SSI (such as the wind climatology of the geographical region through which the storm passes). There is also only a weak link between the SPS metrics and the jet stream types, which means the relationship between the jet and cyclone is not strongly linked to the type of SPS present.

The relationship between the SPS metrics and the PTE categories is significant. The dry, baroclinic growth rates (QGD, SGD) perform relatively well to identify the 'horiz'-type storms, compared to their moist counterparts (QGM, SGM). This can be explained physically because an atmosphere with a horizontal temperature gradient not parallel to the pressure gradient is baroclinic, and this is required for horizontal temperature advection to occur. Another implication of this result is that, despite acting on very different spatial scales, the techniques of measuring the SPS (over the North Atlantic and Europe) and analysing the PTE terms (in a $3^o \times 3^o$ box) are related. This suggests that the large-scale forcing of a storm can also be seen on the smaller-scale, and so the approaches complement each other. The 'diab'-type storms are more strongly linked to the barotropic growth rate and weakly to the two moist, baroclinic growth rates (QGM, SGM). The link with the moist growth rates occurs because diabatic processes such as latent heat release require a moist atmosphere to occur. The link with barotropicity implies that, when baroclinicity is insufficiently strong for horizontal temperature advection to dominate the deepening or for the dry baroclinic growth rates to be exceptionally high, then the horizontal wind shear becomes a more important factor in the SPS, and contributes more strongly towards the deepening.

Finally, the SPSs were studied, in terms of storm predictability. Links were sought between the strength of the SPS and the three quantities of interest of forecast quality: the differences between

forecast and analysis pressure, latitude and longitude. Overall, the strength of an SPS is generally not significantly related to forecast quality.

7.6 Future Work

Exploring the implications of this work in terms of model development is left to future work. Whether uncertainties in the forecast are related to the initial conditions or the model physics is difficult to determine, but this could be unpicked using a different approach. Investigating further the relationship between crossing the jet and forecast quality or spread would be particularly interesting. A method of objectively identifying when the jet is crossed could be devised, which also removes any storms that do not cross the jet ('edge' type) or are associated with more than one jet ('split' type). This could then be applied to a larger set of storms and their deterministic and ensemble forecasts, to identify any model deficiencies. Alternatively, examining the storms selected for the current work in model simulations would allow sensitivity to different parametrisation schemes to be tested, and so shed light on whether it is model error that is leading to the limitations in the forecast. Model simulations could also allow sensitivity to initial conditions to be explored, for example changing the large-scale temperature or humidity field, and analysing the effect on the storm. A third approach could investigate the terms of the PTE in ensemble forecasts would allow exploration of the sensitivity of each term to the initial conditions.

The strong link between SPSs and the selected storms has implication for future work into model development. If SPSs are well forecast but the resulting storms are not, then this would signal a model limitation in extracting the potential energy from the SPS and converting it to the kinetic energy of the storm. However, if the initial SPSs are not well forecast, then this would indicate a flaw in the large-scale forcing that causes an SPS to form. This analysis is left to future work. A statistical approach would be recommended, comparing the distribution of the SPS metrics in reanalysis data to those in forecast data at a variety of lead times. Alternatively, investigating some of the selected storms as case studies would allow deeper analysis and identification of the shortcomings of a particular storm's forecast. Applying these techniques would provide information about the limitations of the simulations of midlatitude cyclones, and so direct model development into upgrading the most important model deficiencies.

A similar question could be asked of climate models. If climate models are to realistically simulate midlatitude cyclones, then they must be able to simulate the SPS. If the models fail to simulate an

SPS, then this indicates a different limitation in the modelling than if the models do simulate an SPS but then do not generate a storm from it. Again, this could be done by examining either case studies or climatology, compared to reanalysis. Previous work has mapped how QGD will evolve under climate change (e.g. Greeves et al., 2007), and compared this to reanalysis data; however, in contrast to the current work, averaging was typically done over a period of years, and not over a domain. Domain averaging means that the evolution of an SPS metric could be explored through time. Furthermore, investigation of the other SPS metrics has not been undertaken, and particularly the moist growth rates could yield interesting results, given the uncertainty in the evolution of the moisture field under future climate. Therefore, a future examination of SPSs in climate models is recommended in the future.

Future work might also include using different measures of storm intensity, such as minimum pressure or maximum vorticity, to select the storms, examine correlation with predictability, or explore the relationship with SPSs. Alternatively, extending the work from using ECMWF data to other centres, such as the UK Met Office or NCEP, could prove a fruitful avenue of investigation. This is not done here, because the tracking algorithm used in this work would require extensive adjustment to cope with lower temporal resolution data, which would make comparing the results to ECMWF data difficult.

Another source of future work could be to use a larger set of storms, because this was not feasible in the current project. If the storms could be categorised automatically, then the forecasts of a large set of storms could be analysed. This would seem more feasible with the PTE categorisation, which is simply based on comparing percentage contributions, rather than the jet analysis, which involves examining plots of the jet stream relative to each storm's track. If automation is not possible, future work could update the results to include severe midlatitude cyclones that have occurred since 2010, as there have been several periods of damaging storms since then, and potentially include investigation of the DIAMET case studies (Vaughan et al., 2014).

Finally, the PTE analysis could also be extended to explore the contribution of different processes to the diabatic heating. Such processes include the release of latent heat by condensation or by melting, absorption of short-wave solar radiation, and absorption or release of long-wave infra-red radiation.

7.7 Summary

Severe midlatitude cyclones were investigated in terms of the different processes that govern their deepening, their predictability, and the large-scale configuration of the atmosphere prior to their development. The approach allowed a set of 31 midlatitude cyclones to be investigated in some depth, representing the spectrum of North Atlantic cyclones. The storms were selected based on the SSI value (Leckebusch et al., 2008b), which quantifies the potential a storm has to inflict damage. However, some high-SSI dates needed to be rejected because they were not associated with midlatitude cyclones, which is the first time that weather phenomena other than such cyclones have been seen to be related to high SSI. The storms were divided in two ways, the first of which was to identify the type of jet stream with which each storm was associated. The second method of categorising the storms used the processes in the PTE, which builds on the work of Fink et al. (2012) to include more storms and to divide them objectively using the percentage contribution of each process.

Analysis of the storms' forecast quality and spread allowed deficiencies in the modelling to be identified. Previous studies have concentrated on statistical or case-study approaches. This investigation of 31 midlatitude cyclones takes the best from both of these, in that statistical tests are performed, but each storm is also examined in depth. Forecast quality is better but forecast spread larger in diabatically-driven storms, compared to those where horizontal temperature advection dominates the deepening. Furthermore, the uncertainty in the forecast decreases earlier in storms that cross the jet stream early in their lifetime compared to storms that cross the storm late, which could mean that there is significant uncertainty associated with the storms' interaction with the jet stream.

Together, these results indicate there are still limitations in modelling baroclinicity and the jet stream interaction as well as diabatic processes; the sources of these problems could stem from model error or from the initial conditions. Improving the model's representation of these processes could make forecasts of storms more reliable. Alternatively, improving the observations or the data assimilation scheme would mean fewer errors in the initial conditions, which can propagate to become large errors in the forecast, particularly when the storm is interacting with the jet and deepening quickly. Determining whether the limitations are mainly due to model error or initial conditions is left to future work.

Identification of the five SPS metrics allowed their relationship with the set of storms to be investigated. Between them, the metrics represent several factors that influence cyclogenesis, including baroclinicity, moisture, vorticity and barotropicity. The barotropic growth rate is used (Gill, 1982, Equation 13.6.12), alongside four baroclinic metrics proposed by three previous studies that took an idealised approach (Emanuel et al., 1987; Lindzen and Farrell, 1980; Whitaker and Davis, 1994). Previous work has examined the statistical picture of these growth rates, or used them in a case study of a midlatitude cyclone, and has concentrated on the dry verstion of the growth rate. The current work includes some effects of moisture in the growth rate, and examines the relationship with a set of real storms. A strong relationship is found: of the 31 storms, 29 were associated with an exceptionally high value of one or more of these metrics. There are many high values of these metrics that are not linked to the selected storms, but these are linked to strong storms that do not track over Europe. The different dynamics of the storms, established by determining which processes dominate the deepening according to PTE analysis, were successfully related to the SPS metrics. This has important implications for identifying limitations in modelling severe, midlatitude cyclones on a range of timescales, from NWP to climate models.

References

Alexander, L. V. and S. F. B. Tett, 2005: Recent observed changes in severe storms over the UK and Iceland. *Geophysical Research Letters*, **32**, L13 704.

Alexandersson, H., H. Tuomenvirta, T. Schmith, and K. Iden, 2000: Trends in the storms of northwestern Europe derived from an updated pressure set. *Climate Research*, **14**, 71–73.

Allen, J. T., A. B. Pezza, and M. T. Block, 2010: Explosive cyclogenesis: A global climatology comparing multiple reanalyses. *Journal of Climate*, **23**, 6468–6484.

Anstey, J. A., et al., 2013: Multi-model analysis of Northern Hemisphere winter blocking, Part I: Model biases and the role of resolution. *Journal of Geophysical Research: Atmospheres*, **118**, 3956?3971.

Azad, R. and A. Sorteberg, 2009: A diagnosis of warm-core and cold-core extratropical cyclone development using the Zwack-Okossi equation. *Atmospheric Science Letters*, **10**, 220–225.

Bader, M. J., G. S. Forbes, J. R. Grant, R. B. E. Lilley, and A. J. Waters, 1995: *Images in weather forecasting : a practical guide for interpreting satellite and radar imagery*. Cambridge University Press.

Beck, C., J. Jacobeit, and P. D. Jones, 2007: Frequency and within-type variations of large-scale circulation types and their effects on low-frequency climate variability in central Europe since 1780. *International Journal of Climatology*, **27**, 473–491.

Beersma, J. J., K. M. Rider, G. J. Komen, E. Kaas, and V. V.Kharin, 1997: An analysis of extratropical storms in the North Atlantic region as simulated in a control and $2 \times CO_2$ time-slice experiment with a high-resolution atmospheric model. *Tellus A*, **49**, 347–361.

Bengtsson, L., S. Hagemann, and K. I. Hodges, 2004b: Can climate trends be calculated from reanalysis data? *Journal of Geophysical Research*, **109**, D11 111.

Bengtsson, L., K. I. Hodges, and S. Hangemann, 2004a: Sensitivity of the ERA40 reanalysis to the observing system: determinations of the global atmospheric circulation from reduced observations. *Tellus A*, **56**, 456–471.

Bengtsson, L., K. I. Hodges, and E. Roeckner, 2006: Storm tracks and climate change. *Journal of Climate*, **19**, 3518–3543.

Beniston, M., et al., 2007: Future extreme events in European climate: an exploration of regional climate model projections. *Climatic Change*, **81 (1)**, 71–95.

Bjerknes, J. and H. Solberg, 1922: Life cycle of the cyclones and the polar front theory of atmospheric circulation. *Geophysical Publications*, **3**, 3–17.

Black, M. T. and A. B. Pezza, 2013: A universal, broad-environment energy conversion signature of nexplosive cyclones. *Geophysical Research Letters*, **40**, 452–457.

Blackmon, M. L., 1976: A climatological spectral study of the 500 mb geopotential height of the northern hemisphere. *Journal of the Atmospheric Sciences*, **33**, 1607–1623.

Blender, R. and M. Schubert, 2000: Cyclone tracking in different spatial and temporal resolutions. *Monthly Weather Review*, **128**, 377–384.

Bolton, D., 1980: The computation of equivalent potential temperature. *Monthly Weather Review*, **108**, 1046–1053.

Booth, J. F., S. Wang, and L. Polvani, 2012: Mid-latitude storms in a moister world: Lessons from idealised baroclinic lifecycle experiments. *Climate Dynamics*, **41**, 787–802.

Brönnimann, S., 2007: Impact of El Niño-Southern Oscillation on European climate. *Reviews of Geophysics*, **45**, RG3003.

Browning, K. A., 2004: The sting at the end of the tail: Damaging winds associated with extratropical cyclones. *Quarterly Journal of the Royal Meteorological Society*, **130**, 375–399.

Browning, K. A. and N. M. Roberts, 1994: Structure of a frontal cyclone. *Quarterly Journal of the Royal Meteorological Society*, **120**, 1535–1557.

Buizza, R. and T. N. Palmer, 1998: Impact of ensemble size on ensemble prediction. *Monthly Weather Review*, **126**, 2503–2518.

Buizza, R., D. S. Richardson, and T. N. Palmer, 2003: Benefits of increased resolution in the ECMWF ensemble system and comparison with poor-man's ensembles. *Quarterly Journal of the Royal Meteorological Society*, **129**, 1269–1288.

Carleton, A. M., 1988: Meridional transport of eddy sensible heat in winters marked by extremes of the North Atlantic Oscillation, 1948/49 – 1979/80. *Journal of Climate*, **1**, 212–223.

Carnell, R. E. and C. A. Senior, 1998: Changes in mid-latitude variability due to increasing greenhouse gases and sulphate aerosols. *Climate Dynamics*, **14**, 369–383.

Catto, J., L. Shaffrey, and K.I.Hodges, 2011: Northern henisphere extratropical cyclones in a warming climate in the HiGEM high-resolution climate model. *Journal of Climate*, **24**, 5336–5352.

Centre for Research on the Epidemiology of Diseases, 2012: The internantional disaster database. URL: www.emdat.be, uRL: www.emdat.be.

Chagnon, J. M., S. L. Gray, and J. Methven, 2012: Diabatic processes modifying potential vorticity in a North Atlantic cyclone. *Quarterly Journal of the Royal Meteorological Society*, **139**, 1270?1282.

Chang, C. B., D. J. Parkey, and C. W. Kreitzberg, 1982: A numerical cate stude of the effects of latent heaheat on a developing wave cyclone. *Journal of the Atmospheric Sciences*, **39**, 1555–1569.

Chang, E. K. M., Y. Guo, and X. Xia, 2012: CMIP5 multimodel ensemble projection of storm track change under global warming. *Journal of Geophysical Research*, **117**, D23 118.

Chang, E. K. M., Y. Guo, X. Xia, and M. Zheng, 2013: Storm-track activity in IPCC AR4/CMIP3 model simulations. *Journal of Climate*, **26**, 246–260.

Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *Journal of Meteorology*, **4**, 135–163.

Christensen, J. H., et al., 2007: Regional climate projections. *Climate change 2007: The physical science basis*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, Eds., Cambridge University Press, chap. 11.

Christensen, J. H., et al., 2014: Climate phenomena and their relevance for future regional climate change. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to*

the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Colucci, S. J., 2010: Stratospheric influences on tropospheric weather systems. *Journal of the Atmospheric Sciences*, **67**, 324–344.

Compo, G. P., et al., 2011: Review article the twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, **137**, 1–28.

Dacre, H. F. and S. L. Gray, 2009: The spatial distribution and evolution characteristics of North Atlantic cyclones. *Monthly Weather Review*, **137**, 99–115.

Davies, H. C., C. Schar, and H. Wernli, 1991: The palette of fronts and cyclones within a baroclinic wave development. *Journal of the Atmospheric Sciences*, **48**, 1666–1689.

Dee, D. P., et al., 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137 (656)**, 553–597.

Della-Marta, P. M., H. Mathis, C. Frei, M. A. Liniger, J. Kleinn, and C. Appenzeller, 2009: The return period of wind storms over Europe. *International Journal of Climatology*, **29**, 437–459.

Deser, C., A. Phillips, V. Bourdette, and H. Teng, 2011: Uncertainty in climate projections: the role of internal variability. *Climate Dynamics*, **38**, 527–546.

Deser, C., R. Tomas, M. Alexander, and D. Lawrence, 2010: The seasonal atmospheric response to projected Arctic sea ice loss on the late twenty-first century. *Journal of Climate*, **23**, 333–351.

Donat, M. G., G. C. Leckebush, J. G. Pinto, and U. Ulbrich, 2010: European storminess and associated circulation weather types: future changes deduced from a multi-model ensemble of GCM simulations. *Climate Research*, **42**, 27–43.

Donat, M. G., D. Renggl, S. Wild, L. V. Alexander, G. C. Leckebusch, and U. Ulbrich, 2011: Reanalysis suggests longterm upward trends in european storminess since 1871. *Geophysical Research Letters*, **38**, L14 703.

Eady, E. T., 1949: Long wave and cyclone waves. *Tellus*, **1(3)**, 33–52.

Emanuel, K., M. Fantini, and A. Thorpe, 1987: Baroclinic instability in an environment of small stability to slantwise moist convection. Part I: Two-dimensional models. *Journal of the Atmospheric Sciences*, **44**, 1559–1573.

Fink, A. H., T. Bruecher, V. Ermert, A. Krueger, and J. G. Pinto, 2009: The European storm Kyrill in January 2007: synoptic evolution, meteorological impacts and some considerations with respect to climate change. *Natural Hazards and Earth System Sciences*, **9 (2)**, 405–423.

Fink, A. H., S. Pohl, J. Pinto, and P. Knippertz, 2012: Diagnosing the influence of diabatic processes on the explosive deepening of extra-tropical cyclones. *Geophysical Research Letters*, **39**, L07 803.

Finn, C. B. P., 1993: *Thermal Physics*. 2d ed., CRC Press.

Fisher, R. A., 1973: *Statistical Methods for Research Workers*. Hafner Publishing Company.

Fourier, J. B. J., 1822: *Théorie Analytique de la Chaleur*. Chez Firmin Didot, père et fils, translation: Freeman, A., 1955: The Analytical Theory of Heat.

Froude, L. S. R., 2009: Regional differences in the prediction of extratropical cyclones by the ECMWF Ensemble Prediction System. *Monthly Weather Review*, **137**, 893–911.

Froude, L. S. R., 2010: TIGGE: Comparison of the prediction of Northern Hemisphere extratropical cyclones by different ensemble prediction systems. *Weather and Forecasting*, **25**, 819–836.

Froude, L. S. R., L. Bengtsson, and K. I. Hodges, 2007a: The predictability of extratropical storm tracks and the sensitivity of their prediction to the observing system. *Monthly Weather Review*, **135**, 315–333.

Froude, L. S. R., L. Bengtsson, and K. I. Hodges, 2007b: The prediction of extratropical storm tracks by the ECMWF and NCEP Ensemble Prediction Systems. *Monthly Weather Review*, **135**, 2545–2567.

Geng, Q. and M. Sugi, 2001: Variability of the North Atlantic cyclone activity analysed from NCEP-NCAR reanalysis data. *Journal of Climate*, **14**, 3863–3873.

Geng, Q. and M. Sugi, 2003: Possible change of extratropical cyclone activity due to enhanced greenhouse gases and sulphate aerosols — study with a high-resolution AGCM. *Journal of Climate*, **16**, 2262–2274.

Gibbons, J. D., 1985: *Nonparametric statistical inference*. M. Dekker.

Gilet, J.-B., M. Plu, and G. Rivière, 2009: Nonlinear baroclinic dynamics od surface cyclones crossing a zonal jet. *Journal of the Atmospheric Sciences*, **66**, 3021–3041.

Gill, A. E., 1982: *Atmosphere-Ocean Dynamics*. Academic Press.

Gómara, I., J. G. Pinto, T. Wollings, G. Masato, P. Zurita-Gotar, and B. Rodriguez-Fonseca, 2012: Rossby wave-breaking analysis of explosive cyclones in the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, **Online first.**, DOI: 10.1002/qj.2190.

Greeves, C. Z., V. D. Pope, R. A. Stratton, and G. M. Martin, 2007: Representation of Northern Hemisphere winter storm tracks in climate models. *Climate Dynamics*, **28**, 683–702.

Gulev, S. K., O. Zolina, and S. Grigoriev, 2001: Extratropical cyclone variability in the Northern Hemisphere from the NCEP/NCAR reanalysis data. *Climate Dynamics*, **17**, 795–809.

Haas, R. and J. G. Pinto, 2012: A combined statistical and dynamical approach for downscaling large-scale footprints of European windstorms. *Geophysical Research Letters*, **39**, L23 804.

Hanley, J. and R. Caballero, 2012: The role of large-scale atmospheric flow and Rossby Wave Breaking in the evolution of extreme windstorms over Europe. *Geophysical Research Letters*, **39**, L21 708.

Harnik, N. and E. K. M. Chang, 2003: Storm track variations as seen in radiosonde observations and reanalysis data. *Journal of Climate*, **16**, 480–495.

Hart, R. E., 2003: A cyclone phase space derived from thermal wind and thermal asymmetry. *Monthly Weather Review*, **131**, 585–616.

Harvey, B. J., L. C. Shaffrey, T. J. Woolings, G. Zappa, and K. I. Hodges, 2012: How large are projected 21st century storm track changes? *Geophysical Research Letters*, **39**, L18 707.

Hawkins, E. and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, **90**, 1095–1107.

Hodges, K., 1994: A general method for tracking analysis and its application to meteorological data. *Monthly Weather Review*, **122**, 2573–2586.

Hodges, K. I., B. J. Hoskins, J. Boyle, and C. Thorncroft, 2003: A comparison of recent reanalysis datasets using objective feature tracking: storm tracking and tropical easterly waves. *Monthly Weather Review*, **131**, 2012–2037.

Holton, J. R., 2004: *An Introduction to Dynamic Meteorology*. 4th ed., Elsevier Academic Press, London.

Hoskins, B. J., 2013: The potential for skill across thje range of the seamless weather-climate prediction problem: a stimulus for our science. *Quarterly Journal of the Royal Meteorological Society*, **139**, 573–584.

Hoskins, B. J. and K. I. Hodges, 2002: New perspectives on the Northern Hemisphere winter storm tracks. *Journal of the Atmospheric Sciences*, **59**, 1041–1061.

Hudson, R. D., 2012: Measurements of the movement of the jet streams at mid-latitudes in the northern and southern hemispheres. *Atmospheric Chemistry and Physics*, **12**, 7797–7808.

James, I. N., 1995: *Introduction to Circulating Atmospheres*. Cambridge University Press.

James, I. N. and L. J. Gray, 1986: Concerning the effect of surface drag on the circulation of a baroclinic planetary atmosphere. *Quarterly Journal of the Royal Meteorological Society*, **112**, 1231–1250.

Jiang, J. and W. Perrie, 2007: The impacts of climate change on autumn North Atlantic midlatitude cyclones. *Journal of Climate*, **20**, 1174–1187.

Joliffe, I. T. and D. B. Stephenson, 2003: *Forecast verification: a practitioner's guide in atmospheric science*. John Wiley and Sons Ltd.

Keeley, S. P. E., R. T. Sutton, and L. C. Shaffrey, 2013: The impact of North Atlantic sea surface temperature errors on the simulation of the North Atlantic European region climate. *Quarterly Journal of the Royal Meteorological Society*, **138**, 1774–1783.

Klawa, M. and U. Ulbrich, 2003: A model for the estimation of storm losses and the identification of severe winter storms in Germany. *Natural Hazards and Earth System Sciences*, **3**, 725–732.

Knippertz, P., U. Ulbrich, and P. Speth, 2000: Changing cyclone and surface wind speeds over the North Atlantic and Europe in a transient GHG experiment. *Climate Research*, **15**, 109–120.

Krueger, O., F.Shenk, R. Feser, and R. Weisse, 2012: Inconsistencies between long-term trends in storminess derived from the 20cr reanalysis and observations. *Journal of Climate*, **26**, 868–874.

Kucharski, F. and A. J. Thorpe, 2000: Upper-level barotropic growth as a precursor to cyclogenesis during FASTEX. *Quarterly Journal of the Royal Meteorological Society*, **126**, 3219–3232.

Kucharski, F. and A. J. Thorpe, 2001: The influence of transient upper-level barotropic growth on the development of baroclinic waves. *Quarterly Journal of the Royal Meteorological Society*, **127**, 835–844.

Lackmann, G. M., L. F. Bosart, and D. Keyser, 1996: Planetary- and synoptic-scale characteristics of explosive wintertime cyclogenesis over the western north atlantic. *Monthly Weather Review*, **124**, 2672–2702.

Lamb, H., 1991: *Historic Storms of Europe, North Sea, British Isles and Northwest Europe*. Cambridge University Press.

Lambert, S. J., 1995: The effect of enhanced greenhouse warming on winter cyclone frequencies and strengths. *Journal of Climate*, **8**, 1447–1452.

Lambert, S. J., 1996: Intense extratropical Northern Hemisphere winter cyclone evets: 1899-1991. *Journal of Geophysical Research*, **101**, 21 319–21 325 (D16).

Lambert, S. J. and J. C. Fyfe, 2006: Changes in winter cyclone frequencies and strengths simulated in enhanced greenhouse warming experiments: results from the models participating in the IPCC diagnostic exercise. *Climate Dynamics*, **26**, 713–728.

Lambert, S. J., J. Sheng, and J. Boyle, 2002: Winter cyclone frequencies in thirteen models participating in the Atmospheric Model Intercomparison Project (AMIP1). *Climate Dynamics*, **19**, 1–16.

Leckebusch, G., D. Renggli, and U.Ulbrich, 2008b: Development and application of an objective storm severity measure for the northeast Atlantic region. *Meteorologische Zeitschrift*, **17 (5)**, 575–587.

Leckebusch, G. and U. Ulbrich, 2004: On the relationship between cyclones and extreme windstorm events over Europe under climate change. *Global and Planetary Change*, **44**, 181–193.

Leckebusch, G. C., B. Koffi, U. Ulbrich, and J. G. Pinto, 2006: Analysis of frequency and intensity of European winter storms from a multi-model perspective, at synoptic and regional scales. *Climate Research*, **31**, 59–74.

Leutbecher, M. and T. N. Palmer, 2008: Ensemble forecasting. *Journal of Computational Physics*, **227**, 3515–3539.

Liberato, M. L. R., J. G. Pintol, I. F. Trigo, and R. M. Trigo, 2011: Klaus – an exceptional winter storm over northern iberia and southern france. *Weather*, **66**, 330–334.

Lindzen, R. and B. Farrell, 1980: A simple approximate result for the maximum growth rate of baroclinic instabilities. *Journal of the Atmospheric Sciences*, **37**, 1648–1654.

Löptien, U., O. Zolina, S. Gulev, M. Latif, and V. Soloviov, 2008: Cyclone life cycle characteristics over the Northern Hemisphere in coupled GCMs. *Climate Dynamics*, **31**, 507–532.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307.

Lynch, A. H. and J. J. Cassano, 2006: *Applied Atmospheric Dynamics*. John Wiley and Sons Ltd.

Ma, H.-Y., et al., 2014: On the correspondence between mean forecast errors and climate errors in cmip5 models. *Journal of Climate*, **Online first:**, http://journals.ametsoc.org/doi/abs/10.1175/JCLI–D–13–00 474.1.

Martin, G. M., S. F. Milton, C. A. Senior, M. E. Brooks, and S. Ineson, 2010: Analysis and reduction of systematic errors through a seamless approach to modeling weather and climate. *Journal of Climate*, **23**, 5933–5957.

Martin, J. E., 2006: *Mid-Latitude Atmospheric Dynamics*. John Wiley and Sons Ltd, Chichester, UK.

McCabe, G., M. Clark, and M. Serreze, 2001: Trends in Northern Hemisphere surface cyclone frequency and intensity. *Journal of Climate*, **14 (12)**, 2763–2768.

McDonald, R., 2011: Understanding the impact of climate change on Northern Hemisphere extratropical cyclones. *Climate Dynamics*, **37**, 1399–1425, doi:{10.1007/s00382-010-0916-x}.

Meehl, G. A., et al., 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis,

K. B. Averyt, M. Tignor, and H. L. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Mehta, C. R. and N. R. Patel, 1983: A network algorithm for performing fisher's exact test in r Œc contingency tables. *Journal of the American Statistical Association*, **78**, 427–434.

Munich Re, 2009: World map of natural hazards. Available on request from MunihcRe Touch.

Murray, R. and S. M. Daniels, 1952: Transverse flow at entrance and exit to jet streams. *Quarterly Journal of the Royal Meteorological Society*, **79**, 236–241.

Murray, R. J. and I. Simmonds, 1991: A numerical scheme for tracking cyclone centres from digital data. Part I: development and operation of the scheme. *Australian Meteorological Magazine*, **39**, 155–166.

Neu, U., et al., 2013: Imilast: A community effort to intercompare extratropical cyclone detection and tracking algorithms. *Bulletin of the American Meteorological Society*, **94**, 529–547.

Office for National Statistics, 2011: Consumer price indices (http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=consumer+price+indices).

Palmer, T. N., J. Barkmeijer, R. Buizza, and T. Petroliagis, 1997: The ECMWF ensemble prediction system. *Meteorological Applications*, **4**, 301–304.

Palmer, T. N., F. J. Doblas-Reyes, A. Weisheimer, and M. J. Rodwell, 2008: Toward seamless prediction: Calibration of climate change projections using seasonal forecasts. *Bulletin of the American Meteorological Society*, **89**, 459–470.

Peings, Y. and G. Magnusdottir, 2014: Response of wintertime Northern Hemisphere circulation to current and projected Arctic sea ice decline: A numerical study with CAM5. *Journal of Climate*, **27**, 244–264.

Phillipp, A., P. M. Della-Marta, J. Jacobeit, D. R. Fereday, P. D. Jones, A. Monberg, and H. Wanner, 2007: Long-term variability of daily North Atlantic-European pressure patters since 1850 classified by simulated annealing clustering. *Journal of Climate*, **20**, 4065–4095.

Pinto, J. G., T. Spangehl, U. Ulbrich, and P. Speth, 2005: Sensitivities of a cyclone detection and tracking algorithm: individual tracks and climatology. *Meteorologische Zeitschrift*, **14**, 823–838.

Pinto, J. G., T. Spangehl, U. Ulbrich, and P. Speth, 2006: Assessment of winter cyclone activity in a transient ECHAM4-OPYC3 GHG experiment. *Meteorologische Zeitschrift*, **15**, 279–291.

Pinto, J. G., U. Ulbrich, G. C. Leckebusch, T. Spangehl, M. Reyers, and S. Zacharias, 2007a: Changes in storm track and cyclone activity in three SRES ensemble experiments with the ECHAM5/MPI-OM1 GCM. *Climate Dynamics*, **29**, 195–210.

Pinto, J. G., S. Zacharias, A. H. Fink, G. C. Leckebusch, and U. Ulbrich, 2009: Factors contributing to the development of extreme North Atlantic cyclones and their relationship with the NAO. *Climate Dynamics*, **32**, 711–737.

Plavcovà, E., 2011: Evaluation of daily temperatures in central Europe and their links to large-scale circulation in an ensemble of regional climate models. *Tellus A*, **63**, 763–781.

R Core Team, 2013: *R: A Language and Environment for Statistical Computing*. Vienna, Austria, R Foundation for Statistical Computing, URL `http://www.R-project.org/`.

Raible, C. C., P. M. Della-Marta, C. Schwierz, H. Wernli, and R. Blender, 2008: Northern Hemisphere extratropical cyclones: A comparison of detection and tracking methods and different reanalyses. *Monthly Weather Review*, **136**, 880–897.

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

Rees, D. G., 2001: *Essential Statistics*. 4th ed., Chapman & Hall / CRC.

Renggli, D., G. C. Leckebusch, U. Ulbrich, S. N. Gleixner, and E. Faust, 2011: The skill of seasonal ensemble prediction systems to forecast wintertime windstorm frequency over the North Atlantic and Europe. *Monthly Weather Review*, **139**, 3052–3068, doi:{10.1175/2011MWR3518.1}.

Riley, K. F., M. P. Hobson, and S. P. Bence, 2006: *Mathematical Methods for Physics and Engineering*. 3d ed., Cambridge Univeristy Press.

Rivière, G. and A. Joly, 2006a: Role of low-frequency deformation field on the explosive growth of extratropical cyclones at the jet exit. part i: Barotropic critical region. *Journal of the Atmospheric Sciences*, **63**, 1965–1981.

Rivière, G. and A. Joly, 2006b: Role of low-frequency deformation field on the explosive growth of extratropical cyclones at the jet exit. part ii: Baroclinic critical region. *Journal of the Atmospheric Sciences*, **63**, 1982–1995.

Rodwell, M. J. and T. N. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quarterly Journal of the Royal Meteorological Society*, **133**, 129–146.

Roebber, P. and M. R. Schumann, 2011: Physical processes governing the rapid deepening tail of maritime cyclogenesis. *Monthly Weather Review*, **139**, 2776–2789.

Sanders, F. and J. Gyakum, 1980: Synoptic-dynamic climatology of 'the bomb'. *Monthly Weather Review*, **108**, 1589–1606.

Sansom, P., D. B. Stephenson, C. A. T. Ferro, and G. Z. dn L. Shaffrey, 2013: Simple uncertainty fframework for selecting weighting schemes and interpreting multi-model ensemble climate change experiments. *Journal of Climate*, **26**, 4017–4037.

Schneider, T., P. A. O'Gorman, and X. J. Levine, 2010: Water vapor and the dynamics of climate changes. *Reviews of Geophysics*, **48**, RG3001.

Schultz, D., D. Keyser, and L. Bosart, 1998: The effect of large-scale flow on low-level frontal structure and evolution in midlatitude cyclones. *Monthly Weather Review*, **126**, 1767–1791.

Schultz, D. M. and G. Vaughan, 2011: Occluded fronts and the occlusion process. *Bulletin of the American Meteorological Society*, 443–466.

Schwierz, C., C. Appenzeller, H. C. Davies, M. A.Linger, W. Műller, T. F.Stocker, and M. Yoshimori, 2006: Challenges posed by and approaches to the study of seasonal-to-decadal climate variability. *Climatic Change*, **79**, 31–63.

Schwierz, C., P. Köllner-Heck, E. Z. Mutter, D. N. Bresch, P.-L. Vidale, M. Wild, and C. Schaer, 2010: Modelling European winter wind storm losses in current and future climate. *Climatic Change*, **101 (3-4)**, 485–514.

Shapiro, M. and D. Keyser, 1990: Fronts, jet streams and the tropopause. *Extratropical Cyclones: The Erik Palmen Memorial Volumne*, C. W. Newton and E. O. Holopaninen, Eds., American Meteorological Society, Boston, 167–191.

Sinclair, M. R. and I. G. Watterson, 1999: Objective assessment of extratropical weather systems in simulated climates. *Journal of Climate*, **12**, 3467–3485.

Spengler, T., J. Egger, and S. T. Garner, 2011: How does rain affect surface pressure in a one-dimensional framework? *Journal of the Atmospheric Sciences*, **68**, 347–360.

Stephenson, D., 2006: The clustering effect. Tech. rep., Willis Research Network.

Stephenson, D. B., V. Pavan, M. Collins, M. M. Junge, R. Quadrelli, and P. C. M. Groups, 2006: North Atlantic Oscillation response to transient greehouse gas forcing and the impact on European winter climate: a CMIP2 multi-model assessment. *Climate Dynamics*, **27**, 401–420.

Stratton, R. A., 2004: Report on aspects of variability in high-resolution versions of HadAM3. Tech. rep., The Met Office. URL `http://www.metoffice.gov.uk/publications/HCTN/HCTN_53.pdf`.

The WASA Group, 1998: Changing waves and storms in the northeast Atlantic? *Bulletin of the American Meteorological Society*, **79 (5)**, 741–760.

Thorncroft, C. D., B. J. Hoskins, and M. E. McIntyre, 1993: Two paradigms of baroclinic-wave life-cycle behaviour. *Quarterly Journal of the Royal Meteorological Society*, **119**, 12–55.

Tilinina, N., S. K. Gulev, I. Rudeva, and P. Kilterman, 2013: Comparing cyclone life cycle characteristics and their interannual variability in different reanalyses. *Journal of Climate*, **26**, 6419–6438.

Tory, K. J., J. D. Kepert, J. A. Sippel, and C. M. Nguyen, 2012: On the use of potential vorticity tendency for diagnosing atmospheric dynamics in numerical models. *Journal of the Atmospheric Sciences*, **69**, 942–960.

Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modemodel global tteleconnection associated with tropical sea surface temperatures. *Journal of Geophysical Research*, **103 (C7)**, 14,291–14,324.

Trigo, I., 2006: Climatology and interannual variability of storm-tracks in the Euro-Atlantic sector: a comparison between ERA-40 and NCEP/NCAR reanalyses. *Climate Dynamics*, **26**, 127–143.

Trigo, I., T.D.Davies, and G.R.Bigg, 1999: Onbjective climatology of cyclones in the mediterranean. *Journal of Climate*, **12**, 1685–1696.

Uccellini, L. W. and P. J. Kocin, 1987: The interaction of jet streak circulations during heavy snow events along the east coast of the United States. *Weather and Forecasting*, **2**, 289–308.

Ulbrich, U., G. C. Leckebusch, and J. G. Pinto, 2009: Extra-tropical cyclones in the present and future climate: a review. *Theoretical and Applied Climatology*, **96 (1-2)**, 117–131.

Ulbrich, U., J. G. Pinto, H. Kupfer, G. C. Leckebusch, T. Spangehl, and M. Reyers, 2008: Changing Northern Hemisphere storm tracks in an ensemble of IPCC climate change simulations. *Journal of Climate*, **21**, 1669–1679.

Ulbrich, U., et al., 2013: Are greenhouse gas signals of northern hemisphere winter extra-tropical cyclone activity dependent on the identification and tracking algorithm? *Meteorologische Zeitschrift*, **22**, 61–68.

Uppala, S. M., et al., 2005: The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131 (612)**, 2961–3012, doi:10.1256/qj.04.176.

Vaughan, G., et al., 2014: Cloud banding and winds in intense european cyclones: Results from the DIAMET project. *Bulletin of the American Meteorological Society*, **in press**, http://dx.doi.org/10.1175/BAMS–D–13–00 238.1.

Wallace, J. M. and P. V. Hobbs, 2006: *Atmospheric Science: An Introductory Survey*. 2d ed., Academic Press, London, UK.

Wang, X., V. Swail, and F. Zwiers, 2006: Climatology and changes of extra-tropical cyclone activity: comparison of ERA-40 with NCEP-NCAR reanalyses for 1958-2001. *Journal of Climate*, **19**, 3145–3166.

Wang, X., Y.Feng, G.P.Compo, V. Swail, F. Zwiers, R.J.Allan, and P.D.Sardeshmukh, 2013: Trends and low frequency variability of extra-tropical cyclone activity in the ensemble of twentieth century reanalysis. *Climate Dynamics*, **40**, 2775–2800.

Wanner, H., S. Bronnimann, C. Casty, D. Gyalistras, J. Luterbacher, C. Schmutz, D. B. Stephenson, and E. Xoplaki, 2001: North Atlantic Oscillation: Concepts and studies. *Survey Geophysics*, **22**, 321–381.

Wernli, H., S. Dirren, M. A. Liniger, and M. Zillig, 2002: Dynamical aspects of the winter storm 'lothar' (24-26 december 1999). *Quarterly Journal of the R*, **128**, 405–429.

Whitaker, J. and C. Davis, 1994: Cyclogenesis in a saturated environment. *Journal of the Atmospheric Sciences*, **51**, 889–907.

Wilbraham, R., 2013: The infulence of the stratosphere on mid-latitude cyclones: A diagnostic approach using the pressure tendency equation.

Williams, K. D., et al., 2013: The transpose-AMIP II experiment and its application to the understanding of Southern Ocean cloud biases in climate models. *Journal of Climate*, **26**, 3258–3274.

Willison, J., W. A. Robinson, and G. M. Lackmann, 2013: The importance of resolving mesoscale latent heat in the North Atlantic storm track. *Journal of the Atmospheric Sciences*, **70**, 2234–2250.

Woolings, T., 2011: Dynamical influences on European climate: an uncertain future. *In Press*.

Woolings, T., A. Hannachi, B. Hoskins, and A. Turner, 2010: A regime view of the North Atlantic Oscillation and its response to anthropogenic forcing. *Journal of Climate*, **23**, 1291–1307.

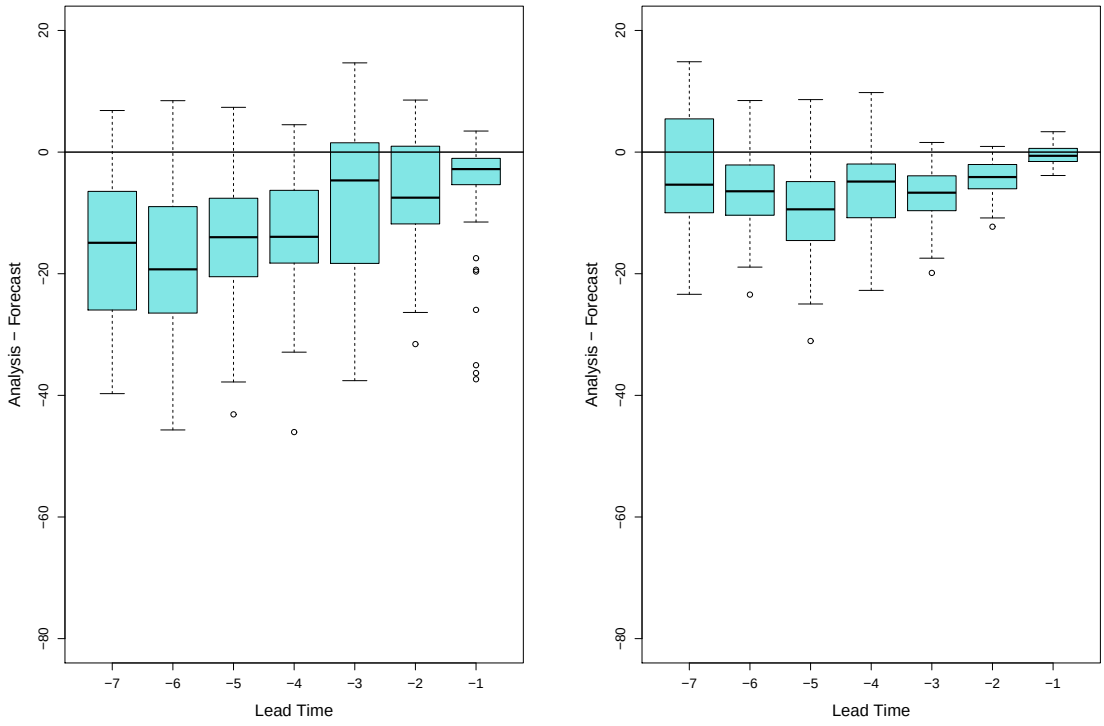
Yip, S., C. A. T. Ferro, D. B. Stephenson, and E. Hawkins, 2011: A simple, coherent framework for partitioning uncertainty in climate predictions. *Journal of Climate*, **24**, 4634?4643.

Zappa, G., L. C. Shaffrey, and K. I. Hodges, 2013: The ability of CMIP5 models to simluate North Atlantic extratropical cyclones. *Journal of Climate*, **26**, 5379?5396.

Appendix A: Additional Ensemble Forecast Plots

(a) Cross Late

(b) Cross Early



(c) Edge

(d) Split

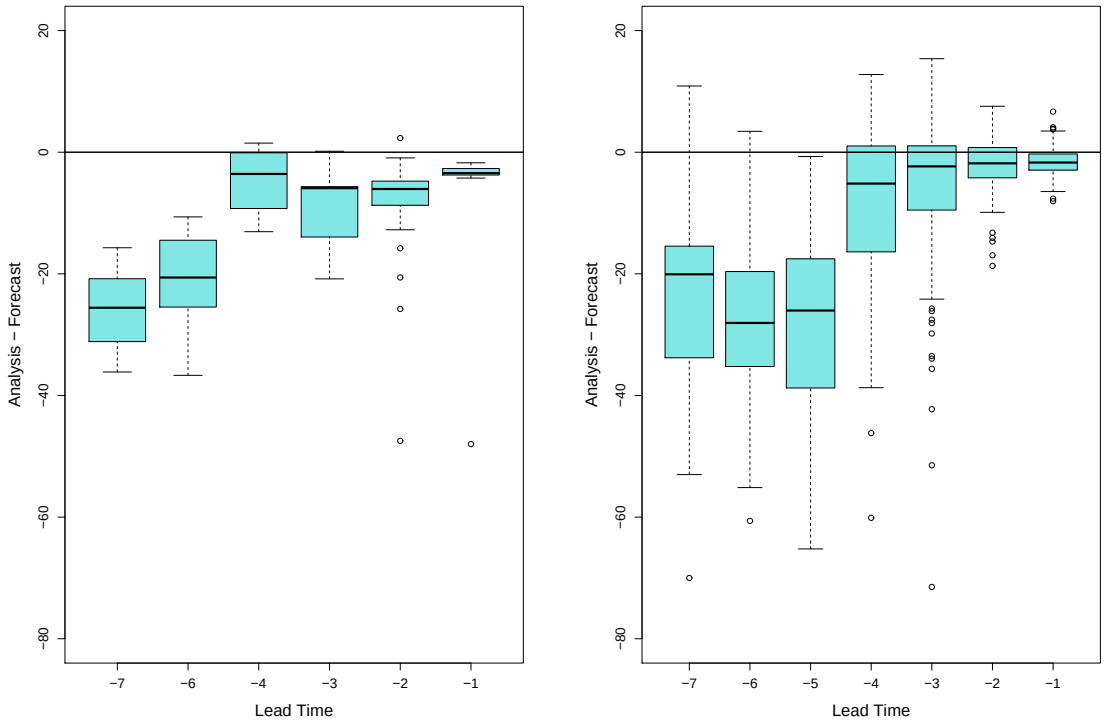
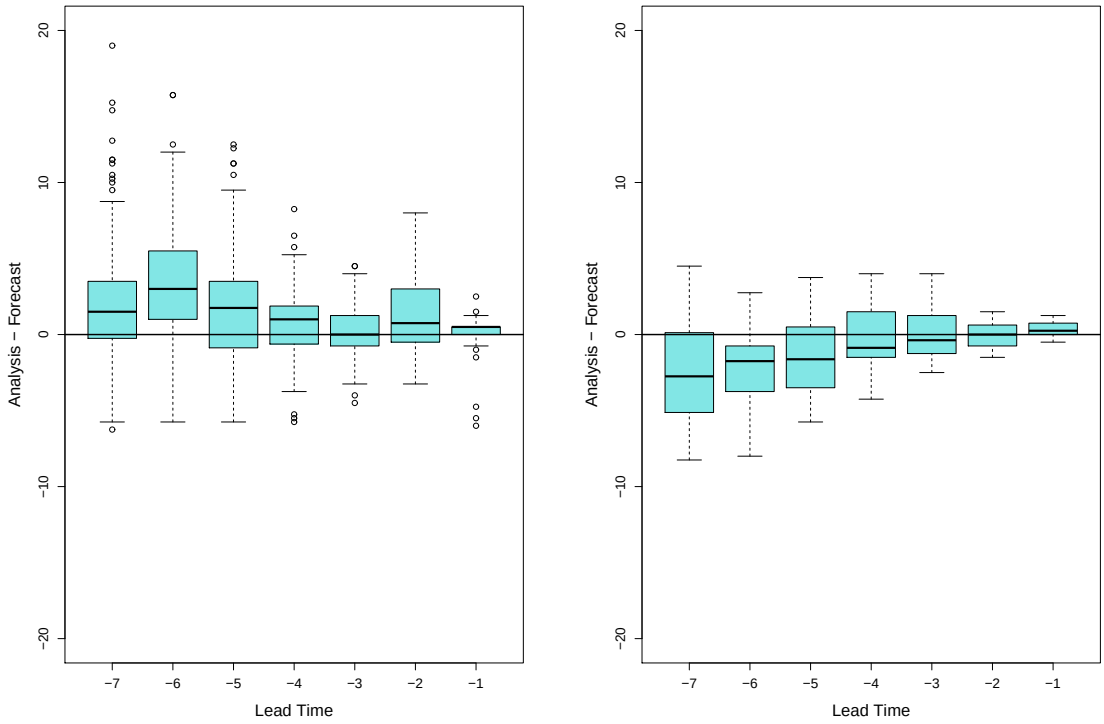


FIGURE 1: Differences between the ensemble forecasts and reanalysis pressure with lead times binned by day, the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split.

(a) Cross Late

(b) Cross Early



(c) Edge

(d) Split

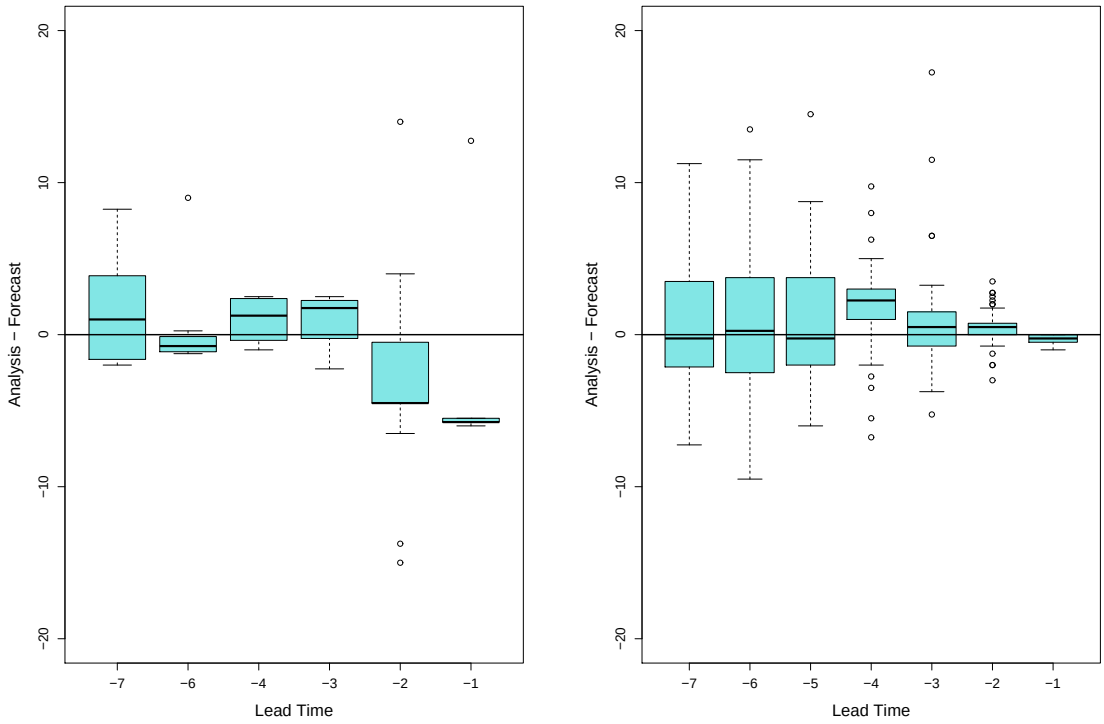
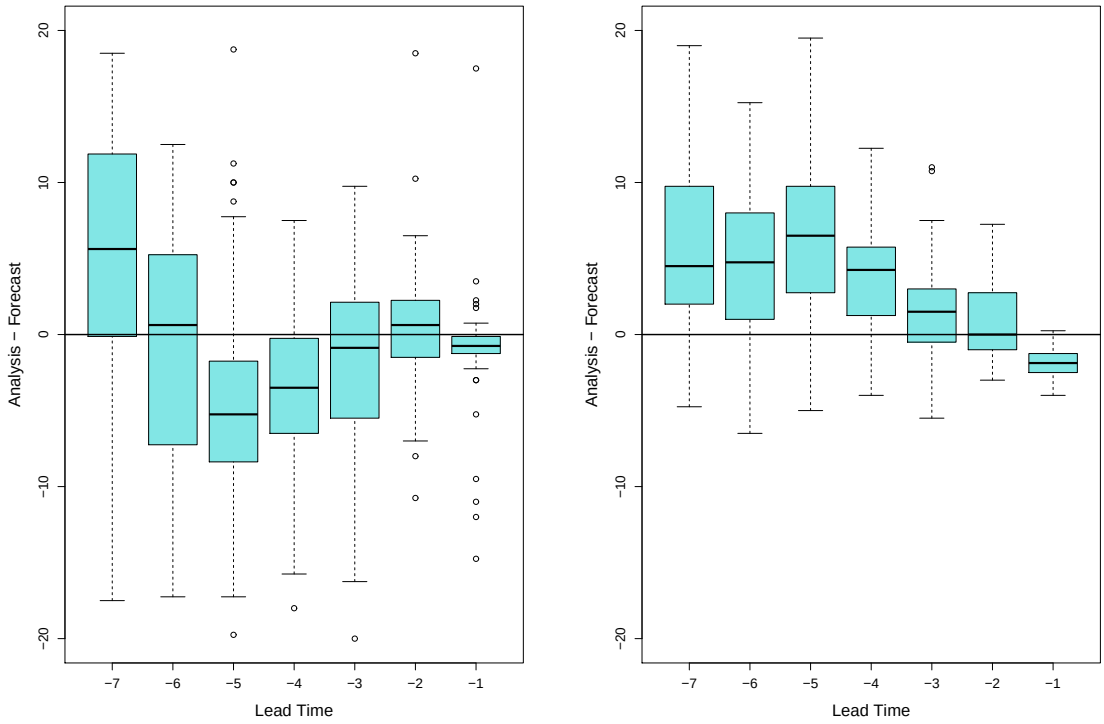


FIGURE 2: Differences between the ensemble forecasts and reanalysis latitude with lead times binned by day, the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split.

(a) Cross Late

(b) Cross Early



(c) Edge

(d) Split

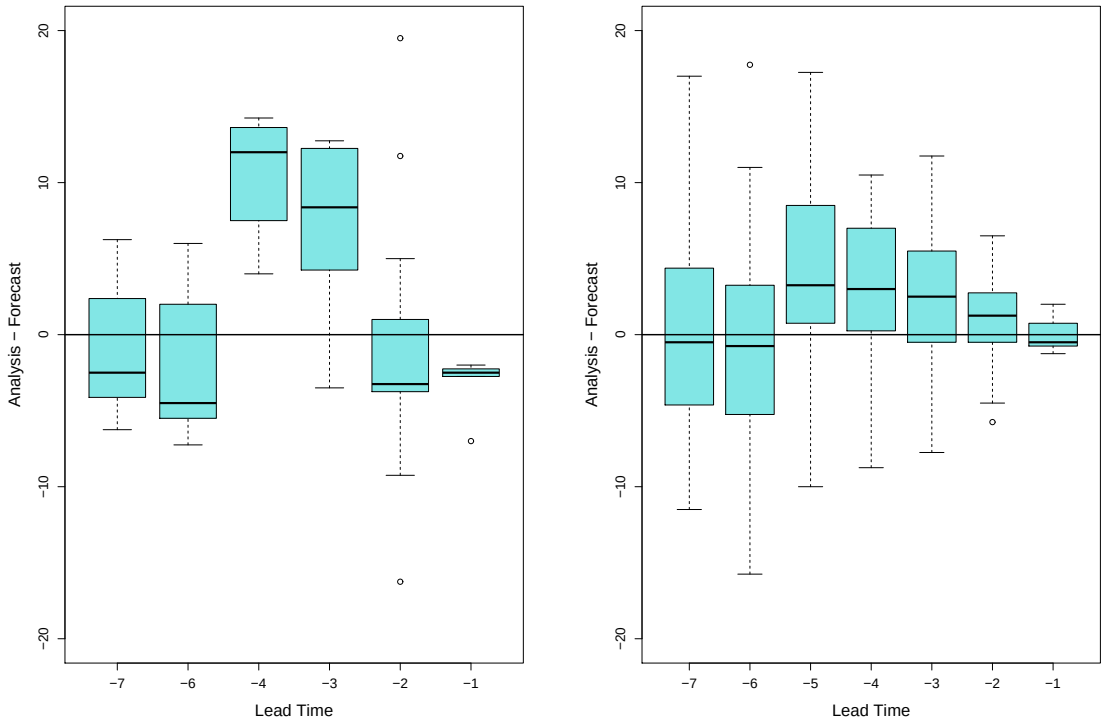


FIGURE 3: Differences between the ensemble forecasts and reanalysis longitude with lead times binned by day, the storms divided by jet stream type: (a) cross late, (b) cross early, (c) edge and (d) split.

(a) Horiz

(b) Diab

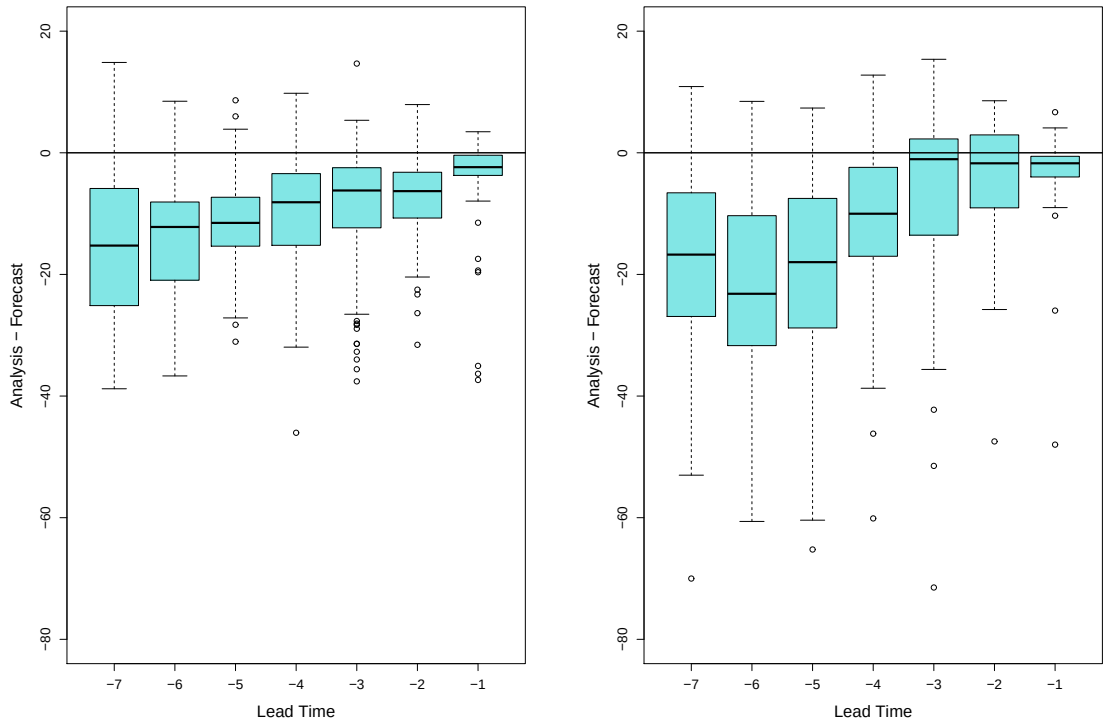


FIGURE 4: Differences between the ensemble forecasts and reanalysis pressure with lead times binned by day, the storms divided by the terms that dominate the PTE: (a) ‘horiz’, (b) ‘diab’.

(a) Horiz

(b) Diab

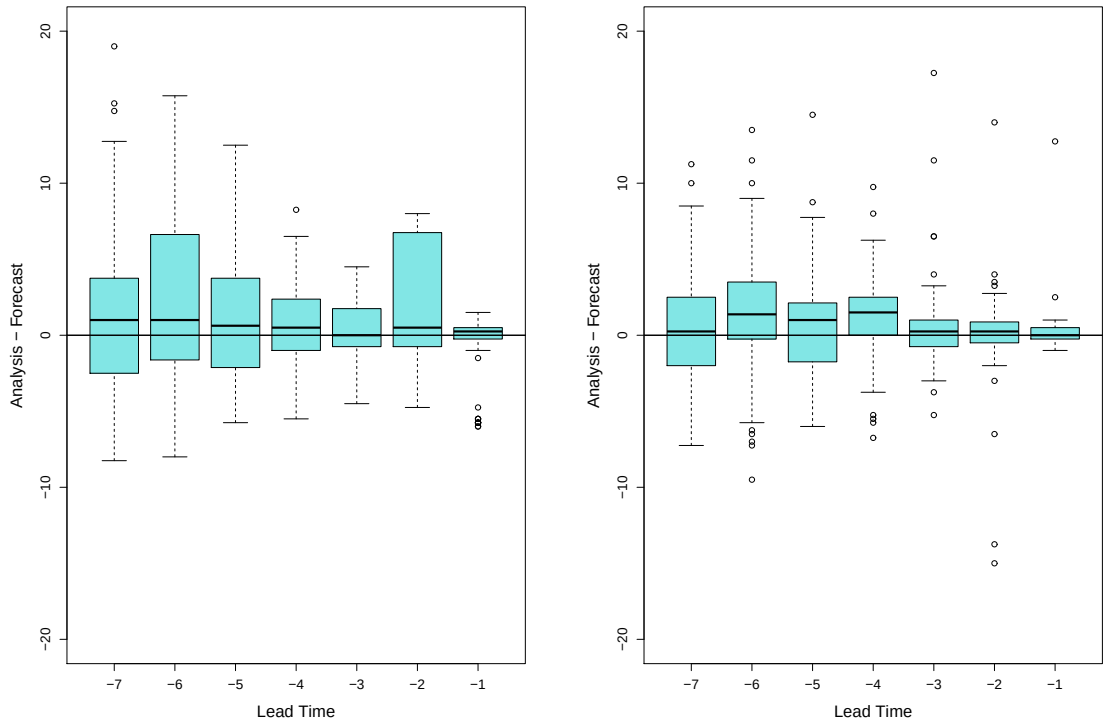


FIGURE 5: Differences between the ensemble forecasts and reanalysis latitude with lead times binned by day, the storms divided by the terms that dominate the PTE: (a) ‘horiz’, (b) ‘diab’.

(a) Horiz

(b) Diab

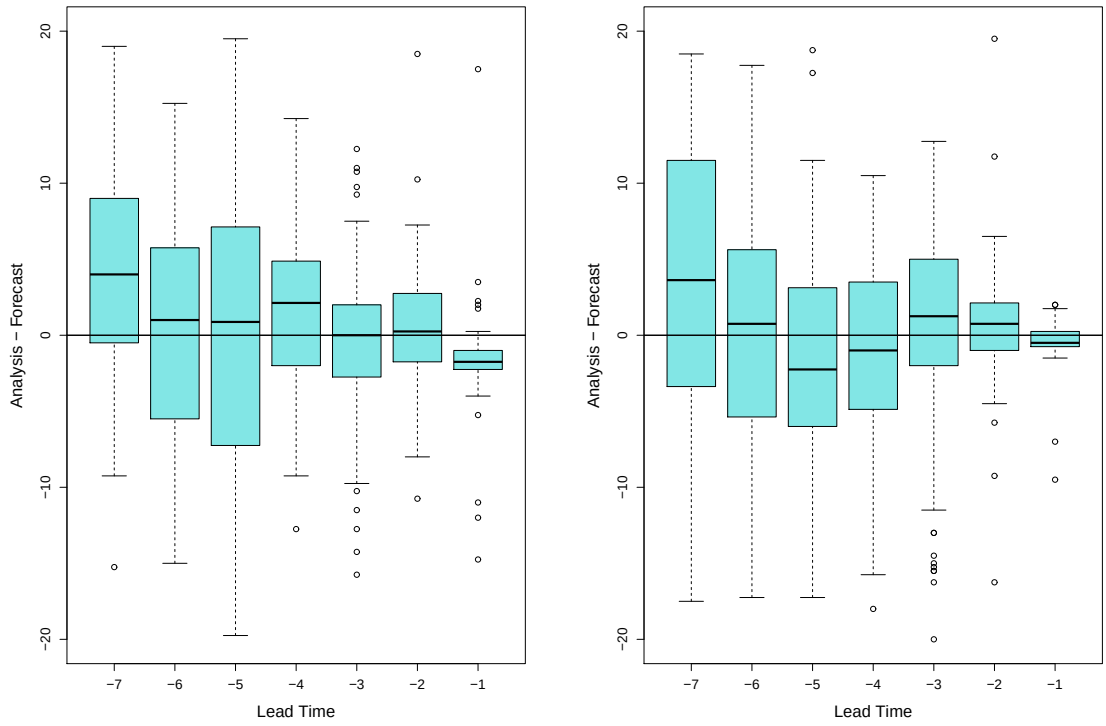


FIGURE 6: Differences between the ensemble forecasts and reanalysis longitude with lead times binned by day, the storms divided by the terms that dominate the PTE: (a) ‘horiz’, (b) ‘diab’.

(a) T_L255 (b) T_L399 (c) T_L639

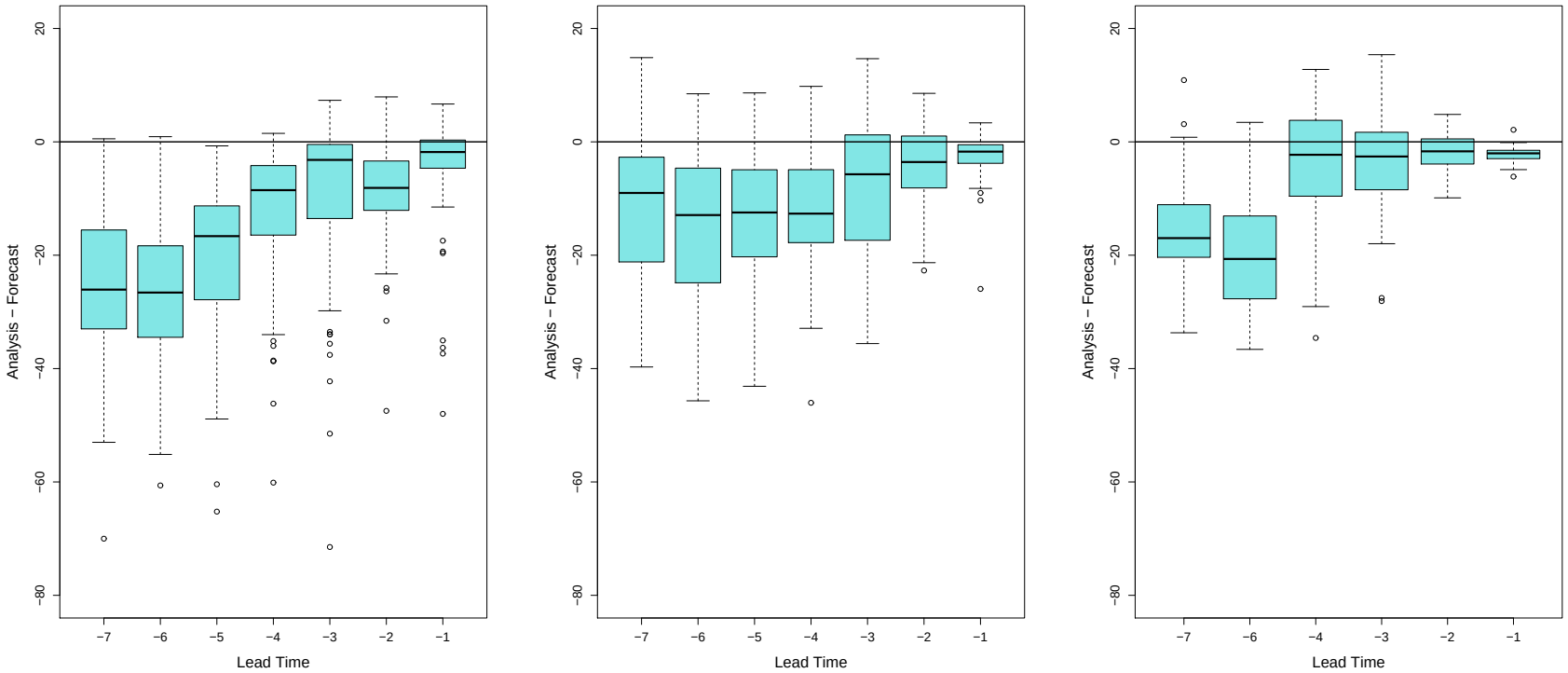


FIGURE 7: Differences between the ensemble forecasts and reanalysis pressure with lead times binned by day, and the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639.

(a) T_L255 (b) T_L399 (c) T_L639

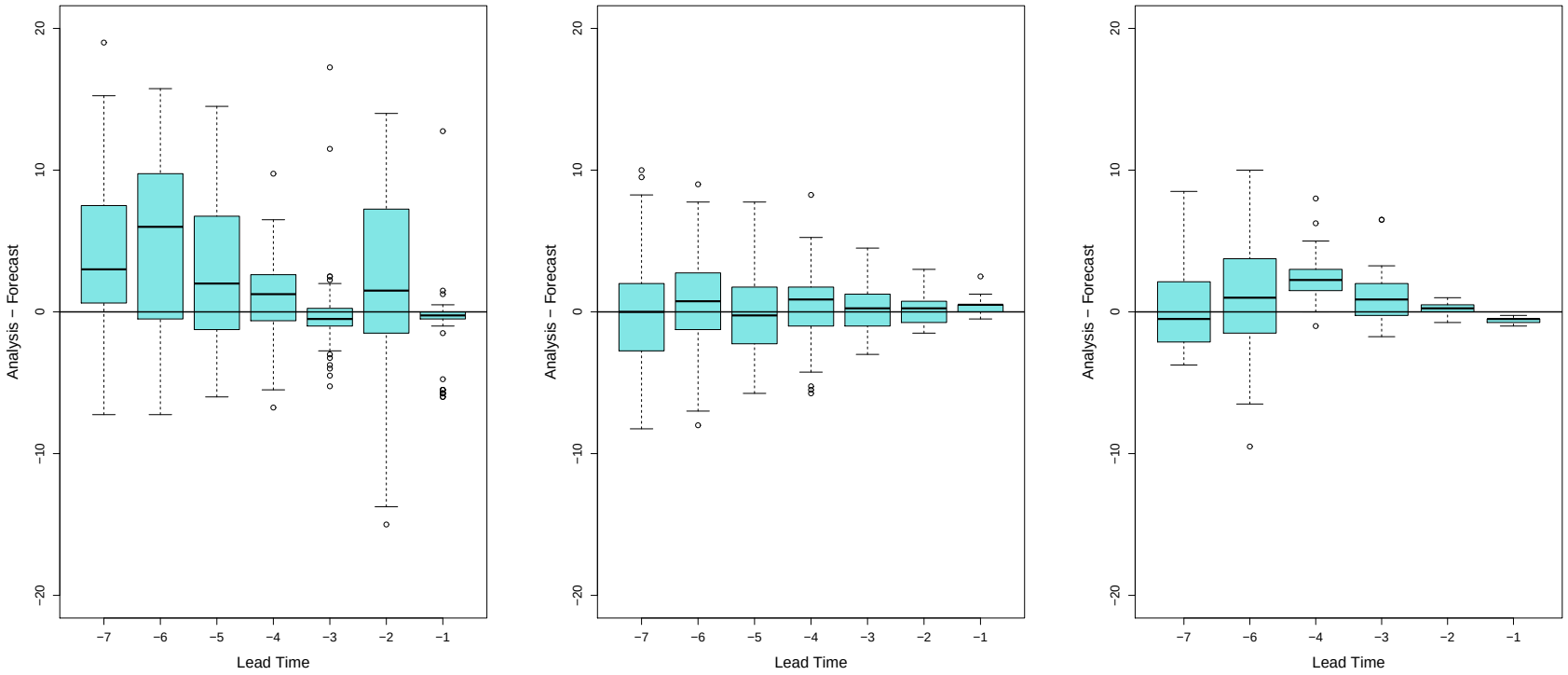


FIGURE 8: Differences between the ensemble forecasts and reanalysis latitude with lead times binned by day, and the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639.

(a) T_L255 (b) T_L399 (c) T_L639

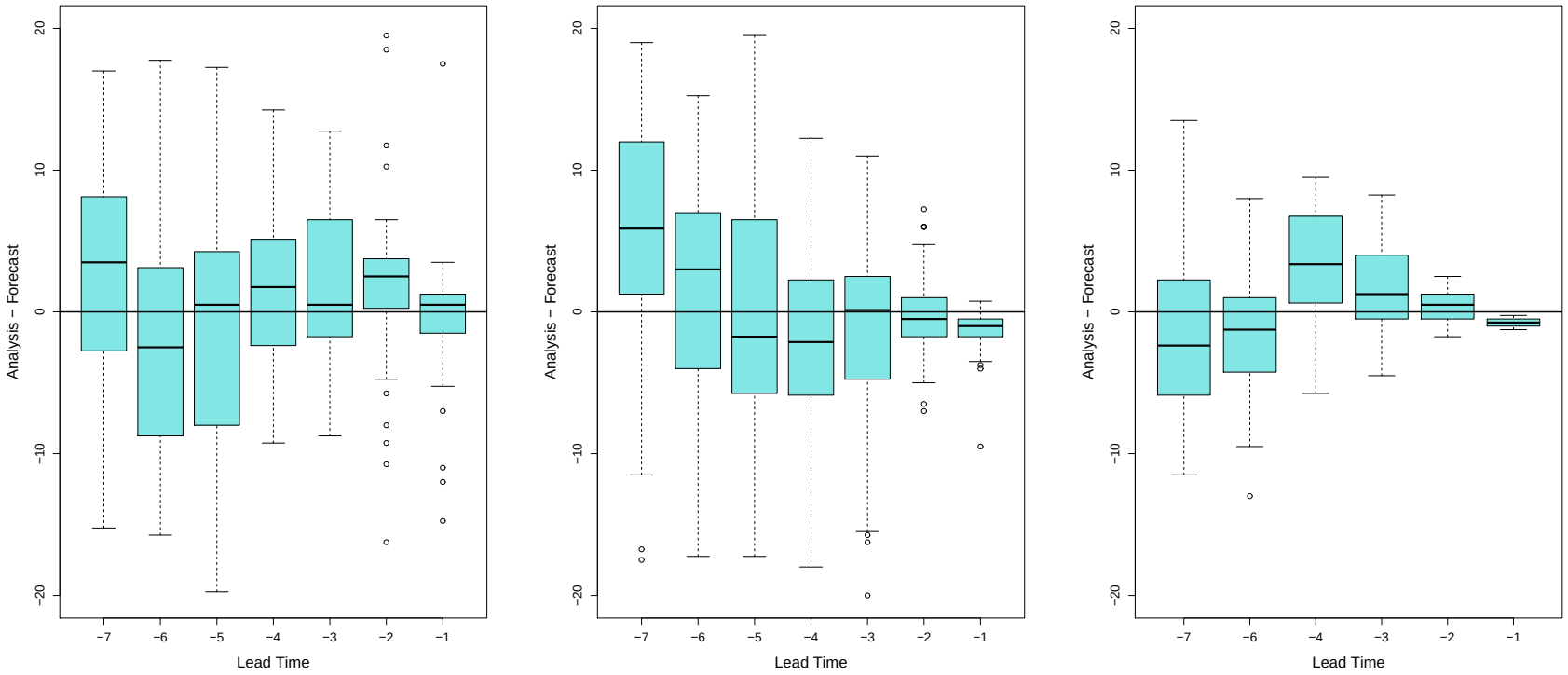


FIGURE 9: Differences between the ensemble forecasts and reanalysis longitude, with the storms divided by the resolution of the forecast: (a) T_L255, (b) T_L399, and (c) T_L639.

Appendix B: Tables of Variables

Γ_d	dry adiabatic lapse rate = $\frac{g}{c_p}$
Γ_m	moist adiabatic lapse rate, $\Gamma_m = g\frac{1+\frac{L_v r}{R_d T}}{c_p+\frac{L_v^2 r\epsilon}{R_d T}}$
ϵ	ratio of R_d to R_v, =0.622
η_g	total vorticity of the geostrophic wind
ζ	vorticity
θ	potential temperature, $\theta = T\frac{p_0}{p}^{\frac{R_d}{c_p}}$
θ_0	potential temperature at the surface
θ_e	equivalent potential temperature (Bolton, 1980)
	$\theta_e = T\frac{p_0}{p}^{0.2854(1.-0.00028r)} exp\left[r\left(\frac{3.376}{t_{lcl}}\right) - 0.00254\right)(1 + 0.00081r)\right]$
ν	number of degrees of freedom
θ_{es}	saturated equivalent potential temperature
ρ	density
σ_{qgd}	Dry, quasi-geostrophic growth rate
σ_{qgm}	Moist, quasi-geostrophic growth rate
σ_{sgd}	Dry, semi-geostrophic growth rate
σ_{sgm}	Moist, semi-geostrophic growth rate
Φ	geopotential

c_p	specific heat of dry air, at constant pressure = 1004.0 J/kg/K
d	the difference between rankings (used in calculation of R_S)
E	evaporation
e	partial pressure of water vapour in an air mass
e_s	partial pressure of water vapour in a saturated air mass $e_s = 6.112exp\left[\frac{17.67T_c}{T_c+243.5}\right]$
f	Coriolis Parameter
g	acceleration due to gravity
L_v	latent heat of vapourisation for water, 2.26MJ
N	Brunt–Väisälä frequency (Equation 6.2)
N_m	moist stability (Equation 6.4)
n	number of members of a population (e.g. data points used to calculate correlation)
P	precipitation
p	pressure (normally mean sea -level pressure)
p_0	reference pressure = 1000.0 hPa
$\dot{Q}$	diabatic heating rate
q	specific humidity
r	mixing ratio, $r = \frac{\epsilon e}{(p-e)}$
R	Pearson's correlation coefficient
R_S	Spearman's rank correlation coefficient
R_d	gas constant for dry air, 287.0 J/kg/K
R_v	gas constant for water vapour, 461.5 J/kg/K
RH	relative humidity = $\frac{e}{e_s}$
T	temperature in Kelvin
T_{LCL}	temperature at the lifting condensation level (Bolton, 1980)
	$T_{LCL} = \frac{2840}{3.5\ln T - \ln e - 4.805} + 55$
T_v	virtual temperature = $\left((1-q) + q\frac{R_v}{R_d}\right) T$ (Lynch and Cassano, 2006, Equation 3.7)
U	total wind speed
u	zonal wind speed
u_g	zonal geostrophic wind speed
v	meridional wind speed
v_g	meridional geostrophic wind speed
w	vertical wind speed
z	geopotential height