88 Wissenschaftliche Berichte des Instituts für Meteorologie und Klimaforschung
des Karlsruher Instituts für Technologie

LEA ANDREA EISENSTEIN

Objective identification and climatology of mesoscale high-wind features within extratropical cyclones

KIT Scientific Publishing

Lea Andrea Eisenstein

Objective identification and climatology of mesoscale high-wind features within extratropical cyclones

Wissenschaftliche Berichte des Instituts für Meteorologie und Klimaforschung des Karlsruher Instituts für Technologie (KIT)
Band 88

Herausgeber: Prof. Dr. C. Hoose
Prof. Dr. P. Knippertz
Prof. Dr. J. G. Pinto

Institut für Meteorologie und Klimaforschung
am Karlsruher Institut für Technologie (KIT)
Kaiserstr. 12, 76128 Karlsruhe

Eine Übersicht aller bisher in dieser Schriftenreihe erschienenen Bände finden Sie am Ende des Buches.

Objective identification and climatology of mesoscale high-wind features within extratropical cyclones

by
Lea Andrea Eisenstein

KIT Scientific Publishing

Karlsruher Institut für Technologie
Institut für Meteorologie und Klimaforschung

Objective identification and climatology of mesoscale high-wind features within extratropical cyclones

Zur Erlangung des akademischen Grades eines Doktors der
Naturwissenschaften von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT) genehmigte Dissertation

von Lea Andrea Eisenstein, M.Sc.

Tag der mündlichen Prüfung: 14. Juli 2023
Hauptreferent: Prof. Dr. Peter Knippertz
Korreferent: Prof. Dr. Joaquim G. Pinto

Impressum

KIT Scientific
Publishing

Karlsruher Institut für Technologie (KIT)
KIT Scientific Publishing
Straße am Forum 2
D-76131 Karlsruhe

www.ksp.kit.edu

Print on Demand 2024 – Gedruckt auf FSC-zertifiziertem Papier

ISSN 0179-5619
ISBN 978-3-7315-1374-2
DOI 10.5445/KSP/1000171924

Preface

The PhD candidate confirms that the research presented in this thesis contains significant scientific contributions by herself. This thesis reuses material from the following publications:

> Eisenstein, L., B. Schulz, G. A. Qadir, J. G. Pinto, and P. Knippertz, 2022a: Identification of high-wind features within extratropical cyclones using a probabilistic random forest – Part 1: Method and case studies. *Weather and Climate Dynamics*, **3**, 1157–1182, doi: 10.5194/wcd-3-1157-2022
>
> Eisenstein, L., B. Schulz, J. G. Pinto, and P. Knippertz, 2023a: Identification of high-wind features within extratropical cyclones using a probabilistic random forest – Part 2: Climatology. *Weather and Climate Dynamics Discussions*, 1–27, doi: 10.5194/wcd-2023-10, in review

The research leading to the results has been accomplished within the "Dynamical feature-based ensemble postprocessing of wind gusts within European winter storms" C5 project of the "Waves to Weather" Transregional Collaborative Research Center (grant no. SFB/TRR 165) funded by the German Science Foundation (DFG). The research proposal for the C5 project was written by Prof. Dr. Peter Knippertz and Dr. Sebastian Lerch.

The development of the objective identification tool described in Chapter 5 based on Eisenstein et al. (2022a) is split into three steps: (1) Subjective labelling, (2) probabilistic random forest and (3) Kriging. Step (1) was performed by the candidate while step (2) was mainly done by Benedikt Schulz and step (3) by Dr. Ghulam A. Qadir. The corresponding sections

5.2.2, 5.2.3 and Appendix A.1 were mainly written by the two colleagues, respectively, but are included in this dissertation for clarification of the method. Section 4.2.3 and the mathematical aspects of the evaluation in Section 5.4.1 were analysed by Benedikt Schulz, while the meteorological interpretation was done by the candidate. The remainder of the manuscript was written by the candidate with advice and corrections from Prof. Dr. Peter Knippertz and Prof. Dr. Joaquim Pinto. The climatology of Eisenstein et al. (2023a) was computed by the candidate, who also wrote the manuscript, with advice and corrections from Prof. Dr. Peter Knippertz, Prof. Dr. Joaquim Pinto and Benedikt Schulz.

The code of the two compared sting-jet identification methods by Gray et al. (2021) and Manning et al. (2022, 2023) in Chapter 7 were adapted to be used with model data from the German Weather Service by Kevin Gramlich in the scope of his Bachelor thesis (Gramlich, 2023) co-supervised by the candidate. The description of these adaptations in Section 4.2.6 is based on his thesis, while being refined by the candidate. The application on an illustrative case study in Chapter 7 is done solely by the candidate.

The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. This copy has been supplied on the understanding that this is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Abstract

Strong winds associated with extratropical cyclones are one of the most dangerous natural hazards in Europe. These high winds are mostly associated with five mesoscale dynamical features: the warm jet (WJ), the cold jet (CJ), cold frontal convection (CFC), strong cold-sector winds (CS) and, at least in some storms, the sting jet (SJ). The timing within the cyclone's life cycle, the location relative to the cyclone core and some further characteristics differ between these features and, hence, likely also the associated forecast errors.

Here, we present RAMEFI (RAndom-forest-based Mesoscale wind Feature Identification), a novel objective identification approach for these high-wind features using a probabilistic random forest (RF) based on each feature's most important characteristics in near-surface wind, rainfall, pressure and temperature evolution. A strength of RAMEFI is that it works flexibly and is independent of location-specific effects and horizontal gradients; thus, it can be applied to irregularly spaced surface observations and to gridded analyses and forecasts of different resolution in a consistent way. As a reference for the RF, we subjectively identify mesoscale wind features in 12 winter storm cases between 2015 and 2020 in hourly surface observations over Europe, using an interactive data analysis and visualisation tool. The RF is then trained on the subjectively set labels and learns physically consistent relations. From the RF, we obtain probabilities of each feature occurring at a given station, which can be interpolated to areal information with a Kriging approach. However, using only near-surface observations, the SJ and CJ are hard to distinguish and are therefore considered together in RAMEFI. The results show a reliable identification for all features, especially for the WJ and CFC, while the distinction of the CJ and CS can be difficult at times, as these features have rather similar meteorological characteristics. The new software RAMEFI is made publicly available for straightforward use by the atmospheric community.

Subsequently, RAMEFI is used to compile a climatology over 19 extended winter seasons (October–March, 2000–2019) based on high-resolution reanalyses of the German Consortium for Small-scale Modelling (COSMO) model over Western and Central Europe. This allows the first long-term objective statistical analysis of the features, their occurrence frequency, geographical distribution and characteristics in wind and humidity parameters. The CS is prominent in most winter storms, while CFC is the least common cause of high winds. However, CFC tends to be the cause of the highest gusts after the CJ, and has the highest gust factor, probably due to convective momentum transport. The WJ produces the weakest winds on average, but affects a larger area than the CJ. Central Europe is more strongly affected by WJ and CFC winds, while the CJ usually occurs farther north over the North and Baltic Seas and surrounding land areas. System-relative composites show that the WJ and CFC tend to occur earlier in the cyclone life cycle than the CJ and CS. Consistently, CS is the most common cause of high winds over Eastern Europe, where cyclones tend to occlude and decay. The WJ mostly occurs within the southeastern quadrant of a cyclone bordering with the narrow CFC in the west. However, the location of CFC relative to the cyclone centre varies greatly between cases. The CS occurs in the southwestern quadrant, while the CJ appears closer to the cyclone centre. This objective climatology largely confirms previous, more subjective investigations but puts these into climatological context and allows a more detailed analysis of feature properties.

Although surface characteristics of the CJ and SJ are similar such that the features are combined in RAMEFI, the origin differs. While the SJ is a descending air stream originating within the cloud head, the CJ keeps at low levels at all times. With the descent the SJ brings high momentum from mid-levels down to the top of the boundary layer or even surface, i.e., commonly creates higher winds and gusts and, thus, is desired to be detected separately. Here, two published SJ (potential) identification methods – one thermodynamic and one kinematic method – are tested on SJ and non-SJ cases using the German Icosahedral Nonhydrostatic (ICON) model to simulate the associated storms and compared to Lagrangian trajectories, the most established and most costly SJ identification method. Results suggest high error rates, such that a combination of parts of the two approaches and RAMEFI is developed into a new method. Based on a simple approach to detect a coherent three-dimensional region of high winds, the method shows first promising results and provides a basis for further refinement and increasing robustness, such that it can be used on different models and resolutions.

This work provides important new insides into the characteristics of mesoscale high-wind features within extratropical cyclones and lays the foundation for a wide range of applications, such as a feature-dependent forecast error analysis, impact research and climate change studies.

Kurzfassung

Starke Winde, die mit außertropischen Zyklonen einhergehen, sind eine der gefährlichsten Naturgefahren in Europa. Diese starken Winde sind meist mit fünf mesoskaligen dynamischen Bereiche verbunden: dem *Warm Jet* (WJ), dem *Cold Jet* (CJ), der kaltfrontalen Konvektion (CFC), starke Winde im Kaltsektor (CS) und, zumindest bei einigen Stürmen, der *Sting Jet* (SJ). Der Zeitpunkt des Auftretens im Lebenszyklus der Zyklone, die Lage relativ zum Tiefdruckkern und einige weitere Merkmale unterscheiden sich zwischen diesen Regionen und daher wahrscheinlich auch die damit verbundenen Vorhersagefehler.

Hier stellen wir RAMEFI (*RAndom-forest-based Mesoscale wind Feature Identification*) vor, einen neuartigen objektiven Identifizierungsansatz für diese Starkwindbereiche unter Verwendung eines probabilistischen *Random Forests* (RF), der auf den wichtigsten Merkmalen jedes Bereichs in der bodennahen Wind-, Niederschlags-, Druck- und Temperaturentwicklung basiert. Eine Stärke von RAMEFI ist, dass es flexibel und unabhängig von lokalen Effekten und horizontalen Gradienten funktioniert. Daher kann es auf unregelmäßig verteilte Bodenbeobachtungen und auf gegitterte Analysen und Vorhersagen mit unterschiedlicher Auflösung in konsistenter Weise angewendet werden. Als Referenz für den RF identifizieren wir subjektiv die Starkwindbereiche in 12 Winterstürmen zwischen 2015 und 2020 in stündlichen Bodenbeobachtungen über Europa mit Hilfe eines interaktiven Datenanalyse- und Visualisierungstools. Der RF wird dann auf dieser Klassifizierung trainiert, der physikalisch konsistente Beziehungen lernt. Von dem RF erhalten wir Wahrscheinlichkeiten für das Auftreten jedes Starkwindbereichs an den einzelnen Stationen, die mit Hilfe von Kriging in räumlich nahtlose Informationen interpoliert werden können. Da jedoch nur bodennahe Beobachtungen verwendet werden, werden der SJ und CJ in RAMEFI zusammen betrachtet. Die Ergebnisse zeigen eine zuverlässige Identifizierung für alle Bereiche, insbesondere für den WJ und CFC, während sich die Unterscheidung von CJ und CS manchmal schwierig gestalten kann, da die Merkmale ziemlich ähnliche meteorologische Eigenschaften haben. Die neue Software RAMEFI wird der Öffentlichkeit zur Verfügung gestellt, damit sie von der atmosphärischen Gemeinschaft unkompliziert genutzt werden kann.

Anschließend wird RAMEFI für eine Klimatologie über 19 Wintersaisons (Oktober-März, 2000-2019) auf der Grundlage vom deutschen COSMO-Model (*Consortioum for Small-scale Modelling*) über West- und Mitteleuropa verwendet. Dies ermöglicht die erste langfristige objektive statistische Analyse der Windbereiche, ihrer Häufigkeit des Auftretens, der geografischen Verteilung und der Eigenschaften in Wind- und Feuchteparametern. Der CS ist bei den meisten Winterstürmen vorherrschend, während CFC die am wenigsten häufige Ursache für Starkwinde ist. CFC ist jedoch tendenziell die Ursache für die stärksten Böen nach dem CJ und hat den höchsten Böenfaktor, wahrscheinlich aufgrund des konvektiven Impulstransports.

Der WJ erzeugt im Durchschnitt die schwächsten Winde, betrifft aber ein größeres Gebiet als der CJ. Mitteleuropa ist stärker von den WJ- und CFC-Winden betroffen, während der CJ weiter nördlich über der Nord- und Ostsee und ihrer ländlichen Umgebung auftritt. System-relative Komposita zeigen, dass WJ und CFC tendenziell früher im Lebenszyklus einer Zyklone auftreten als CJ und CS. Konsistent dazu ist CS die häufigste Ursache für Starkwinde über Osteuropa, wo Zyklonen dazu neigen, zu okkludieren und sich aufzulösen. Der WJ tritt meist im südöstlichen Quadranten einer Zyklone auf und grenzt im Westen an schmale CFC. Allerdings variiert die Lage von CFC relativ zum Tiefdruckkern von Fall zu Fall stark. Der CS tritt im südwestlichen Quadranten auf, während der CJ näher am Tiefdruckkern auftritt.

Diese objektive Klimatologie bestätigt weitgehend frühere, eher subjektive Untersuchungen, stellt diese aber in einen klimatologischen Kontext und ermöglicht eine detailliertere Analyse der Eigenschaften der Windbereiche.

Obwohl die bodennahen Eigenschaften von CJ und SJ ähnlich sind, sodass die Bereiche in RAMEFI kombiniert werden, sind die Ursprünge unterschiedlich. Während der SJ ein absinkender Luftstrom aus dem Wolkenkopf ist, bleibt der CJ in niedrigen Höhenniveaus. Mit dem Absinken transportiert der SJ hohen Impuls in die bodennahe Grenzschicht und verursacht damit oft höhere Winde und Böen und sollte daher separat identifiziert werden. Hier werden zwei veröffentlichte Methoden zur Identifizierung von SJ(-Potential) – eine thermodynamische und eine kinematische – an SJ- und Nicht-SJ-Fällen getestet, wobei das deutsche ICON-Modell (*ICOsahedral Nonhydrostatic*) zur Simulation der Stürme verwendet und mit Trajektorien verglichen wird, welche die etablierteste und rechenaufwendigste SJ-Detektionssmethode ist. Die Ergebnisse deuten auf hohe Fehlerquoten hin, sodass eine Kombination aus Teilen der zwei Ansätze und RAMEFI zu einer neuen Methode entwickelt wird. Auf der Grundlage eines einfachen Ansatzes zur Erkennung eines dreidimensionalen Starkwindgebiets zeigt die Methode erste vielversprechende Ergebnisse, und es wird eine Strategie zur weiteren Entwicklung und Erhöhung der Robustheit diskutiert, sodass sie bei verschiedenen Modellen und Auflösungen angewendet werden kann.

Diese Arbeit liefert neue Erkenntnisse über die Eigenschaften von mesoskaligen Starkwindgebieten in außertropischen Zyklonen und legt den Grundstein für ein breites Spektrum von Anwendungen, wie z.B. eine regime-abhängige Vorhersagefehleranalyse, Forschung der Auswirkungen und Studien zum Klimawandel.

Contents

1 Introduction

Storms come and go,
and with every passing storm
there is a new lesson to glean from.
– Pastor Jimmie Duran,
Peace of the Midst of the Storm

In the mid-latitudes, extratropical cyclones can produce some of the most severe natural hazards, especially during wintertime. These winter storms can cause high wind speeds, heavy precipitation, storm surges and, thus, considerable damage. Prominent examples for Central Europe are the storms *Lothar* (December 1999; Wernli et al., 2002), *Kyrill* (January 2007; Fink et al., 2009) and more recently storm *Eunice* (February 2022; Volonté et al., 2023a,b). From October to March, weather services regularly warn against the risks such wind storms can bring. Figure 1.1 shows an example of the German Weather Service (DWD) from the *WarnWetter* app[1] for storm *Eunice* with orange warnings (stage 2 – 65 $\mathrm{km\,h^{-1}}$ to 104 $\mathrm{km\,h^{-1}}$) for Southern Germany, red warnings for Northern and Central Germany (stage 3 – 105 $\mathrm{km\,h^{-1}}$ to over 120 $\mathrm{km\,h^{-1}}$) and dark red warnings along the North Sea coast (stage 4 – regionally over 140 $\mathrm{km\,h^{-1}}$). While these warnings are mainly concerning high wind speeds, heavy precipitation associated with the fronts was significant

---

[1] More information about the DWD *WarnWetter* app at `https://www.dwd.de/EN/ourservices/warnwetteapp/warnwetterapp.html` (last access 21 May 2023)

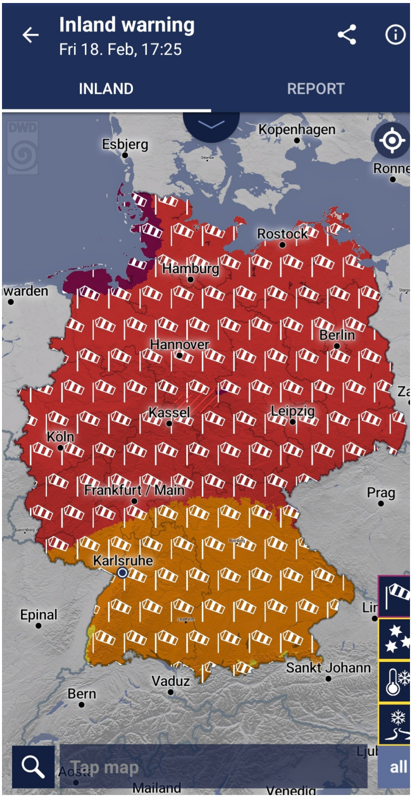


Figure 1.1: Official inland warning against storm *Eunice* by the DWD within the DWD Warn-Wetter app. Screenshot taken on 18 February 2022, 17:25 UTC+1. ©Deutscher Wetterdienst

as well. Early warnings enable the population to prepare for weather-related disasters and stay safe, as also stated in a recent press release by the World Meteorological service (WMO)[2].

---

[2] See official press release from the WMO published on 22 May 2023 (press release number: 22052023): `https://public.wmo.int/en/media/press-release/economic-costs-of-weather-related-disasters-soars-early-warnings-save-lives` (last access 23 May 2023).

A cyclone's impact is dependent on a number of factors, such as its size, storm track, intensity and occurring wind features. High winds are typically associated with four mesoscale features within the synoptic-scale cyclone: the warm conveyor belt jet or, in short, warm jet (WJ); the cold conveyor belt jet or, in short, cold jet (CJ); cold frontal convection (CFC); and the sting jet (SJ). Furthermore, high wind speeds are often detected within the cold sector (CS) without the formation of a distinct mesoscale feature. As the cold front itself, the cold sector is usually convectively active, leading to downward momentum transport in the vicinity of showers or even thunderstorms.

All features can cause damage due to strong gusts; thus, it is important to accurately forecast them and their associated wind fields and in doing so to increase the quality and lead time of warnings to prevent casualties and damages. A widely employed approach to improve forecasts (not only of wind) is statistical post-processing (Vannitsem et al., 2021), during which the model output is corrected on the basis of past forecast errors. The performance of various ensemble post-processing methods for wind gusts has been discussed in a recent paper by Schulz and Lerch (2022), who found that approaches ranging from classical statistical methods to novel neural-network-based techniques significantly improve forecast reliability and accuracy. However, Pantillon et al. (2018), who applied one of the classical statistical methods to ensemble forecasts of wind gusts and analysed the performance for several winter storms over Germany, found that post-processing can actually considerably worsen the forecast in some cases, as it did, for example, for storm *Christian* (October 2013), which is known to have developed an SJ (Browning et al., 2015). Similar results were also found for storm *Friederike* in January 2018 (not shown). This can generally come from a "misprediction" of the cyclone track or intensity but could also indicate that the characteristics of individual mesoscale high-wind features are not well represented. If that was the case, a feature-dependent post-processing could lead to further improvement, as it could consider the specific dynamical characteristics and how they are treated in the forecast model. By developing an objective identification algorithm for the wind features, this work lays the foundation for further exploring this idea.

As all of the previous approaches are purely subjective, relatively time-consuming and, thus, hard to automate or are missing one or more of the introduced features, this work aims to develop an objective analysis of the different mesoscale wind features that can flexibly be applied to station and gridded data and, therefore, serve as a basis for forecast evaluation and post-processing development, as well as climate change studies. The strategy that we follow is to start with a subjective identification (as in previous studies) but to use the results to then train a probabilistic random forest (RF) to develop an objective procedure that can be applied to cases outside of the training data set. The identification is designed to be independent of horizontal gradients (and hence resolution) and can principally be applied to observations from a single weather station. In addition, the identification is based on tendencies over 1 h only, making it applicable to time series with gaps. The output of the RF are feature probabilities rather than binary identification, which allow for an evaluation of how well individual data points fit the typical feature characteristics as well as the identification of hybrid features or transition zones. Our newly developed method is referred to as RAMEFI (RAndom-forest-based MEsoscale wind Feature Identification) and is made publicly available (Eisenstein et al., 2022c).

Once trained, the RF can also be used on different data sets and enable the first objective long-term climatology of the WJ, the CJ and SJ, CFC and CS winds in a consistent way. This climatology is compiled for 19 extended winter seasons to expand and complement existing shorter and/or more general climatologies and focuses on the occurrence frequency and characteristics over western and Central Europe.

As RAMEFI is used for surface observations only, the CJ and SJ are taken together for the method development and the compiled climatology. Nevertheless, the SJ is a feature with high wind risk that has been extremely underestimated in the past as mentioned above. A low-cost detection of the SJ comparable to the detection of the other features in RAMEFI is ambitious using only surface parameters. However, two previously published SJ (potential) identification methods – an instability-based and a kinematic approach – are tested in the scope of this word. Lastly, a strategy for an extension of RAMEFI to detect SJs is discussed. Given that the provision of such a feature-dependent post-processing tool can enhance the forecasts of strong winds and wind gusts, it can potentially contribute towards better weather warnings and impact forecasting of such events (e. g., Merz et al., 2020).

This thesis is structured as follows: First, literature about extratropical cyclones and their associated wind features are summarised in Chapter 2. Research questions tackled in the scope of this work are given in Chapter 3, and used data and methods are introduced in Chapter 4. The results are split into three chapters: The newly developed objective identification method RAMEFI is discussed in detail using example case studies in Chapter 5. In Chapter 6, RAMEFI is used to compile a climatology over western and Central Europe. Chapter 7 focuses on an objective SJ identification based on published methods and provides some preliminary results and a strategy for an extension of RAMEFI. Lastly, Chapter 8 concludes the findings of this work, and Chapter 9 gives an outlook.

2 Background theory

Extratropical cyclones belong to the most severe natural hazards in Europe, especially during wintertime. They can cause strong and destructive winds and heavy precipitation leading to considerable damage. Strong winds are commonly associated with low-level jets accompanying extratropical cyclones and convection along the front and in the cold sector.

This chapter discusses previous literature with the focus on extratropical cyclone evolution, including important processes such as contributing instabilities in the atmosphere, and the causes of high wind speeds associated with cyclones.

2.1 Evolution of extratropical cyclones

While extratropical cyclones can be classified by various measures, such as dynamical forcings, air flows and clouds (Catto, 2016), the focus here is on conceptual cyclone models. This sector focuses on the development and evolution of an imbalance to an extratropical cyclone and mesoscale processes contributing.

2.1.1 Baroclinic instability

The development of an extratropical cyclone is primarily associated with baroclinic instability. The Earth's atmosphere is characterised by a gradual meridional temperature gradient between the equator and the poles caused by radiative heating. In mid-latitudes, cold polar air masses and warm subtropical air masses collide along a frontal zone with a strong temperature gradient called polar front and driving the polar jet stream. The jet stream is a band of high wind speeds at around 200 hPa to 300 hPa circling the globe parallel to the isotherms. This front is enhanced during winter time, when less to no solar radiation reaches the polar region. Diabetic heating, e. g., also sensible heat fluxes, is further influenced by different responses to solar heating over land and ocean. The dependency of diabatic heating on latitude, the surface and air motions often leads to an occurrence of imbalance. As warm air is less dense than cold air, the temperature gradient creates a pressure gradient. This in turn leads to air moving from regions of high pressure to regions of low pressure, i. e., a wind flow. Baroclinicity describes the state of the atmosphere when the temperature gradient and pressure gradient are misaligned. Following the thermal wind balance, this leads to a vertical shear of the zonal wind and rotating of the air masses further affected by the Coriolis effect. In this unstable flow, a small disturbance can quickly intensify and create cyclones and anticyclones.

In the atmosphere, rotation is generally measured by the vorticity $\zeta = \nabla \times \vec{v}$, with the three-dimensional wind speed vector $\vec{v}$. In general, ζ considers deformation, stretching and shear. Absolute vorticity ζ_a further considers the Earth's rotation by summing ζ and the Coriolis frequency $f = 2\Omega \cos\phi$, where Ω is the rotation rate of the Earth and ϕ is latitude. Hoskins et al. (1985) introduced the so-called "potential vorticity thinking", with potential vorticity (PV) combining the absolute vorticity and static stability (see Section 2.1.3):

$$PV = -g \cdot (\zeta + f) \cdot \frac{\partial \theta}{\partial p}, \tag{2.1}$$

with the gravitational acceleration $g \approx 9.81\,\mathrm{m^2 s^{-2}}$, potential temperature θ and pressure p. It is usually measured in PVU, where $1\,\mathrm{PVU} = 10^{-6}\,\mathrm{m^2\,s^{-1}\,K\,kg^{-1}}$. PV is conserved in adiabatic flow and, hence, a local change in PV is associated with advection only. If PV is not conserved, changes can be caused by diabatic processes that influence the large-scale flow.

2.1.2 Conceptual cyclone models

The development of extratropical cyclones is often associated with two conceptual cyclone models. Over 100 years ago, Bjerknes (1919) proposed the first-ever cyclone model introducing a front as the moving boundary between two air masses, i. e., a warm and moist subtropical and a cold and dry arctic air mass, which is still used today and commonly called the Norwegian cyclone model. The cyclone model consists of four stages as shown in Figure 2.1a for the Northern Hemisphere:

I) A disturbance at low levels in the frontal region between colder northern and warmer southern air masses causes the formation of an initially small wave.

II) The disturbance spreads to higher levels and the angle between the fronts begins to sharpen, leading to the formation of a warm sector in-between. Furthermore, the cyclone begins to rotate counter-clockwise.

III) The warm sector reaches its maximum size, which is accompanied by a maximum conversion of potential to kinetic energy. However, the cold front moves faster than the warm front, causing the warm sector to narrow.

IV) The fronts start to occlude when the cold front reaches the warm front, causing the air from the warm sector to rise. At this stage the axis of the cyclone becomes vertical and the intensification slows down until the fronts are fully occluded before the cyclone begins to dissipate.

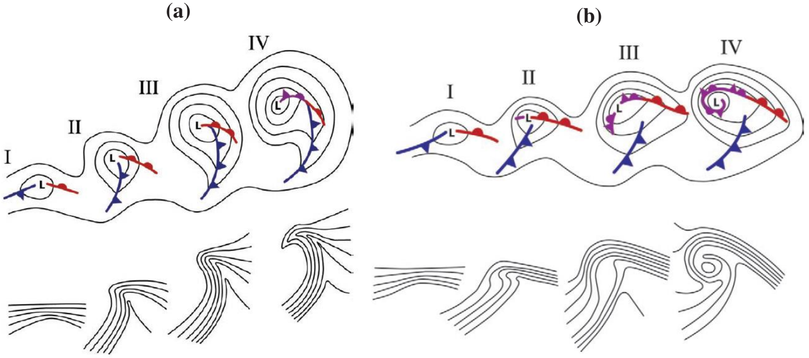


Figure 2.1: Conceptual models of extratropical cyclones: (a) Norwegian cylone, (b) Shapiro-Keyer cyclone. Top: lower-tropospheric (e. g., 850 hPa) geopotential height and fronts (blue – cold front, red – warm front, violet – occluded front). L marks the cyclone centre. Bottom: θ in the lower troposphere. The stages in the respective cyclone evolutions are separated by approximately 6—24 h . The characteristic scale of the cyclones based on the distance L, to the outermost geopotential height contour in stage IV is 1 000 km. Figure reprinted from (Schultz and Vaughan, 2011, their Fig. 2 and 12). © American Meteorological Society. Used with permission.

While the occlusion process was typically described as a "catch-up" in earlier works, Schultz and Vaughan (2011) suggest viewing it as a "wrap-up" of the frontal wave around the cyclone centre and a reduction of the warm sector in the process. This implies that the catch-up is not the cause of the occlusion but rather a result of the wrap-up.

However, not all extratropical cyclones follow the evolution of a classical Norwegian cyclone as shown in Shapiro and Keyser (1990), who simulated several idealised and real cases, ultimately leading to a second cyclone model, the so-called Shapiro-Keyser cyclone model. In analogy to the Norwegian model, it is described in four stages shown in Figure 2.1b.:

I) As in the case of a Norwegian cyclone, a wave is forming along the front due to a disturbance in low levels.

II) A frontal fracture between the warm and cold fronts occurs. The cold front is typically inactive, straight and weak compared to Norwegian cyclones (Schultz et al., 1998).

III) The warm front bends backwards around the centre, resulting in a T-bone structure. Therefore, the warm front is often referred to as "bent-back front".

IV) While the bent-back front wraps around the cyclone centre, and hence is displayed as an occluded front in Figure 2.1b after Schultz and Vaughan (2011), warm air is secluded as shown by the isotherms in the lower panel of Figure 2.1b.

In both conceptual models, the extratropical cyclone undergoes an intensification called cyclogenesis and a decay called cyclolysis. Although both develop an occluded front, the wrap-up in case of Shapiro-Keyser cyclones is not associated with a catch-up of the fronts. Research on when and why cyclones follow a certain model is still on-going. A possible factor is the importance of surface friction (Schultz and Zhang, 2007). Weaker surface friction over the ocean compared to land favours Shapiro-Keyser cyclones, which are therefore also referred to as marine cyclones (e. g., Hines and Mechoso, 1993; Rotunno et al., 1998). Idealised simulations further analysed across-jet shear (Davies et al., 1991; Thorncroft et al., 1993; Shapiro et al., 1999) showing that developments within regions without added cyclonic barotropic shear usually result in Shapiro-Keyser cyclones and developments within regions of added barotropic shear in Norwegian cyclones, where barotropic shear describes a change in wind not associated with temperature or density changes. Schultz et al. (1998) propose that Norwegian cyclones usually develop in diffluent upper-tropospheric background flows, while Shapiro-Keyser cyclones favour confluent background flows. This importance of along-jet shear results in developments of Shapiro-Keyser cyclones being more common in the western North Atlantic and Norwegian cyclones in the eastern North Atlantic and western Europe (Clark and Gray, 2018). More recent work by Coronel et al. (2016) shows that idealised cyclones initialised on the jet axis develop into Norwegian cyclones, while Shapiro-Keyser cyclones emerge from initialisations on the warm side of the jet.

Intense cyclones are often accompanied by dry intrusions (Browning, 1997), a dry descending air stream originating from the lower stratosphere (see Section 2.2.5). Moreover, they are associated with explosive cyclogenesis defined by Sanders and Gyakum (1980) as a pressure drop of at least

$$\frac{24\,\mathrm{hPa}\sin\phi\,/\,\sin 60^\circ}{24\,\mathrm{h}}. \tag{2.2}$$

For intense cyclones, a prominent hooked cloud head is often visible in satellite images. Pirret et al. (2017), who analysed factors contributing to the evolution, suggest that on average around 90 % of a cyclone's cyclogenesis stage is dominated by baroclinic and diabatic processes. Consistently, further studies (e. g., Grønås, 1995; Gray et al., 2011) suggest that latent heat release plays an important role, as for example in the case of the Great Storm of 1987 (Shutts, 1990). Intense and explosive cyclones are commonly associated with crossing the jet stream from southwest to northeast (Gilet et al., 2009; Pinto et al., 2009).

2.1.3 Mesoscale instabilities

While synoptic-scale processes are fundamental for the cyclone development, mesoscale processes can also play an important role in the intensification and dissipation of the system. Mesoscale instabilities are defined for a steady-state environment, which can generally be described as stable, neutral or unstable. If a parcel is displaced, it will return to its initial state in the case of stable conditions. However, in an unstable environment the parcel will accelerate away from its initial position and, in a neutral environment, it will remain in its new position.

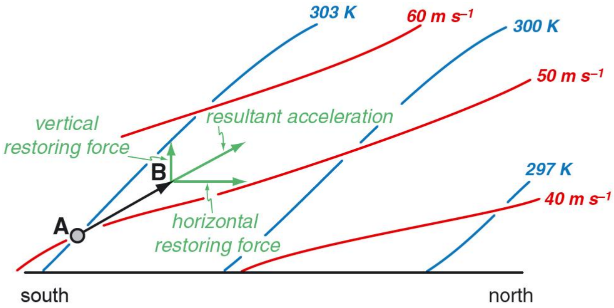


Figure 2.2: Schematic meridional cross-section of isentropic (blue) and geostrophic momentum surfaces (red) in a symmetrically unstable atmosphere. A tube of parcels that is displaced from position A toward position B experiences a resultant acceleration that is directed away from the original equilibrium position. Figure and caption reprinted from Markowski and Richardson (2010), their Fig. 3.9 © 2010 John Wiley & Sons, Ltd.

Inertial instability describes an air parcel being unstable with respect to horizontal displacements arising from wind shear and the Coriolis force. It is diagnosed by the vertical component of the absolute vorticity $\zeta_z < 0$. Analogously, static instability considers vertical motions and is diagnosed by the vertical gradient of the buoyancy $\partial_z b < 0$. For saturated conditions, instability with regard to vertical motions is referred to as conditional instability and identified using $\partial_z \theta_E^* < 0$ with θ_E^* being the saturated equivalent potential temperature. Unlike conditional instability, where the focus is on the displacement of an air parcel, convective – or potential – instability considers the vertical displacement of an atmospheric layer under saturation.

When a parcel is inertially and statically stable, a slantwise displacement can result in acceleration and is therefore symmetrically unstable. This is illustrated in Figure 2.2 with θ surfaces in blue and geostrophic momentum M_g surfaces in red. If a parcel is experiencing a slantwise displacement as shown from position A to B, the new environment around the parcel is both cooler and faster. Being inertially and statically stable, this results in two restoring forces – one upward and one northwards – leading to the slantwise acceleration. Symmetric instability is thereby diagnosed by $PV < 0$.

Accordingly, conditional symmetric instability (CSI) describes the state of the atmosphere when a slantwise unstable air parcel achieves saturation during lifting and is generally diagnosed by the saturated moist *PV* (MPV^*):

$$MPV^* = -g \cdot \zeta_\mathrm{a} \cdot \nabla_\mathrm{p} \theta_\mathrm{E}^* < 0 \tag{2.3}$$

Furthermore, to identify CSI, inertial stability ($\zeta_\mathrm{z} > 0$) and conditional stability ($\partial_\mathrm{p} \theta_\mathrm{E}^* < 0$) have to be present following previous literature (e. g., Parton et al., 2009; Gray et al., 2011; Baker et al., 2014).

Note that in extratropical cyclones, however, mesoscale instabilities should be treated with caution as a steady-state environment is not given and timescales of changes in the background and instability release cannot be properly separated. Nevertheless, an instability analysis can be helpful to give an indication of unstable reactions in the atmosphere, even though they might not be released as expected in a steady-state environment. To resolve mesoscale instabilities – and their release – in simulations, a high horizontal resolution of at least 10 km to 12 km as well as high vertical resolution is necessary Clark et al. (2005).

2.2 Characteristics of high-wind features

High winds within extratropical cyclones can have several origins. Here, the most common high-wind features, already mentioned in Chapter 1 are discussed in more detail. They are schematically depicted in Figure 2.3. These are in particular the WJ (Section 2.2.1), the CJ (Section 2.2.2), the SJ (Section 2.2.3), CFC (Section 2.2.4) and CS (Section 2.2.5). CS winds can include post-CFC, dry intrusions and further causes of high winds and are not attributed to a single mesoscale feature. Hewson and Neu (2015) gathered characteristic of WJ, CJ and SJ and summarised the occurrence relative to the cyclone centre and life cycle, footprint in Figure 2.4. Note that the authors included CFC in the WJ category, but did not involve CS in their analysis.

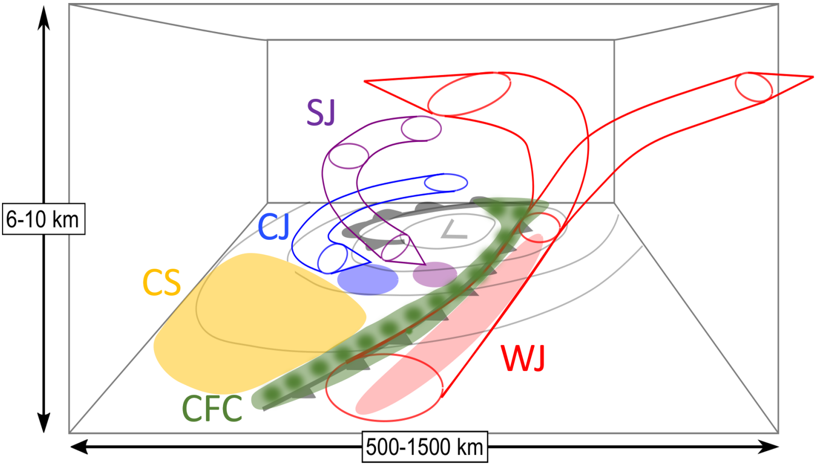


Figure 2.3: Conceptual model of the 3D structure of a Shapiro-Keyser cyclone showing the WJ (red), CJ (blue) and SJ (magenta). In each case, the region of strong surface winds is indicated by an ellipse. Figure as shown in Eisenstein et al. (2022a) adapted from Clark and Gray (2018) (their Figure 7) to include CFC (green) and CS (gold).

Characteristics of these dynamical features in extratropical cyclones are discussed in the following.

2.2.1 Warm conveyor belt and warm jets

An important feature in extratropical cyclones is the warm conveyor belt (WCB; Wernli and Davies, 1997; Eckhardt et al., 2004; Madonna et al., 2014), which is characterised by warm, moist air moving polewards ahead of the cold front. It starts near the surface in the equatorward portion of the warm sector and later ascends above the surface cold front. During the ascent of typically more than 600 hPa (Wernli and Davies, 1997), the WCB splits into a cyclonic and anticyclonic branch as seen by the red tubes in Figure 2.3. While the cyclonic part forms the cloud head and usually causes heavy precipitation along a narrow region, the anticyclonic part rises above the warm front and brings more moderate precipitation over a wider area.

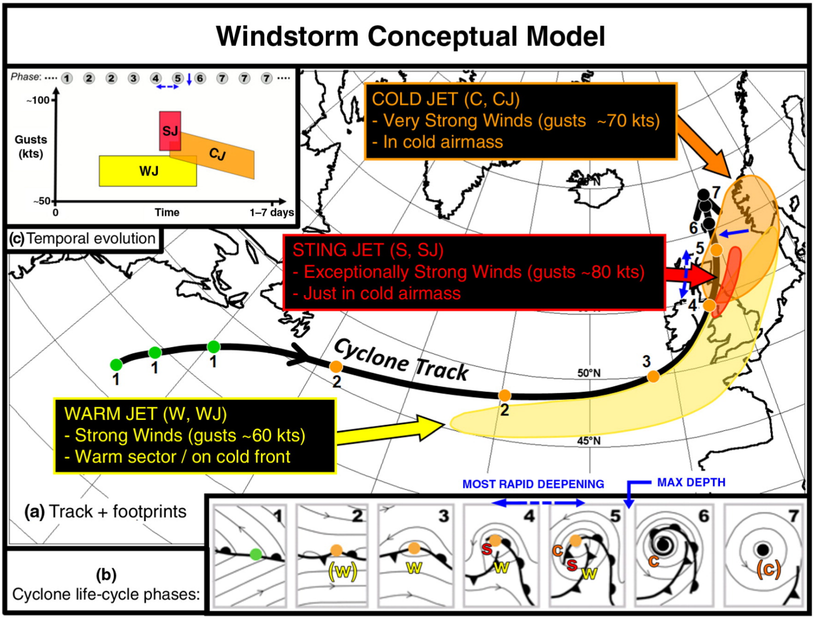


Figure 2.4: Conceptual model of an extratropical cyclone: (a) cyclone track (black) with WJ (yellow), CJ (orange) and SJ (red) footprints, (b) synoptic-scale evolution of fronts and isobars around the cyclone, (c) temporal evolution of gust strength for each jet. The dashed blue line denotes the period of most rapid deepening, whilst the solid blue arrow shows the time of maximum depth. Figure reprinted from Hewson and Neu (2015), their Fig. 1.

Overall, the WCB is the main cause for long-lasting precipitation (Catto, 2016) and it is visible through elongated cloud bands above the frontal zone in satellite images (Schemm and Wernli, 2014). Furthermore, the WCB can be the cause of strong convection along the cold front (Hewson and Neu, 2015). Here, we focus on the early stages of the WCB while it is still near the surface and can cause high winds there and refer to it as the WJ. Contrary to Hewson and Neu (2015), we define the WJ as the region ahead of the cold front and its convection, hence ahead of the CFC feature (Section 2.2.4), as displayed by the red shaded ellipse in Figure 2.3. Located in the warm sector of the cyclone between the two typically rain-active fronts, the WJ is usually characterised by positive temperature anomalies, decreasing pressure with time and little or no precipitation. The WJ is usually associate with the first strong winds starting in an early stage of a cyclone (Figure 2.4). Maximum gusts of around 25 $\mathrm{m\,s^{-1}}$ are typical (Hewson and Neu, 2015). Since the WCB is an ascending air stream, the winds at the surface weaken and disappear in later stages when the WCB no longer affects the boundary layer. The jet is long-lived with a duration of 24 h to 48 h and can cause a large surface wind footprint with a width of 200 km to 500 km and length of up to 1 000 km (Hewson and Neu, 2015). While the predictability was evaluated to be good with a relatively high coherence in space and time, Hewson and Neu (2015) found that the occurrence of very high winds within the warm sector are rather unusual, while Parton et al. (2010) associates 40 % of strong wind events with winds within the warm sector. Due to generally stable conditions in the warm sector, wind speeds above the boundary layer are usually much higher than surface gust speeds. Compared to the CJ and SJ, the WJ is the most long-lasting, but – as already mentioned – typically does not cause the most destructive winds (Figure 2.4).

2.2.2 Cold conveyor belt and cold jets

The CJ is associated with the main airflow of the cold conveyor belt (CCB) that turns cyclonically around the centre of the low (blue tube in Figure 2.3). At first, the CJ moves around the northwestern flank behind the occluded or bent-back front beneath the cloud head (CCBa in Earl et al., 2017; Gentile and Gray, 2023). Since it is travelling against the motion of the low-pressure system, the CJ is hard to see in Earth-relative winds until it wraps around the cyclone centre, except for extreme cases (Earl et al., 2017). Once wrapped around (CCBb in Earl et al., 2017; Gentile and Gray, 2023), the CJ can cause strong surface gusts near the tip of the front as shown in Figure 2.3 by the blue shaded ellipse. This usually happens around the time the maximum intensity is reached (Figure 2.4). The CJ weakens when the low decays or shortly before that. The jet mainly stays close to the ground, i. e., below 850 hPa (Smart and Browning, 2014), or is ascending slightly during its life cycle (Martínez-Alvarado et al., 2014). With typical maximum gusts around 30 m s^{-1}, the CJ is stronger than the WJ but with a typical lifetime of 12 h to 36 h does not last as long (Hewson and Neu, 2015). The impacted area expands with time, while the CCB, and hence the CJ, wraps around the cyclone centre when finally the footprint can reach a width of around 100 km to 800 km and a length of up to 2 500 km (Hewson and Neu, 2015). Although forming later in the life cycle of the parent cyclone than the SJ, both jets can coexist (Figure 2.4). A damaging CJ is more common than an SJ or WJ over Europe (Hewson and Neu, 2015).

2.2.3 Sting jets

A third potential airflow is the SJ, which can only occur in Shapiro-Keyser cyclones, where it descends from mid-levels within the cloud head into the frontal fracture region (magenta tube in Figure 2.3; Clark and Gray, 2018). The first time strong winds associated with the bent-back front of a Shapiro-Keyser cyclone but not with the CJ were mentioned, was in the work of Grønås (1995) analysing the *New Year's Storm* of 1992. In his paper, the author referred to the bent-back front as "the poisonous tail" in analogy to a scorpion. In their analysis of the *Great Storm* of 1987, Browning (2004) and Clark et al. (2005) found a descending air stream from mid-levels to be the cause of strong surface winds. Keeping the terminology by Grønås (1995), they referred to this phenomenon as "the sting at the end of the tail", i. e., a sting jet.

An SJ first appears in stage II of a Shapiro-Keyser cyclone (Figure 2.4; Section 2.1.2) at the tip of the bent-back front. It grows in size with time and moves eastwards with the wrap-up of the front. When the CJ wraps around the low, the SJ can be replaced by the CJ or merge with it. Hewson and Neu (2015) suggest an average surface footprint of less than 100 km in width and

up to 800 km in length. SJs usually last just a few hours but can be active up to 12 h in extreme cases and can cause wind gusts stronger than that of a CJ with typically $35\,\mathrm{m\,s^{-1}}$ although being less frequent (Figure 2.4).
A summary of consistent characteristics in both real cases and idealised simulations can be found in Baker et al. (2014) and the review of Clark and Gray (2018). The SJ originates within the cloud head and usually shows a drop in relative humidity RH when exiting it. However, this is not the case for storm *Egon* (January 2017; Eisenstein et al., 2020), possibly due to strong evaporative cooling. It descends on slanted surfaces of constant wet-bulb potential temperature θ_W while accelerating. If reaching the surface, it can cause strong near-surface wind gusts. Already in the first SJ paper (Browning, 2004), evaporative cooling and CSI were both suggested to be important contributions in the development of an SJ. Since then, various studies found agreeing results (e. g., Gray et al., 2011; Martínez-Alvarado et al., 2012; Volonté et al., 2018; Eisenstein et al., 2020; Volonté et al., 2020), while others deemed it not to be important (e. g., Smart and Browning, 2014; Slater et al., 2015; Coronel et al., 2016). Similarly, some studies consider evaporative cooling to contribute (e. g., Clark et al., 2005; Browning et al., 2015; Eisenstein et al., 2020) or to be negligible (e. g., Baker, 2009; Baker et al., 2014; Smart and Browning, 2014; Coronel et al., 2016). Schultz and Sienkiewicz (2013) discuss the importance of frontal dynamics, more specifically frontolysis at the tip of the bent-back front where the SJ occurs, and suggest mesoscale moist processes to be less relevant, allowing for SJ analysis in coarser resolution models. Slater et al. (2015, 2017) and Coronel et al. (2016) support this idea in their case studies. Volonté et al. (2018) analysed storm *Tini* (February 2014) in two model simulations with a coarser grid spacing of 24 km and a finer grid spacing of 12 km. While a descending air stream could be found in both simulations, the mechanisms and intensity differed. The lower resolution showed a weaker wind maximum and frontal dynamics to be the most important contributor.

In higher resolution simulations mesoscale instabilities can be resolved, such that trajectory analysis revealed the importance of CSI as a contributing factor to the intensity of an SJ.

While over a dozen SJ cases were analysed and simulated over the years, they mostly focused on the descent to the top of the boundary layer and neglected the momentum transport down to the surface. One study addressing this momentum transport within the boundary layer is done by Rivière et al. (2020) by running idealised simulations with varying horizontal resolution and surface roughness focusing on the downward transfer within the boundary layer. While winds at the surface increase with an increase in resolution and surface roughness, they hardly differ above the boundary layer. The 1 km simulation with typical land-surface roughness shows an SJ stronger than the CJ and downward transfer of high-momentum organised by along-wind moist convective rolls. The descending branch is correlated with the maximum surface gusts, while the ascending branch is located near the minimum of horizontal zonal wind speeds and peaks in precipitation at the top of the boundary layer. Passive tracers confirm that the convective rolls, which are initiated and strengthened by cooling caused by sublimation and melting of ice hydrometeors below the SJ, are accompanied by rapid downward transfer to a few hundred meters above the surface and that turbulent fluxes due to convective instability are responsible for the mixing within the last hundred meters. While the strong vertical mixing found in this study is consistent with results in observational studies (Parton et al., 2009; Browning et al., 2015), previous simulations had problems to resolve it (Parton et al., 2009; Baker et al., 2014; Coronel et al., 2016), suggesting that although a horizontal grid spacing of 10 km to 12 km and vertical grid spacing of 200 m to 300 m Clark and Gray (2018) might be sufficient for simulating an SJ above the boundary layer, it is not enough to resolve downward mixing.

The most common way to identify SJs is to compute Lagrangian trajectories (e. g., Volonté et al., 2018; Eisenstein et al., 2020, see Figure 2.6), for which a high resolution is required, horizontally and vertically, but also in time. Recent work has tried to identify SJs with low-cost approaches. For example, Gray et al. (2021) introduced an instability-based precursor tool, while Manning et al. (2022) developed a kinematic approach looking for reversals in the vertical gradient of horizontal wind speed along streamlines. These approaches are discussed further in Section 2.4.

While previous studies agree on the SJ being a descending air stream, the contributions of different mechanisms, e. g., mesoscale instabilities, frontal dynamics and turbulence, are still debated as they change from case to case and between different model configurations and resolutions. So far, capturing all proposed contributions and causes has turned out to be challenging, such that the development of SJs remains a topic of current research.

2.2.4 Cold-frontal convection

The passage of cold fronts is often accompanied by heavy precipitation, sometimes in form of convective lines, which, in turn, can cause strong wind gusts associated with the downward transport of high momentum from above the boundary layer. CFC is displayed in Figure 2.3 as an elongated line along the front with several darker spots representing individual convective events. Cold frontal rain, snow and graupel is responsible for around 28 % of extreme precipitation events in the mid-latitudes (Catto and Pfahl, 2013).

The frontal zone is characterised by a marked change in wind direction and a decrease in temperature (Clark, 2013). In extreme cases, tornadoes can occur in association with CFC, causing even more hazardous winds (e. g., *Kyrill*; Fink et al., 2009). Earl et al. (2017) suggest a separation of CFC into convective lines and pseudo-convective lines. The latter do not strictly satisfy identification criteria for convective lines but show characteristics of organised, strong convection and may fulfil the criteria earlier or later. A convective line shows a clear signal in radar imagery at 3 km to 10 km height (Parton et al., 2010). In case of a kata cold front, CFC occurs at and ahead of the surface cold front (Lackmann, 2011). Although high winds might then strictly speaking occur within the warm sector, we have decided not to associate it with the WJ due to the physically distinct characteristics. In case of an ana cold front, CFC occurs at and behind the cold front, and the distinction is therefore more straightforward. Shapiro-Keyser cyclones usually have a rather weak cold front, hence high winds are less likely to be associated with CFC (Schultz et al., 1998).

2.2.5 Cold-sector winds

The cold sector is the region behind the cold front (gold shading in Figure 2.3). In this area high winds can be caused by post-cold-frontal convection and also by dry intrusions. This region of a cyclone is generally known for its instability and turbulent behaviour. A dry intrusion is, as an SJ, a descending air stream, but in this case one that originates near the tropopause or even the lower stratosphere. This way it brings dry air down to the middle and lower troposphere moving towards the cyclone centre (Raveh-Rubin, 2017), where it leads to a dry slot behind the front later in the cyclone's life cycle (Thorncroft et al., 1993). Overrunning the CCB, it increases the potential instability leading to severe convection in the dry slot (Catto, 2016). Dry intrusions – and later in the cyclone's life cycle the dry slot near the cyclone centre – can often be seen in water vapour satellite imagery. Furthermore, it is visible in isentropic *PV* maps in the middle troposphere due to the increased *PV* values brought down from the stratosphere or upper-troposphere compared to usually weaker *PV* at these heights. As mentioned above, explosive cyclones are commonly accompanied by dry intrusions (Browning, 1997). Raveh-Rubin (2017) state that dry intrusions can cause destabilisation and increased wind gusts. Furthermore, cold fronts show a stronger temperature gradient, higher winds and more precipitation when accompanied by a dry intrusion (Catto and Raveh-Rubin, 2019; Raveh-Rubin and Catto, 2019).

Most publications consider all high-winds on the colder side of a cyclone to be a CJ or SJ and do not distinguish these features from other cold-sector winds (e. g., Manning et al., 2022; Gentile and Gray, 2023). Given the relatively large area of the cold sector in many cyclones and the appearance of discernible substructures, we decided to specifically separate out CFC, CJ and SJ, and label the remaining strong winds as CS.

2.3 Climatologies

Given that extratropical cyclones are such a dominant feature of the mid-latitudes, several objective algorithms have been developed for identifying cyclones and their tracks from digital data, either reanalysis or climate model data (see Ulbrich et al., 2009; Neu et al., 2013, for an overview). Depending on the perception on what a cyclone is, various variables can be used for tracking (Hoskins and Hodges, 2002), but the most common are mean sea level pressure p_{msl} and 850 hPa relative vorticity ζ. During winter, three local maxima of cyclone frequency are found over the Northern Hemisphere, namely over the North Atlantic, North Pacific and Mediterranean as displayed in Figure 2.5a (Ulbrich et al., 2009). While the two first features are identified by all methods, the maximum over the Mediterranean is dependent on the resolution of the used data set and methodology. Neu et al. (2013) compared 15 cyclone tracking methods and found significant differences in life cycle characteristics but a large consistency for long-living, intense cyclones. Concentrating on only the strongest systems (top 5 %), the track density shows a maximum from the East coast of Canada along Greenland to the Norwegian Sea (Figure 2.5b). Dacre et al. (2012) describe the development and compilation of an extratropical cyclone atlas using 200 extreme North Atlantic cyclones over a 20-year period. The atlas includes composites of horizontal and vertical cyclone structure, multiple parameters (e.g., cloud cover, wind and relative humidity) and cyclone evolution, while also identifying the warm and cold conveyor belts and dry intrusions.

A climatology focusing on near-surface winds can be found in Laurila et al. (2021). They focus on the North Atlantic and Europe by defining an extreme wind factor, which is the monthly 98th percentile divided by the monthly mean wind speed. While they found no linear trend between 1979 and 2018, they showed that the strongest winds are mostly connected with storm tracks in the winter season. This is consistent with the review paper of Feser et al. (2015), who concluded that decadal variability is the dominant feature of storminess over the region in the last 100–150 years, and only regional and short term trends can be identified. As expected from the surface and boundary layer characteristics, Laurila et al. (2021) also identified a distinct land-sea contrast in the 10 m wind speed. The same is true for wind gusts, given for example the very different gust factors typically found for offshore/inland areas (Wieringa, 1973).

To the best of our knowledge, the first climatology focusing on different mesoscale wind regions within cyclones is Parton et al. (2010), who differentiate cold frontal events, warm-sector events, tropopause folds/warm fronts, SJs and unclassified events within data from a wind-profiling radar in Wales

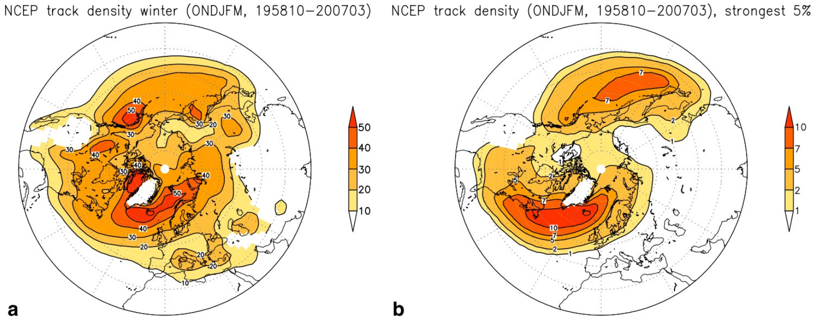


Figure 2.5: Track density for the extended winter seasons (October–March) 1958/59–2006/07 in the Northern Hemispherea using the NCEP-NCAR reanalysis. Grid points with an altitude over 1 500 m are removed: (a) all systems, adapted from Pinto et al. (2005), their Fig. 6a; (b) 5 % strongest systems with respect to the Laplacian of p_{msl}, Figure reprinted from Ulbrich et al. (2009), their Fig. 1a,b.

over a 7-year period. According to them, warm-sector events are the most common cause of strong winds with around 40 %, while cold-frontal events make up around 24 % over the investigated area, which may not be representative of cyclones in general. A study by Rivière et al. (2015) suggests that in early stages a cyclone is dominated by the WJ and later by the CJ. As mentioned above, Hewson and Neu (2015) analysed observations and reanalysis data of 29 wind storms with a focus on the three low-level jets, that is, WJ, CJ and SJ. They included CFC in their WJ analysis instead of treating it as an independent feature. They created an idealised conceptual model for the timing relative to the cyclone life cycle (Figure 2.4), the location relative to the cyclone centre and their strength, while also suggesting differences in further characteristics, such as instability/stability, vertical gradient of horizontal wind speed, etc. Furthermore, Earl et al. (2017) looked at the most common causes of high surface gusts in UK extratropical cyclones and, besides WJ, CJ and SJ, included several convection-induced high-wind features. They based the identification of SJs on satellite images, the location of the gusts within the cyclone and the deepening rate. Since no confirmation with Lagrangian trajectory analysis was done, the identified features are referred to as "potential SJs". Consistent with other studies, Earl et al. (2017) found that although WJs and CJs are the most common causes when looking at the highest 1 % of daily maximum wind gusts, CFC and potential SJs cause the highest 0.1 % of daily maximum gusts. However, their focus lays solely on winter storms over the UK similar to Parton et al. (2010).

Recently, some first objective approaches to identify mesoscale wind features have been developed by Manning et al. (2022) and Gentile and Gray (2023). Both studies use the strong thermal (and moisture in the latter) gradient to detect fronts and define high winds on the warm side as WJ. While Manning et al. (2022) identify SJs using a kinematic objective identification (see Section 2.4.2) and define all further high winds on the cold side of the fronts as CJ, Gentile and Gray (2023) distinguish between the CJ travelling against the system motion (named CCBa) and the CJ wrapping around the cyclone centre (CCBb) using the wind direction. Thus, the latter resembles our definition of a CJ merged with the CS. Although Manning et al. (2022) focus on future changes and Gentile and Gray (2023) on a 9-year climatology, both works conclude that winds in the cold sector to the west and south of the cyclone centre cause higher wind speeds than in the warm sector. Furthermore, Gentile and Gray (2023) analyse atmosphere-ocean-wave coupling based on ocean stations and find that the CCBb is the most common cause of high winds with an increasing proportion of CCBa to the north-east of the UK.

2.4 Objective identification approaches for sting jets

The first sign of an occurring SJ is often said to be a banded structure within the cloud head visible in satellite images and at times also radar data. This banded structure is an indication for slantwise circulation and release of CSI associated with an SJ. According to Browning (2004), each band is associated with a band of high winds at the surface.

As mentioned previously, the confirmation of SJs is typically done using Lagrangian trajectories (see Section 4.2.5), following the movement of an air parcel backwards for several hours. This approach allows for a proper analysis of the descent itself, but also of intensification and behaviour of parameters traced along the trajectories, such as a decrease in RH or constant θ_{W} and the contribution of mesoscale instabilities and evaporative cooling as discussed in Section 2.2.3. Furthermore, the trajectories can be compared to other low-level jets as shown exemplarily for storm *Egon* (January 2017) in Figure 2.6. However, for a proper trajectory analysis, at least hourly data is needed, and a 15 min resolution has proven even more beneficial in recent case studies by Volonté et al. (2018) and Eisenstein et al. (2020). This makes Lagrangian trajectories computational costly and time-consuming, such that

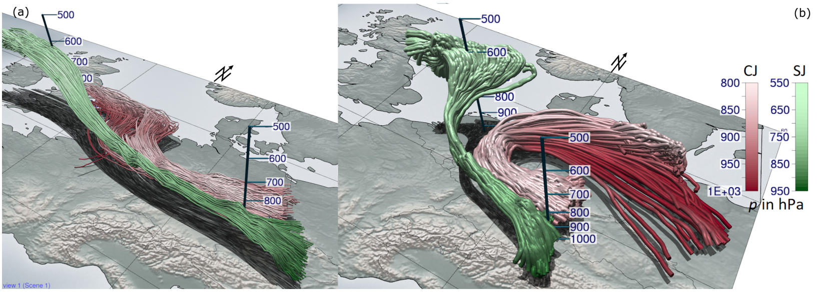


Figure 2.6: Three-dimensional visualisation of sting jet (SJ; green) and cold jet (CJ; red) trajectories started at 0200 UTC on 13 January 2017 in (a) an Earth- and (b) a system-relative framework in the REF simulation. Shading indicates altitude in hPa. Imagery produced using Met.3D. Both SJ and CJ trajectories must exceed a wind speed of $37\,\mathrm{m\,s^{-1}}$ at starting time. The pressure of SJ trajectories must decrease by at least 150 hPa in 8 h, while it has to remain below 800 hPa for CJ trajectories at all times. Figure and caption reprinted from Eisenstein et al. (2020), their Fig. 7a,b.

an operational use is challenging. However, since SJs are associated with hazardous winds, a proper forecast is essential to estimate the wind risk and avoiding casualties, which has motivated other approaches to identify SJs.

In recent years, two approaches appeared to be promising: an instability-based precursor tool and a kinematic-based method. While for a detailed analysis of SJ cases, trajectories are still necessary, these methods can help to detect SJs in longer time period data sets or in operational use, making them important for forecasters and climatologies.

2.4.1 Instability-based precursor tool

One possibility to detect SJs is to observe the occurring instabilities, more precisely CSI within the cloud head (Section 2.1.3). Martínez-Alvarado et al. (2013) developed a diagnostic for sting-jet precursors in low-resolution data sets. Since mesoscale instabilities need a resolution of at least 10 km to 12 km (Clark et al., 2005), they used the downdraught slantwise convective available potential energy (*DSCAPE*) as an indication for CSI. This is the maximum kinetic energy that is available for a slantwise descending air parcel, such as is the case for an SJ, if CSI is released. They define the criterion for *DSCAPE* as:

$$DSCAPE = \int_{p_{\text{top}}}^{p_{\text{bottom}}} R_{\text{d}} \left(T_{\text{v,e}} - T_{\text{v,p}}\right) d\ln(p) \geq 200\,\text{J}\,\text{kg}^{-1}, \qquad (2.4)$$

where p_{top} and p_{bottom} are the top and bottom pressure levels of a hypothetical downdraught, R_{d} is the dry air gas constant and $T_{\text{v,p}}$ and $T_{\text{v,e}}$ are the parcel and environmental virtual temperatures, respectively. High values imply slantwise instability. High *DSCAPE* can occur within the dry slot or in the WCB region where CSI cannot be released due to missing moisture, such that further criteria are needed to ensure the location of CSI within the cloud head close to the cyclone centre.

This method was refined in Martínez-Alvarado et al. (2012), who analysed 100 intense extratropical cyclones in ERA-Interim data by adding the following thresholds:

1. $RH_{\text{ice}} \geq 80\,\%$ to ensure enough moisture for release of CSI and thus a likely location within the cloud head
2. $|\nabla \theta_{\text{W}}| \geq 10^{-5}\,\text{K}\,\text{m}^{-1}$ to ensure a certain proximity to a front
3. $(\vec{v}_{\text{h}} \cdot \nabla \theta_{\text{W}}) \geq 10^{-4}\,\text{K}\,\text{s}^{-1}$, with the horizontal wind velocity $\vec{v}_{\text{h}}$ indicating the temperature advection to ensure this front to be a cold front. Positive values – as used here – indicate cold air flowing towards warm air.

Furthermore, CSI points are only retained within 100° to 300° relative to the cyclone's motion (where 0° is the direction of motion) to exclude signals in the warm sector and surroundings. An evaluation using trajectory analysis of 15 randomly chosen storms shows high consistency. Overall, the authors found that 23 % to 32 % of the storms presumably developed an SJ depending on the chosen thresholds. The precursor tool was used again in a study by Martínez-Alvarado et al. (2018) for climate simulations, showing an increase of potential SJs by 60 % and a shift towards Europe, whereas the proportion of explosive cyclones does not show a significant trend.

The successful application leads to a – now operational – tool used by the UK Met Office to give forecasters an indication of SJ potential and enables possible adjustments of warnings and risk management, which was introduced in Gray et al. (2021) based on Martínez-Alvarado et al. (2012). The original code was translated from MATLAB to Python and some adjustments regarding input data and algorithm speed-up were made. The precursor tool is used for the global ensemble forecast model MOGREPS-G (Met Office Global and Regional Ensemble Prediction System – Global; Bowler et al., 2008) with a native horizontal grid spacing of 20 km regridded to 0.5°. However, the tool is not used as a hard classification as in previous studies but rather as a qualitative guide for forecasters. By applying it to every ensemble member, an SJ potential is obtained by assessing the detected CSI points in each member as seen in Figure 2.7 for the example of storm *Brandon* (January 2020).

Cyclones are tracked using the Hodges (1994) tracking algorithm. Within 1 000 km of the cyclone centre, a grid point is considered a CSI point if the above defined criteria are satisfied. The centroid of an area of adjacent CSI points has to be located within a radius of 700 km around the cyclone centre within the cloud head region, where moisture is available for instability release.

Figure 2.7 shows two panels available for forecasters at the UK Met Office. The cyclone track and isobars based on the control ensemble member are displayed by the red line and black contours, respectively. The plus symbols indicate the cyclone centre of each member. While Figure 2.7a shows an ensemble-based probability of at least 20 CSI points, Figure 2.7b displays the centroids of precursor regions for each individual member. The size is proportional to the number of CSI points.

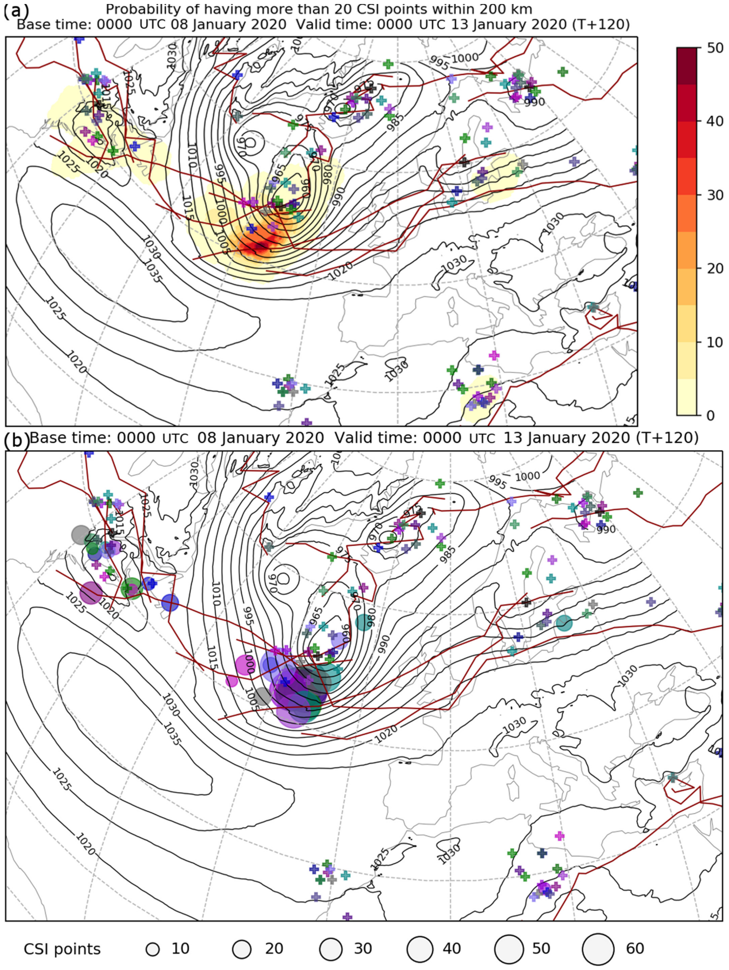


Figure 2.7: Example of graphical SJ precursor output for storm *Brandon* (5-day forecast valid at 13 January 2020, 00 UTC). (a) Ensemble-based probability for SJ precursor for constitute regions of more than 20 CSI points within a 1 000 km radius around a cyclone centre and the corresponding centroids within 700 km of a centre. (b) Centroid locations of CSI points (coloured circles; size corresponding to the number of connecting points). The red lines represent the cyclone tracks of identified storms and black contours show p_{msl} of the control ensemble member. Cyclone centres are indicated by the plus symbols (colour-coded by ensemble members matching the circles in (b)). The plus symbols indicate the locations of the cyclone centres. Figure reprinted from Gray et al. (2021), their Fig. 2.

2.4.2 Kinematic approach

While the precursor tool has been developed for coarse resolution forecast data, Manning et al. (2022) use a kinematic method to detect SJs in low and high resolution climate data. The method first examines if a cyclone is a Shapiro-Keyser cyclone by detecting the warm seclusion in the vicinity of the cyclone centre developing in stage IV (Section 2.1.2). The SJ detection follows, only if a Shapiro-Keyser cyclone development is confirmed.

To detect a warm seclusion, cyclones are first tracked using the algorithm by Hoskins and Hodges (2002) and the area of 1 000 km north, east, south and west of the centre is extracted. Each latitudinal section through this area is examined for a peak of θ_W. A threshold $\theta_\mathrm{W,thresh}$ is set as the value of θ_W of the first grid point of a consecutive number of grid points showing an increase. The peak has to exceed $\theta_\mathrm{W,thresh} + 2\,\mathrm{K}$. This step is repeated in longitudinal direction. Coinciding peaks in both directions are considered potential warm seclusion points. Furthermore, a p_msl trough is detected analogously, where the core has to be lower than $p_\mathrm{msl,thresh} - 6\,\mathrm{hPa}$. If the number of potential warm seclusion points identified within the p_msl core exceeds 500 (in case of the 2.2 km resolution of their used simulations), a warm seclusion is detected. Note that the values of 2 K and 6 hPa are chosen subjectively. For further discussion see Manning et al. (2022).

Figure 2.8 illustrates how an SJ is extracted by the example of storm *Anatol* (December 1999). In Figure 2.8a the cloud head is indicated by RH_ice at 500 hPa exceeding 80 % (black contour) and blue contours depict the frontal boundaries and warm seclusion. A wind speed maximum at the tip of the cloud head is evident (shading). The basis of this method is the descent indicated by a reversal of the vertical gradient of horizontal winds $\Delta_\mathrm{z} v$ between 700 hPa and 850 hPa ($v_{700} - v_{850}$) along streamlines. The streamlines are computed at 850 hPa originating at grid points where the wind speed exceeds the 98th percentile of the winter season (December – February; e. g., gold dot in Figure 2.8b). SJ streamlines have to show a consecutive number of

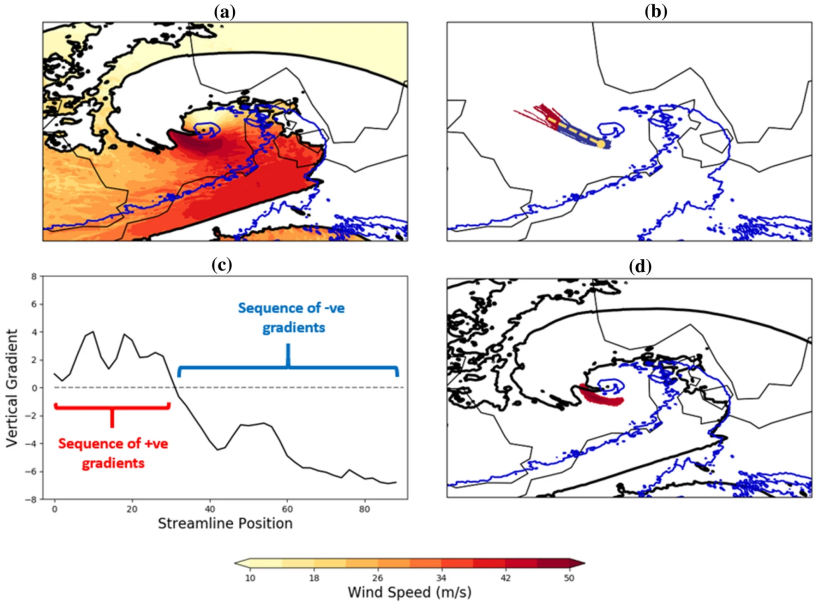


Figure 2.8: Illustration of the methodology for SJ detection by the example of storm *Anatol* (December 1999). (a) Shading corresponds to v_{850}, no shading enclosed by the black line shows $RH > 80\,\%$ at 500 hPa and the blue contour indicates the frontal boundary and warm seclusion by $\theta_{\mathrm{W,thresh}}$. (b) streamlines showing gradient reversals of $\Delta_z v$ (red – positive $\Delta_z v$ (+ve), blue – negative $\Delta_z v$ (-ve). (c) example series of $\Delta_z v$ along a streamline as indicated by the gold line in (b). (d) the extracted SJ at 850 hPa, which corresponds to the blue area in (b) with a wind speed of at least $4\,\mathrm{m\,s^{-1}}$ less than the maximum wind speed. Figure reprinted from Manning et al. (2022), their Fig. 2.

points with positive gradients followed by a sequence with negative gradients as shown in Figure 2.8c. To account for the origin within the cloud head, the median of RH_{ice} within the sequence of positive gradients has to exceed $80\,\%$. Furthermore, within the sequence of the negative gradient, the median of RH_{ice} has to be at least 50 percentage points (RH_{param}) lower to indicate that the potential SJ is leaving the cloud head (Baker et al., 2014) and maximum wind speed has to be at least $6\,\mathrm{m\,s^{-1}}$ higher (v_{param}). Overall, using the native grid spacing of 2.2 km, at least 50 streamlines have to satisfy these

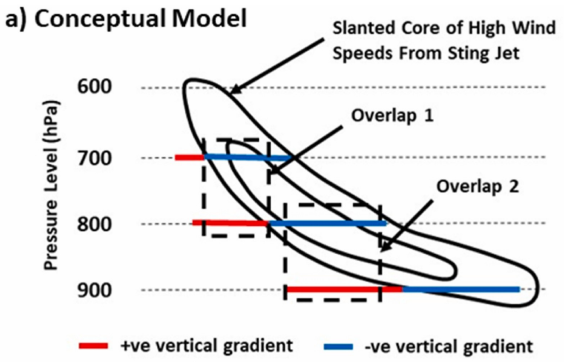


Figure 2.9: Illustration of components of the SJ algorithm by Manning et al. (2023) displaying an idealised vertical cross-section showing an SJ. Red lines represent regions where wind speeds are higher in the level above and blue lines where they are weaker. Figure reprinted from Manning et al. (2023), their Fig. 1a.

criteria to confirm the presence of an SJ. The thresholds were chosen to optimise performance, and the reader is referred to Manning et al. (2022) for further discussion. The authors evaluated their method by applying it to a hindcast covering most documented SJ storms in the last decades. The only SJ not detected was that of storm *Egon* (Eisenstein et al., 2020) but otherwise agreement was large.

In a more recent study, Manning et al. (2023) adapted their method to be more objective, as the rather subjectively chosen thresholds seem to be overtuned for the data set used in Manning et al. (2022). Instead of using only the difference in wind speed between 700 hPa and 850 hPa, the difference between four levels are used as shown in Figure 2.9 (upper: 600 – 700 hPa, middle: 700 – 800 hPa, lower: 800 – 900 hPa). The middle pair needs to satisfy the previously mentioned RH_{ice} and maximum wind speed criteria (RH_{param} and v_{param}, respectively) from Manning et al. (2022). Furthermore, the middle pair must overlap with the lower or upper pairs, that is at least 20 % of the positive sequence in the middle pair must overlap with the negative sequence of the upper pair, or, analogously, 20 % of the positive sequence of the lower pair has to overlap with the negative sequence of the middle pair (Figure 2.9). This allows for coherent features and the subjective threshold of at least 50 streamlines satisfying SJ conditions in Manning et al. (2022) can be abandoned.

3 Research questions

As discussed, extratropical cyclones are often associated with high winds causing immense damage at times, especially during wintertime. The forecast of these events is therefore critical to prepare for and prevent casualties and losses. To improve the forecast, statistical post-processing is often used to correct biases and forecast errors of model output (Vannitsem et al., 2021), as mentioned in Chapter 1. However, Pantillon et al. (2018) showed that post-processing of wind gusts can sometimes lead to forecast busts. Under the hypothesis that these forecast busts are associated with dynamical features within the winter storm, this work aims to identify the high-wind features introduced in Section 2.2 in an objective manner, extend the knowledge of their characteristics and lay the foundation to analyse forecast errors and possibly develop a feature-dependent post-processing method.

Previous work discussed in Chapter 2 suggests various characteristics of the features differing in timing relative to the cyclone life cycle, location relative to the cyclone centre, but also atmospheric stability, intensity and gustiness. However, most approaches were based on subjective identification (e. g., Hewson and Neu, 2015; Earl et al., 2017) or a simplified identification based on the location in the warm or cold side of a front (e. g., Manning et al., 2022; Gentile and Gray, 2023). Furthermore, many studies do not focus on CFC or CS, as they are included in their definition of the WJ or the CJ, respectively (e. g., Hewson and Neu, 2015; Manning et al., 2022; Gentile and Gray, 2023). However, CFC, for example, shows fundamentally different behaviour compared to WJ, despite occurring back-to-back. While the width of a cold front is narrow and the embedded convection spotty, CFC is associated with a high gustiness (Earl et al., 2017) and, hence, with a high wind risk. To account for the different nature of each feature – and the hypothesised different forecast errors – it is important to consider each feature separately.

This work is divided into three main result topical areas: identification of different high wind features, a climatology of these features and, lastly, the detection of SJs. Research questions for each chapter are formulated below.

Chapter 5 concentrates on the question of how high-wind features can be objectively identified. The primary objective is to develop an objective identification of all introduced wind features by exploiting the distinct characteristics using as few parameters and constraints as possible, addressing the following questions:

RQ1a Can the introduced wind features – WJ, CJ, SJ, CFC and CS – be detected reliably using only surface data?

RQ1b Is it possible to develop a method independent of horizontal organisation of input data, hence, applicable to both station observations and gridded data without additional adaptations?

RQ1c Can a machine learning approach learn distinctive characteristics to distinguish the features?

Being able to detect the different wind features in a flexible way enables the computation of a long-term climatology, both verifying discussed characteristics from mostly subjective studies, adding to them and possibly solving seemingly contradicting suggestions from them. Chapter 6 focuses on the question of how the high-wind features are characterised. For this, a climatology over Europe for 19 extended winter seasons (October to March, 2000–2019) – using both an observational and a reanalysis data set – is performed addressing the following aspects:

RQ2a Are results obtained by an objective approach consistent with previous (subjective) literature?

RQ2b What new aspects can an objective long-term climatology bring?

RQ2c Which European regions are commonly more affected by which feature?

The disadvantage of focusing on surface parameters for the objective identification is not being able to distinguish the SJ from the CJ in Shapiro-Keyser cyclones. Up until recently, the only way to distinguish SJ and CJ was the computation of Lagrangian trajectories, which, however, are computationally costly. Therefore, Chapter 7 discusses which low-cost approach can be used to identify SJs. In particular, the advantages and disadvantages of the two introduced alternative SJ (potential) detection methods and how they could be applied to an operational framework using the ICON (ICOsahedral Non-hydrostatic) model of the DWD are discussed. The following questions are addressed:

RQ3a How robust are the tools in regard to model choice and resolution?

RQ3b How do they perform over land, especially in regions with complex topography?

RQ3c What recommendations can be given to the DWD for the development of an operational tool based on the current operational system?

By tackling these research questions, we will be able to identify the different high-wind feature in an objective and consistent manner for the first time and help extending our knowledge of them. In the long run, this work will help to analyse possibly varying forecast errors and improve the overall forecast and, hence, warning of wind gusts caused by extratropical cyclones.

4 Data and methods

This section introduces the data sets used for the training, evaluation and first applications of RAMEFI, as well as model simulations used for the comparison of SJ detection methods. Furthermore, it is described how to evaluate and filter RAMEFI output and a way to visualise cyclones in a three-dimensional framework including the frontal structure. Lastly, Lagrangian trajectories and necessary adaptations to the introduced precursor tool to detect SJs are discussed.

4.1 Data

In the scope of this dissertation, three data sets are used. While an observational and a reanalysis data set are used for RAMEFI, Chapter 7 focuses on the SJ identification using model simulations that are discussed here.

4.1.1 Surface observations

The main basis of the analysis of 12 chosen case studies and the training of the RF is a data set of hourly surface observations available from 2001 to mid-2020. This includes mean sea level pressure (p_{msl}), 2 m air temperature ($T_{2\mathrm{m}}$), wind speed at 10 m ($v_{10\mathrm{m}}$), wind direction at 10 m (d) and precipitation amount (RR). Using $T_{2\mathrm{m}}$ and p_{msl}, we further compute the surface pressure using the barometric height formula to then calculate the potential temperature ($\theta_{2\mathrm{m}}$). Our focus is on Europe, more specifically stations within the area of $-10\,^{\circ}$E to $20\,^{\circ}$E, $40\,^{\circ}$N to $60\,^{\circ}$N (stations included in the green box in Figure 4.1). Around 1700 stations are included of which, however, fewer than 400 stations on average observe all five parameters. For the training of the RF (Section 5), we focus on stations that measure at least three of the five parameters. The most frequent missing parameter in the hourly data is *RR*, as many stations only measure 3- or 6-hourly precipitation. However, many stations, especially over Germany, measure *RR* only and, hence, are not usable for the training of the RFs (Section 5.2.2) but still helpful to inform our subjective labelling (Section 5.2.1). In addition, we exclude mountain stations, that is, those with a station height above 800 m, as we suspect these to be dominated by orographic influences that may blur the feature characteristics we want to identify. This leaves around 750 stations per time step. The distribution of stations measuring three of the five parameter is displayed in Figure 4.1 for a random time step.

In order to take into account the diurnal and seasonal cycles as well as location-specific characteristics (e. g., exposed stations in coastal regions) in $\theta_{2\mathrm{m}}$ and also $v_{10\mathrm{m}}$, we decided to normalise these parameters by their climatology. For $\theta_{2\mathrm{m}}$, this means $\tilde{\theta} = \theta_{2\mathrm{m}}/\theta_{50}$, where θ_{50} is the median for the specific location, time of day and day of the year ±10 days. A comparison with ±30 days for an exemplary month showed no significant differences, hence, ±10 days were used to save computational cost. This is done analogously for $v_{10\mathrm{m}}$ using the 98th percentile ($\tilde{v} = v_{10\mathrm{m}}/v_{98}$), as we are mostly interested in high winds in this work. The 98th percentile is used in analogy to standard high-wind quantities such as the Storm Severity Index (SSI), which is computed from stations where measured gusts exceed the local 98th percentile and provides an integral indication for the strength of the cyclone and the associated potential damage (see Section 4.2.1). Both θ_{50} and v_{98} are computed for the time period 2001 to 2019. Moreover, we are interested in temporal tendencies of p_{msl}, $\tilde{\theta}$ and d, here represented simply by the difference between the current and the prior time step (Δp_{msl}, $\Delta\tilde{\theta}$ and Δd, respectively). All parameters and their descriptions are listed in Table 4.1.

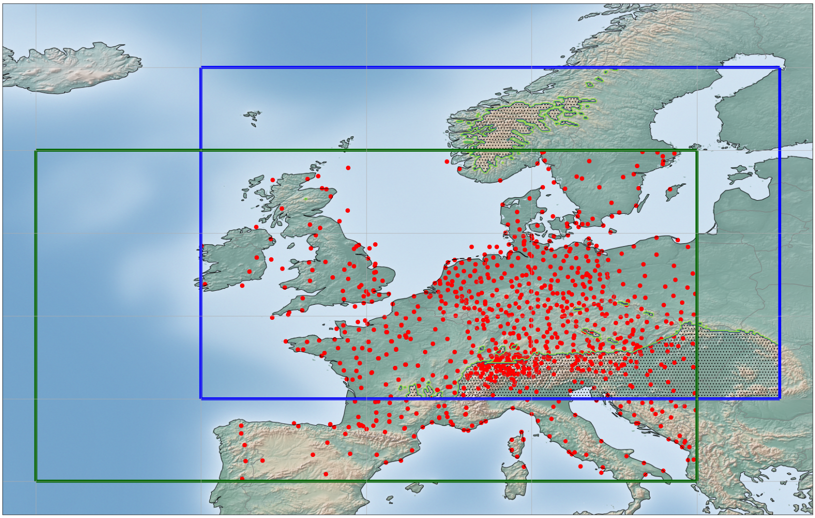


Figure 4.1: Distribution of stations measuring at least three of the parameters used for training (exluding tendencies, i. e., $\tilde{v}$, p_{msl}, $\tilde{\theta}$, RR, d) for a random time step. Shading corresponds to the altitude. The blue box marks the area used for the training of RAMEFI and the compiled climatology (−10 °E to 25 °E and 45 °N to 65 °N), where hatching indicates excluded areas with an altitude over 800 m and areas impacted by orographic effects in the southeast, i. e., the Balkans, while the green box marks the ICON-LAM domain (−20 °E to 20 °E and 40 °N to 60 °N).

4.1.2 COSMO-REA6

As an example for a gridded data set, we use COSMO-REA6 data from the Hans-Ertel-Centre for Weather Research, which is a reanalysis based on the formerly operational COSMO (Consortium for Small-scale Modeling) model from the DWD covering the European CORDEX domain[1] with a grid spacing

---

[1] See `https://cordex.org/domains/cordex-region-euro-cordex/` (last accessed 23 May 2023).

Table 4.1: Overview of the variables considered for the objective identification using the probabilistic RF (upper part) and added parameters from COSMO-REA6 for the climatology (lower part). The fourth column indicates whether the variable is used as an input variable for the final version of the RF. The associated percentiles and medians are computed with respect to the location, time of day and day of the year ±10 days. Tendencies are calculated as $\Delta x = x - x_{-1}$, where x_{-1} is the observation of the previous time step.

Variable	Description	Unit	RF	Derivation
$v_{10\mathrm{m}}$	10m wind speed	$\mathrm{m\,s^{-1}}$	–	Station observation.
$\tilde{v}$	Normalised v	1	✓	$\tilde{v} = v_{10\mathrm{m}}/v_{98}$, where v_{98} is the 98th percentile
d	Wind direction	$^\circ$	✓	Station observation
Δd	Tendency of d	$^\circ\,\mathrm{h^{-1}}$	✓	$\Delta d = d - d_{-1}$
p_{msl}	Mean sea level pressure	hPa	✓	Station observation
Δp_{msl}	Tendency of p	$\mathrm{hPa\,h^{-1}}$	✓	$\Delta p = p - p_{-1}$
$T_{2\mathrm{m}}$	2m air temperature	K	–	Station observation
$\theta_{2\mathrm{m}}$	Potential temperature	K	✓	Derived from T and p
$\tilde{\theta}$	Normalised $\theta_{2\mathrm{m}}$	1	✓	$\tilde{\theta} = \theta_{2\mathrm{m}}/\theta_{50}$, where θ_{50} is the median
$\Delta\tilde{\theta}$	Tendency of $\tilde{\theta}$	$\mathrm{h^{-1}}$	✓	$\Delta\tilde{\theta} = \tilde{\theta} - \tilde{\theta}_{-1}$
RR	Precipitation	$\mathrm{mm\,h^{-}}$	✓	Station observation
v_{gust}	10 m wind gust speed	$\mathrm{m\,s^{-1}}$	–	parameterised
g_{v}	Gust factor	1	–	$g_{\mathrm{v}} = v_{\mathrm{gust}}/v_{10\mathrm{m}}$
$q_{2\mathrm{m}}$	2 m specific humidity	$\mathrm{g\,kg^{-1}}$	–	
$RH_{2\mathrm{m}}$	2 m relative humidity	%	–	
cc	Total cloud cover	%	–	

of 0.055°, i. e., roughly 6 km (Bollmeyer et al., 2015). Several parameters, including wind, temperature, humidity and pressure, are assimilated using a

nudging technique. Observations from radiosondes, aircraft, wind profilers and surface stations are used for the nudging. However, precipitation is not assimilated, which can lead to larger deviations between the reanalysis and observations by rain gauges or radar (Bach et al., 2016; Hu and Franzke, 2020).

The reanalysis is available from 1995 to mid-2019. The same surface parameters as mentioned in Section 4.1.1 are used. The data set contains p_{msl}, $T_{2\mathrm{m}}$, RR and the zonal and meridional surface wind components, from which we can compute $v_{10\mathrm{m}}$ and d. Again, we further calculate $\tilde{\theta}$ and the temporal tendencies Δp_{msl}, $\Delta\tilde{\theta}$ and Δd. Due to computational cost, we compute θ_{50} and v_{98} for the 10-year time period 2005 to 2015 only, but this should have a negligible effect on the final outcome. In addition to the parameters mentioned above, COSMO-REA6 allows us to include further variables, such as 10 m wind gusts (v_{gust}), 2 m specific humidity ($q_{2\mathrm{m}}$), 2 m relative humidity ($RH_{2\mathrm{m}}$) and total cloud cover (cc) (see lower part in Table 4.1). Using v_{gust} and $v_{10\mathrm{m}}$, we can also calculate the gust factor $g_{\mathrm{v}} = v_{\mathrm{gust}}/v_{10\mathrm{m}}$ to account for gustiness of the features.

The data, originally on a rotated grid, are regridded to a latitude-longitude grid with a grid spacing of 0.062 5°, i. e., roughly 6.5 km analogously to the resolution of the deterministic ICON-EU forecast (see Section 4.1.3), for the area of $-10\,^{\circ}$E to $30\,^{\circ}$E, $40\,^{\circ}$N to $65\,^{\circ}$N. However, for the climatology, we focus on the blue box in Figure 4.1 marking the area of $-10\,^{\circ}$E to $25\,^{\circ}$E, $45\,^{\circ}$N to $65\,^{\circ}$N.

4.1.3 ICON model

The ICON model from the DWD is the current operational forecast model in Germany following the COSMO model. Since 2015 it has been run both globally and over Europe (ICON-EU), while a regional run over Germany became operational in early 2021 (ICON-D2). ICON is run on an icosahedral-triangular Arakawa C grid having the advantage of a homogeneous grid

Table 4.2: Initialisation times of ICON-LAM simulations for Shapiro-Keyser cyclones.

Storm	Date	Time
Egon	12 January 2017	00 UTC
Xavier	04 October 2017	12 UTC
Friederike	17 January 2018	12 UTC
Eberhard	09 March 2019	12 UTC
Diana II	01 March 2020	00 UTC
Eunice	17 February 2022	06 UTC

spanning on the entire globe in nearly constant distance between grid points. Hence, the model is independent of latitude and the singularity at the poles, which other models have to deal with, is not an issue. ICON and ICON-EPS (Ensemble Prediction System) have grid spacings of 13 km and 26 km, respectively, while the nest over Europe is run on 6.5 km and 13 km for ICON-EU and ICON-EU-EPS, respectively[2]. ICON-D2 and ICON-D2-EPS have a convection-permitting grid spacing of 2.2 km with explicit deep convection and parametrised shallow convection. For a detailed model description, the reader is referred to Zängl et al. (2015).

---

[2] On 23 November 2022 the grid spacing was upgraded for ICON-EPS and ICON-EU-EPS from 40 km and 20 km to 26 km and 13 km, respectively. For further information see `https://www.dwd.de/DE/fachnutzer/forschung_lehre/numerische_wettervorhersage/nwv_aenderungen/_functions/DownloadBox_modellaenderungen/icon_eps/pdf_2022/pdf_icon_eps_23_11_2022.pdf?__blob=publicationFile&v=2` (last access 21 May 2023)

For the confirmation of SJ storms (Chapter 5) and comparison of SJ detection approaches in Chapter 7, we use the ICON limited area mode (ICON-LAM) to simulate the storm and get a deterministic output. The ICON-LAM simulations are computed analogously to Eisenstein et al. (2020) over Europe (40 °N to 60 °N, 20 °W to 20 °E; see green box in Figure 4.1) with a 0.062 5° grid spacing translating to roughly 6.5 km and 90 vertical levels (MIDRES in Eisenstein et al., 2020), hence, fine enough for resolving an SJ and contributions of mesoscale instabilities. As initial and boundary conditions the deterministic ICON-EU forecast is used. Initialisation times for each case are listed in Table 4.2. The hourly output covers 48 time steps, thus a simulation time of two days and includes surface, model and pressure level data.

4.2 Methods

In the following, various methods used throughout this dissertation are discussed. Firstly, the SSI (Section 4.2.1) and visualisation tool (Section 4.2.2) used for the compilation of case studies for the training of RAMEFI and the approach to evaluate the method (Section 4.2.3) are introduced followed by filtering RAMEFI output for the climatology (Section 4.2.4). The last two sections focus on the SJ detection, i. e., Lagrangian trajectories (Section 4.2.5) and adaptations done to the SJ precursor tool (Section 4.2.6).

4.2.1 Storm Severity Index

The SSI was originally developed to estimate windstorm-related damage to buildings and infrastructure (Klawa and Ulbrich, 2003). With this aim, daily wind gust maxima ($v_{\mathrm{g,max}}$) for DWD stations were first scaled with its 98th percentile ($v_{\mathrm{g,98,s}}$) to take local conditions into account. Next, the exceedances above $v_{\mathrm{g,98,s}}$ are cubed to account for the wind destructiveness and are weighted with the population density as a proxy for the insured values. Later developments introduced formulations for grid-based data, i. e., reanalysis data and climate models, definition of affected areas ("windstorm footprints") and feature tracking (e. g., Pinto et al., 2007; Leckebusch et al., 2008). Following Pantillon et al. (2018), we use the SSI formulation for station observations over Germany considering only the meteorological impact (no population weighting):

$$\mathrm{SSI} = \sum_{\text{station s}} \left\{ \left(\frac{v_{\mathrm{g,max,s}}}{v_{\mathrm{g,98,s}}} - 1 \right)^3 \right\}_{v_{\mathrm{g,max}} > v_{\mathrm{g,98}}} \tag{4.1}$$

Daily maximum wind gusts over Germany are available from the DWD surface network[3]. The SSI is calculated for each single day and displayed in Table 5.1 for the 12 selected cases. If a storm affects the region for several days, the maximum SSI of the daily values is selected. As the aggregated SSI value is dependent on the number of stations, possible changes in the surface network need to be taken into account for when analysing longer time series. For the period of our case studies (2015–2020; see Section 5.1), the number of stations remained stable, and thus a comparison of the SSI values is largely fair. Still, the obtained SSI values only serve as a comparison for the selected case studies in this work and should not be compared numerically to SSI values from other publications.

4.2.2 Three-dimensional visualisation and front detection

For an overview of winter storm cases, the three-dimensional visualisation tool Met.3D is used (Rautenhaus et al., 2015). Originally developed for a field campaign using EPS forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF), it has been continuously advanced and

---

[3] Data available at `https://cdc.dwd.de/portal/` (last access 05 May 2023)

extended with new tools and made applicable for other data sets. The user can interactively analyse a case and choose between horizontal and vertical cross-sections, as well as isosurfaces. Furthermore, Met.3D can either compute Lagrangian trajectories on the fly along, for example, a movable vertical pole or horizontal cross-sections, or visualise pre-computed trajectories (see Section 4.2.5) making it suitable for SJ or WCB visualisation.

A recently added feature is the three-dimensional front detection by Beckert et al. (2023) allowing for a more in-depth analysis of extratropical cyclone cases. This detection follows a gradient-based approach by Hewson (1998) and Kern et al. (2018) and is able to distinguish warm and cold fronts. In addition to the common two-dimensional front detection, 3D fronts allow analysing the direction of tilting of a front, the presence of upper level fronts and more. While various detection parameters can be chosen, here θ_W is used.

The tool is used for qualitative analyses of case studies, as it provides a quick and easy way to get a detailed look at the development over time, the location of fronts and high wind speeds and helps the user to find the right time steps and locations for further detailed analysis.

4.2.3 Assessing probability predictions for multiple wind features

As the output of RAMEFI is of probabilistic nature, a simple evaluation using error rates, such as false alarm and miss rate, is not applicable. Furthermore, the predictions include five different features. Probability predictions of three or more classes, such as the wind features, are typically evaluated by downscaling to two-class problems, of which the one-against-all and all-pairs approaches are two well-known examples (Zadrozny and Elkan, 2002). While the one-against-all approach compares the occurrence of one wind feature against all others grouped together, the all-pairs approach considers the conditional probabilities for each pair of classes, for example, the conditional probabilities of the WJ and the CJ when one of the two features materialises. The one-against-all approach is used to evaluate how well one specific wind feature is forecast, the all-pairs approach to evaluate the ability to discriminate between two wind features.

The probabilities are evaluated based on the paradigm of Gneiting et al. (2007) that a prediction should aim to maximise sharpness subject to calibration. Calibration refers to the consistency of the prediction and the observation, while sharpness is a property of the prediction alone and refers to the associated uncertainty. In a nutshell, a probability f is called calibrated if the conditional event probability (CEP; conditional on f) matches f, for example, if a 20 % forecast is issued 100 times, the event should occur about 20 times. Further, a probability prediction is said to be sharper, the more confident the prediction is, that is, the closer to 0 or 1. Both calibration and sharpness can be assessed qualitatively via reliability diagrams, which display the calibration curve (Sanders, 1963; Wilks, 2011). The calibration curve is a plot of the CEP dependent on the probability f, which is close to the diagonal if the predictions are calibrated. In addition to the calibration curve, the frequency of the probabilities is illustrated by a histogram. The more U-shaped the histogram is, the closer the predictions are to 0 and 1 and thus the sharper. The reliability diagrams shown in this paper are based on the novel CORP-approach (Consistency, Optimality, Reproducibility, PAV-algorithm based) of Dimitriadis et al. (2021), which yields optimal calibration curves and eliminates the need for implementation decisions for the calculation of the calibration curve.

Quantitatively, calibration and sharpness can be assessed using the Brier score (BS; Brier, 1950). The lower the score, the better the prediction. Here, we compare our RF probabilities with the class frequencies observed in the training data of the RF using the Brier skill score (BSS), which denotes the improvement of the RF over a prediction based on the class frequencies in the RF training data, where a negative percentage corresponds to worse predictions, 0% to no improvement and 100% to an optimal prediction. Details on the practical implementation of the evaluation of probability predictions are provided in Appendix A.1.1. For details on the mathematical formulation, the reader is referred to the Appendix of Eisenstein et al. (2022a).

4.2.4 Filtering for cyclone tracks

In general, RAMEFI (see Section 5) can be used without any filters. However, for a meaningful climatology we aim to exclude high winds not associated with extratropical cyclones and, hence, the mesoscale wind features are targeted here. So to filter the gained probabilities and also to compile a storm-relative analysis, objectively determined cyclone tracks are used. The cyclones are identified and tracked on the basis of ERA5 (Hersbach et al., 2020) p_{msl} data using an objective tracking algorithm (Murray and Simmonds, 1991; Pinto et al., 2005). The algorithm primarily searches for the minimum p_{msl} in the vicinity of a $\nabla^2 p_{\mathrm{msl}}$ maximum to find potential cyclone centres within a radius of 750 km. To filter out weak and thermal lows, or cyclones over high orography, we follow the criteria from Pinto et al. (2009). The method settings used for ERA-Interim (Neu et al., 2013) were slightly adapted to handle the higher spatial resolution data of ERA5, while the time resolution was kept at 6 h intervals. The resulting cyclone tracks are then interpolated linearly to gain hourly information. The track density, i. e., number of cyclones passing over a grid point (Ulbrich et al., 2009), of all cyclones within the 19 years can be seen in Figure 4.2. As expected given the selected study area (blue box in Figure 4.2), cyclone tracks corresponding to the identified features typically travel over the British Isles and the North Sea towards the Baltic Sea.

We further remove time steps, where fewer than 5 % of all grid points during a time step are associated with one of the features (excluding NF), i. e., weak cyclones. Time steps with a cyclone moving through and at least 5 % windy conditions are referred to as "stormy time steps". In addition, we extract the surroundings of a cyclone that is the area of $\pm 15^\circ$ in zonal and -15° to 5° in

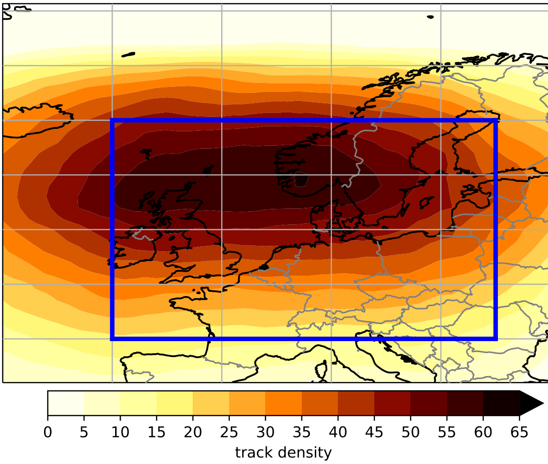


Figure 4.2: Cyclone track density (average number of tracks per year per (° lat)^2) of all cyclones within the examined time period (2000–2019). The blue box represents the study area.

meridional direction around a cyclone centre. A reduced distance in northern direction is chosen, as the examined wind features typically occur south to southwest of a cyclone.

4.2.5 Lagrangian trajectories

As discussed in Section 2.4, to be certain an SJ occurred, Lagrangian trajectories have been an important tool in the previous literature (e.g., Gray et al., 2011; Baker et al., 2014; Volonté et al., 2018; Eisenstein et al., 2020). Trajectories are computed analogously to Eisenstein et al. (2020) using LAGRANTO (LAGRangian ANalysis TOol; Wernli and Davies, 1997; Sprenger and Wernli, 2015). As coherent air streams cannot be followed when turbulence occurs, this tool should not be used in the boundary layer. Therefore, starting points are defined by grid points in the vicinity of the cyclone centre, hence in the frontal-fracture region of a Shapiro-Keyser cyclone, in 850 hPa where wind speeds exceed a certain threshold depending on the strength of the

SJ. These grid points are then extended to vertical layers between 900 hPa to 800 hPa following previous studies. Note that winds at lower pressure levels might be affected by orography over land. While a frequency of 15 minutes is supposed to be beneficial for a detailed trajectory analysis, a frequency of 1 h is sufficient to confirm SJs (e. g., Volonté et al., 2018; Eisenstein et al., 2020) and used here for a comparison to the introduced detection approaches.

Several starting times are chosen to identify time steps and the overall duration of the SJ occurrence. Backward trajectories over a time period of 12 h are used to identify the descending nature of an SJ. Furthermore, forward trajectories are computed for a period of 6 h to detect further descent to the boundary layer and merged with the backward trajectories. In addition to the v_{850} threshold to focus on high winds, trajectories have to exceed a Δp threshold within a certain time, here 8 h, indicating the descent to extract the SJ from the computed trajectories and separate it from the CJ.

4.2.6 Adaptation of the SJ precursor approach

The precursor tool by Gray et al. (2021) is based on the method by Martínez-Alvarado et al. (2013) and was developed for low-resolution models, where mesoscale processes and, hence, SJs cannot be resolved. Here, however, we apply it to high-resolution model simulations with a grid spacing of 0.062 5°. Although mesoscale instabilities should be resolved (comparing to the 10–12 km grid spacing deemed necessary in Clark et al., 2005), the method should still be cheaper computationally compared to Lagrangian trajectories and is therefore attractive for operational use. Nonetheless, the regridding of higher resolution brings higher computational costs compared to the 20 km ensemble forecast used at the UK Met Office and increased noise, potentially disrupting the coherency of an area of detected CSI points. Furthermore, our focus is on continental Europe and less on the North Atlantic and the UK alone, such that orographic effects may distort the signals significantly.

The following adaptations to the code[4] were made:

- θ_W was computed over a recursive loop, stopping when a specific threshold was reached. This subfunction was exchanged by an approximation of θ_W following Bolton (1980) and Davies-Jones (2008):

$$T_\mathrm{LCL} = \frac{1}{\frac{1}{T-55} - \frac{\ln(RH/100)}{2840}} + 55 \tag{4.2}$$

$$\begin{aligned}\theta_\mathrm{E} = T \cdot \left(\frac{1000}{p}\right)^{\left(0.2854\cdot(1-0.28\times10^{-3}\cdot m)\right)} \\ \cdot \exp\left(\left(\frac{3.376}{T_\mathrm{LCL}} - 0.00254\right)\cdot m \cdot (1+0.81\times10^{-3}\cdot m)\right)\end{aligned} \tag{4.3}$$

$$257\,\mathrm{K} < \theta_\mathrm{E} < 377\,\mathrm{K}:$$

$$\theta_\mathrm{W} = 45.114 - 51.489\left(\frac{C}{\theta_\mathrm{E}}\right)^{\lambda} + C \tag{4.4}$$

$$377\,\mathrm{K} \leq \theta_\mathrm{E} < 674\,\mathrm{K}:$$

$$\begin{aligned}\theta_\mathrm{W} = 43.38 - 51.489\left(\frac{C}{\theta_\mathrm{E}}\right)^{\lambda} + 0.0609\left(\frac{\theta_\mathrm{E}}{C}\right)^{\lambda} \\ - 0.01005\left(\frac{\theta_\mathrm{E}}{C}\right)^{2\lambda} + C,\end{aligned} \tag{4.5}$$

---

[4] The original code is available at `https://github.com/omartineza/csisounding/tree/79cca7d` (version from 28 June 2019 – last access 10 May 2023)

where T_{LCL} is the temperature at the lifting condensation level, m is the saturation mixing ratio, $C = 273.15\,\mathrm{K}$ and $\lambda = c_{\mathrm{pd}} R_{\mathrm{d}}^{-1} \approx 3.504$ (c_{pd}: the specific heat of dry air; R_{d}: the gas constant). This exchange speeds up the script from several minutes per time step to mere seconds in case of 0.062 5° grid spacing.

- grid points with an altitude of over 800 m are removed
- the fields are smoothed using a Gaussian filter
- The original data to which the code is applied to does not include RH such that it is computed using the specific humidity and the saturation vapour pressure. This is done analogously for RH_{ice}. In case of storm *Friederike*, RH_{ice} seemed to have unreasonable low values (never reaching 80 %) for all time steps using the original function (Gramlich, 2023). As RH is given in the output of the ICON-LAM simulation used here, the following formulas were tested as used in Eisenstein et al. (2020) (following Murphy and Koop, 2005):

$$p_{\mathrm{ice}} = \exp\left(9.550426 - \frac{5723.265}{T} + 3.53068 \cdot \ln T - 0.00728332\right) \tag{4.6}$$

$$\begin{aligned} \ln p_{\mathrm{liq}} \approx 54.843 - \frac{6763.22}{T} - 4.210 \ln T + 0.000367 \cdot T \\ + \tanh\left(0.0415 \cdot (T - 218.8)\right) \\ \cdot \left(53.878 - \frac{1331.22}{T} - 9.445 \ln T + 0.014 \cdot T\right) \end{aligned} \tag{4.7}$$

$$RH_{\mathrm{ice}} = RH \cdot \frac{p_{\mathrm{liq}}}{p_{\mathrm{ice}}}, \tag{4.8}$$

with water vapour pressure over ice p_{ice} and liquid p_{liq}. Based on positive results, we decided to keep these formulas for the computation of RH_{ice} here.

- As *DSCAPE* is computed using a recursive loop similarly to θ_W, it takes a lot of time. Instead of computing it for every grid point, we only calculate *DSCAPE* for grid points satisfying the other criteria ((2)–(4); Section 2.4)

5 Identification of high-wind features using a probabilistic random forest

As discussed in Chapter 2, several previous studies focused on the characteristics of different wind features (e. g., Hewson and Neu, 2015; Earl et al., 2017). However, as all of these approaches are purely subjective, only consider some of the introduced wind features, or are relatively time-consuming and, thus, hard to automate, we aim to develop an objective analysis of the different mesoscale wind features that can be applied flexibly to station and gridded data and, therefore, serve as a basis for climatological studies, forecast evaluation, post-processing development and impact research.

The strategy that we follow is to start with a subjective identification (as in previous studies) but to use the results to then train a probabilistic random forest (RF) to develop an objective procedure that can be applied to cases outside of the training data set. The identification is designed to be independent of horizontal gradients (and hence resolution) and can principally be applied to observations from a single weather station. In addition, the identification is based on tendencies over 1 h only, making it applicable to time series with gaps. Our newly developed method is referred to as RAMEFI (RAndom-forest-based MEsoscale wind Feature Identification). Given that the provision of such a feature-dependent post-processing tool can enhance the forecasts of strong winds and wind gusts, it can potentially contribute towards better weather warnings and impact forecasting of such events (e. g., Merz et al., 2020).

This chapter will show examples using surface stations (Section 4.1.1) and COSMO reanalysis data (Section 4.1.2). A full long-term climatology is the focus of Chapter 6. The output of the RF are feature probabilities, rather than binary identification, which allow for an evaluation of how well individual data points fit the typical feature characteristics as well as the identification of hybrid features or transition zones.

5.1 Case studies

In this work we focus on 12 winter storm case studies between the years 2015 and 2020 listed in Table 5.1. The selection was based on their SSI over Germany (see Section 4.2.1), caused damage and impacted area. This includes the eight winter storms with the highest SSI during this time period plus four subjectively chosen more moderate storms to capture a healthy diversity of cyclones and features. The selected cases occurred during the extended winter half year between the end of September and end of March.

They vary in terms of their cyclone tracks (Figure 5.1) and occurring high-wind features, and include both Norwegian and Shapiro-Keyser cyclones. We also include two storms, namely *Herwart* and *Sabine*, with an exceptional large pressure gradient leading to a stronger background wind field, such that it is more difficult to distinguish the features and the contribution of them to the storm's wind footprint. Further, *Sabine* stands out to be an extremely deep cyclone with a minimum core pressure of 944 hPa during its lifetime.

The selected case studies include four Shapiro-Keyser cyclones, namely *Egon*, *Xavier*, *Friederike* and *Eberhard*. To confirm the presence of a possible SJ, ICON simulations (Section 4.1.3) and Lagrangian trajectories using LAGRANTO (see Section 4.2.5; Sprenger and Wernli, 2015) were computed analogously to *Egon* in Eisenstein et al. (2020). Descending trajectories ($v > 35\,\mathrm{m\,s^{-1}}$, $\Delta p > 125\,\mathrm{hPa}$) were found for *Friederike*, but not for *Xavier* nor *Eberhard*. In first tests, we further included *Diana II* (01–02 March 2020), which developed an SJ over France and southwestern Germany.

Table 5.1: Selected winter storm cases from 2015 to 2020 over Central Europe (names as given by FU Berlin), date, maximum observed gust speed (location), SSI over Germany and associated high-wind features including exceptional large pressure gradients (SPG). The cyclone tracks are displayed in Figure 5.1.

Case	Date	max. gust [km h^{-1}]		SSI	Features
		above 800 m	below 800 m		
Niklas	31 Mar 2015	192 (Zugspitze, D, 2 964 m)	148 (Weinbiet, D, 553 m)	20.8	WJ, CFC, CJ
Susanna	09 Feb 2016	158 (Patscherkofel, AT, 2 247 m; Pilatus, CH, 2 106 m)	158 (Île de Groix, FR, 46 m)	3.6	WJ, CFC, CJ
Egon	12-13 Jan 2017	150 (Fichtelberg, D, 1 231 m)	148 (Weinbiet, D, 553 m)	5.9	WJ, SJ, CJ
Thomas	23-24 Feb 2017	158 (Brocken, D, 1 134 m)	152 (Capel Curig, UK, 216 m)	3.0	WJ, CJ
Xavier	05 Oct 2017	202 (Sněžka, CZ, 1 602 m)	141 (Saint-Hubert, BE , 563 m)	6.3	WJ, CJ
Herwart	29 Oct 2017	176 (Fichtelberg, D, 1 231 m)	144 (List/Sylt, D, 26 m)	15.2	WJ, CFC, SPG
Burglind	03 Jan 2018	217 (Feldberg, D, 1 490 m)	150 (Wädenswil, CH, 463 m)	15.2	WJ, CFC, CJ
Friederike	18 Jan 2018	204 (Brocken, D, 1 134 m)	144 (Hoek Van Holland, NL, 7 m)	18.3	WJ, SJ, CJ
Fabienne	23 Sep 2018	141 (Feldberg, D, 1 490 m)	158 (Weinbiet, D, 553 m)	4.6	WJ, CFC
Bennet	04 Mar 2019	181 (Cairngorn, UK, 1 245 m)	151 (Cape Corse, FR, 106 m)	5.1	WJ, CFC
Eberhard	10 Mar 2019	194 (Sněžka, CZ, 1 602 m)	141 (Weinbiet, D, 553 m)	10.1	WJ, CJ
Sabine	09-10 Feb 2020	195 (Sněžka, CZ, 1 602 m)	219 (Cape Corse, FR, 106 m)	20.0	WJ, CFC, SPG

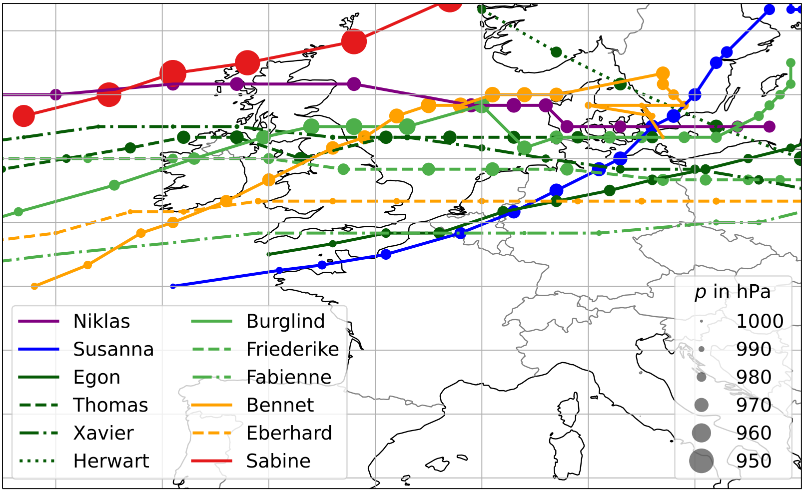


Figure 5.1: Cyclone tracks of case studies from Table 5.1. Colour indicates the year (purple = 2015, blue = 2016, dark green = 2017, green = 2018, orange = 2019, red = 2020), linestyle indicates the order in that year (solid - dashed - dashdot - dotted). The size of the markers correspond to the minimum mean sea level pressure p_{msl} – here p.

Although it caused gusts of over 100 km h^{-1} in the plains, e. g., 107 km h^{-1} in Rheinstetten, D (116 m), it only affected a very small area, also noticeable with an SSI of only 0.59. Hence, after we decided to combine the SJ and CJ (see Section 5.4.1 and Appendix A.2), we decided to remove the storm from our set.

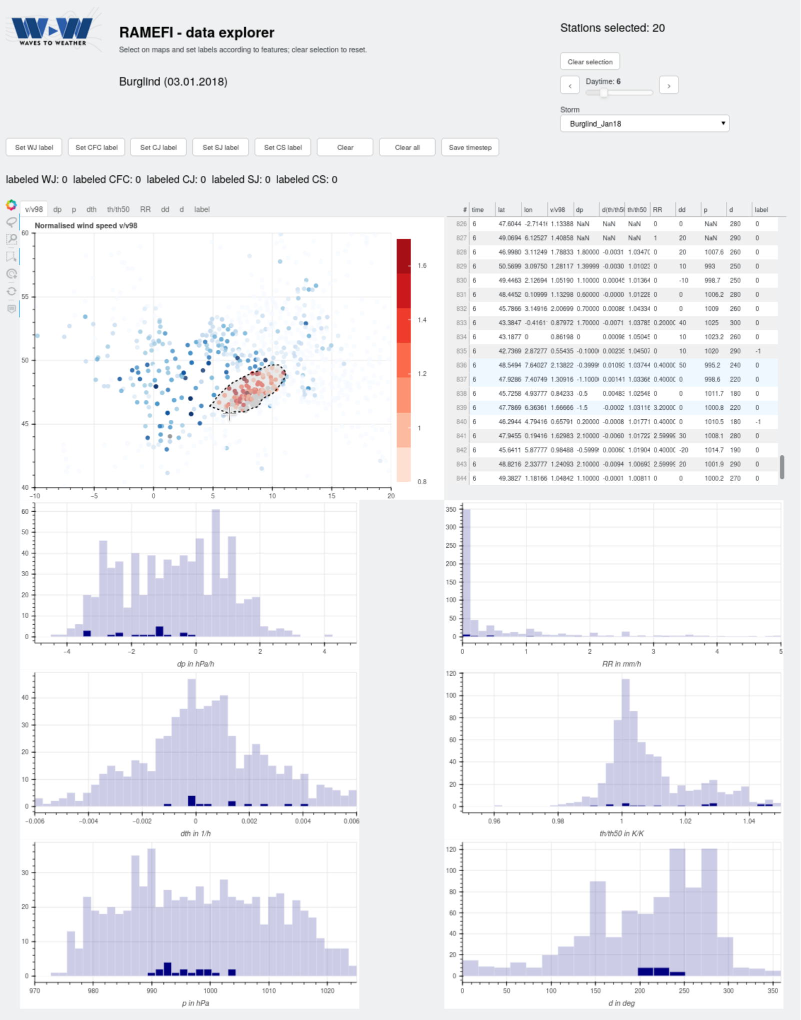
5.2 RAMEFI

Our new method RAMEFI focuses on strong but not only exceptionally high wind speeds. The latter are usually indicated by the 98th percentile. To obtain a sufficiently large storm area and to base its definition on a widely used reference, we decided to include stations reaching 80 % of their 98th percentile, i. e., $\tilde{v} \geq 0.8$. To capture usually narrow and fast-moving features such as CFC, RAMEFI requires hourly data. All used parameters are independent of the location of the station/grid point and horizontal gradients, such that, in principle, the approach can be applied to a single station and data sets with differing horizontal resolution. The approach evaluates each 1 h interval independently.

RAMEFI includes three steps described in the following subsections. First, we identify the features subjectively in surface observations in 12 selected case studies, such that each station is assigned to a specific feature. These labels are then used to train RFs for feature prediction on the basis of a cross-validation approach. In a final step, we obtain forecasts on a grid by interpolating the predicted probabilities using a Kriging approach. For the COSMO-REA6 data, the features are identified analogously. Instead of training separate RFs, we apply the RFs trained on the surface observations. As the COSMO-REA6 forecasts are already grid-based, the Kriging step is obsolete.

5.2.1 Subjective labelling using an interactive tool

Given the sometimes unclear distinction between the high wind features of interest in realistic cases, we decided to base our algorithmic development on how experienced meteorologists would identify the features on the basis of a wide range of parameters and their evolution in time and space. To accelerate and facilitate the subjective labelling of high-wind features, we developed an interactive tool using the open-source data visualisation package `bokeh` for Python (Bokeh Development Team, 2021), where one can switch between the available parameters (Table 4.1) in a graphical display and select an area to set labels using a mouse-controlled lasso tool as shown in Figure 5.2. Furthermore, histograms of the various parameters are shown for all and selected data points only. The values and set labels for each station/grid point are displayed in a data table, which can be saved before selecting a new time step or case.



RAMEFI - data explorer

Select on maps and set labels according to features; clear selection to reset.

Burglind (03.01.2018)

Stations selected: 20

Daytime: 6

Storm: Burglind_Jan18

labeled WJ: 0 labeled CFC: 0 labeled CJ: 0 labeled SJ: 0 labeled CS: 0

#	time	lat	lon	v/v98	dp	d(th/th5	th/th50	RR	dd	p	d	label
826	6	47.6044	-2.7141	1.13388	NaN	NaN	NaN	0	0	NaN	280	0
827	6	49.0694	6.12527	1.40858	NaN	NaN	NaN	1	20	NaN	290	0
828	6	46.9980	3.11249	1.78833	1.80000	-0.0031	1.0347	0	20	1007.6	260	0
829	6	50.5699	3.09750	1.28117	1.39999	-0.0030	1.0102	0	10	993	250	0
830	6	49.4463	2.12694	1.05190	1.10000	0.0004	1.01364	0	-10	998.7	250	0
831	6	48.4452	0.10999	1.13298	0.60000	-0.0000	1.0122	0	0	1006.2	280	0
832	6	45.7866	3.14916	2.00699	0.70000	0.0008	1.04334	0	0	1009	260	0
833	6	43.3847	-0.4161	0.87972	1.70000	-0.0071	1.0378	0.2000	40	1025	300	0
834	6	43.1877	0	0.86198	0	0.0009	1.0504	0	10	1023.2	260	0
835	6	42.7369	2.87277	0.55435	-0.1000	0.0023	1.04507	0	10	1020	290	-1
836	6	48.5494	7.64027	2.13822	-0.3999	0.0109	1.03744	0.4000	50	995.2	240	0
837	6	47.9286	7.40749	1.30916	-1.1000	0.00141	1.0336	0.4000	0	998.6	220	0
838	6	45.7258	4.93777	0.84233	-0.5	0.0048	1.0254	0	0	1011.7	180	0
839	6	47.7869	6.36361	1.66666	-1.5	-0.0002	1.0311	3.2000	0	1000.8	220	0
840	6	46.2944	4.79416	0.65791	0.20000	-0.0008	1.01771	0.4000	0	1010.5	180	-1
841	6	47.9455	0.19416	1.62983	2.10000	-0.0060	1.0172	2.5999	30	1008.1	280	0
842	6	45.6411	5.87777	0.98488	-0.5999	0.0006	1.01904	0.4000	-20	1014.7	190	0
843	6	48.8216	2.33777	1.24093	2.10000	-0.0094	1.0069	2.5999	20	1001.9	290	0
844	6	49.3827	1.18166	1.04842	1.10000	-0.0001	1.00811	0	0	1000.2	270	0

Figure 5.2: Screenshot of the interactive visualisation and labelling tool as part of RAMEFI (Eisenstein et al., 2022c) using the Python package `bokeh` (Bokeh Development Team, 2021). The user can switch between timesteps and loaded winter storms. Top row shows a map of stations and their parameters, here $\tilde{v}$, where the user can select an area by mouse using a lasso or clicking on single stations. Labels for the introduced features can be set for the selected stations. A table includes all data points and parameters. Histograms are shown for several parameters for the whole region (light blue) and currently selected stations (dark blue).

The guiding principles for the labelling were extracted from the scientific literature (Section 2.2) and are mainly based on the location relative to the cold front and cyclone core. For this, analysis charts of the DWD and the UK Met Office were used for orientation. For a more detailed analysis, we further used the three-dimensional front detection of the interactive visualisation software Met.3D (Beckert et al., 2023, Section 4.2.2). In our surface parameters, a cold front is then mostly identified through the characteristic change of the sign of Δp_{msl}. It is labelled CFC, if a larger area of precipitation along it is observed, while high winds ahead of the front within the warm sector are labelled WJ. The CJ is mostly detected through its hook-shaped wind footprint at the tip of a wrapped-around occlusion or bent-back front as well as through its proximity to the cyclone centre. An SJ is labelled, where model-based trajectories analogous to Eisenstein et al. (2020) confirm a descending air stream. The area behind the cold front that is not associated with the CJ or SJ is labelled as the CS. An example of the reasoning behind the labelling is described for one time step of storm *Burglind* in Section 5.3.

The subjective labelling was done for the introduced 12 extratropical cyclone case studies (Section 5.1). In total 282 time steps have been analysed. As mentioned in Section 4.1.1, we excluded mountain stations and stations, where less than three of the given parameters were measured. This leaves around 750 stations per time step for the subjective labelling. Overall for the 12 case studies, we have 77.517 data points where $\tilde{v} \geq 0.8$, of which 19 200 (24.77 %) are not associated with a feature (NF), 21 809 (28.13 %) were labelled as CS, 19 501 (25.16 %) as WJ, 11 705 (15.1 %) as CJ, 3 800 (4.9 %) as CFC and 1 502 (1.94 %) as SJ. However, the SJ is a small, short-lived and rare feature and the characteristics of SJs and CJs in surface parameters are very similar due to the proximity in both time and space. A first training with SJ and CJ as separate features showed that a clear distinction is not possible with the information at hand, and that the SJ is often detected as CJ. However, an SJ is also detected when no SJ occurred, as further discussed in Section 5.4.1. Note that the first training was based on a different set of storms including a third SJ storm and more predictors. For more details, see Appendix A.2. Following these results, we decided to include it in the more frequent CJ feature, increasing the values for CJ to 13 207 data points (17.04 %).

The features were further labelled in all case studies using the interactive tool for COSMO-REA6 data – except for *Sabine* that occurred outside of the reanalysis time period. These labels are used to evaluate the predictions generated by the station-based RFs for a grid-based data set (Section 5.2.2 and 5.4.1). For computational reasons, i. e., as labels are set for every grid point rather than an area, we downsampled the COSMO grid to every third grid point in the zonal and meridional directions, resulting in a grid spacing of 0.187 5°, i. e., around 21 km. Moreover, we excluded ocean grid points to avoid uncertainty in the cause of differences in the output, as the characteristics of the high-wind features might be different from land due to different surface friction, surface heat fluxes, etc. Regions with high wind speed not directly associated with a winter storm, especially over Italy and the Balkans, were not labelled.

5.2.2 Probabilistic Random Forest

First approaches to identify the wind features were based on a point scale for several characteristics fulfilled at a given station, similar to the ones used for labelling the features as discussed in Section 5.2.1. As this approach was dependent on rather subjective thresholds, we decided to test a machine learning approach trained on the set labels.

An RF (Breiman, 2001) is a popular, robust machine learning method for classification and regression problems that does not rely on parametric assumptions but instead is based on the idea of decision trees (Breiman, 1984). Given data for which we want to generate predictions, a decision tree operates by assigning each sample to one of its so-called leaves, which is done by subsequent queries at the so-called nodes, where we check whether one of the input variables exceeds a certain threshold. Each leaf is associated with that subset of the training data that satisfies the criteria of the corresponding nodes. Tailored to the specific task, a prediction is then derived from the associated subset. First tests using a hard classification were not successful in separating the wind features. This is to be expected given the case-to-case variability and uncertainty, thus a probabilistic RF is used. Here, we obtain probability forecasts by using the frequencies of the observed wind features among the samples in the corresponding leaf. The queries at the nodes, that is, the specific input variables and corresponding thresholds, are chosen automatically such that the terminal leaves are as diverse as possible. The maximal depth, the minimal node size and the diversity criterion are tuning parameters and can be chosen by the user. An RF builds a randomised ensemble of decision trees, where the generation of each tree is based on a different subsample of the training data and that of each node on a different subset of input variables. To obtain a final prediction from the ensemble, the individual predictions of the decision trees are aggregated. For further details on RFs, we refer to Breiman (2001) and Hastie et al. (2009). In a meteorological context, probabilistic RFs have already been applied to predict damaging convective winds (Lagerquist et al., 2017) and severe weather (Hill et al., 2020), but also in a general form for a wide range of applications such as ensemble post-processing of surface temperature and wind speed (Taillardat et al., 2016).

Machine learning methods such as the RF are often referred to as black boxes due to a lack of interpretability, although there exist several techniques to understand what the models have learnt and how the predictions are related to the input variables, typically referred to as predictors (McGovern et al., 2019). We will apply two of these predictor importance techniques, one to find the most relevant predictors and one that illustrates the effect of the predictor values on the RF probabilities. The first is the permutation importance of a predictor (Breiman, 2001). Proceeding separately for each predictor, the values of that predictor are shuffled randomly within the test data in space and time such that the physical relation to the observed wind feature is broken. Then, based on these permuted predictor values, new predictions are generated and compared to those obtained with the original data. The worse the predictions become (with respect to an evaluation measure), the more important the predictor. Here, we measure predictive performance with the BS that was introduced in Section 4.2.3, the importance measure is referred to as BS permutation importance. The second technique is given the partial dependence plot (PDP; Greenwell, 2017), which illustrates the effect of a predictor on the prediction. Given a fixed predictor, a PDP shows the expected RF probability dependent on the value of the predictor variable while averaging out the effects of the other predictors. Hence, a PDP illustrates how the RF probabilities depend on the value of a specific predictor variable, on average. For more details, we refer to McGovern et al. (2019).

In this study, we apply RFs to generate probabilities of the wind features presented in Section 2.2. The input variables used are listed in Table 4.1. For the station-based observations, we use a cross-validation scheme on the different winter storm cases, that is, for each winter storm the predictions are generated by an RF that is trained on the data of the remaining 11 winter storms. Training RFs in a similar cross-validation scheme for the COSMO-REA6 data becomes computationally infeasible as the underlying data sets become too large. Since the underlying processes should coincide for both the station- and model-based data, we instead apply the station-based RFs in the same scheme to generate probabilities using the COSMO-REA6 data. Details on the implementation including the choice of the tuning parameters can be found in Appendix A.1.2.

Due to normalising $\theta_{2\mathrm{m}}$ and $v_{10\mathrm{m}}$, the trained RF is fairly independent of location-specific information, such that it can hopefully be applied successfully to other mid-latitude regions around the world affected by extratropical cyclones. However, before doing that we recommend a thorough sanity check, particularly when using it over the ocean and mountainous regions.

5.2.3 Kriging

As it is difficult to envision a coherent area of a certain wind feature from probabilities at single stations that are distributed irregularly over the study area, we interpolate the station-based probabilities to a regularly spaced grid in order to visualise the results. In geostatistics, this is generally achieved by Kriging (Matheron, 1963). In principle, the Kriging predictions (here on the grid) are the weighted averages of the input data (here the station data), where the specification of the weights is driven by the covariance of the underlying random process. Under the assumption of Gaussianity (Rasmussen and Williams, 2005), Kriging provides the optimal full predictive distribution. The key requirement for the implementation of Kriging in the context of Gaussian processes is the specification of the mean and the covariance function.

In this study, we perform univariate Kriging to obtain probability maps for each wind feature, where we specify the mean and covariance function by a constant mean function and the stationary Matérn covariance function (Matérn, 1986; Guttorp and Gneiting, 2006). For the estimation, we resort to the method of maximum likelihood estimation for Gaussian processes. However, since the input data are, in our case, probabilities and thus deviate from the Gaussianity assumption, we perform a data transformation for approximate Gaussianity. For the production of probability maps, we independently perform Kriging on each of the class probabilities, hence univariate Kriging, and normalise the resulting probabilities for each grid point such that, across the multiple wind feature, the probabilities sum to 1. Note that the Kriging predictions are only obtained for areas over land, where our winter storms occurred and where a sufficient amount of data was available for a reliable interpolation. More details regarding the implementation of the Kriging approach are provided in Appendix A.1.3. For details on the mathematical formulation, the reader is referred to the Appendix of Eisenstein et al. (2022a).

5.3 Illustrative case study: Storm Burglind

In this section a full case study for Storm *Burglind* is presented to illustrate the functionality of the new feature detection method RAMEFI. *Burglind* is relatively close to a "textbook" cyclone and shows a feature evolution largely in concordance with the literature.

5.3.1 Synoptic evolution

Storm *Burglind* (also known elsewhere as $Eleanor$) developed as a secondary cyclone on 02 January 2018 over the North Atlantic and reached the British Isles at the end of that day. The core pressure dropped by more than 27 hPa in 24 h and, thus, exceeds the criteria for an explosive cyclogenesis after Sanders and Gyakum (1980) (see Equation 2.2). The minimum pressure occurred just east of the English North Sea coast around 03 UTC on 03 January. The cyclone then tracked mostly eastward across the North Sea and Baltic Sea before heading north-eastward in later stages (Figure 5.1). The cold front crossed France and Germany in the first half of 03 January and caused high winds due to CFC. Ahead of the front, high winds were associated with the WJ. Later, when the occlusion front wrapped around the cyclone centre, the CJ dominated as the cause of high winds in addition to CS winds further away from the cyclone core.

5.3.2 Application of feature identification

Figure 5.3 shows the most important parameters, namely p_{msl}, RR, $\Delta\tilde{\theta}$, to distinguish the high-wind features for one selected time step, i. e., 03 January 2018, 06 UTC, to illustrate the process of the subjective labelling as described in Section 5.2.1. The cyclone centre was located to the east of the UK over the North Sea (red x in Figure 5.3). The cold front stretched from northwestern Germany to France (see Figure 5.4 in the Appendix). Highest winds were observed ahead of and along the cold front and over western France (contours

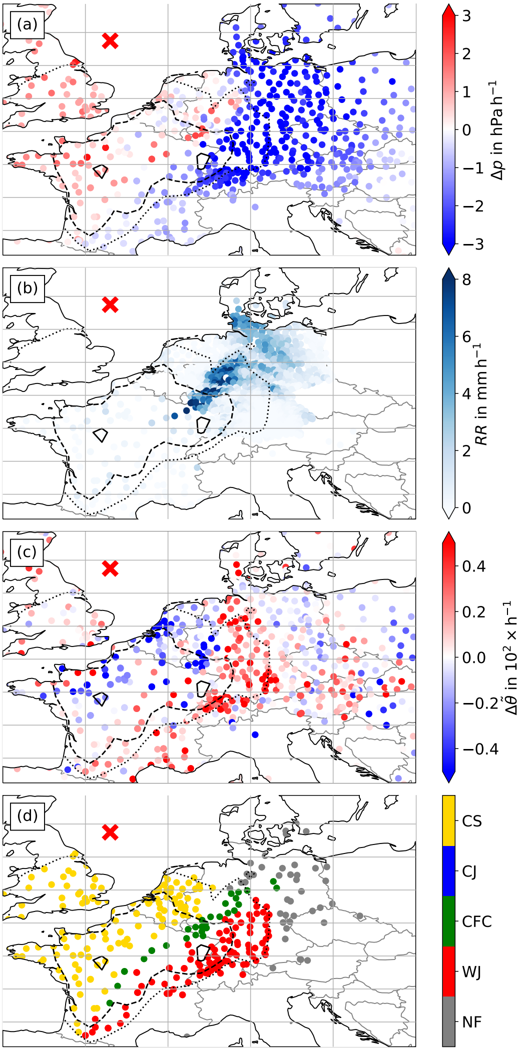


Figure 5.3: Storm *Burglind* on 03 January 2018, 06 UTC. Scattered dots show station observations for (a) Δp_{msl} – here Δp, (b) *RR*, (c) $\Delta\tilde{\theta}$ and (d) subjectively identified wind features for stations where $\tilde{v} \geq 0.8$ and at least three of the five initial parameters are measured (Section 5.2.1). The contours show the interpolated $\tilde{v}$ (dotted $\tilde{v} = 0.8$, dashed $\tilde{v} = 1$, solid $\tilde{v} = 1.2$) to display the regions where the highest winds occurred. The red x indicates the cyclone centre

in Figure 5.3). Figure 5.3a shows p_{msl} decreasing strongly with values of mostly above $3\,\mathrm{hPa\,h^{-1}}$ over southern and Eastern Germany ahead of the cold front. The increase after the front is a lot weaker with some stations even still showing a weak decrease, which can for example be caused by small-scale processes like convection. Nevertheless, coinciding with the location of the front, several stations observe a p_{msl} increase of around $2\,\mathrm{hPa\,h^{-1}}$. This region also coincides with heavy rain with values of around $5\,\mathrm{mm\,h^{-1}}$ to more than $8\,\mathrm{mm\,h^{-1}}$ (Figure 5.3b). Slightly lower amounts were observed along the occluded front and northern part of the warm front. Furthermore, we note a change in $\tilde{\theta}$ (Figure 5.3c) with an abrupt decrease in the frontal region and a shift of d from south-westerly to westerly winds (not shown). While $\Delta\tilde{\theta}$ indicates large tendencies over northern France and Western Germany, it shows noisier behaviour further away from the highest winds. Following this, we set labels for a WJ in the region of negative Δp_{msl}, positive $\Delta\tilde{\theta}$ and ahead of the high values of *RR*. CFC is labelled in the frontal region, where the heavy precipitation occurred. As the occluded front is not wrapped around the core yet (see Figure 5.4), implying that a CJ is not yet occurring at this point in time, the region behind the cold front is entirely labelled as CS. All set labels are displayed in Figure 5.3d.

The forecasts of feature probabilities by the RF, which is trained on the 11 other cases (Table 5.1), for 06 UTC and other time steps in a 3 h interval are shown in Figure 5.5 after Kriging was applied to generate a gridded field of probabilities. Moreover, $\tilde{v}$ and p_{msl} contours are added for orientation. An animation for the entire lifetime of *Burglind* is provided at Eisenstein et al. (2022b).

Comparing Figure 5.3d (subjectively identified features) and Figure 5.5b (RF predictions) shows the features mostly in consistent areas with high confidence. CFC is identified in a smaller region, which is partly due to missing precipitation observations, since this is the most important variable to predict CFC as will be discussed in Section 5.4.2.

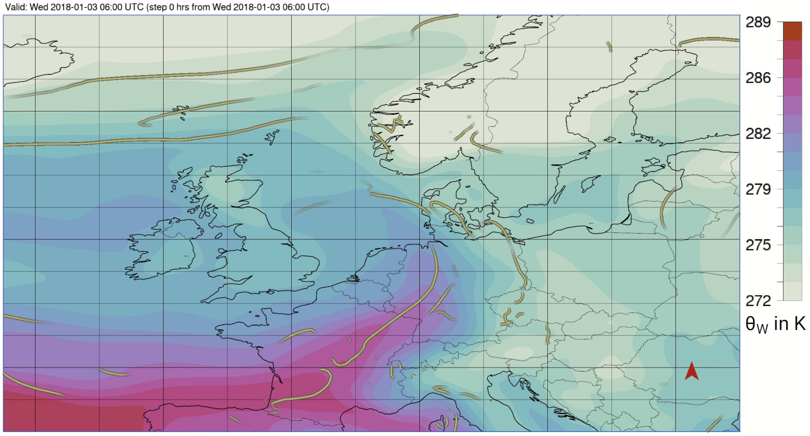


Figure 5.4: Detected fronts at 850 hPa (green tubes) and θ_W (shading) for storm *Burglind* on 03 January 2018, 06 UTC using the visualisation software Met.3D (Beckert et al., 2023; Rautenhaus et al., 2015). Data source: ERA5.

As stated, the WJ is the first feature to occur during the lifecycle of a cyclone, as it is the case for *Burglind*. Figure 5.5a shows 03 January 2018, 03 UTC where high probabilities of a WJ are predicted for most of Central France. This region is followed by a smaller region of CFC along the front and CS behind it. As the cyclone evolves further, the cyclone centre and the identified features coherently move further east, while the area affected by the WJ diminishes (Figure 5.5b). 3 h later, the WJ dissolves north of the Alps, while still being followed by a line of identified CFC (Figure 5.5c). At this time step, which is also the time of minimum core pressure, the CJ is identified at the coasts of Belgium and the Netherlands, while the CS is detected further away from the cyclone centre. Highest $\tilde{v}$ are observed along the CFC and remainder of the WJ. The WJ and CFC vanish completely until 12 UTC, when the CJ and also the region of high $\tilde{v}$ extends to western Germany (Figure 5.5d) and moves further east following the cyclone centre during the next 3 h (Figure 5.5e). At 18 UTC, when the storm and $\tilde{v}$ weaken, the CJ starts to diminish and the probabilities of both CJ and CS decrease as well.

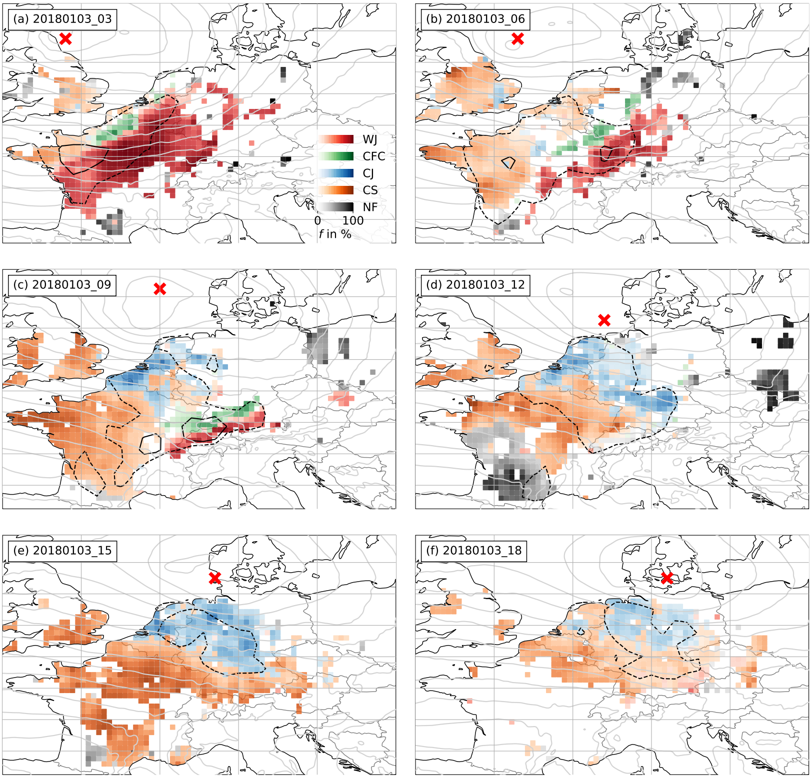


Figure 5.5: RF probabilities f for each high-wind feature after Kriging for Storm *Burglind* on 03 January 2018: (a) 03 UTC, (b) 06 UTC, (c) 09 UTC, (d) 12 UTC, (e) 15 UTC and (f) 18 UTC. The dashed contour shows $\tilde{v} = 1$, the solid contour $\tilde{v} = 1.2$). The red x indicates the cyclone centre. Light grey contours show COSMO-REA6 p_{msl} in a 4 hPa-interval.

Even though the RF is only dependent on few meteorological parameters and their development over the last hour, looking at all time steps together, the features are largely coherent, both spatially and temporally, and behave as described in the literature (see Section 2.2). While the WJ and CFC appear in earlier stages of the lifecycle (Figure 5.5a-c), they disappear in later stages, when CJ and CS dominate as the cause of high wind speeds. NF is mostly

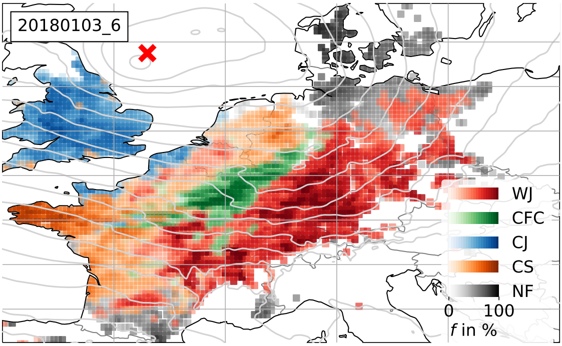


Figure 5.6: RF-derived probabilities f for each feature for *Burglind*, 03 January 2018, 06 UTC as Figure 5.5b but for COSMO-REA6.

detected at the peripheries of a cyclone or not connected to one at all during all time steps, indicating that most of the high wind speeds in the vicinity of the cyclone are in fact associated with the introduced wind features and identified accurately.

5.3.3 Comparison to gridded data

One interesting application of our new algorithm is the comparison of gridded data sets with station observations. To do this successfully, we need to ascertain that such data sets can be fairly compared.

To provide a visual impression of the differences between station- and reanalysis-based results, Figure 5.6 shows the example of *Burglind* again, such that it can be directly compared to Figure 5.5b. Results are mostly consistent with even higher probabilities most of the time in the reanalysis. Note that COSMO-REA6 provides complete information on a dense regular grid in contrast to the irregularly distributed stations that have to be interpolated. This leads to more coherent areas here. Particularly CFC covers a larger region with higher probabilities. This is due to the spatially higher resolved *RR* field in COSMO-REA6, the most important parameter for the detection of the CFC feature. Over the UK, high probabilities of a CJ are predicted by the RF. This contrasts with the identified CS in station data and the subjective analysis, where a CS was labelled, since no hook-shaped structure of high winds was discernible yet. However, the occurrence of a hook-shaped structure cannot be accounted for in the spatially independent approach of RAMEFI making it difficult to distinguish these otherwise similar features. This is consistent with the more common overforecast of CJ in the COSMO reanalysis than in observations as will be discussed in Section 5.4.1.

5.4 Statistical evaluation

In Section 4.2.3, we described how we evaluate probability predictions for the wind features. Here, we first apply this concept to the RF probabilities for the station data and the COSMO reanalysis. At the end of the section, we investigate the relationship between the predictors and the RF probabilities.

5.4.1 Evaluation of the RF probabilities

The evaluation of the station-based RF probabilities is split into three parts. First, we quantitatively compare the RF forecasts with the class frequencies in the training data, then we assess how well the RFs predict the individual wind features in the one-against-all approach, and finally we check how well the predictions distinguish two features with the all-pairs approach. For each storm that we predict, the class frequencies of the other 11 storms are used as a benchmark prediction. As expected, we find that the RF probabilities outperform the benchmark in terms of the BS for the prediction of each winter storm. The overall improvement is 24.7 %, while for the different storms it ranges from 11.8 % to 34.7 % with 11.8 % being the skill for *Xavier*, which is discussed in some detail in Section 5.5.4.

Figure 5.7 shows the reliability diagrams of the RF probabilities in the one-against-all approach for the occurrence of NF and the four specific wind features (WJ, CFC, CJ and CS). We observe that the probabilities are in

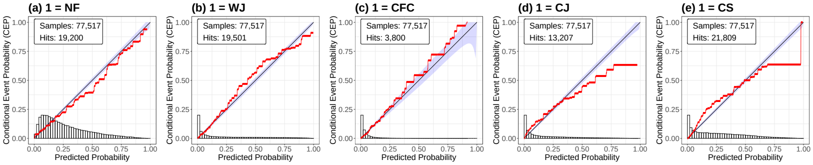


Figure 5.7: CORP reliability diagrams of the RF probabilities for the individual wind features in the one-against-all approach including all 12 storms.

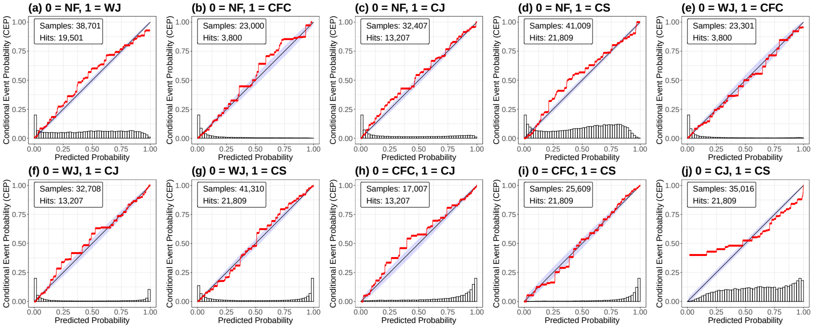


Figure 5.8: CORP reliability diagrams of the conditional RF probabilities comparing two wind features in the all-pairs approach including all 12 storms.

general well-calibrated for all five cases, as the calibration curves closely follow the diagonal. The predictions are generally reliable, especially for small probabilities, which are most frequent in this setting, as the peaks of the histograms illustrate. Therefore, the RFs identify the non-occurrence of a specific wind feature with high confidence (Figure 5.7a). For larger probabilities, the predictions of NF, the WJ and the CFC are well-calibrated, as the calibration curves stay reasonably close to the diagonal (Figure 5.7a–c), while for the CJ and CS (Figure 5.7d,e) larger deviations are evident. In both cases, the RF overforecasts the events, that is, the predicted probability is generally too large.

The reliability diagrams of the all-pairs approach are displayed in Figure 5.8, which show that the RFs yield well-calibrated probabilities for the distinction of all feature pairs but one. When the RF predicts that the CJ is more likely to occur than the CS (in case one of those two materialises), the RFs overforecast the CJ, meaning that the CS occurs more often than predicted (Figure 5.8j). This is consistent with the results from the one-against-all approach, where we found that the CJ and CS predictions were not well-calibrated for high probabilities, indicating that the RF fails to distinguish them for large conditional probabilities of the CJ. Further, the histogram of this pairwise comparison shows that the RF cannot discriminate between the two features with high confidence. This issue can be seen best for the storms *Herwart* and *Sabine*, which both did not develop a CJ, however a CJ was identified by the RFs (Section 5.5.2). The main meteorological reason for this problem is the general similarity of the two features and that the hook-shaped structure, which is used for the subjective identification of a CJ, cannot be considered in the RF, such that the distinction is mainly based on p_{msl}, as will be discussed in Section 5.4.2. Other than that, the calibration curves of the other pairs follow closely the diagonal. Moreover, we note that the WJ is distinguished well from the CJ and CS, as the U-shaped histograms of the probability distributions show (Figure 5.8f,g).

For first test runs with fewer cases used for training (Section 5.2.2; Appendix A.2), we also included the SJ and post-CFC as separate features. However, due to the small number of occurrence of post-CFC of just over 100 data points, a training was not feasible and post-CFC was included in CS. While the SJ included more data points, its characteristics in surface parameters are too similar to the CJ. The results showed high uncertainty between the SJ and CJ. While the SJ region often showed higher probabilities for the CJ, cases with no SJs showed probabilities of SJ occurrence in the CJ region at times. Figure A.1a,b in Appendix A.2 shows reliability diagrams for the SJ and SJ vs. CJ. Further discussion can be found in Appendix A.2. Following this, the SJ was merged with the CJ for following trainings.

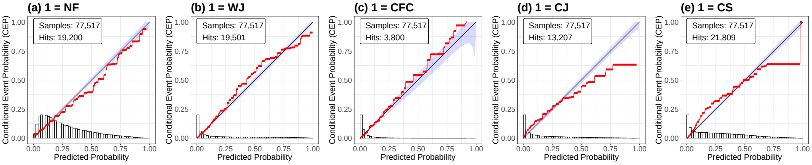


Figure 5.9: As Figure 5.7 but based on COSMO-REA6 data.

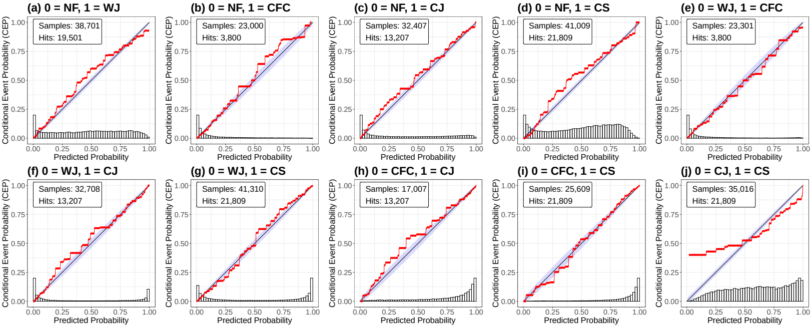


Figure 5.10: As Figure 5.8 but based on COSMO-REA6 data.

Using the original set of case studies, the reliability diagrams showed on average slightly worse performance than the final version discussed above (not shown). Especially the distinction between CJ and CS was more difficult with *Sabine* and *Herwart* carrying much weight. Replacing these two cases resulted in lower predictability for CFC, such that we decided to include them while also keeping the new cases. Furthermore, *Diana II* was replaced as it only affected a very small area, while the rest of Europe was affected by *Diana I* and later *Diana III* and an inclusion was not advantageously after we decided to merge the SJ with the CJ feature.

For the predictions derived from the COSMO-REA6 data, the RF probabilities are also well able to distinguish the features, although the RFs used were trained on station-based data. The predictions exhibit similar characteristics and perform, as expected, only slightly worse than for the station data. As before, the skill of the BS is calculated with respect to a benchmark prediction based on the class frequencies and is 19.6 % for all storms. For eight of the selected storms, we observe improvements ranging from 11.0 % to 37.5 %, however for *Herwart* and *Susanna* the skill scores are $-0.8\,\%$ and $-11.6\,\%$, respectively, indicating a decrease in predictive performance. For *Susanna*, this is due to a larger high-wind region ahead of but not directly connected to the cyclone for multiple time steps. While the predictions for *Herwart* look consistent in both data sets at first sight, fewer stations over Poland, where the CJ was overforcast (Section 5.5.2), are available, such that the overforcast in the gridded data carries more weight compared to the station data.

Further, we find at times high probabilities of mostly WJ in COSMO-REA6 in regions where winter storms are uncommon and where no features were labelled at all, as e. g., Italy and the Balkans (not shown). However, on the synoptic scale, the trough still affects some parameters in the region, that is, decreasing Δp_{msl} on the eastern side of the core and d as the wind follows the isobars (not shown), which are the most important parameters to distinguish NF and WJ (Figure 5.12). Therefore, high winds caused by, e. g., mountainous effects, such as the foehn effect or land-sea breeze, might be falsely identified as a WJ. Thus, the RF should only be applied to regions affected by extratropical cyclones. As these regions have not been labelled (Section 5.2.1), we excluded them from our evaluation.

The reliability diagrams of the one-against-all approach for the COSMO-REA6 data (see Figure 5.9) show that the calibration curves deviate more from the diagonal than for the station-based data (Figure 5.7) but are still reasonably close to calibrated. For the WJ and the CJ, we observe slight overforecasting (Figure 5.9b,d), whereas we observe underforecasting for the CFC (Figure 5.9c). For the CS, we observe a similar calibration curve to the station-based data (Figure 5.9e). The distinction of the individual features, which we assess via the all-pairs approach in Figure 5.10, results in mostly well-calibrated probabilities. The largest deviations from calibration are observed again for the distinction of the CJ and the CS, as discussed above, and for the distinction of the WJ and the CFC (Figure 5.10e), where the WJ is identified more frequently than observed. This might be caused by a spatially extended area of precipitation further into the warm sector at times due to missing data assimilation for that parameter (Section 4.1.2). Overall, the predictions based on the COSMO-REA6 data are satisfactory considering that the RF models were trained on data from the station observations.

5.4.2 Predictor importance

To identify the predictors most relevant for the prediction of the wind features and the discrimination between two features, we calculate the BS permutation importance for the one-against-all and all-pairs approach. The BS permutation importance in the one-against-all approach is displayed in Figure 5.11. In general, Δp_{msl} is the most important predictor variable, especially for the WJ. Only for CFC it is not an important predictor, as it can occur slightly ahead of the cold frontal pressure trough, hence in a region of positive Δp_{msl}, as described in Section 2.2.4. On the other hand, the absolute p_{msl} values seem to be of less importance for WJ and NF, which occur further away from the cyclone centre than CJ and CS, for which p_{msl} indicates the proximity to the cyclone centre. For CFC, we find instead that RR is the most relevant predictor variable as expected, while being less important for WJ, CJ and CS. d seems to be relevant for most features, as it is a characteristic for the location relative to the cyclone centre. This also leads to a high importance for NF occurring more frequently north or west from the cyclone centre. However, d is not important for CFC, probably as convection leads to a more variable wind direction and due to the characteristic jump in d at cold fronts. To the contrary, Δd is of minor relevance for all features as well as $\Delta\tilde{\theta}$. A more important temperature-based predictor seems to be $\tilde{\theta}$, although again being less relevant for CFC. Lastly, $\tilde{v}$ shows its highest importance for NF, as higher wind speeds are less likely to be found at the boundary of a cyclone.

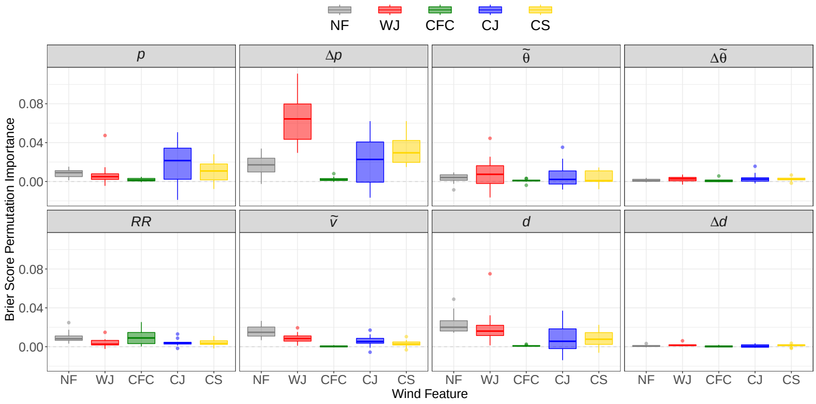


Figure 5.11: Boxplots of the BS permutation importance of the RF probabilities for the individual wind features and predictor variables in the one-against-all approach. The boxplots are calculated over the individual winter storms. Note that p corresponds to p_{msl} here.

In the all-pairs approach (Figure 5.12), we can attribute the importance of the predictor variables more accurately. The key to distinguish the WJ from all other features is Δp_{msl}, especially from the CJ and CS. This is consistent with the one-against-all discussion above. The large outlier in Δp_{msl} in WJ vs. CJ is related to storm *Herwart* as further discussed in the following section. Of secondary importance is d, particularly when compared to CJ, CS and NF. Temperature also plays some smaller role in the distinction of the WJ. For CFC the by far most important predictor is RR but when compared against the CJ p_{msl}, Δp_{msl}, $\tilde{\theta}$, $\Delta\tilde{\theta}$ and d also contribute. The positive outlier in RR is related to storm *Fabienne* (not shown). The distinction of the CJ to other features is more complex. p_{msl} is relevant in all CJ-pairs, as already discussed. The distinction of CJ from NF additionally hinges upon Δp_{msl}, $\tilde{\theta}$ and d.

The shortcomings of the RFs to distinguish CS and CJ is also reflected in Figure 5.12 by partly negative values for p_{msl}, Δp_{msl} and $\tilde{\theta}$. A negative value indicates that the RF probabilities perform better, when we break the link to the target variable by randomly permuting the predictor values. As discussed

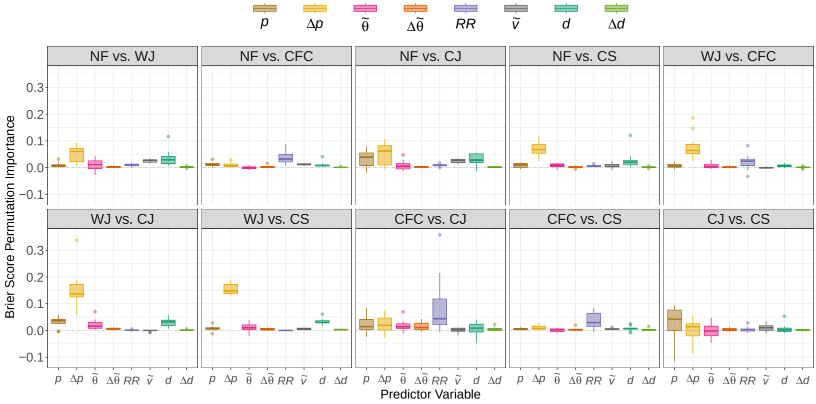


Figure 5.12: Boxplots of the BS permutation importance of the RF probabilities comparing two wind features for the predictor variables in the all-pairs approach. The boxplots are calculated over the individual winter storms. Note that p corresponds to p_{msl} here.

further in the following section, this is mostly due to storm *Sabine*, which reached an unusually low minimum core pressure of less than 950 hPa over the Norwegian Sea (see Figure 5.1). Because of this, p_{msl} values in the CS over continental Europe were similar to values typical of a CJ.

Early tests included further predictors, namely $v_{10\mathrm{m}}$, latitude, longitude and altitude. As $v_{10\mathrm{m}}$ is our target variable and strongly dependent on the location of the station, we decided to exclude it as a predictor. While the location-specific predictors longitude and latitude showed a certain predictor importance (as for example visible in Figure A.1c in Appendix A.2), it was merely an indication of the occurrence relative to the cyclone life cycle and the cyclone centre and would affect the identification of features for storms with an unusual cyclone track. On the other hand, altitude showed no substantial predictor importance. For the sake of flexibility, we removed all location-specific predictors.

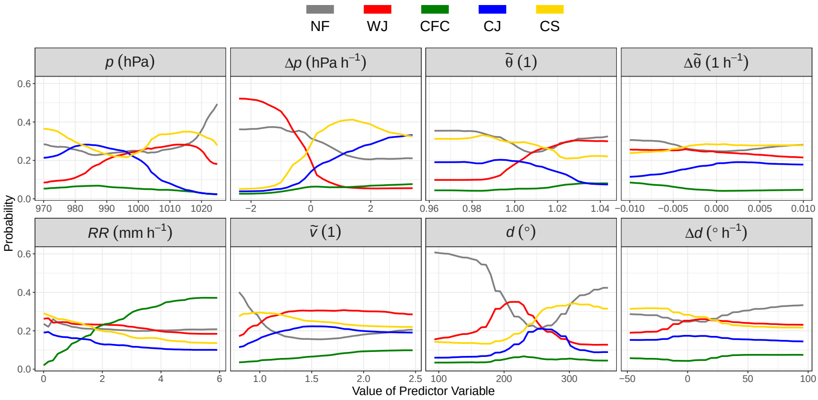


Figure 5.13: Partial dependence plots for the predictor variables and wind features. Note that p corresponds to p_{msl} here.

We do not only want to identify the most relevant predictors, but also investigate their effect on the predictions, which is illustrated for the eight predictor variables by the PDPs in Figure 5.13. Again, the largest impact is found for Δp_{msl}. The probability of observing a WJ is largest for small values of Δp_{msl} and declines rapidly as the tendency increases and switches signs, while the probabilities of the CS and CJ increase. Probabilities for NF decrease slightly, while changes for CFC are small. For little *RR* the probability of a CFC is close to zero, but consistently increases with increasing precipitation. In turn, probabilities for other features slightly decrease with increasing precipitation. In general, CJ and CS show high probabilities for low p_{msl} values consistent with their occurrence during the most intense stage of a cyclone (Section 2.2). However, surprisingly CS shows higher probabilities than CJ between 970 hPa to 980 hPa, although the CJ is usually closer to the cyclone centre. This is again associated with the unusual behaviour of storm *Sabine* with its deep pressure minimum but no subjectively identified CJ. As such intense cyclones are rare, we are confident that the RF performs well in most more ordinary cases. As discussed previously, d is dependent on the location relative to the cyclone centre. As the introduced features are all located south to west of the cyclone, we focus on values from 90° to 360° only. Within the WJ d values mostly show south-westerly winds and do not change drastically. Probabilities for CFC increase with a positive wind shift, leading to more westerly and north-westerly winds for CFC but also following features, i. e., CJ and CS. Δd shows almost no change in probabilities for all features consistent with its low BS permutation importance. $\tilde{\theta}$ shows an increasing trend for the WJ, while the probabilities decrease for the other features, most strongly for the CJ, as one would expect. For $\Delta\tilde{\theta}$ we see indications of the air mass change at the cold front and thus higher probabilities in CFC for negative values. The CJ shows a slightly positive trend, while all the others are flat.

Overall, investigating the importance of the predictor variables on the predictions, we find that the RFs largely learn physically consistent relations as described in Section 2.2 and Section 5.2.1.

5.5 Discussion

Section 5.3 showed a successful application of the introduced method to storm *Burglind* both in station- and grid-based data sets. However, extratropical cyclones rarely follow the textbook Norwegian cyclone model exactly. For example, sometimes high winds are mainly related to an exceptionally strong synoptic-scale pressure gradient rather than associated with the mesoscale features that we have developed an objective identification algorithm for. As described in Section 5.5, we selected 12 wind storm case studies to train the RF based on surface observations with the diversity of cyclones and features in mind, such that the RF is representative for a climatology over a longer time period.

This section discusses how the trained RF deals with some well-known deviations from idealised cyclone models such as double fronts and convergence lines (Section 5.5.1), large background pressure gradients (Section 5.5.2), as well as the specific characteristics of Shapiro-Keyser cyclones and SJ (Section 5.5.3). In addition, we further discuss advantages and disadvantages of not using spatial dependencies in the feature identification (Section 5.5.4). A complete set of results for all case studies can be accessed at Eisenstein et al. (2022b).

5.5.1 Double fronts and convergence lines

Real-world frontal structures of extratropical cyclones can differ considerably from idealised conceptual models, e. g., in terms of complex vertical structure, strong tilts, or a secondary frontal zone parallel to the main front. Here, we are interested in synoptic systems with double cold fronts and convergence lines with high winds. While a cold front is associated with a second low pressure trough, a convergence line develops where two airflows collide and can occur independently of a cyclone. The area between a primary and secondary cold front can have characteristics of a warm and/or cold sector, and thus high-wind features are predicted with higher uncertainty by the RF. This can also be the case for the area between a cold front and a convergence line.

The example from our selection of 12 cyclones that illustrates this best is *Bennet*. On 04 March 2019 at 12 UTC the cyclone centre was located over the North Sea to the west of Denmark (Figure 5.14a). The primary front is located at the north-eastern border of Germany and Bavaria and has already weakened (see Figure 5.15). A secondary strong temperature gradient can be found over northwestern Germany, Luxembourg and France. However, it is uncertain, if this should be classified as a front or convergence line. While synoptic charts from DWD show a convergence line, UK Met Office charts show an upper-level cold front at 06 UTC and an occluded front 6 h later. As this feature shows characteristics typical of CFC (see Section 2.2.4), it was ultimately labelled as such by the first author and is referred to as a secondary front.

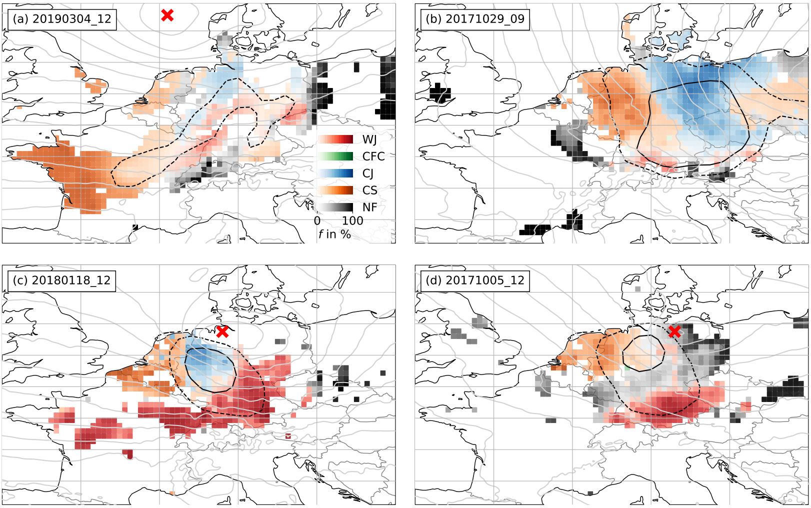


Figure 5.14: As Figure 5.5 but for (a) storm *Bennet*, 04 March 2019, 12 UTC; (b) storm *Herwart*, 29 October 2017, 09 UTC; (c) storm *Friederike*, 18 January 2018, 12 UTC; and (d) storm *Xavier*, 05 October 2017, 12 UTC. Note that at the time shown, storm *Herwart* had already exited the plot area to the east.

Figure 5.14a shows the mesoscale wind features identified by the RF. At this time *Bennet* causes its highest winds slightly ahead of the secondary front and a smaller region behind the primary front (see dashed line in Figure 5.14a). For the area between the fronts, the RF predicts both CS and WJ with medium confidence. This is to be expected, as some parameters, especially Δp_{msl}, $\tilde{\theta}$ and $\Delta\tilde{\theta}$, show behaviour of both features with a tendency towards CS. RF predictions along the secondary front show only low probabilities for CJ and CFC. However, at earlier and later time steps, that is, ahead of the primary cold front and behind the secondary one, the prediction of the WJ, CJ and CS are accurate, as can be seen in the full animation provided at (Eisenstein et al., 2022b). So looking at the entire lifetime of the storm, satisfactory identification can be obtained from RAMEFI.

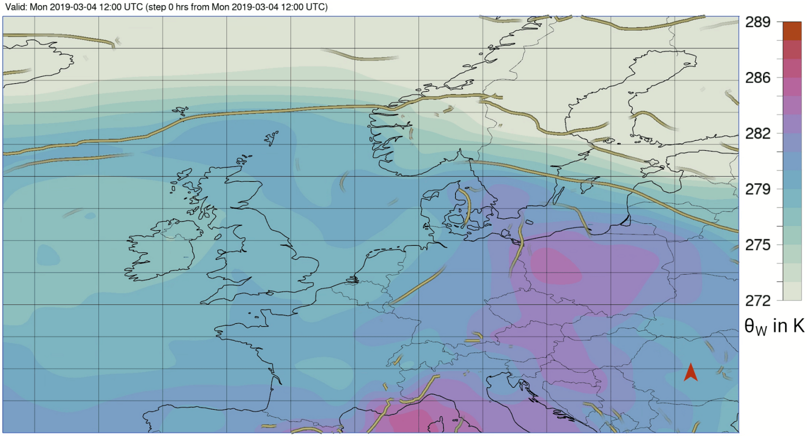


Figure 5.15: As Figure 5.4 but for storm *Bennet* on 04 March 2019, 12 UTC.

5.5.2 Strong background pressure gradients

Very intense cyclones are often accompanied by a strong large-scale pressure gradient (e. g., Fink et al., 2009), which in turn causes high wind speeds unconnected to one of the four mesoscale wind features under study but can be enhanced by them. With an underlying strong wind field, the detection and distinction of the features might be more complicated. Good examples from our list of case studies to illustrate this are the storms *Herwart* and *Sabine*. Figure 5.14b shows *Herwart* in a late development stage on 29 October 2017 at 09 UTC when the pressure gradient is already weakening (see light grey contours). Around this time, the cyclone centre travelled over Poland (outside the area shown in the figure). An occurrence of a CJ seems unlikely in that region, since a typical hook-shape structure cannot be seen in wind observations (not shown). The occluded front was rather weak and did not fully wrap around the cyclone centre (see Figure 5.16). So ultimately, the already high wind speeds caused by a strong background pressure gradient over Germany (black lines in Figure 5.14b) make the subjective labelling of additionally occurring features quite challenging. The RF shows high

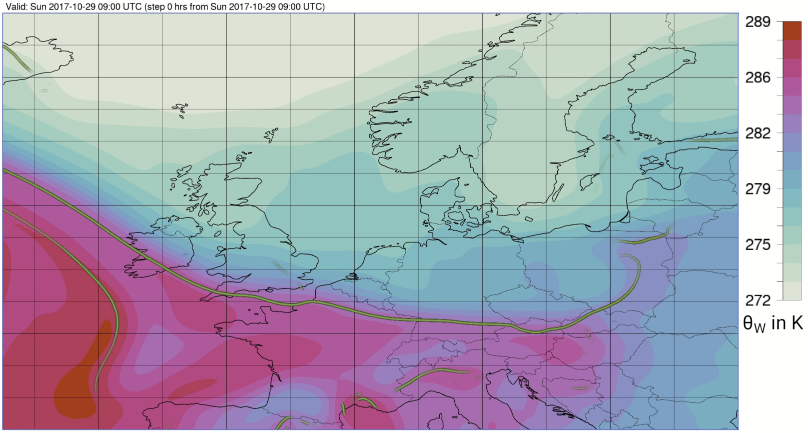


Figure 5.16: As Figure 5.4 but for storm *Herwart* on 29 October 2017, 09 UTC.

probabilities of a CJ for several hours in this region, although it was originally labelled to be CS. The main reason for this is that the proximity to the cyclone centre, reflected in low p_{msl}, is the most important predictor to distinguish the CJ and CS (Figure 5.12). Nevertheless and even in this unusual case, the prediction by the RF is still reasonable, since both a subjective and objective identification of the two features here is ambiguous in surface observations alone.

In the case of *Sabine* the cyclone centre did not cross continental Europe but moved through the Norwegian Sea (Figure 5.1). The minimum pressure reached less than 950 hPa. Stations over Central Europe still observed p_{msl} below 970 hPa, which is lower than the cyclone centres of most of our other case studies, making *Sabine* a quite unusual case. As discussed already in Section 5.4, this causes difficulties to distinguish the CJ and CS of *Sabine*, somewhat similar to *Herwart*. Although a CJ is identified in an area of low p_{msl} values, this region is not in the vicinity of the cyclone centre, as it was the case for *Herwart*. An animation of the feature identification for all time steps of *Sabine* is provided at (Eisenstein et al., 2022b). In this case the CJ-CS distinction issue could have been avoided to some extent by including spatial dependencies in the identification algorithm at the expense of losing the capability for flexible application as discussed in Section 5.2. We will return to this issue in Section 5.5.4.

5.5.3 Shapiro-Keyser cyclones and sting jets

As already explained in Section 2.2, cold fronts of Shapiro-Keyser cyclones are usually weaker than those of Norwegian cyclones, such that CFC wind features hardly occur. A good example for this is storm *Friederike*. Figure 5.14c shows 12 UTC on 18 January 2018, when the cyclone centre just reached the North Sea coast of Germany. High probabilities of a WJ occur over Central Germany, while a CJ is identified by the RF over northeastern Germany. CFC probabilities are very low along the entire cold front. Lagrangian trajectories confirmed an SJ in the region where the CJ is identified during that time. As the CJ and SJ are considered together in this work (Section 5.2.1), the area was labelled as CJ and, hence, identified accurately. This area also coincides with the highest wind speeds associated with *Friederike* (see black lines in Figure 5.14c). West of the identified CJ region, a CS feature is detected. This shows that CJ and SJ indeed show similar characteristics in surface parameters and can be considered as one feature in this context.

5.5.4 Spatial independence

The decision not to use spatial (nor temporal beyond 1 h) dependencies in the identification algorithm makes our method highly flexible in its application, but the local approach can also cause issues where features deviate from their stereotypical characteristics. One example for the problem is the CJ of *Sabine* as discussed in Section 5.5.2. Another example is storm *Xavier*, where for several hours many points within the vicinity of the cyclone show the highest probability for NF rather than for any of the mesoscale wind features (Figure 5.14d). The main reason for this appears to be that *Xavier* was characterised by unusually cool $\tilde{\theta}$ and high p_{msl} (not shown), generally two of the most important parameters to distinguish features (Figure 5.11). While one predictor behaving in an unusual way could be compensated, as e. g., in the case of *Fabienne* (see Eisenstein et al., 2022b), two anomalous behaviours unsurprisingly result in considerably greater uncertainty.

A possible solution to the issues described here on the basis of *Sabine*, *Xavier* and *Fabienne* is to not only regard anomalies from diurnal and seasonal cycles but also to include some kind of spatial background, e. g., by normalising p_{msl} by the core pressure to detect the region close to the cyclone centre, or comparing $\tilde{\theta}$ to the mean state over Europe during the period of the storm to detect the warm sector. However, such a step would bring its own set of problems. Any spatial mean would require an arbitrary decision about the considered area, which may vary greatly from cyclone to cyclone. Moreover, spatial means computed from surface observations are not representative due to the irregular spacing of the stations. Essentially, as the features identified by the RF still occur in the expected areas as described in Section 2.2, we conclude that a flexible local approach offers more advantages than disadvantages overall.

5.6 RAMEFI as a near-real time product

Due to its flexibility, RAMEFI can also be applied to forecast data using it as a near-real time tool. For now, this is done using the deterministic ICON-EU forecast, but it is planned to expand the analysis to ensemble data (ICON-EU-EPS) as well. However, for the normalisation of $v_{10\mathrm{m}}$ and $\theta_{2\mathrm{m}}$, not enough ICON-EU data has been assembled yet, such that the climatology of COSMO-REA6 is used, which covers the same area with a comparable resolution. As discussed in Section 4.1.2, the data are regridded to the same resolution as the ICON data for this work. The usage of a reanalysis to normalise forecast data might cause some discrepancies, as systematic biases

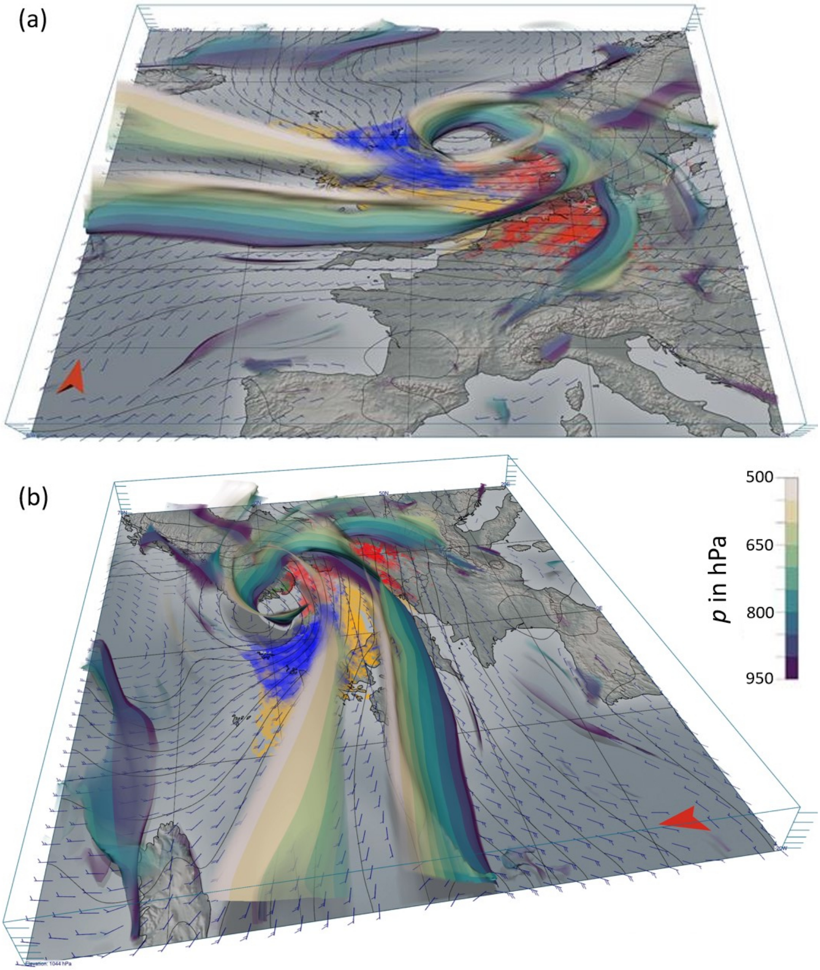


Figure 5.17: RAMEFI probabilities and 3D fronts of storm Ulf on 17 February 2023, 11 UTC. Fronts are shaded by p, RAMEFI probabilities are shaded in red for WJ, green for CFC, blue for CJ and orange for CS. The red arrow indicates north. Image produced using Met.3D.

such as overestimation of wind speed in forecasts (e. g., Hess, 2020; Schulz and Lerch, 2022) may differ between the data sets. Nevertheless, these differences are mostly negligible for the probabilistic identification, as the predictor dependency of $\tilde{v}$ and $\tilde{\theta}$ is rather low compared to Δp_{msl}, p_{msl} and RR anyway.

Furthermore, RAMEFI probabilities can be combined with the three-dimensional front detection developed within Met.3D (Beckert et al., 2023, see Section 4.2.2) to visualise the essential characteristics and structure of an extratropical cyclone. An example is shown in Figure 5.17 for storm Ulf (17 February 2023), which developed as a Norwegian cyclone, from different angles to illustrate the benefit of an interactive three-dimensional visualisation of a cyclone. At 11 UTC, Ulf is located over the North Sea. An occluded front is visible, indicating that the cyclone is in stage IV of a Norwegian cyclone life cycle (Section 2.1.2; Figure 2.1a). As expected, while the WJ is located mostly in the warm sector between the two fronts (Figure 5.17a), the CS and CJ are within the cold sector. The three-dimensional fronts reveal a secondary upper-level cold front, which seems to coincide with the area of transition between CJ and CS (Figure 5.17b). The cold front is tilted backwards, suggesting CFC to be located slightly behind the surface cold front. While this is not visible here, an interactive analysis with changing angles and the ability to turn fronts off and on enables a more detailed look at the storm. Note that NF is not displayed in this case, as the interest is on the introduced features.

5.7 Conclusions

High wind and gust speeds can be caused by distinct mesoscale features within extratropical cyclones, which occur during different stages of the cyclone lifecycle, in varying regions relative to the cyclone centre and have distinctive meteorological characteristics (e. g., Hewson and Neu, 2015). These differences likely imply differences in hazardousness, forecast errors and, hence, risk to life and property.

To better understand, monitor and predict these mesoscale features, we developed RAMEFI, a first-ever objective identification method that is able to reliably distinguish the four most important features, that is, the WJ, the CJ, CFC and CS. The rare and often short-lived SJ is included in the CJ category, as their surface characteristics are often rather similar and 3D trajectories are required for a clean distinction (Gray et al., 2021).

The first step was to build a browser-based, interactive tool to subjective label surface stations over Europe for 12 selected winter storm cases between 2015 and 2020. Based on the outcome, we trained a probabilistic RF based on the eight predictors $\tilde{v}$, p_{msl}, Δp_{msl}, $\tilde{\theta}$, $\Delta\tilde{\theta}$, RR, d and Δd. We note that we set a $\tilde{v}$ threshold of 0.8 to focus on high wind areas. However, we do not expect the RF to be sensitive to small changes in the threshold and, in principle, the RF can be applied to wind speeds below this. Being independent of spatial behaviour or gradients, the approach is very flexible and can be applied to single stations or grid points and various data sets with differing grid spacing. However, due to the fast movement of meteorological features in stormy situations hourly resolution is required, making the algorithm inapplicable to some climate data sets. To obtain areal information from irregular station data, Kriging was applied on the station-based probabilities generated by the RF.

The trained RFs are generally well-calibrated. Merely, the distinction between CJ and CS is more challenging, since the two features show similar characteristics in most parameters except for lower p_{msl} in the CJ being located nearer to the cyclone centre. Overall, the RFs learn physically consistent relations reflected in the importance of individual predictors. For example, while Δp_{msl} appears to be most important for WJ, CJ and CS, RR is substantial for the identification of CFC.

A detailed analysis of the RF feature probabilities for the selected cases shows a high consistency with the subjectively set labels with only few disagreements, mostly in cases of large deviations from standard cyclone models. While the identification of WJs has the highest confidence, the identification of CFC is least certain due to relatively few surface stations reporting hourly precipitation and thus less training data. Even the distinction between the relatively similar CS and CJ works well in most cases and time steps. In some cases, however, high probabilities of CJs are predicted by the RF in areas where no CJ was identified subjectively due to a missing hook-shaped structure and occlusion front, or too large distance from the cyclone centre (e. g., *Herwart*, *Sabine*). Despite the spatial independence of the method, putting the predicted probabilities together on a horizontal map and following the storm evolution in time shows a high degree of coherence for each feature, demonstrating the success of our method.

The station-based RFs are also applied to COSMO reanalysis data without any adaptations to the new data set. Nevertheless, the obtained results are mostly consistent and only slightly less calibrated. This demonstrates that the method could be readily applicable to other analysis and forecast datasets. Although applying RAMEFI over regions other than that used in the training has not been examined yet, relying on location-independent predictors suggests that it should be possible with no or only little modification.

Now that the RAMEFI method is fully developed, it enables a number of exciting follow-on studies. The next step is to use the objective identification approach to compute a long-term climatology over Europe based on station observations and COSMO reanalysis data in Chapter 6.

6 Climatology of high-wind features over Europe

Although previous literature discussed different causes of winds within extratropical cyclones, their climatologies were based on more subjective categorisations for a limited sample size (e. g., Hewson and Neu, 2015; Earl et al., 2017). RAMEFI will for the first time allow a statistically substantiated analysis of the characteristic of the mesoscale wind features in terms of size, lifetime, position relative to the cyclone core, occurrence relative to the life cycle of the cyclone and wind characteristics.

The goal of this chapter is to expand and complement existing shorter and/or more general climatologies using RAMEFI, the first tool to objectively distinguish WJ, CFC, CJ + SJ and CS. With this aim, a high-resolution regional reanalysis data set for 19 extended European winter seasons is used. Other observational data sets are used for specific aspects or for comparison.

The chapter is structured as follows: First, we shortly describe necessary adaptations to RAMEFI and the gained output (Section 6.1). Section 6.2 focuses on the occurrence of the identified high-wind features, i. e., frequency, relative to the cyclone centre and cyclone life cycle, while Section 6.3 discusses the different characteristics of the features.

6.1 Application of RAMEFI

Here, we apply RAMEFI to station observations (Section 4.1.1) and COSMO-REA6 (Section 4.1.2) data under windy conditions during the extended winter months, regardless of whether a storm occurred or not. However, we later filter the output for cyclone occurrence as discussed in Section 4.2.4. Note that the 12 cases used for training (see Section 5.1) are also included in the data used for the climatology. The reasoning behind the cross-validation approach in Chapter 5 was to evaluate whether the RF is able to reliably identify the features in unseen data. Here, we want to generate a climatology of the high-wind features rather than testing the method. Hence, it is unproblematic to apply RAMEFI to the same data it was trained on. Instead, we obtain an identification that mirrors the subjective identification within in these storms, and is still consistent with the entire climatology due to the same model underlying the identification.

In contrast to Chapter 5, RAMEFI is also applied to ocean grid points of COSMO-REA6, where it has not been systematically evaluated such that results should be treated with some caution there. Due to varying friction, orography, heating of the surface and more, the wind distribution over the ocean has a fundamentally different form compared to land (e. g., Wieringa, 1973; Born et al., 2012), as displayed for exemplary locations in Figure 6.1. The right tail of the distribution shows considerably stronger winds leading to a higher number of time steps exceeding 80 % of the 98th percentile. Overall, the threshold is exceeded around 45 % more often over the ocean compared to land. To nevertheless allow a fair comparison, we normalise the occurrence by the number of time steps with windy conditions. Characteristics in other parameters are discussed in Section 6.3. Nevertheless, looking at various cases over the 19-year period, the ocean and land do not seem to behave fundamentally different with respect to feature detection and their probability distributions. Exemplary cases can be accessed at Eisenstein et al. (2023b).

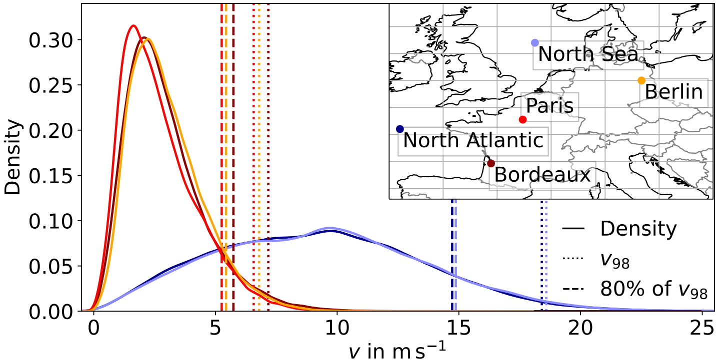


Figure 6.1: Distribution of $v_{10\mathrm{m}}$ for two ocean grid points in the North Atlantic (dark blue) and the North Sea (light blue) and three land grid points close to Bordeaux (dark red), Paris (red) and Berlin (orange), respectively. Dotted lines mark the 98th percentile, dashed lines 80 % of the 98th percentile. Note that v corresponds to $v_{10\mathrm{m}}$ here.

The obtained probabilities are filtered by cyclone occurrence, and we concentrate specifically on the area of ±15° in zonal and −15° to 5° around a cyclone centre as discussed in Section 4.2.4.

As RAMEFI provides a probabilistic identification, each feature is assigned a probability from 0 to 1. The distribution of the probabilities for each feature are shown in Figure 6.2 for both data sets. While WJ and CS show a similar distribution in both data sets with peaks around 48 %, the maximum of CJ slightly differs and is lower at around 43 % for COSMO-REA6 and 40 % for observations. The highest uncertainty in the feature detection can be seen for CFC, which shows overall lower probabilities with a peak around 33 %. The biggest difference in the data sets is found for NF. COSMO-REA6 shows a peak at 50 %, but a plateau between 50 % and 73 % in the observations. This will be discussed further in Section 6.2.1. This probabilistic information is used in two different ways: Firstly, we assign the feature with the highest probability to a given time and grid point, ignoring all other probabilities (referred to as MAXP hereafter). Secondly, we exploit

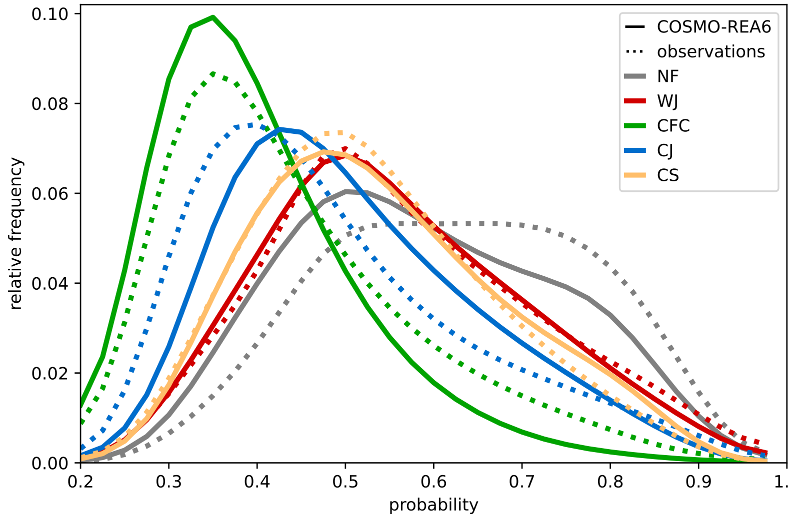


Figure 6.2: Distribution of RAMEFI probabilities for each feature over the 19 extended winter seasons using COSMO-REA6 (solid) and station observation data (dashed). The density is calculated based on smoothed histograms.

the probabilistic nature of the identification by interpreting the (accumulated) feature probabilities as the expected number of features (referred to as ACCP hereafter). The calibration of the feature probabilities, which was checked in Section 5.4.1, is a critical condition for this approach. The second approach is particularly important for features with less confident detection, which might be underrepresented in the first approach[1], e. g., the CFC as shown in Figure 6.2 (see also Section 6.2.1).

---

[1] As an example, consider an identification of 75 % for rain versus 25 % for no rain. Although we always detect rain via the first approach, rain was actually observed, on average, every fourth case (given the probabilities are calibrated).

6.2 Occurrence of high-wind features

One of the main aspects of this climatology is the occurrence of the mesoscale wind features in time and space. As mentioned in Section 2.2, the features develop during different times in a cyclone life cycle and in different areas of the cyclone. In addition to the overall frequency of the features as well as diurnal, seasonal, and yearly variations (Section 6.2.1), RAMEFI further gives us the possibility to obtain the occurrence of the wind features both in an Earth-relative (Section 6.2.2) and a system-relative framework, i. e., relative to the cyclone centre and relative to the cyclone life cycle (Section 6.2.3).

6.2.1 Relative occurrence frequency

Overall occurrence

Figure 6.3 shows the relative frequency of the wind features for both observations (first row) and COSMO-REA6 (lower three rows) over all time steps and grid points satisfying stormy time step conditions (Section 4.2.4). While the left two columns show all features, including NF, the right two columns neglect NF to focus on the identified mesoscale wind features. Furthermore, the frequencies for the most probable feature (MAXP) is displayed next to the accumulated probabilities (ACCP), i. e., the expectation. For better comparison between the observations and reanalysis, COSMO-REA6 is displayed for land grid points, ocean grid points and all grid points.

As displayed in Figure 6.3a, of the four mentioned features CS shows the highest proportion with 21.5 % as the most probable feature, followed by the WJ with just under 15 %. The CJ reaches merely 4 %, while the least common feature is CFC with under 1 %. However, NF has a proportion of almost 60 %, thus almost three times as much as the CS. This might be caused by NaN values within the data set, which are replaced by the mean values of the variable. This complicates the distinction between the features as less parameters include information about the current conditions, leading

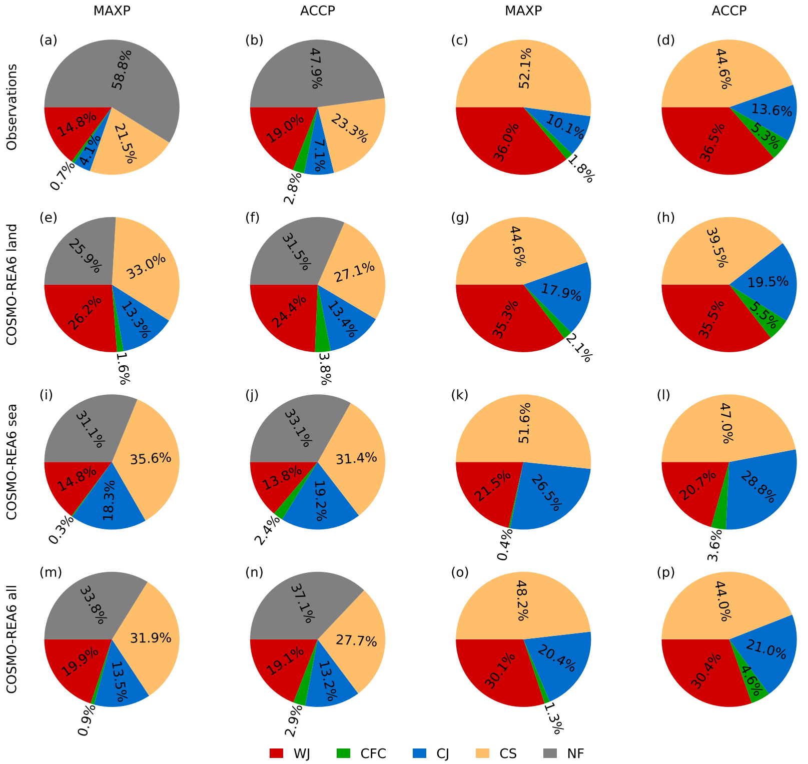


Figure 6.3: Relative frequency of features for station observations (a–d), COSMO-REA6 land grid points (e–h), COSMO-REA6 ocean grid points (i–l) and all COSMO-REA6 grid points (m–p) for the most probable feature (MAXP; first and third column) and accumulated probabilities (ACCP; second and fourth column). The first two columns include NF, while it is neglected in the latter two due to its high frequency.

to higher probabilities for NF and causing NF to be the most probable feature more often. Indeed, Figure 6.3b shows that the proportion of NF reduces by over 10 percentage points if probabilities for all features and not only the most probable one are taken into account. Although this leads to an increase of all mesoscale features, it is not to the same amount. While the features with overall higher probabilities, namely WJ and CS as shown in Figure 6.2, increase by around 10 % to 30 %, CJ shows an increase of 73 %. The CFC, which shows the highest uncertainty, increases by 300 % demonstrating the gain by using the assigned probabilities. Nevertheless, CFC is by far the least common cause for high winds.

Neglecting NF in Figure 6.3c,d draws the focus on the ratio of the mesoscale features themselves. In the MAXP perspective, CS and WJ are the cause of high winds in over 50 % and 36 %, respectively, while the more damaging features CJ and CFC (e. g., Hewson and Neu, 2015; Earl et al., 2017) only show a proportion of 10.1 % and 1.8 %, respectively. However, these are also the features with lower certainty (Figure 6.2). Hence, in the ACCP perspective, CJ and CFC come to a total of around 19 %, i. e., an increase of around 60 %. While the WJ also increases slightly, the proportion of CS decreases by almost 8 percentage points. This suggests that the CJ and CFC mostly lose against the CS to be the most probable feature.

In contrast to the station observations, the proportion of NF is considerably lower in COSMO-REA6 data as seen in Figure 6.3e,f. This supports the hypothesis that the high proportion in observations is due to NaN values, as the data are of course complete for all parameters here. As the overall certainty of NF is lower for COSMO-REA6 and closer to the density of CS and WJ (solid lines in Figure 6.2), the proportion in Figure 6.3f actually increases, while the CS and WJ proportion decreases. As the proportion of NF substantially affects also the proportions of the other features, we compare observations and gridded data without NF from here on (Figure 6.3g,h). So apart from that, the largest difference between the two data sets can be seen in the percentage of the CJ with an increase of almost 8 percentage points for land grid points in COSMO-REA6 and over 16 percentage points for ocean grid points for MAXP. This is due to the fact that the CJ commonly occurs in northern continental Europe, over the sea and in Scandinavia, i. e., regions where less station observations are available in our data set (see Section 6.2.2 and Figures 6.6e and 4.1). On the other hand, CS mostly occurs further south, i. e., farther away from the cyclone centre, such that a higher percentage in CJ leads to lower percentages for CS in the gridded data over land. The WJ, which occurs more over land than over sea (comparing Figure 6.3e–h and i–l), shows similar percentages on both observations and COSMO-REA6 over land. With similar probability distributions for the mesoscale features for both data sets (Figure 6.2) it is again apparent that CFC shows a higher percentage in ACCP. Comparing sea and land grid points (Figure 6.3e–h vs. i–l) finally shows that CJ and CS occur more often and in a wider area over the ocean compared to CFC, which almost exclusively occurs over land, where friction is higher and static stability lower during daytime. As mentioned before, WJ is more common over land with around 15 percentage points more in both MAXP and ACCP.

Finally, Figure 6.3m–p show the proportions for all grid points. Note that the number of land grid points is around 35 % higher than the number of ocean grid points. Overall, almost half of windy conditions are caused by CS, followed by the WJ and CJ with around 30 % and 21 %, respectively. Again, for CFC the difference between MAXP and ACCP shows a considerable difference from 1.3 % to almost 5 %.

As the overall frequencies are similar and differences plausible in both data sets and COSMO-REA6 has the advantage of a homogeneous field without missing parameters, we focus on the gridded data set from here on.

To examine how robust these numbers are, we computed three subsets of nine randomly chosen winter seasons. A figure analogous to Figure 6.3m–p for each set can be found in Appendix B (Figure B.1). The proportions vary just slightly with an average of around 2 %, as to be expected considering fluctuations between winter seasons (see discussion of Figure 6.5 below) and suggesting overall consensus of the feature occurrence.

Figure 6.4 shows the proportions of each feature for MAXP analogously to Figure 6.3m,o (COSMO-REA6 all) but only including grid points where $\tilde{v} \geq 1.2$ (left columns) and $\tilde{v} \geq 1.4$ (right columns). Note that this is only the case for around 1.5 % and 0.1 %, respectively, of the previously included data points. Considering only $\tilde{v} \geq 1.2$, the proportion of NF reduces by over

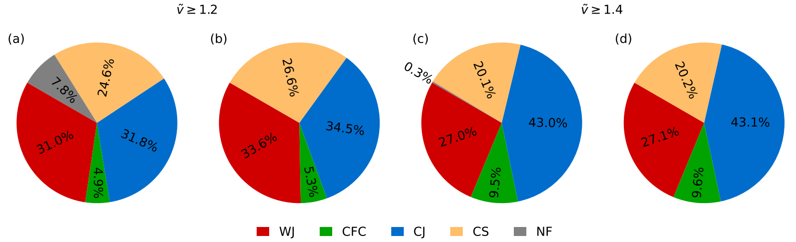


Figure 6.4: As Figure 6.3m,o (COSMO-REA6 all; MAXP) but for grid points with (a),(b) $\tilde{v} \geq 1.2$ and (c),(d) $\tilde{v} \geq 1.4$.

75 % for MAXP (Figure 6.4a), while it decreases to almost 0 % if only grid points with $\tilde{v} \geq 1.4$ are included (Figure 6.4c). This suggests that higher winds within the vicinity of a cyclone are mostly associated with one of the introduced features. When NF is neglected, the proportion of CS is reduced by 45 % to 58 % (Figure 6.4b,d) suggesting that while the CS affects a large area, it is less common for it to be the cause of extreme winds. In contrast, the more damaging features CJ and CFC (e. g., Earl et al., 2017) show an increased proportion. CJ shows an increase of about 69 % for $\tilde{v} \geq 1.2$ and 111 % for $\tilde{v} \geq 1.4$. CFC shows an even stronger increase with over 300 % and over 600 %, respectively. These results are consistent with the wind characteristics, as will be discussed in Section 6.3. Meanwhile, the proportion of WJ decreases by about 10 % as it usually causes weaker winds compared to the CJ and CFC (e. g., Hewson and Neu, 2015; Earl et al., 2017).

Temporal occurrence

The discussed proportions in Figure 6.3 depend not only on the occurrence of the feature but also on its size. Figure 6.5 shows the seasonal and interannual evolution of stormy time steps and occurrence of each feature, respectively. This is computed as the sum of stormy time steps during which a certain feature is detected as the most probable feature at at least 100 grid points over

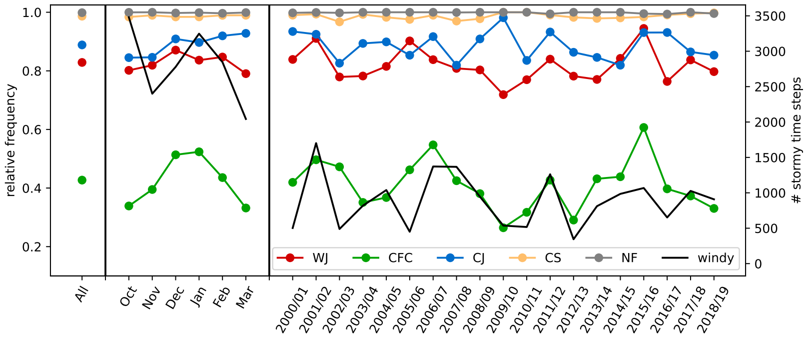


Figure 6.5: Number of stormy time steps (black line and right axis) and number of time steps with the feature occurring divided by the number of stormy time steps (left axis) for all time steps (left), each month (middle) and each winter season (right).

the investigation domain, normalised by the number of stormy time steps. This measure is thus independent of the number of grid points, where the feature is identified, i. e., the size of an affected area.

Looking at all stormy time steps of the investigation period (left side of Figure 6.5), it is evident that NF (grey dot) occurs somewhere in the domain practically for every moment in time. The same holds for CS (orange dot) with only marginally lower frequencies. Both CJ (blue) and WJ (red) occur in over 80 % of stormy time steps with slightly higher values for CJ. Together with the lower proportion of CJ compared to WJ as seen in all panels of Figure 6.3 except (i)–(l) (over the ocean) this suggests that the CJ is on average a smaller feature than the WJ. The least frequent feature with around 43 % is CFC (green) contributing to the low proportions in all panels of Figure 6.3.

With respect to the mean seasonal cycle, the black line in the middle section of Figure 6.5 shows that November and March are the least stormy months during our investigation period 2000 to 2019 with a peak in January in between. This is consistent with the cyclone track density plots shown in Figure B.2 of Appendix B. Somewhat surprisingly, October shows the highest number of stormy time steps of all months. On long-term average, the majority of storms occur between December and February, and thus the 98th percentile of wind speeds is highest for that period and lower towards autumn and spring, consistent with Feser et al. (2015); Laurila et al. (2021). The on average lower wind speeds result in a lower threshold of $\tilde{v}$. The recent two decades show a larger number of noteworthy storms in October over Central Europe – e. g., *Jeanett* (2002), *Christian* (2013), *Xavier* (2017), *Herwart* (2017) – compared to no noteworthy storms in November, leading to a larger number of stormy time steps in the 19 years. The higher frequency is consistent with October showing a slightly higher cyclone density than November (Fig, B.2a,b). With respect to the individual features, the rather rare CFC has a marked seasonal cycle with an apparent peak in December and January. The WJ has a smaller relative peak during peak winter months. While the relative frequency of CJ is only slightly higher than for WJ, its frequency increases with the winter passing leading to a maximum of over 90 % in March. Recall that the $\tilde{\theta}$ predictor is a normalised parameter, such that a cooling Arctic with progressing winter and possibly more cold-air outbreaks, i. e., lower values of $\tilde{\theta}$, might lead to higher numbers of high winds being allocated to the CJ – and CS. Both NF and CS are so frequent that an annual cycle is not evident in this analysis.

Finally, the right section of Figure 6.5 shows the interannual evolution of stormy time steps and wind features. Overall, lower numbers of stormy time steps, such as 2002/03, 2005/06, 2009/10, 2010/11 and 2012/2013 are consistent with negative values of the North Atlantic Oscillation (NAO; e. g., Wanner et al., 2001), which describes the large-scale circulation over the North Atlantic and originally represents the pressure difference between Iceland and the Azores (Hurrell, 1995). For our study area, slightly positive NAO values facilitate stormy conditions, as the typical cyclone paths for such NAO conditions corresponds to tracks over the British Isles, the North and the Baltic Sea. Accordingly, most peaks are associated with positive NAO phases. Furthermore, quieter winter seasons consistent with literature such as 2010/2011 (e. g., Santos et al., 2013; Laurila et al., 2021) can be found.

With 2009/10 being a particular cold winter season (Wang et al., 2010), the occurrence of detected WJs is lower compared to other winters, while CJ shows a peak. The peaks of WJ and CFC in 2015/16 are consistent with the winter season being particularly wet and warm as discussed in McCarthy et al. (2016). Again, NF and CS occur too frequently to detect an interannual cycle. Overall, all features have no to very weak positive correlation with the number of stormy time steps (0 % to 10 %). The coefficient of variance is lowest for CS and stormy time steps with 41 % and highest for CFC with 54 %.

With respect to long-term trends, a slight decline is evident consistent with the overall decrease in the number of winter storms in a warming climate (Priestley and Catto, 2022). However, given that our investigation period covers only 19 years, a Mann-Kendall test (significance level of 0.05; Hussain and Mahmud, 2019) did not indicate statistical significance in any of the time series and the decadal variability is dominating (Feser et al., 2015).

6.2.2 Earth-relative statistics

An Earth-relative framework enables us to learn which regions are commonly affected by which feature. Figure 6.6 shows a geographic distribution of the relative frequency of NF and the four mesoscale wind features (MAXP) as well as of overall windy conditions. An analogous plot for ACCP can be found in Appendix B showing overall similar properties (Figure B.3). Results are displayed on $0.5^\circ \times 0.5^\circ$ boxes, thus aggregating over 16 grid points.

The number of windy conditions, i. e., $\tilde{v}$ exceeding 0.8, in proximity to the cyclone centre is displayed in Figure 6.6a. This criterion leads to the highest numbers over the North Sea, Denmark, northern Germany and the Baltic Sea, i. e., south of the maximum track density shown in Figure 4.2. Furthermore, choosing a threshold of 80 % of the 98th percentile results in more windy conditions over the ocean (Figure 6.1), hence, the absolute frequencies of each feature are normalised by the number of exceedances for each grid point.

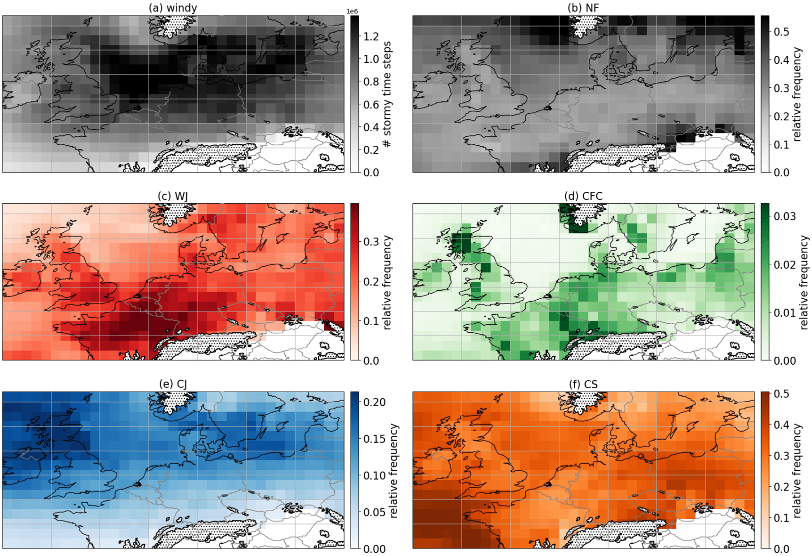


Figure 6.6: Relative frequency of (a) windy conditions, (b) NF, (c) WJ, (d) CFC, (e) CJ and (f) CS. (b)–(f) are normalised by (a). Results are displayed on 0.5° × 0.5° boxes. Hatching indicates grid points with an altitude above 800 m

Due to orographic effects and higher noise in these regions, we exclude grid points above 800 m (hatching in Figure 6.6) and the area east of the Alps, including Hungary, Slovenia and the Balkans. Note that the frequencies of NF, WJ, CFC, CJ and CS add up to 1.

Figure 6.6b shows that NF is most common with a relative frequency of around 50 % in the periphery of the area, i. e., north and further south of the most common cyclone paths and wind footprints. Although removing areas of high orography itself, their effects can still be seen upstream, when the mostly westerly winds encounter mountain barriers such as the Scandinavian mountains, Western Alps and Carpathians leading to higher frequencies of NF in these regions. Over Western and Central Europe, high winds are usually closer to the cyclone centre, such that they are mostly associated with one of the mesoscale features.

The WJ occurs mostly over Western Europe with relative frequencies of almost 40 % and decreases over Germany (Figure 6.6c) down to 20 % over Poland. This is consistent with the occurrence early in the life cycle when the cyclone is still in the western regions. The higher frequency east of the Scandinavian mountains should be treated with caution as they might be caused by orography such as the foehn effect when a cyclone crosses the mountains. Moreover, a land-sea contrast is visible, which is weaker in ACCP (Figure B.3). A possible explanation are the different thermal characteristics of land and ocean, such as daytime heating leading to differences in $\tilde{\theta}$. Hence, just trained over land, RAMEFI might have difficulties distinguishing the WJ from the CS over sea in ambiguous situations. However, looking at exemplary winter storm cases, where features are well-developed, a difference in the detection over ocean and land grid points is not evident (Eisenstein et al., 2023b).

CFC shows low frequencies under 4 % for MAXP (Figure 6.6d), while values are twice as high for ACCP (Figure B.3d) due to the lower certainty of the feature (Figure 6.2) as discussed in Section 6.2.1. However, a distinct land-sea contrast is visible in both, where CFC seems to be almost exclusively detected over land. This might be due to land effects, such as frictional convergence and land surfaces being heated up more strongly than over ocean during the day, leading to destabilisation of the atmosphere. Moreover, CFC develops slightly later than the WJ when the cold front intensifies (Section 6.2.3 and Figure 6.7b). Note the patchy behaviour over land, possibly caused by local small-scale effects due to, among other things, surface roughness and orography. Distinct maxima are found east of or over mountainous regions, such as the Scottish Highlands and the Scandinavian mountains. Here, again, results given by RAMEFI should be treated with caution. As orography can induce convection, CFC might be detected without the occurrence of a cold front or even cyclone.

As expected, the occurrence of the CJ (Figure 6.6e) maximises in the northern half of the domain, much farther north than for WJ (Figure 6.6c), with a distinct footprint over the northern British Isles, the North Sea and the Baltic Sea. Over the British Isles, where intense cyclones are more frequent than over the North Sea (not shown), the CJ shows a maximum with over 20 %. Overall, the CJ occurs mainly over the sea and coastal areas (although RAMEFI was trained over land).

CS shows high values in the Bay of Biscay, where other features are rarely detected, and a rather abrupt drop over France in the east (Figure 6.6f). Together with the opposite patterns for WJ this suggests a possible false detection in some cases. As explained above, we suspect some systematically different behaviour between land and ocean to be at least partly responsible for this. A second peak in CS can be found over Eastern Europe, where most other features have already weakened at that late stage in the cyclone life cycle. Overall, the CS occurs further south than the CJ and thus farther away from the cyclone centre.

6.2.3 System-relative statistics

As before, we concentrate on the area within ±15° in zonal direction, $-15°$ and $+5°$ in meridional direction of the cyclone centre. This translates to around ±1 073 km in zonal direction at 50° latitude and 1 670 km and 557 km in southern and northern direction, respectively. Figure 6.7 shows a composite over the 19 extended winter seasons from 2000–2019 relative to the cyclone centre (a) and life cycle (b).

With respect to the mean spatial distribution, the WJ mostly occurs within the southeastern quadrant of a cyclone consistent with conceptual models (Figure 2.3). As shown in Figure 6.7a, the WJ usually has a distance from 250 km to 1 500 km from the centre. The CFC occurs around 3° to 5° farther to the west, i. e., upstream with respect to a westerly flow and also slightly shifted to the north, closer to the cyclone centre. Since CFC is a relatively small

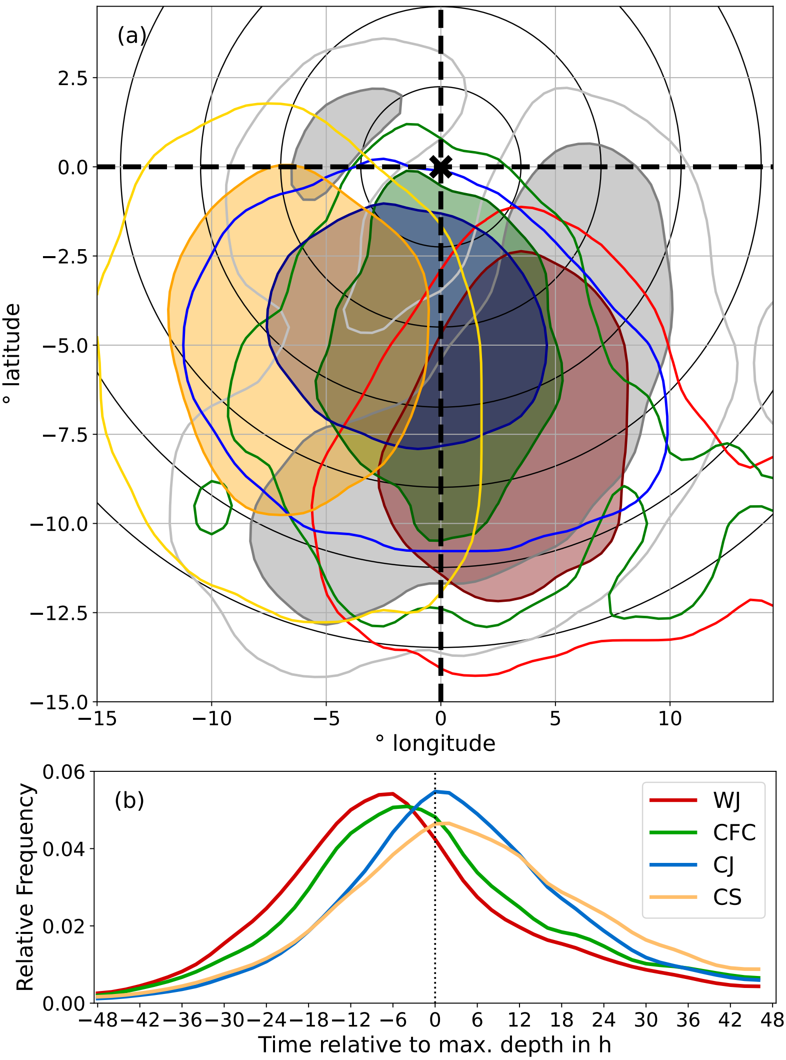


Figure 6.7: Occurrence of the identified mesoscale wind features (a) relative to the cyclone centre – including NF in grey – and (b) relative to the cyclone life cycle. The contours in (a) show the area with most feature occurrences, including 25 % (filled contours) and 50 % (outer contours) of the detected features. Black circles show the distance to the cyclone centre in 250 km increments using 50°N as a reference latitude.

elongated and narrow feature (as is the front itself), the location is harder to pinpoint over so many cases and the location varies the most from case to case compared to the other features. Thus, although CFC overlaps with other features statistically, this should usually not be the case for individual cyclones. The CJ is situated to the southwest to south of the cyclone centre with some statistical overlap with CFC. It occurs closer to the centre than the CS, which dominates the southwest quadrant and can have a distance of up to 1 500 km from the cyclone centre. The 25 % shaded area for NF is split into two patches. The majority of NF is detected to the southeast to south of the cyclone centre, mostly coinciding with the WJ but extending its reach to the north and southwest. The northern part of this patch is located in the area of the warm front and is possibly connected with the CCBa as introduced in Gentile and Gray (2023). A second, smaller patch can be found to the northwest of the cyclone centre. In this region, the CJ usually occurs before it is wrapped around the cyclone centre. As the CJ follows the bending of the front, the wind direction differs from the CJ later on, when it wraps around the centre, such that RAMEFI does not identify this part of the CJ. However, high wind speeds in this area are usually only caused for very strong CJs, as the relative movement of the air in this area is against the cyclone motion, weakening the Earth-relative wind speed, particularly in fast-moving weaker cyclones (Eisenstein et al., 2020).

Figure 6.7b shows the relative frequency of the four mesoscale wind features throughout the lifetime of the parent cyclone from 2 days before until 2 days after the time of maximum depth, i. e., the deepest pressure minimum during a cyclone's life cycle (marked as 0 in Figure 6.7b). The WJ is the first feature to develop with a maximum at -6 h. 50 % of detected WJ grid points occur between -18 h and the time of maximum depth. With a small offset of around 2 h to 3 h the CFC follows, consistent with its more western location in Figure 6.7a. The peak is slightly lower at around -4 h and 50 % of CFC points are detected between -16 h and 6 h. Contrary to Hewson and Neu (2015), the CJ develops already several hours before the time of maximum depth. However, the peaks of CJ and also CS are around 0, even though the peak of CS is lower. The WJ and CFC are already decreasing at that time, as the warm air begins to occlude consistent with Hewson and Neu (2015). While the CJ occurrence decreases faster, the CS stays on for longer. Overall, 50 % of CJ and CS occur within −8 h and 10 h and −8 h and 14 h, respectively.

When combining both composites, we can see how the features occur in different locations during the cyclone's life cycle. This is displayed for four exemplary time slots in Figure 6.8. An animation showing all time slots from 24 h before to 24 h after time can be accessed at Eisenstein et al. (2023b). 24 h before the time of maximum depth, only WJ and CFC appear in the composite, the round maximum of WJ to the south-southeast of the cyclone centre and the CFC with a more north–south elongated maximum closer to the cyclone centre (Figure 6.8a). CJ and CS develop in the following hours to the southwest of the cyclone centre, while the size of the WJ increases as shown 12 h later (Figure 6.8b). The area of CFC also increases, however the increase in width is probably rather due to the variation in location than an increase in size. Around the time of maximum depth, as shown in Figure 6.8c, the WJ area decreases and shifts farther away from the centre in southeastern direction, now with a stronger west–east orientation. The CJ has increased in size and now stretches across both southern quadrants, whereas the CS fills most of the southeastern quadrant. Furthermore, NF covers most of the area

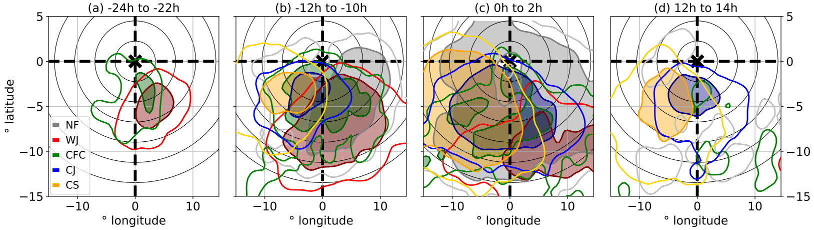


Figure 6.8: As Figure 6.7a but for different times relative to the cyclone life cycle, i. e., time of maximum depth: (a) -24 h to -22 h, (b) -12 h to -10 h, (c) 0 h to 2 h and (d) 12 h to 14 h.

overlapping with all other features. As mentioned above, the area northeast to north of the cyclone centre, which does not overlap with any of the mesoscale features, corresponds to the CCBa as described in Gentile and Gray (2023). 12 h after time of maximum depth (Figure 6.8d), WJ and CFC have mostly vanished, while CJ and CS are much diminished in size.

Overall, these results are mostly consistent with idealised schematics and conceptual models in the literature (e. g., Hewson and Neu, 2015, Figure 2.4). However, the very large set of differing cyclone developments in our comprehensive data set is able to show a larger variety. Since our study domain is too small to cover the whole development of the investigated cyclones, especially the early stages of a feature might be missed, such that an analysis of feature duration and comparison with the literature (e. g., Hewson and Neu, 2015) is not meaningful. Over the investigation domain, the WJ, CJ and CS have a broadly similar duration, with an average of around 20 h (not shown). In contrast, the smaller and rarer CFC appears only half as long with an average of around 11 h (not shown).

6.3 Characteristics of high-wind features

Using RAMEFI over a 19-year time period also allows us to analyse the distributions of selected meteorological parameters for each feature, i. e., finding characterisations of meteorological conditions. By construction, the eight parameters used for the training of RAMEFI (upper part of Table 4.1) behave as already documented in Section 5 and are therefore not displayed here, but only briefly discussed. Instead, we concentrate on wind speed (v_{10m}) and gusts (v_{gust}), the gust factor (g_{v}), specific and relative humidity (q and RH_{2m}, respectively) and total cloud cover (cc). Figure 6.9 shows boxplots for these parameters for all grid points, while a distinction between land and sea grid points can be found in Figure 6.10.

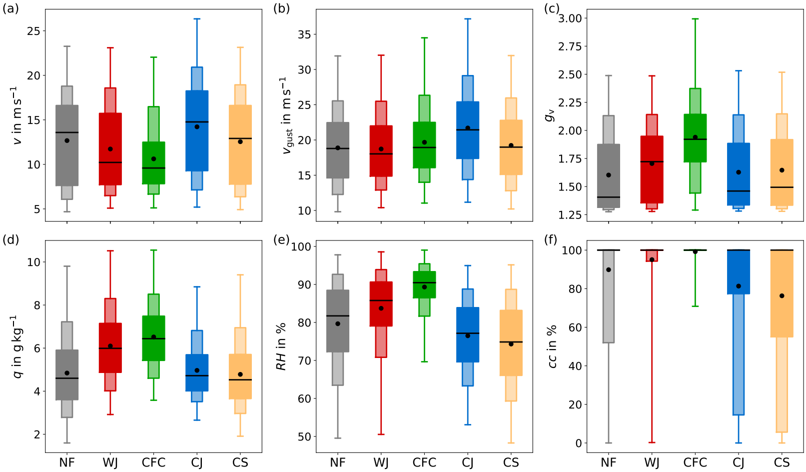


Figure 6.9: Boxplots of each feature for (a) v, (b) v_{gust}, (c) g_{v}, (d) q, (e) RH and (f) cc. Broader boxes show the 25th to 75th percentile, thinner ones the 10th to 90th percentile and whiskers the 1st and 99th percentile. Lines and dots indicate the median and the mean, respectively. Note that v, RH and q correspond to v_{10m}, $RH_{2\text{m}}$ and $q_{2\text{m}}$, respectively, here.

With respect to p_{msl} (not shown), the CJ has the deepest pressure being closest to the cyclone centre, while WJ and CS show higher values. In contrast to Figure 5.13, the CS has no second peak at low p_{msl}, which was due to the exceptionally deep storm *Sabine* (February 2020) included in the training. This is not the case here, and even if it was included, it would carry much less weight in a composite of almost 20 years. Being usually ahead of the cold front, WJ shows falling pressure, while pressure rises in the CS and CJ areas. Furthermore, WJ and CFC show warmer $\tilde{\theta}$ compared to CJ and CS (not shown). RR values over 1 mm h^{-1} are only common for CFC. CFC is characterised by slightly positive Δd values, while hardly any wind shift is found for WJ, CJ and CS.

As seen in Figure 6.9a, $v_{10\text{m}}$ is usually highest for the CJ with the median being around 15 m s^{-1} and the 99th percentile over 25 m s^{-1}, making it the most common cause for high winds (e. g., Hewson and Neu, 2015; Gentile

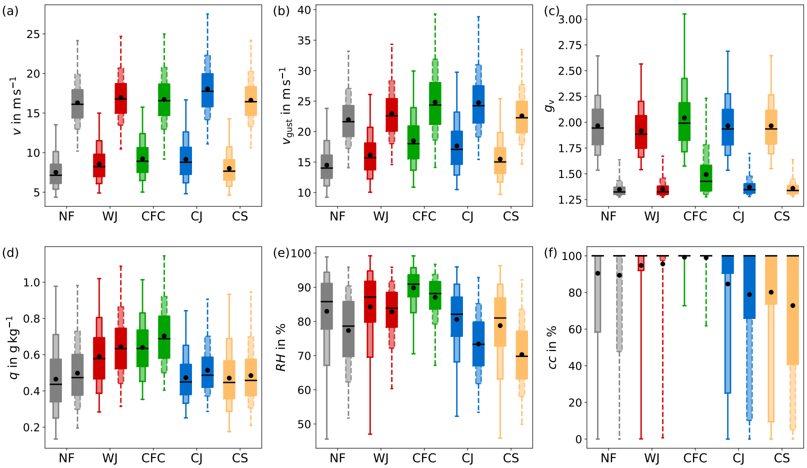


Figure 6.10: As Figure 6.9 but with land (solid boxplots) and sea (dashed boxplots) grid points separated.

and Gray, 2023). This is also consistent with the increasing proportion with a higher $\tilde{v}$ threshold, as discussed in Section 6.2 (Figure 6.4). While CS and WJ show similar 99th percentiles at around 23 m s^{-1}, the median of WJ at around 10 m s^{-1} is about 3 m s^{-1} lower than the median of CS. NF has a slightly higher median than CS, but a similar mean value. The lowest values are found for CFC with a median of under 10 m s^{-1}. Naturally, $v_{10\mathrm{m}}$ over the ocean is considerably higher than over land (Figure 6.10) affecting the overall $v_{10\mathrm{m}}$ depending on how often they occur over land or ocean (Figure 6.3e–l and Figure 6.6). v_{gust} shows similar behaviour for CJ, CS, WJ and NF. However, CFC shows the second highest gusts with up to 35 m s^{-1} (Figure 6.9b). This leads to the highest g_v, which simply displays the ratio between v_{gust} and $v_{10\mathrm{m}}$ and reaches 3 in case of CFC (Figure 6.9c). This is not surprising, as convection is associated with high instability and turbulence. Although both wind and gust speeds are much higher over the ocean, the gust factor differs

significantly between land and sea (Born et al., 2012) with less friction and other causes of turbulence over the sea leading to a weaker increase of gust speeds compared to wind speeds (Figure 6.10a–c).
With respect to moisture and cloud variables, CFC shows the highest values of specific and relative humidity, followed by the WJ in the warm sector, CJ and lastly CS (Figure 6.9d,e). CS may also include high winds caused by dry intrusions (Raveh-Rubin and Catto, 2019; Catto and Raveh-Rubin, 2019), leading to overall drier conditions for this feature. Moreover, especially the SJ, which is here included in the CJ feature, occurs in the dry slot area. Following the Clausius-Clapeyron relation, it is also intuitive that warmer temperatures enable higher values of $q_{2\mathrm{m}}$. Consistently, all features have lower q values and higher $RH_{2\mathrm{m}}$ values over land. Figure 6.9f shows the total cloud cover. While stratocumulus and stratus clouds are common in the warm sector ahead of the cold front, cloudless areas can still be found for the WJ in contrast to CFC. Both CJ and CS show a wide distribution. While the CJ appears at the tip of the cloud head, partly below and slightly ahead of it, allowing for both cloudy and cloudless conditions, the CS is often associated with post-CFC, and thus mixture between cloudless skies and showers. At times, NF is detected along the warm front (Figure 6.7), which is characterised by cloudy conditions, but can also show lower cc values in other areas, e. g., in the warm sector.

6.4 Conclusions

Damaging winds accompanying extratropical cyclones can be caused by several mesoscale features with different characteristics and, thus, also differing forecast errors and potential for damage. To analyse these differences, we developed a novel objective and flexible probabilistic identification tool called RAMEFI (RAndom-forest based MEsoscale wind Feature Identification) introduced in Chapter 5. The method is trained on the basis of surface observations for 12 storm cases but due to spatial independence and removing location-specific effects, once trained, it can be applied to gridded data without any modification. Here RAMEFI is used to compile a – to the best of our knowledge – first-ever long-term objective climatology of the four wind features WJ, CJ, CS and CFC based on station observations and a high-resolution reanalysis data set (COSMO-REA6) for a time period of 19 extended winter seasons, i. e., October to March. Using the reanalysis data also allows investigating ocean grid points. Although a systematic validation as done for land in Section 5.4.1 is not performed for ocean areas, a subjective inspection of several cases during the analysed time period did not reveal fundamental differences. However, due to a different shape of the wind speed distribution, the threshold of $\tilde{v} > 0.8$ used to define windy conditions results in more frequent occurrence over the ocean than over land.

The considered area includes Western and Central Europe, however, excluding grid points above 800 m altitude and the Balkans. Focusing on grid points with windy conditions within the vicinity of a cyclone centre, i. e., $\pm 15°$ in zonal and $-15°$ and $5°$ in meridional direction, we determined the relative frequency of mesoscale wind features in both an Earth-relative and cyclone-relative framework. Furthermore, distinctive characteristics in wind (wind speed v_{10m}, gust speed v_{gust}, gust factor g_{v} and humidity parameters (specific humidity $q_{2\mathrm{m}}$, relative humidity $RH_{2\mathrm{m}}$, cloud cover cc) were investigated. The main findings of the climatological analysis for the individual features are:

Warm jet (WJ)

- Characterised by decreasing p_{msl}, warm temperatures, almost no precipitation and mostly south-westerly winds
- First to occur within the southeastern quadrant of a cyclone, peak around 6 h before the time of maximum depth
- Detected mostly over land in more than 80 % of stormy time steps
- Most common over southern UK, France, Benelux states and Germany

- Mostly cloudy conditions; warm sector allows for rather humid conditions

Cold-frontal convection (CFC)

- Associated with heavy precipitation, a shift in wind direction and cooling temperatures
- Narrow feature along the cold front, least common, location varies considerably from case to case
- Occurs almost exclusively over land where, e. g., daytime heating and frictional convergence, can strongly enhance the development of convection along the cold front, and is also detected particularly around mountainous areas (sometimes orographic triggering independent of cold front)
- Highest g_v with rather low v_{10m} and high v_gust up to $35\,\mathrm{m\,s^{-1}}$ and higher in extreme cases
- Highest values of q_2m, RH_2m and cc connected with the convection

Cold jet (CJ)

- Shows increasing but overall deepest p_msl, westerly winds and cold temperatures
- More frequent over sea than land in over 80 % of stormy time steps mostly affecting northern UK, the North and Baltic Seas, Scandinavia and northern Germany
- Smaller feature than WJ, occurring close to the cyclone centre, first to the southwest and later to the south of it
- Usually cause of highest winds and gusts
- Cloudy conditions below cloud head and drier at the tip of it in dry-slot region

Cold sector (CS)

- Associated with cold temperatures, westerly winds, increasing and higher p_{msl} compared to CJ
- Occurs in almost all winter storms and time steps affecting a relatively large area in the southwestern quadrant of a cyclone
- Last to decay, hence most common cause of high winds over Eastern Europe
- Sunnier conditions with patches of post-cold frontal convection
- Overall drier conditions due to dry intrusion

The locations of the features relative to the cyclone centre found in this climatology are mostly consistent with conceptual models based on case studies or subjective identification (see Figure 2.3). While previous literature suggests a cyclone to be first dominated by the WJ and second by the CJ (e.g., Hewson and Neu, 2015; Rivière et al., 2015), this climatology further revealed the occurrence of CFC in early development stages and the dominance of CS in later ones. Also further characteristics in wind and humidity parameters show mostly consistent behaviour to previous studies (e. g., Hewson and Neu, 2015; Earl et al., 2017). The large number of storms investigated helped to reveal the large variability in the location of CFC in a system-relative framework, similar to the blurring of frontal boundaries in composites discussed in Dacre et al. (2012). Other differences to the literature include the time of occurrence of the CJ already several hours before the time of maximum depth in contrast to Hewson and Neu (2015). Overall, RAMEFI allows for a more objective and more thorough analysis and description of the mesoscale wind features. This climatology demonstrates the applicability of RAMEFI for longer time periods and data it was not trained on. The new data set can serve the community as a climatological reference for case studies or in combination with other objective climatologies (e. g., Sprenger et al., 2017).

7 Identification of sting jets

RAMEFI enables us to identify several key wind features within extratropical cyclones. However, since it is solely based on surface parameters, the distinction of CJ and SJ is challenging and the two features are combined. As an SJ can cause higher gusts and, thus, damage (Hewson and Neu, 2015), it is desirable to detect SJs or at least the potential of possible SJs in an operational framework.

Until recently, the SJ occurrence was mostly validated a posteriori using Lagrangian trajectories (e. g., Volonté et al., 2018; Eisenstein et al., 2020, see Section 2.4). Due to their high computational cost, efforts were made to develop a cheaper method suitable for operational use or computing climatologies. Two methods – an instability-based precursor tool by Gray et al. (2021) and a kinematic-based detection by Manning et al. (2022) – are introduced in Section 2.4. Here, a comparison of these methods with Lagrangian trajectories is discussed using the example of storm *Eunice* (February 2022), which is known to have developed an SJ over the UK (Mühr et al., 2022; Volonté et al., 2023a,b). Furthermore, a Shapiro-Keyser cyclone – storm *Xavier* (October 2017; see also Chapter 5) – which does not show an SJ in trajectory analysis, is used to evaluate how the tools handle a non-SJ case. Lastly, a strategy and first results of an own method based on ideas from the introduced tools are discussed, possibly complementing the RAMEFI method.

7.1 Shapiro-Keyser cyclone case studies

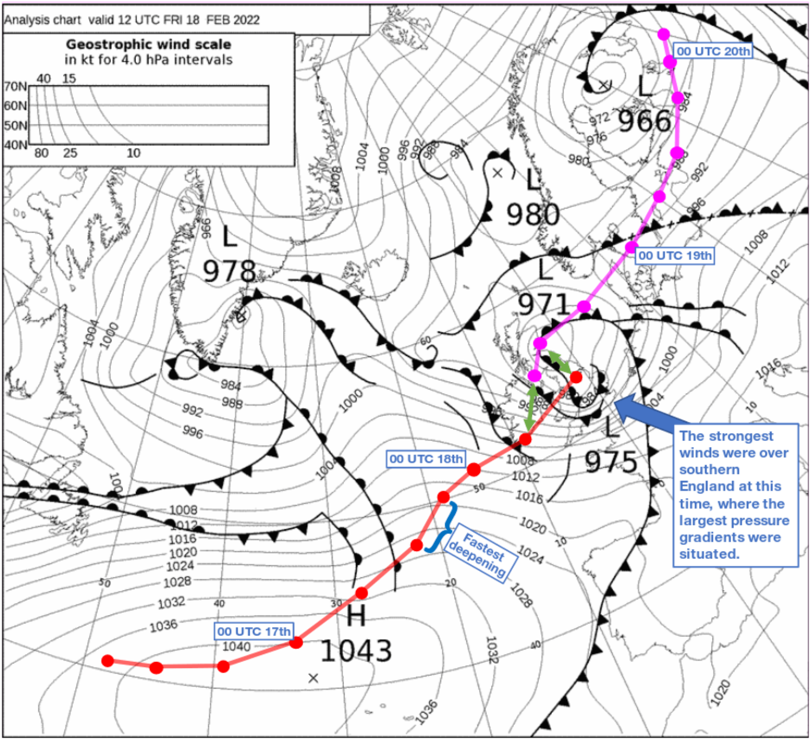
7.1.1 SJ storm Eunice

In February 2022 a storm series of three intense extratropical cyclones affected Europe in just a few days (Mühr et al., 2022): *Ylenia* (named *Dudley* by the UK Met Office; 16–17 February 2022), *Zeynep* (*Eunice*; 18–19 February 2022) and *Antonia* (*Franklin*; 20–21 February 2022). This series caused a total of EUR 3 851 m of insured damage over Europe[1]. The storms developed in association with a strong and almost zonally frontal zone accompanied by a strong jet stream.

In this chapter, the focus is on storm *Zeynep/Eunice* as its development is representative of a Shapiro-Keyser cyclone with an occurring SJ. Hereafter, we refer to the storm as *Eunice* following recent literature (Hewson et al., 2022; Volonté et al., 2023a,b). *Eunice* was named by the UK Met Office on 14 February and received its name *Zeynep* by the FU Berlin on 16 February, when it built along the cold front of *Ylenia/Dudley*. It moved over the North Atlantic to the UK, where its centre split in two, as shown in Figure 7.1 by the red and magenta lines (Hewson et al., 2022) and satisfies the threshold for an explosive cyclogenesis with a pressure drop of roughly 35 hPa within 24 h (see equation 2.2). The minimum core pressure was below 970 hPa in the evening hours of 18 February while located over the North Sea. The storm broke the record of highest wind gust in England with 196 $\mathrm{km\,h^{-1}}$ at the station The Needles, Isle of Wight (Mühr et al., 2022), possibly caused by an SJ. A banded structure within the cloud head was visible in satellite images through the night and early morning hours of 18 February (Volonté

---

[1] According to the final loss report (21 February 2023) from `https://www.perils.org` (last access 03 May 2023)



The strongest winds were over southern England at this time, where the largest pressure gradients were situated.

Figure 7.1: UK Met Office synoptic chart for 18 February 2022, 12 UTC (©Crown Copyright) and track of *Eunice* in 6 h intervals. At this time, *Eunice* had two centres (marked as 971 hPa and 975 hPa lows). The tracks of the two centres are shown in red and magenta. Green arrows connect the centres that correspond to the same valid times. Figure reprinted from Hewson et al. (2022).

et al., 2023a). The later confirmed SJ was mainly affecting the North Atlantic, the UK and the North Sea, while high wind gusts and damages over Germany in later hours were mostly caused by CFC.

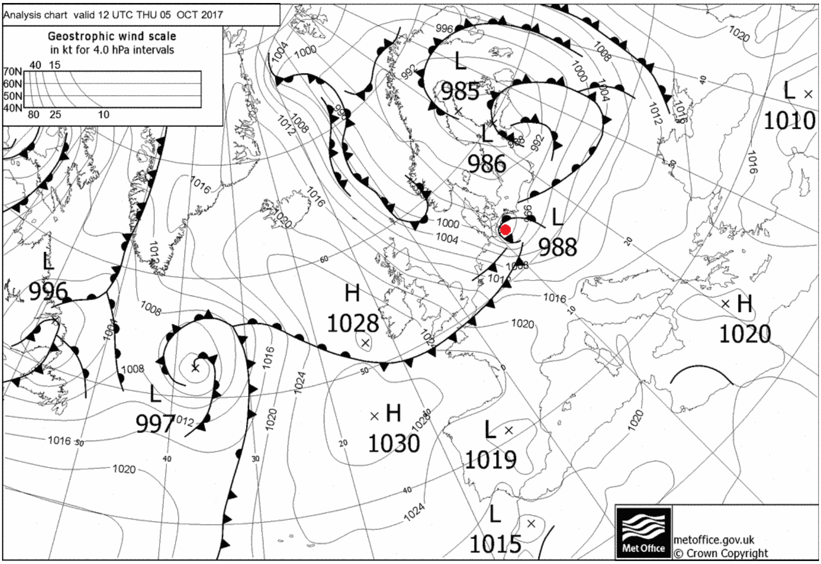


Figure 7.2: UK Met Office synoptic chart for 12 UTC on 05 October 2017, when *Xavier* was situatied over Germany. Cyclone centre marked in red ©Crown Copyright

7.1.2 Non-SJ storm Xavier

Storm *Xavier* was part of the case studies used for RAMEFI (Chapter 5) mainly affecting the North Sea region, Germany, Poland and the Czech Republic on 05 October 2017 (see Figure 5.1). It was a fast-moving cyclone with a minimum core pressure of around 986 hPa at 15 UTC on 05 October accompanied by high gusts exceeding 120 $\mathrm{km\,h^{-1}}$ at several stations in the plains in Northern Germany. Figure 7.2 shows the storm over Northern Germany with a core pressure of 988 hPa at 12 UTC on 05 October, resembling stage III of a Shapiro-Keyser cyclone shortly before the time of maximum depth (Figure 2.1b; Section 2.1.2). *Xavier* caused 9 casualties and a total of EUR 324 m of insured losses[2]

The occurrence of an SJ was hinted by a prominent cloud head, however, could not be confirmed by Lagrangian trajectories as discussed in Section 7.2. As *Xavier* was investigated as one of the 12 RAMEFI case studies (see Sections 5.1 and 5.5.4) and was a Shapiro-Keyser cyclone without an occurring SJ, it makes a suitable case study for this chapter.

7.2 Trajectory analysis

Following Eisenstein et al. (2020), Lagrangian trajectories were computed for *Eunice* and *Xavier* for several starting times. Naturally, the number of trajectories satisfying the SJ criteria is dependent on the choice of thresholds for v and Δp as evident in Table 7.1 for *Eunice*. The first threshold of $v > 37\,\mathrm{m\,s^{-1}}$ and $\Delta p > 150\,\mathrm{hPa}$ were chosen in analogy to the reference criteria discussed in Eisenstein et al. (2020). For *Eunice* up to almost 8000 trajectories satisfy these criteria at noon on 18 February. In comparison, *Egon* showed a maximum of just over 200 trajectories, suggesting that *Eunice* developed an extraordinarily intense SJ. However, some of these trajectories might also be part of a dry intrusion, as suggested by the trajectory analysis of Volonté et al. (2023b) on the eastern flank of starting positions. As these numbers are quite high, more extreme thresholds were selected as well, showing a stronger decrease in number in the earlier time steps with a higher v criterion ($v > 42\,\mathrm{m\,s^{-1}}$). In contrast, increasing only the Δp criterion ($> 200\,\mathrm{hPa}$) shows a stronger decrease in number in the later hours. This

---

[2] According to the final loss report (05 October 2018) from `https://www.perils.org` (last access 03 May 2023.

suggests that the SJ experienced a stronger decrease at first, while causing higher winds in later time steps. Even with a v criterion of $42\,\mathrm{m\,s^{-1}}$ and a descent of more than 200 hPa the maximum number of trajectories remains above 3000 at noon. Eisenstein et al. (2020) showed for storm *Egon* that the proportion of CSI is higher for trajectories satisfying higher thresholds, suggesting that the CSI contribution actually strengthens the SJ. This is, however, not analysed in the scope of this chapter. For storm *Xavier* only a maximum of 10 trajectories satisfy the conditions of $v > 35\,\mathrm{m\,s^{-1}}$ and $\Delta p > 125\,\mathrm{hPa}$ at 15 UTC on 05 October, i. e., being deemed insufficient to be considered an SJ.

For storm *Eunice*, two exemplary time steps – 08 UTC and 11 UTC – are shown in Figure 7.3 including the frontal structure and SJ trajectories satisfying weak criteria (see left column of Table 7.1). It is evident that the descending trajectories are split into multiple streams located west and southeast of the tip of the bent-back front instead of being one coherent stream. The western stream originates from farther north and shows a southward travel, while the southern stream originates from the west and travels eastwards. While the western streams are stronger in the earlier time steps (Figure 7.3a), the number of trajectories decreases later on when the other stream becomes more prominent (Figure 7.3b). Considering the number of trajectories with increasing thresholds (Table 7.1) suggests that the western stream shows stronger descent, while the southern stream developed higher wind speeds.

Figure 7.4 gives a closer look at the characteristics with the Earth-relative and system-relative path and several parameters traced along the trajectories satisfying the stronger criteria of $v > 42\,\mathrm{m\,s^{-1}}$ and $\Delta p > 200\,\mathrm{hPa}$ at 11 UTC. This time is chosen because it corresponds to the time of the new gust speed record at The Needles. Note that the plot here shows the merged backwards and forwards trajectories. With these stronger criteria, the western trajectories completely disappear as shown in Figure 7.4a. However, these trajectories are also split in two - one part ascending before descending (shading in Figure

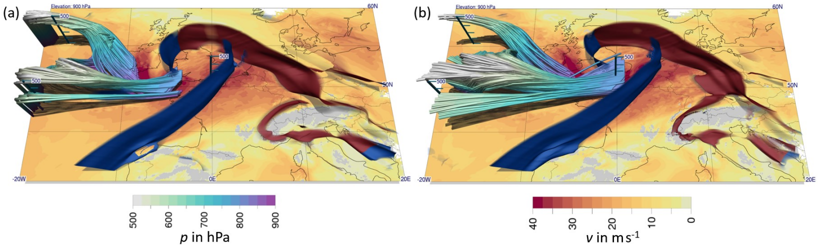


Figure 7.3: Backward trajectories of *Eunice* with starting times at (a) 08 UTC and (b) 11 UTC on 18 February 2022 shaded by p. The frontal structure shows both the warm front in red and the cold front in blue. The horizontal cross section is shaded by v at 900 hPa. Images produced using Met.3D (Beckert et al., 2023; Rautenhaus et al., 2015).

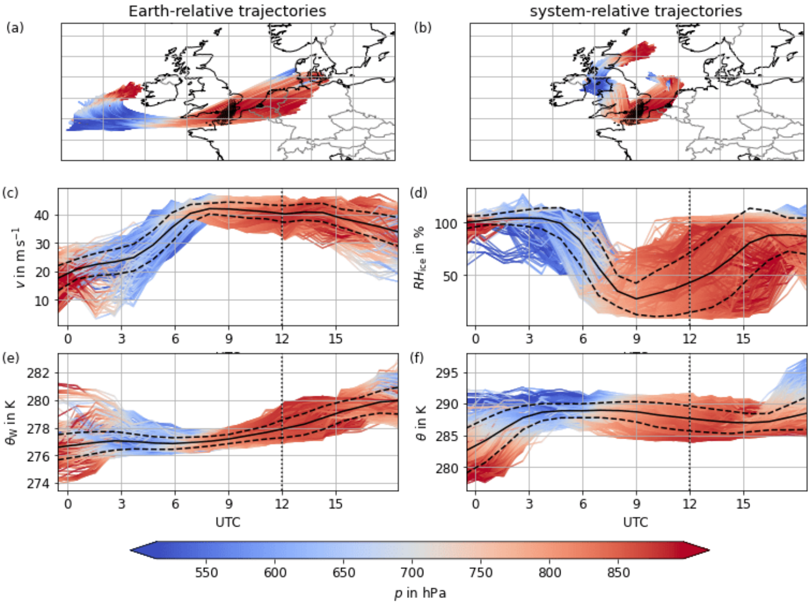


Figure 7.4: Trajectories started on 18 February 2022, 11 UTC and computed backwards for 12 h and forwards for 6 h: (a) Earth-relative path, (b) system-relative path, (c) v, (d) RH_{ice}, (e) θ_{W} and (f) θ. Selection criteria of $v > 42\,\mathrm{m\,s^{-1}}$ and $\Delta p > 200\,\mathrm{hPa}$ leave 2577 trajectories, of which every fourth is plotted here. Shading corresponds to p, black points in (a) and (b) indicate starting positions and the black dotted line in the remaining panels the starting time. The black solid line shows the mean of trajectories and the dashed lines ± 1 standard deviation.

Table 7.1: Number of SJ trajectories of storm *Eunice* using different thresholds for multiple starting times (day/hour). In comparison, storm *Egon* showed a maximum of 206 trajectories satisfying $v > 37\,\mathrm{m\,s^{-1}}$ and $\Delta p > 150\,\mathrm{hPa}$.

Time	$v > 37\,\mathrm{m\,s^{-1}}$, $\Delta p > 150\,\mathrm{hPa}$	$v > 37\,\mathrm{m\,s^{-1}}$, $\Delta p > 200\,\mathrm{hPa}$	$v > 42\,\mathrm{m\,s^{-1}}$, $\Delta p > 150\,\mathrm{hPa}$	$v > 42\,\mathrm{m\,s^{-1}}$, $\Delta p > 200\,\mathrm{hPa}$
18/04	877	65	132	5
18/05	1752	437	278	20
18/06	1863	667	284	9
18/07	1839	667	435	46
18/08	2724	917	944	338
18/09	4557	1790	2117	1036
18/10	5887	3794	3257	2022
18/11	5482	3190	3445	2577
18/12	7805	3990	5276	3315
18/13	7457	3315	5400	2670
18/14	4586	1662	3558	1374
18/15	2590	562	1903	524
18/16	1553	133	1244	126
18/17	762	19	659	10

7.4), which originates from farther north and one part descending from higher levels. Figure 7.4b finely shows the movement along the bent-back front in a system-relative framework. However, a small deformation can be seen over the UK due to the split cyclone centre. Although both air streams accelerate during their descent to around $40\,\mathrm{m\,s^{-1}}$ (Figure 7.4c), the latter shows lower RH_{ice} values in the early stages consistent with a dry intrusion as suggested by Volonté et al. (2023b) (Figure 7.4d). The descent is characterised by almost constant θ_W as typical for an SJ (Figure 7.4e; Section 2.2.3). Figure 7.4 shows only minimal cooling in θ suggesting no considerable contribution of evaporative cooling.

The trajectory analysis of storm *Eunice* shows an extraordinarily intense – in both descent and wind speed – and long-lived SJ split in two air streams. The western stream is more intense in the early hours, while the southern stream increases in size and strength later on when the western stream slowly vanishes. Furthermore, some trajectories seem to belong to a dry intrusion with rather dry RH_{ice} values at all time steps consistent with Volonté et al. (2023b).

7.3 Instability-based precursor tool

The precursor tool by Gray et al. (2021) is used on the MOGREPS-G by the UK Met Office operationally with a native grid spacing of around 20 km regridded to 0.5°. The output for storm *Eunice* that was available to forecasters of the UK Met Office using the MOGREPS-G-EPS was recently discussed in Volonté et al. (2023a) and shows potential of SJ occurrence at midnight of 18 February already 60 h prior, albeit with low probabilities due to ensemble spread concerning the location.

Here, we apply the approach to an ICON-LAM simulation (Section 4.1.3) and, hence, lose the advantage of a probabilistic output. However, Lagrangian trajectories confirm an SJ in the simulation as shown in Section 7.2, while it would possibly not be present in all ensemble members of an EPS forecast. Therefore, the tool should be able to detect enough CSI points indicating an SJ using *DSCAPE* and additional criteria as outlined in Section 2.4. Furthermore, the used ICON-LAM simulation is – as the name LAM implies – not global as the originally used data. This might result in boundary effects, but should not be relevant for the analysis over the UK. The grid spacing is with around 6.5 km almost four times as fine and, hence, shows smaller-scale structures and more noise, also due to the better resolved orography.

The resolution is high enough to resolve mesoscale instabilities in contrast to MOGREPS-G, which might affect the amount of detected CSI points diagnosed by *DSCAPE*.

The method is applied to three different resolutions to compare the benefits and disadvantages of any chosen resolution: (1) the native 0.062 5° grid-spacing of the ICON-LAM simulation, (2) regridded to 0.2° similar to the 20 km grid spacing of MOGREPS-G and (3) regridded to 0.5°, which is the original grid spacing the precursor tool is applied to. Compared to the lowest resolution, applying the method to 0.2° grid spacing takes about twice, and to 0.062 5° grid spacing over ten-times as much computational time for a single time step, respectively. Originally, the output shows a probability for more than 20 CSI points in each ensemble member. This threshold translates to 125 and 1280 CSI points for a grid spacing of 0.2° and 0.062 5°, respectively. Figure 7.5 shows grid points satisfying the precursor criteria shaded by *DSCAPE* for multiple time steps applied to ICON-LAM for storm *Eunice* (upper three rows) and storm *Xavier* (lower two rows). Red circles indicate a centroid if the number of CSI points is above the introduced thresholds within 700 km of the cyclone centre and blue circles if it is farther away. The figure further includes CSI detected using MPV^* as discussed in Section 2.1.3 and used for high-resolution data in previous literature (e. g., Gray et al., 2011) (black dots). For more clarity, grid points meeting these criteria are referred to as MPV points hereafter. For 00 UTC on 18 February 2022, the storm is located east of Ireland over the North Atlantic. The high resolution in 7.5a clearly shows CSI points along the cloud head as expected. However, also a blue circle in the area of the cold front is visible, although criteria to exclude fronts are set in the precursor tool. A cold front is characterised by instability release, which is resolved at such high resolution and possibly leads to high *DSCAPE* values. Only a few MPV points are detected at this stage. Regridding the data to 0.2° still shows many CSI points along the cold front (Figure 7.5b), while they are absent for 0.5° grid spacing (Figure 7.5c). 6 h later, however, when the storm moved to southern Ireland, the

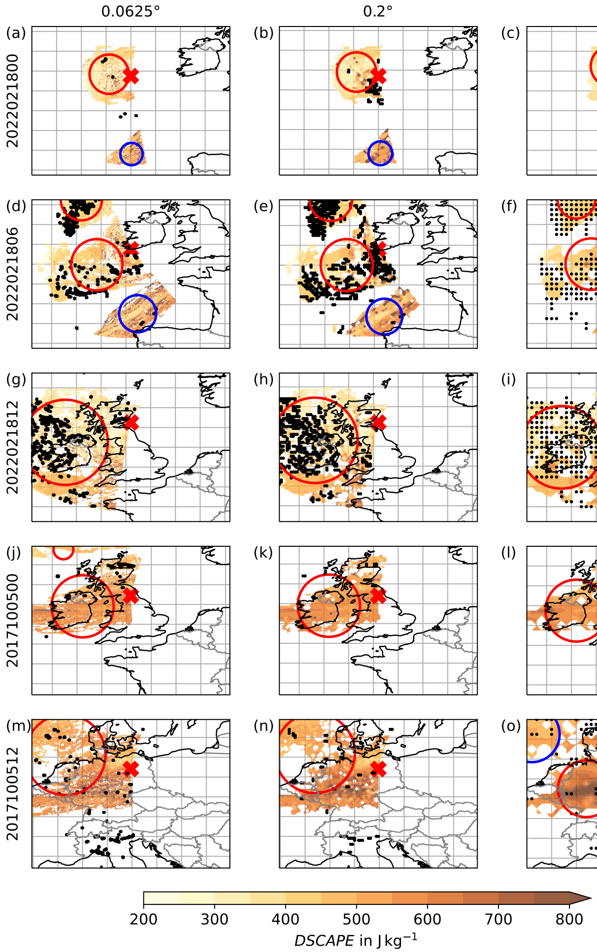


Figure 7.5: CSI points for three resolutions: left column - model resolution 0.0625°, middle column MOGREPS-G-EPS resolution 0.2° and left column SJ precursor resolution 0.5°. The precursor tool was applied to three time steps on 18 February 2023 (storm *Eunice*) - (a)-(c) 00 UTC, (d)-(f) 06 UTC, (g)-(i) 12 UTC - and two time steps on 05 October 2017 (storm *Xavier*) – (j)-(l) 00 UTC and (m)-(o) 12 UTC. Shading indicates *DSCAPE* values for points satisfying all four SJ precursor criteria. The red cross marks the cyclone centre and the red circles mark the centroid of CSI points within the vicinity of 700 km of the cyclone centre and blue circles if farther away. Black dots show MPV points indicating CSI by satisfying MPV^* and further criteria outlined in Section 2.1.3.

blue circle is visible at all three resolutions. Nevertheless, as its centroid is too far away from the centre, it is not considered by the algorithm (Figure 7.5d–f). Furthermore, a second area in the vicinity of the cyclone centre appears northwest of it at all three resolutions, probably still associated with the cloud head, merely not connected to the other area. More MPV points are detected around the cyclone centre as well, confirming the potential for CSI. For this time, trajectories indicate an SJ (Table 7.1), suggesting that instability may have been released. However, at noon, almost twice as many CSI points are diagnosed (Figure 7.5g–i). While the area of CSI points becomes more fragmented with a decrease in resolution, the number of points remains similar. In the evening hours, when the SJ slowly vanishes according to trajectories, a significant number of CSI points is still visible at all three resolutions, when CSI should have been released (not shown). In theory, the precursor tool should diagnose more CSI points before the actual descent of the SJ and less in later hours when CSI has been released contributing to the SJ descent. However, this does not seem to be the case here. Overall, detected MPV points show similar results to CSI points for storm *Eunice* for all time steps.

Although an SJ was not confirmed for storm *Xavier* by trajectory analysis (Section 7.2), Figure 7.5j–o shows diagnosed CSI points for two time steps 12 h apart in the vicinity of the cyclone centre. Moreover, the number of CSI points for single time steps is at times even higher than for storm *Eunice*. In contrast, almost no MPV points are found in the evolution of *Xavier*, suggesting no potential for CSI and an overestimation by the precursor tool.

The similarity between the different resolutions suggests that the data can easily be regridded to coarser resolution for the precursor tool, while still diagnosing CSI points with less computational cost. However, the number of CSI points might be higher in high-resolution simulations or forecasts even if regridded to the same grid operationally used at the UK Met Office, as mesoscale instabilities (release) can be resolved and, hence, processes that are influenced by these instabilities, also on a larger scale. Therefore, it is wise to only use the tool on low-resolution data for which it was originally developed. At the DWD, a suitable model would be the global ensemble model ICON-EPS with a nest over Europe (ICON-EU-EPS) with grid spacings of 26 km and 13 km, respectively. However, a proper evaluation would be necessary, as 13 km is close to the limit of resolving mesoscale instabilities release (10–12 km; Clark et al., 2005).

In general, CSI does not only occur with SJs as shown in the case of storm *Xavier* and not all SJs show CSI contribution. Furthermore, the precursor tool seems to overestimate CSI. Together, this suggests a possibly high false alarm rate. Nonetheless, the precursor tool is able to bring the possibility of intense storms with a likelihood of SJ occurrence to the attention of a forecaster, who can then keep an eye out for further signs and if necessary adapt warnings.

7.4 Kinematic-based streamline approach

The kinematic-based method by Manning et al. (2022) uses the difference of the horizontal wind speed between 700 hPa and 850 hPa. Here, the refined method using three pairs (upper, middle, lower; Manning et al., 2023) is applied to storm *Eunice*. As the approach is supposed to be an approximation of trajectories, the tool is only used on time steps, when SJ criteria are met (see Section 7.2). As in Section 7.2, the time of the highest gust speed in the UK at 11 UTC is used as an example for this section.

Figure 7.6 shows the vertical gradient of the horizontal wind speed $\Delta_z v$ for the three pairs. A region of negative values, i. e., higher wind speeds in the lower level of a pair, can be seen in all pairs, while most evident in the upper pair (Figure 7.6a). This coincides with the SJ region as expected. Especially in the upper and middle pairs (Figure 7.6a,b) a distinct positive gradient occurs along the cold front, indicating the location of the WCB. Overall, the values in the lower pair (Figure 7.6c) are lowest. Note that the lower level at 900 hPa might be affected by orography over land, especially in the vicinity of a low pressure system, where the geometric height of the 900 hPa surface can be considerably reduced.

The gradient reversal is visible along computed streamlines as displayed in Figure 7.7. All pairs show several small coherent regions of streamlines in the southwestern quadrant of the cyclone. While the streamlines in the upper pair (Figure 7.7a) over the coast of southwestern England and southern Wales roughly coincide with the location of the trajectories of the southern air stream, the streamlines in the middle and lower pair (Figure 7.7b,c) located near Scotland and northern England coincide with the trajectories of the south-western air stream (Figure 7.3). As the streamlines are subject to noise in the high resolution data as apparent here, but probably even more so in the 2.2 km grid spacing originally used, Manning et al. (2022) included two filters. Firstly, the maximum wind speed in the negative sequence of $\Delta_z v$ along a streamline (blue shading in Figure 7.7) has to be higher than a threshold v_{param} compared to the positive sequence (red shading in Figure 7.7) and, secondly, the relative humidity within the negative sequence has to decrease by specific percentage points (RH_{param}) compared to the positive sequence. Both in Manning et al. (2022) and Manning et al. (2023), these

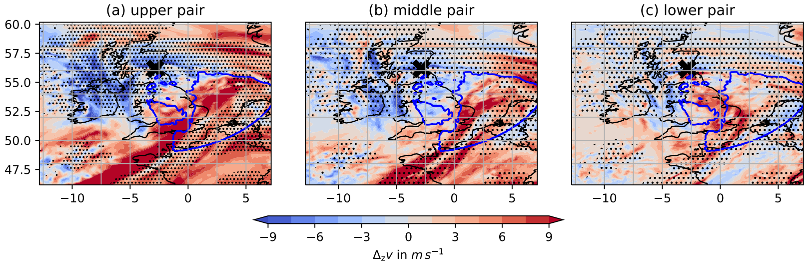


Figure 7.6: Difference in wind speed between $\Delta_z v$ (a) $600-700$ hPa, (b) $700-800$ hPa and (c) $800-900$ hPa at 18 February 2022, 11 UTC. Dotted areas correspond to $RH_{\text{ice}} > 80\,\%$, the blue line indicates the warm seclusion and frontal structure through $\theta_{W\text{thresh}}$ and the black x marks the cyclone centre.

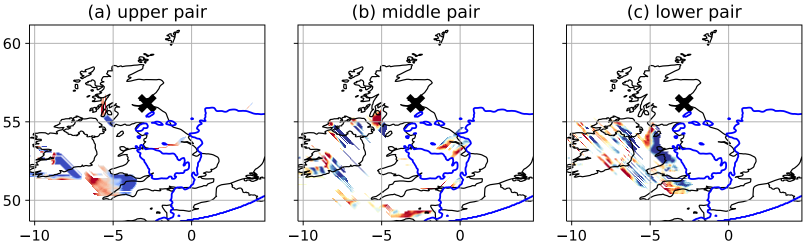


Figure 7.7: Streamlines at (a) 700 hPa, (b) 800 hPa and (c) 900 hPa shaded by $\Delta_z v$ as seen in Figure 7.6 for the respective pair showing a reversal in $\Delta_z v$ (blue – negative, red – positive). The blue line indicates the warm seclusion and frontal structure through $\theta_{W,\mathrm{thresh}}$ and the black x marks the cyclone centre. Note that the streamlines here are not filtered by the additional criteria of RH_{param} and v_{param}.

thresholds are $v_{\mathrm{param}} = 6\,\mathrm{m\,s^{-1}}$ and $RH_{\mathrm{param}} = 50$ percentage points. Applying these filters removes all streamlines visible in Figure 7.7. Although *Eunice* has an exceptionally strong SJ as discussed in Section 7.2, no SJ is detected with these thresholds using absolute values. Reducing either the v_{param} or RH_{param} threshold by half allows for only a few streamlines to remain in the lower and middle pair, respectively. This suggests that – although filtering the noise might be necessary – these parameters seem not adequate to do so.

An addition to the method discussed in Manning et al. (2023) is the condition of overlapping as illustrated in Figure 2.9. At least 20 % of the positive sequence of the middle pair has to overlap with the negative sequence of the upper pair, or, analogously, at least 20 % of the negative sequence of the middle pair has to overlap with the positive sequence of the lower pair. Applying this to the chosen time step shows no overlap areas and, hence, an SJ is not detected (not shown). Although some overlap takes place, if many streamlines show a gradient reversal, the threshold of at least 20 % overlap is quite high and not met in any time step. Furthermore, the overlap is supposed to be an indication of the descent and should be seen in vertical cross-sections. However, in an Eulerian point of view, an SJ is often a more vertical region of high wind speeds and not necessarily a diagonal one as shown by the example

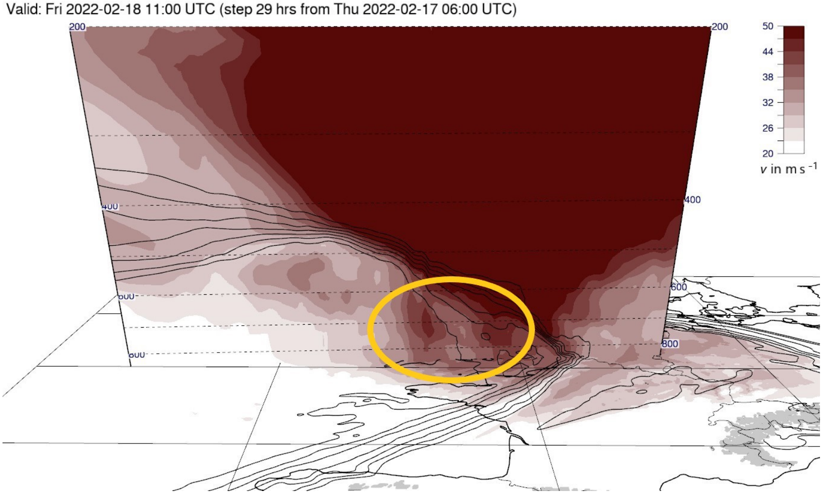


Figure 7.8: Horizontal and vertical cross-section shaded by v for storm *Eunice* on 18 February 2022, 11 UTC. Black contours indicate the frontal structure by θ_{E}. The ellipse marks the SJ area. Image produces using Met.3D.

of *Eunice* in Figure 7.8 and in further published case studies (e. g., Eisenstein et al., 2020; Gray et al., 2021; Volonté et al., 2023b, their Fig. 7, Fig. 4 and Fig. 2, respectively).

The method was evaluated for a higher horizontal resolution (2.2 km) and included mostly SJ storms over the North Atlantic. However, it might not be applicable to a – even just slightly – finer grid spacing of 6.5 km and over land, where the 900 hPa pressure level might be influenced by orography and the boundary layer as mentioned prior. The dependency of horizontal resolution should be tested further. While the detection of Shapiro-Keyser cyclones using the warm seclusion before confirming an SJ is a useful way to reduce computational cost and decrease false alarms, it makes the method dependent on the full evolution of a cyclone since the SJ usually occurs earlier than the development of the warm seclusion (Section 2.2.3). Furthermore, for temporally low resolution data, the warm seclusion and, hence, a possible SJ, might be missed. The SJ detection is further dependent on – partly subjectively chosen – thresholds, which seem quite ambitious. Even with lowered thresholds, the SJ of storm *Eunice* was not detected, despite trajectory analysis showing a large and intense descending air stream. Similar results were also found for storm *Egon* and storm *Friederike* (not shown). While the overall idea and reasoning behind each step are comprehensible and seem promising, these subjectively chosen thresholds and further mentioned downsides seem to lead to a high miss rate, at least for the data set used here.

7.5 Approach for an own SJ identification

While the discussed methods might be good indicators of SJ occurrence for the data sets they were developed for, they do have limitations for wider applications as discussed prior. This motivates to investigate how some ideas of these methods can be used to compliment RAMEFI in identifying an SJ. In the scope of this section, a strategy for an own method is discussed, and first results are shown.

Firstly, the region around the cyclone centre is extracted following both Manning et al. (2022) and Gray et al. (2021) with a radius of 700 km. Secondly, analogously to Manning et al. (2023), the lower, middle and upper pair of $\Delta_z v$ are calculated for regions where $v_{850} > 30\,\mathrm{m\,s^{-1}}$, similar to the threshold for starting positions of trajectories. However, the computation of streamlines was abandoned as they are not deemed necessary. Regions of adjacent grid points with negative $\Delta_z v$ need to have a certain size to reduce noise. Furthermore, instead of the overlap of positive sequences with negative ones as done in Manning et al. (2023), an overlap of regions with a negative gradient is sought. This is supposed to identify areas of high winds. No threshold is set for the proportion of overlap for a single contiguous area, i. e., it is detected if some part of an area in one pair overlaps with an area in another pair. However, at least 10 % of considered grid points have to overlap

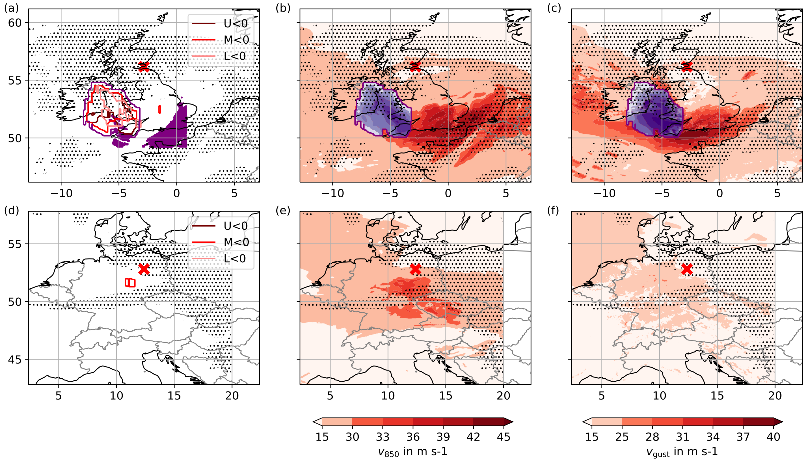


Figure 7.9: SJ identification using areas of negative $\Delta_z v$ in three level pairs (U – upper pair, M – middle pair, L – lower pair) analogously to Manning et al. (2023) for (a) – (c) storm *Eunice* (18 February 2022, 11 UTC) and (d) – (f) storm *Xavier* (05 October 2017, 12 UTC). The first column shows smoothed areas of negative $\Delta_z v$ for each pair in different shadings of red and an enveloping contour in purple if an SJ is detected. Purple dots are SJ trajectory starting positions (see Section 7.2). The second and third columns show the wind speed at 850 hPa and gust speed in 10 m, respectively, in red shading for non-SJ grid points and purple shading where an SJ has been detected. Black dots indicate $RH_{\text{ice}} > 80\,\%$ and the red cross marks the cyclone centre.

with at least one other pair. This also allows for more diagonally located regions of high wind speed. To ensure these regions extend up to 600 hPa at least 5 % has to be detected in the upper pair. In contrast, a CJ stays at low levels and a region of high winds usually does not exceed 800 hPa.

As in Gray et al. (2021), signals in the warm sector or along the frontal structure have to be filtered out. While the precursor tool uses two criteria involving θ_W, here, RAMEFI is used. RAMEFI considers CJ and SJ together, such that the SJ is expected in regions of high probabilities for CJ. However, as discussed in Sections 5.4.1 and 5.5, the CS and CJ cannot always be distinguished with certainty. Hence, grid points with high probabilities for both CJ and CS are considered, while WJ and CFC are excluded.

Figure 7.9a shows the resulting areas of negative sequences for each pair in red contours. If they meet the overlap criteria mentioned above, an SJ is detected and the area encircled as seen here by the purple contour. The northern part of this area, which is supposed to be the origin of the SJ, coincides with high values of RH_{ice}, hence is located within the cloud head. The tip of the area also coincides with starting points of trajectories. Note that the area of the starting points is quite small compared to the identified SJ here due to the different v_{850} criteria used – while here we use $v_{850} > 30\,\mathrm{m\,s^{-1}}$, a higher threshold of $37\,\mathrm{m\,s^{-1}}$ is used for the trajectories to focus on the core of the SJ and reduce the computational cost. A second area of starting positions farther east is not identified here. However, while high winds occur at 850 hPa (Figure 7.9b), the high momentum does not completely reach the surface as visible in the surface gusts in Figure 7.9c. As suggested by Volonté et al. (2023b) and Section 7.2, parts of these starting positions are a dry intrusion rather than an SJ. However, the contour does not seem to include the whole area of the SJ. Looking at Figure 7.8 shows that the region that is not detected as an SJ here – but is detected using trajectories – is connected to high wind speeds from the upper troposphere. Therefore, it is not a distinct region of high winds with a negative $\Delta_z v$, but a region showing increasing v with height. Nevertheless, the method was able to detect a large portion of the SJ and is doing so for every time step (not shown).

To ensure this method is actually detecting SJs and not CJs, it was also applied to storm *Xavier*. While some time steps show regions of negative gradients in some of the pairs, the additionally chosen criteria ensure that the CJ of *Xavier* is not identified as an SJ. An example time step – 12 UTC – is shown in Figure 7.9d–f. Although the middle pair shows regions of negative $\Delta_z v$ (Figure 7.9d), it is not overlapping with another pair, and is thus not identified as an SJ.

The tool was also tested on further SJ and non-SJ storms, such as storms *Egon* (Eisenstein et al., 2020), *Friederike* and *Eberhard* (see Section 5.1) and an exemplary time step for each case can be found in Appendix C (Figure C.1). As envisaged, the method detects an SJ for *Egon* and *Friederike* but none for *Eberhard*. Although these are promising results, the method has to be properly evaluated using data from further storms, also including Norwegian cyclones, as they are not excluded by the identification of a warm seclusion as in the method of Manning et al. (2022, 2023) in order to be independent of time and the full evolution of a cyclone. A downside of this approach is that CJ and SJ cannot be separated when they are merged, as is often the case in the last stages of an SJ. For these cases, trajectories are still needed to confirm their origin. However, the method shows higher temporal accuracy as the precursor tool, which only considers the region within the cloud head.

7.5.1 Possible strategies for further refinement and evaluation

While the method shows first promising results, the criteria are quite simple and might cause a high false alarm rate. Some further ideas to prevent this are discussed here. However, they have not been tested yet and might also lead to misses and thus may need tuning using a sufficient number of cases.

- **Origin within the cloud head:** The SJ has its origin within the cloud head and usually shows a decrease in RH (and RH_{ice}) during its descent. This is also used in the methods by Manning et al. (2022, 2023) and Gray et al. (2021). As RH does not always show a strong decrease, e. g., in the case of strong evaporative cooling as happened in storm *Egon* (Eisenstein et al., 2020), an RH criterion might cause a higher miss rate. However, a criterion to ensure the origin within the cloud head is reasonable and should be included. For this, the northern/northwestern part of the identified area, which is associated with the originating region of an SJ, has to be defined and should meet a criterion of $RH_{\text{ice}} > 80\,\%$ at a certain level, e. g., 500 hPa, or layer (taking the average).

- **Momentum transfer to the surface:** Most methods – including trajectory analysis – only identify SJs in the lower troposphere above the boundary layer. Due to turbulence, trajectories cannot be used reliably within the boundary layer. Another lower pair of pressure levels would probably not show a negative gradient due to surface friction and the lowest pressure level in models would also not be entirely trustworthy, similarly to the comments on the 900 hPa layer. However, Hewson and Neu (2015) suggest that for SJs, which actually reach the surface, $v_{\text{gust}} \approx v_{850}$. In contrast, for the CJ they state that v_{gust} is usually weaker than v_{850}. This difference in the two features can be used here.

- **Probabilistic identification:** Similar to RAMEFI, a probabilistic identification could bring the benefit of handling forecast uncertainty and estimating the risk of an SJ. However, the number of SJ storms is quite low, such that training of an RF as for RAMEFI might not yield satisfactory results. A simple idea would be the contribution of various criteria to a point scale leading to a probability. While this idea would bring many benefits, a proper test of different methods and evaluation will be quite challenging.
- **Mesoscale instabilities:** As discussed, not all SJ storms are equally affected by CSI. Nonetheless, it could be used as an additional criterion. This could for example be built into a probabilistic detection similar to Gray et al. (2021). Furthermore, conditional instability could indicate an occurring momentum transfer down to the surface (e. g., Volonté et al., 2018; Eisenstein et al., 2020).
- **Excluding dry intrusions:** As seen in the example of storm *Eunice*, dry intrusions can bring strong winds from the stratosphere or upper troposphere down to the lower troposphere. With only two criteria for the selection of SJ trajectories (v and Δp), the dry intrusion sometimes cannot be distinguished. However, a dry intrusion is usually characterised by high PV values of 2 PVU or higher (e. g., Raveh-Rubin, 2017). This can be used to detect the dry intrusion as a separate feature – also within the scope of this method.

- **Dependency on resolution:** So far, the method was only tried on 6.5 km grid spacing, which corresponds to the grid spacing of the deterministic ICON-EU forecast at DWD. However, an SJ is a small-scale and short-lived feature, and its development and possible momentum transfer to the surface can easily change with small disturbances. Hence, considering weather forecasts, it would be of interest to get an SJ potential similar to Gray et al. (2021), which provides a spatial SJ probability based on the SJ detection in each ensemble member of an EPS forecast. The ICON-EU-EPS has a grid spacing of 13 km, while ICON-D2-EPS has one of 2.2 km. After a proper evaluation of the final method, it should further be tested for robustness to resolution changes and, hence, the usage of one or both EPS to gain a probability of occurrence by detecting possible SJs in each member.

- **Choosing thresholds:** Here, several thresholds concerning the size of regions and wind speed were chosen for preliminary tests. These thresholds have to be evaluated and adjusted to be more objective and prevent over-tuning to the cases and data set used here.

7.6 Conclusions

Storm *Eunice* was an exceptionally intense winter storm in February 2022 that developed an SJ and is therefore an interesting recent case to test and compare different SJ (potential) identification methods. Section 7.2 confirmed the occurrence of an SJ for several hours over the North Atlantic, the UK, the North Sea and Benelux states using an ICON-LAM simulation with a grid spacing of 6.5 km. The trajectory analysis further revealed a dry intrusion on the eastern flank of detected SJ trajectories. The same simulation was used for the application of the two alternative approaches. It remains to be seen if the methods can be used on data sets other than the original ones. While the precursor tool by Gray et al. (2021) uses a global ensemble forecast model with a grid spacing of 20 km, the kinematic approach by Manning et al. (2022) was developed for a very fine grid spacing of 2.2 km. The methods are fundamentally different in their used resolution and, hence, can work with different characteristics of an SJ.

The ICON simulation used here has a much finer resolution than what the precursor tool was built for. Even if regridded to lower resolutions, the mesoscale instability release resolved naturally affects the dynamics during the evolution of the storm. This leads to consistent regions of CSI points in all resolutions, suggesting that regridding the data before application does not lead to substantial losses, but rather has the big advantage of saved computational cost. While enough CSI points meet the precursor criteria for storm *Eunice*, they are also met for storm *Xavier*, which did not develop an SJ. Moreover, the conditions for *Eunice* are met for a longer time period than the SJ actually occurred according to the trajectory analysis. Comparing the detected CSI points to MPV points shows a substantial overestimation of CSI within the cloud head. While many MPV points are found for storm *Eunice*, almost none appear during the life cycle of storm *Xavier*. On one hand, this agrees with the main idea that CSI can be used as an indicator for SJs. On the other hand, it suggests that *DSCAPE* is not a suitable choice to indicate CSI or that the other chosen criteria are too weak. Overall, the results imply a high false alarm rate. However, it should be noted that these results might be due to the high resolution. For a proper analysis, the tool should be tested on the ICON-EPS or ICON-EU-EPS forecasts in an operational framework.

In contrast to the precursor tool, the kinematic approach seems to have a high miss rate, probably due to over-tuning to the used data set and subjective choices for thresholds of most criteria. While streamlines showing a reversal in $\Delta_z v$ are detected at first, they mostly disappear after filtering them using RH_{ice} and v differences between positive and negative sequences. Furthermore, the last criterion of overlap between the different levels added in Manning et al. (2023) does not result in a detection of an SJ in any time step, even when using the streamlines not filtered by RH_{param} and v_{param}. The assumption of the SJ being a diagonal region in an Eulerian view does not seem valid in several storms – at least at resolutions lower than the 2.2 km used in their study. Hence, an overlap of positive and negative sequences from different levels cannot be matched. Again, as the resolution used here is different from the original paper, this analysis would benefit from an analysis using the matching resolution, which in this case is given by ICON-D2-EPS. Although both approaches are based on promising ideas, they have disadvantages if used as a definite detection tool for SJs in an operational framework. Further evaluation has to be done using the respective resolutions to clarify if the methods are too data-dependent to be used flexibly. However, some concerns were raised regarding some criteria leading to too high false alarm or miss rates. To get a reliable, low-cost identification of SJ (potential), it might be beneficial to combine ideas from both methods into a new tool, which might suit as an extension of the RAMEFI method.

The basis of the new tool is the idea from Manning et al. (2022) to search for regions of negative $\Delta_z v$. However, the computation of streamlines and the overlap of positive and negative sequences are abandoned. Instead, a partial overlap of negative sequences is sought. RAMEFI is used to exclude signals in WJ and CFC regions, making the θ_W criteria and location within 100° to 300° of the cyclone in Gray et al. (2021) unnecessary. Thus far, the tool was tested on three SJ storms and two non-SJ storms, detecting all SJs with no false alarms. However, a full evaluation is yet to be done including the application on different models, resolutions and areas. Furthermore, a strategy for implementing further criteria based on typical SJ characteristics was proposed to make the tool more robust. Nevertheless, the method shows promise as a suitable addition to RAMEFI to identify all introduced wind features. This is of particular interest as it would also enable a forecast error analysis for the SJ, an aim which inspired this work.

8 Conclusions

Strong winds associated with extratropical cyclones can cause severe damage and casualties in the mid-latitudes, especially in winter time. It is therefore of interest to forecast them accurately and give adequate warnings such that people in affected areas can prepare and stay safe. However, the past has shown that forecast busts can happen, as for example during storm *Friederike* in January 2018, when the forecast quality varied for different subregions within the storm. An extratropical cyclone can be accompanied by several mesoscale features: the warm jet (WJ), the cold jet (CJ), the sting jet (SJ) and cold-frontal convection (CFC). In addition, high winds usually also occur within the cold sector (CS) caused for example by dry intrusions and post-CFC. As all of these features have distinctive characteristics, such as their occurrence relative to the cyclone centre and to the cyclone life cycle (e. g., Hewson and Neu, 2015), it is hypothesised that they also have different forecast error characteristics. The main aim of this work was to lay a solid foundation to further investigate this hypothesis through the development of new tools and climatological analysis.

Previous studies that concentrated on wind features within extratropical cyclones used subjective approaches, did not separate all introduced features and/or only covered a short time period. However, to properly analyse characteristics, forecast errors and possibly develop a feature-dependent post-processing, an objective identification tool is needed. Chapter 5 discusses the development of such a detection method using a probabilistic random forest (RF) called RAMEFI (RAndom-forest-based MEsoscale wind Feature

Identification). RAMEFI is trained on subjectively set labels for each feature within 12 winter storm cases and includes only eight surface parameters based on pressure, temperature, precipitation and wind as predictors. It is spatially independent and, additionally, only requires the current and prior time step due to the tendency predictors (Δp, $\Delta\tilde{\theta}$, Δd); a temporal evolution over several hours is not needed. This ensures a great flexibility of the tool.

Chapter 5 tackled the overarching research question on how high wind features can be objectively identified and lead to the following answers to the individual research questions posed in Chapter 3:

RQ1a | Can the introduced wind features – WJ, CJ, SJ, CFC and CS – be detected reliably using only surface data?

RAMEFI is based on an observational data set with hourly resolution including only a few surface parameters, namely mean sea level pressure p_{msl}, normalised wind speed at 10 m $\tilde{v}$, precipitation amount RR, normalised potential temperature $\tilde{\theta}$, wind direction d and the temporal tendencies Δp_{msl}, $\Delta\tilde{\theta}$, Δd. The evaluation of the method showed a satisfactory differentiation between WJ, CFC, CJ and CS and overall good agreement with the subjective identification of the different wind features only using surface data. However, the distinction between CJ and SJ using only surface parameters remains challenging, as characteristics such as occurrence relative to the cyclone life cycle and centre are partly coinciding. Therefore, the two features are considered together in the scope of RAMEFI. Furthermore, RAMEFI shows at times uncertainty in the distinction between CJ and CS, as again, their characteristics in the available parameters are similar, both being located within the cold sector of a cyclone.

RQ1b | Is it possible to develop a method independent of horizontal organisation of the input data, hence, applicable to both station observations and gridded data without additional adaptations?

To eliminate location-specific effects and data-set-specific biases, $v_{10\mathrm{m}}$ and $\theta_{2\mathrm{m}}$ are normalised by their 98th percentile and median, respectively, and location-based predictors, such as latitude, longitude and altitude, are not considered. Both of these properties enable the application of RAMEFI to other data sets, as shown by the example of a gridded reanalysis data set. However, the application over the ocean has not been properly evaluated, and the wind distribution differs significantly. Nevertheless, looking at various cases shows no substantial obvious misclassifications. The flexibility of RAMEFI brings many advantages, however, also some limitations. For example, the observational data set used for training is distributed inhomogeneously over Europe making it impossible to compute a horizontal temperature gradient that could separate the warm and cold sector of a cyclone. Instead, $\tilde{\theta}$ is used as an indication of the sectors. Although leading to a less clear separation of warm and cold sector, the WJ shows warmer $\tilde{\theta}$ values compared to CJ and CS, hence, the parameter sufficiently helps to identify the features. Furthermore, the proximity to a cyclone centre, which is an important factor to distinguish CJ and CS, is also only given indirectly by p_{msl}. Nevertheless, it yields satisfactory results for most cases and time steps. Double fronts or convective lines cause high uncertainty – indicated by overall low probabilities – as regions in between fronts might show characteristics of both cold and warm sector. However, these situations are not clearly defined and meteorologists themselves would have problems identifying high wind features in such a region.

RQ1c | Can a machine learning approach learn distinctive characteristics to distinguish the features?

RAMEFI is based on a probabilistic RF trained on subjective labels of 12 case studies. Only using the above mentioned parameters, the RF learns meteorological consistent characteristics also reflected in the predictor importance. For example, the RF suggests a higher probability of WJ occurrence with negative Δp_{msl} as expected ahead of the cold front and positive values behind it. Furthermore, the RF expects higher precipitation rates for CFC compared to all other features and the CJ to occur closer to the cyclone centre than the adjacent CS. While RAMEFI is able to distinguish the features based on many properties used for the subjective labelling, some characteristics, such as the hook-shaped structure of a CJ, cannot be learned due to the desired spatial independence. Yet, RAMEFI is able to distinguish the features satisfactorily based on the meteorological nature of each feature, and a big advantage of this approach is the probabilistic output, which also takes into account case-to-case variability and, hence, uncertainties of a given dynamical situation.

A flexible objective method such as RAMEFI with low computational cost, can be easily applied to a long-term data set if hourly data for the necessary parameters are provided, and $v_{10\mathrm{m}}$ and $\theta_{2\mathrm{m}}$ can be normalised by their 98th percentile and median, respectively, to remove location-specific effects and seasonal and diurnal cycles.

In Chapter 6 we applied RAMEFI to a high-resolution reanalysis data set (COSMO-REA6) over Western and Central Europe over 19 extended winter seasons (October – March) to tackle the second set of research questions on how the high-wind features are characterised. Occurrence frequencies were also compared to the observational data set for the same time period. RAMEFI probabilities were extracted around the core of a tracked cyclone

(±15° in zonal and $-15°$ to 5° in meridional direction) to ensure the focus on winter storms. Earth-relative and system-relative occurrences of the features – both temporal and spatial – were analysed, as well as characteristics in other parameters than the ones used as predictors for training. Overall, the objective climatology in Chapter 6 provides the following answers to the posed research questions:

RQ2a | Are results obtained by an objective approach consistent with previous (subjective) literature?

The results are mainly consistent with previously proposed characteristics (e. g., Hewson and Neu, 2015; Earl et al., 2017; Gentile and Gray, 2023). The features occur in the expected regions relative to the cyclone centre as proposed in the used idealised schematic displayed in Figure 2.3. However, the location of the CFC associated with the cold front varies from case to case and with the stage of a cyclone. As suggested by Hewson and Neu (2015), the WJ occurs before the CJ. The CJ commonly causes the highest winds and gusts, while WJ and CS show weaker wind speeds but affect a broader area. As expected, CFC shows a high gustiness consistent with its turbulent nature.

RQ2b | What new aspects can an objective long-term climatology bring?

In contrast to many previous studies, CFC and CS are considered separate features and are not merged with WJ and CJ, respectively, and RAMEFI enables us to extract distinct characteristics of these features. Furthermore, RAMEFI allows us to allocate robust numbers of occurrence. The most common cause of high winds is the CS followed by WJ. CFC, on the other hand, is a narrow feature only affecting a small area at a certain time step and, hence, is less common. However, it shows a strong gustiness. It is shown that CFC occurs a little later than the WJ in a cyclone's life cycle. Furthermore, CFC mostly occurs over land. This might be due to differences in surface heating over ocean and land among other things. However, these result should be treated with some caution, as RAMEFI has not been properly evaluated over ocean grid points.

RQ2c | Which European regions are commonly more affected by which feature?

The most common features to cause high wind in a region are dependent on the cyclone track. Cyclones are typically eastward-moving systems. Reaching Western Europe, they can be in any stage of a cyclone's life cycle. Hence, all features can occur in this region. In contrast, Eastern Europe mostly experiences cyclones in their final stage, when most features have vanished, such that this region is mostly affected by CS. Just south of the highest track density over Scotland, the North Sea, southern Scandinavia and northern Central Europe, the CJ is most common as it is the feature occurring closest to the cyclone centre.

RAMEFI allows for a broad overview of the features in a consistent way. Only the SJ cannot be analysed in a climatology yet. Although the SJ is a rather rare and small-scale feature with characteristics similar to a CJ at the surface, it is suggested that the SJ creates a higher wind risk with higher gust speeds (e. g., Hewson and Neu, 2015). Therefore, a suitable detection algorithm is sought for operational application but also for a climatology to improve our understanding of the feature and how it might change in the future.

In recent years, two objective SJ (potential) detection methods proved to be promising: a mesoscale-instability-based precursor tool (Gray et al., 2021) and a kinematic streamline-based approach (Manning et al., 2022, 2023). While they have been developed for different applications – one for forecasters using a low resolution and the other for convection-permitting, i. e., high-resolution, climate simulations –, both seemed suitable for our intentions. The basis of the detection of the precursor tool is *DSCAPE* as an indication of CSI, which is known to commonly contribute to SJs. In contrast, the kinematic method is looking for a reversal in the vertical gradient of the horizontal wind speed in several levels to detect a slanted region of high wind speeds.

Chapter 7 investigated the set research questions on which low-cost approach can be used to identify SJs by comparing the two introduced methods. However, the results were unsatisfactory for our used data set, which led us to combine ideas from both tools and further literature but also including RAMEFI. So far, the method is based on a very simple approach of finding three-dimensional areas of high wind speed in the lower troposphere, similar to the main idea of Manning et al. (2023). A strategy for further possible implementations was outlined, and a proper evaluation is needed to confirm the promise of such a simple and low-cost method complimenting RAMEFI. The work presented in Chapter 7 leads to the following answers to the set research questions:

RQ3a | How robust are the two SJ identification approaches in regard to model choice and resolution?

The precursor tool, originally used on ensemble forecasts with 20 km grid spacing – and regridded to 0.5° –, seems to overestimate the number of CSI points for both SJ and non-SJ storms. The results suggest a high false alarm rate for high-resolution data, even after regridding, in comparison with conventional CSI detection (here MPV points) for model simulations with 6.5 km grid spacing. This grid spacing is high enough to resolve mesoscale instability, hence, the amount of *DSCAPE* might be increased. Therefore, the tool is not applicable for higher-resolution data. However, it might be more suitable for the ICON-EPS or ICON-EU-EPS forecasts (26 km and 13 km, respectively). Yet, a high false alarm rate can still be expected due to the differences between detected CSI and MPV points.

For the streamline-based detection, the opposite is the case. While the first step results in a promising amount of streamlines showing a reversal in the vertical gradient of horizontal wind speed, the candidate points are subsequently removed by further criteria, suggesting a high miss rate.

RQ3b | How do they perform over land, especially regions with complex topography?

The refined streamline approach by Manning et al. (2023) is based on four pressure levels to calculate three pairs of wind speed gradients with the lowest level being 900 hPa. Considering a low-pressure system and orography over land, this level might already be affected by boundary layer effects or the surface itself. Hence, streamlines in the lower pair are likely unreliable at times and the detection of SJs might be affected.

In contrast, the precursor tool does not show substantial differences between ocean and land, as its focus is on the cloud head, which is barely affected by land effects. Hence, it can be equally used over ocean and land.

RQ3c | What recommendations can be given to the DWD for the development of an operational tool based on the current operational system?

Using the precursor tool with the right resolution, it can serve as an indication of SJ occurrence for forecasters. However, the region affected by the SJ after its descent cannot be determined and no information on downward transport of momentum to the surface can be provided. It is solely useful to alert forecasters that the storm might cause exceptionally high gusts that are often underestimated in forecasts. On the other hand, the streamlines from the kinematic approach are able to give a more specific location of the SJ itself. Nevertheless, the presented results in this work showed that the region with detected streamlines in each level are rather spotty and not a coherent air stream, making pinpointing the location more difficult. Overall, a combination of the methods and using further characteristics of SJs could offer the most favourable strategy while retaining a low cost compared to Lagrangian trajectories, which provide the most detailed analysis of SJ but have high computational costs.

In conclusion, this dissertation presented a novel method to objectively identify different high-wind features within extratropical cyclones – the WJ, the CJ+SJ, CFC and CS winds – in a probabilistic manner. This method enabled the first-ever objective long-term climatology of these features. The results provide extended knowledge on the characteristics of each feature obtained in a consistent way. While the SJ is part of the CJ feature in RAMEFI, a strategy to extend the tool by a separate SJ identification algorithm was provided with first promising results. Overall, this work lays the foundation to further investigate the different wind features and their effects in various ways, from improving the forecast of wind gusts to analysing possible changes in a warming climate.

9 Outlook

With an objective identification tool like RAMEFI – and a reliable SJ-detection extension in the near future – a variety of new developments can be pursued, first and foremost a feature-dependent forecast error analysis. Since the discussed features are prone to different dynamics and effects within a cyclone, it is hypothesised that these differences result in varying forecast errors. Especially features characterised by small-scale dynamics and instabilities that might not be resolved in a forecast, such as CFC or the SJ, are believed to have higher forecast errors compared to larger-scale features like the WJ. If this turns out to be true, a next step would be a feature-dependent post-processing method. As winds are commonly overestimated (e. g., Hess, 2020), many post-processing methods are trained to reduce the winds. However, this might lead to forecast busts, as seen in the example of storm *Christian* (October 2013; Pantillon et al., 2018) and storm *Friederike* (January 2018) over Germany. Both of the storms were accompanied by an SJ and at least for *Friederike* the forecast busts have been confirmed to be associated with the SJ (not shown).

RAMEFI output can be used in different ways for a post-processing method. Either the feature with the highest probability is assigned to a grid point, or the probabilities themselves are used. For example, the probabilities can be used as additional predictors, allowing the use of a single post-processing method for all grid points. Another approach would be to train a post-processing method for each feature independently and use it for the respective detected feature. In any case, note that RAMEFI should not be applied to the ensemble mean, as features like CFC characterised by a narrow line will not be represented well in the mean. Instead, RAMEFI should be applied to the single members. From there, it can either be tested, if the occurrence probability taking all members together is suitable or if a post-processing method should be used that works for each member individually instead of only the ensemble mean.

Additional to improving the forecast through post-processing, RAMEFI can be used to assist forecasters in their decision-making, similar to the SJ precursor tool (Gray et al., 2021) at the UK Met Office. While first approaches were already shown in Section 5.6 using Met.3D in combination with three-dimensional front detection (Beckert et al., 2023) for a deterministic forecast, the usage of ensemble forecasts might be of particular benefit. Besides purely weather-forecast-related applications, RAMEFI can also be used for impact research (Merz et al., 2020), possibly linking certain damages and socioeconomic impacts to the different mesoscale wind features and helping with risk assessment and management.

Furthermore, RAMEFI was built to be a flexible tool, aiming to be independent of location-specific effects. Therefore, it can in theory be used for other regions in the mid-latitudes that are commonly affected by extratropical cyclones. However, this was not tested in the scope of this work. Applying RAMEFI to different regions of the world might show differences in the development of cyclones and occurrence of features. Moreover, a detailed analysis of data over the ocean is needed. Although results for several case studies do not show a substantial difference in feature detection between land and ocean (Eisenstein et al., 2023b), the wind distribution is significantly different and, hence, the threshold of $\tilde{v} > 0.8$ used here allows for more detections over the ocean compared to land grid points. Furthermore, $\tilde{\theta}$ might behave differently as land surfaces heat up more strongly than ocean surfaces during the day. A detailed analysis of the differences will help to learn more about the dynamics of the features and might suggest an additional training for ocean grid points. Data and software have been published (Eisenstein et al., 2022c) to allow other researchers to easily apply the new developments to their specific regions and questions.

RAMEFI can further be helpful in the SJ identification, as discussed in Chapter 7. The presented approach is based on a simple and computationally cheap idea following previous SJ detection methods. While the tool shows first promising results, many ideas were outlined to make the tool more robust, and a proper evaluation of the method is planned. These include the consideration of further criteria, such as RH_{ice} and mesoscale instabilities, and trying to include the downward transport of momentum to the surface. Once the tool shows reliable results, a climatology similar to the one described in Chapter 6 can be produced. Furthermore, a first global SJ climatology can be compiled tackling the question if SJs also occur in other regions of the world and if so, whether they behave differently. It is of further interest to evaluate the increased wind risk of the SJ, especially compared to the CJ, in an objective way.

Additionally, the already published approaches by Gray et al. (2021) and Manning et al. (2022, 2023) should be evaluated on further data sets. The results presented here might – partly – be caused by the horizontal resolution, which differs from the ones originally used. It is suggested to further explore both methods using the ICON-EPS or ICON-EU-EPS for the precursor tool and ICON-D2(-EPS) for the kinematic method. While feature-dependent post-processing might not be feasible for training on SJs due to their rather rare occurrence, a reliable SJ identification would be beneficial as a guidance for forecasters, as discussed for RAMEFI above.

Being able to detect the WJ, the CJ, CFC, CS and possibly the SJ covers the most important wind hazards associated with winter storms. However, investigating further dynamical features such as dry intrusions is of interest, too. Similar to the SJ, they can descend to the top of the boundary layer – as seen in storm *Eunice* – and in some cases also to the surface, causing high wind gusts. Furthermore, Section 6.2.3 showed that parts of NF, which was detected north to northeast in the proximity of the cyclone centre, could be regarded as the CJ before it wraps around the cyclone centre, which is referred to as CCBa in recent literature (e. g., Earl et al., 2017; Gentile and Gray, 2023). Although it is uncommon for the CJ to cause high wind speeds while travelling against the cyclone motion, this study showed a distinct occurrence frequency, which makes it desirable to treat this region as a separate feature.

Lastly, climate change studies can benefit from RAMEFI. Previous studies suggested that the changes in wind speed are associated with winds within the warm sector (Priestley and Catto, 2022) and the number of SJs (Martínez-Alvarado et al., 2018; Catto et al., 2019). While the overall number of cyclones is shown to decrease in the future climate, the number of extreme cyclones is suggested to increase (Priestley and Catto, 2022). As changes in, for example, moisture and temperature might also affect the dynamics of the wind features, the occurrence frequency and intensity might also change.

In summary, this work presents a great opportunity for exploration and advancement in various meteorological fields, such as weather forecasting, impact research and climate change studies. The flexibility of RAMEFI facilitates its application to a wide range of data sets, while the probabilistic output it generates is of great benefit for diverse analytical pursuits.

A Appendix to Chapter 5

A.1 Implementation details

In this section, we provide technical details on the evaluation of the RF probabilities, the probabilistic RFs and the Kriging.

A.1.1 Evaluation of the RF probabilities

The CORP reliability diagrams were generated using the `reliabilitydiag`-package (Dimitriadis et al., 2021), the PDPs using the `pdp`-package (Greenwell, 2017). Each reliability diagram is based on the RF probabilities and observations of all storm cases. For the PDPs, one partial dependence curve has to be calculated for each RF generated in a fold of the cross-validation, that is, for each winter storm. The final curves are then obtained by a weighted average depending on the sample size of the folds.

A.1.2 Random Forests

RF is implemented via the `ranger`-package (Wright and Ziegler, 2017) in `R` (R Core Team, 2021). Table 4.1 summarises the predictors used, Table A.1 the chosen tuning parameters. One question in the implementation is the handling of missing values, which an RF cannot process. The station-based samples frequently miss values of one or more predictor variables, especially precipitation is affected. We tried different strategies to handle missing values such as leaving out instances with missing values or replacing the missing values with a mean value and found similar results. Therefore, we decided to replace the missing values in order to use the largest sample

Table A.1: Overview of the tuning parameters of the probabilistic RF.

Tuning Parameter	Value
Number of trees	1,000
Number of predictors considered at each split	2
Minimal node size	10
Maximal depth	Unlimited
Splitting criterion	Gini

size possible, which is desirable for the evaluation and the Kriging step. In each fold of the employed cross-validation scheme, the missing values (both in the training and test set) are replaced by the mean value of the associated predictor variable in the training set.

A.1.3 Kriging

In our practical implementation of Kriging, we employ the `R`-packages `fields` (Douglas Nychka et al., 2017), `mvtnorm` (Genz et al., 2021), `maps` (Original S code by Richard A. Becker and Allan R. Wilks. R version by Ray Brownrigg. Enhancements by Thomas P Minka and Alex Deckmyn., 2018) and `maptools` (Bivand and Lewin-Koh, 2021). Additionally, we transform the probabilities by using the `bestNormalize` package (Peterson, 2021) to achieve approximate Gaussianity, which automatically chooses a suitable transformation from a set of commonly used transformations. The probabilities on the grid generated via the univariate Kriging need to be normalised such that they sum up to 1. However, at some grid cells distant from the cyclone track, the predicted probabilities are small for all of the wind feature and normalisation results in unrealistic predictions. Thus, we only perform the normalisation at grid cells where the accumulated probability is larger or equal to 20 %. For the visualisation, we further drop the grid cells where the largest normalised probability is smaller than 20 % (which includes the grid cells for which no normalisation was performed).

A.2 First test trainings

First tests included the SJ as a separate feature to be identified. Note that the RF was trained on more parameters, that is $v_{10\mathrm{m}}$ and the location-specific parameters longitude, latitude and altitude. We excluded $v_{10\mathrm{m}}$ as the absolute wind speed is our target variable for post-processing and the location-specific parameters to be spatially independent and be able to use RAMEFI on other regions of the world as well. Furthermore, the runs included a slightly different set of case studies, in which *Niklas*, *Susanna* and *Thomas* were missing (Table 5.1). However, storm *Diana II* (01–02 March 2020), which developed an SJ over France and Southwest Germany, was included. Ultimately, the storm was excluded in later runs as it only affected a very small area, while the rest of Europe was affected by *Diana I* and later *Diana III* and an inclusion was not advantageously after we decided to merge the SJ with the CJ feature.

From 60.572 data points, only 1 833 were labelled as an SJ. As mentioned in Section 5.2.1, an SJ was labelled when and where trajectories confirmed a descending air stream. In comparison, around 7 858 data points were labelled as CJ.

Figure A.1a shows a reliability diagram for the SJ in the one-against-all approach analogously to Figure 5.7. The probabilities are not calibrated and demonstrate an overforecast of the feature. Looking at the individual storms, it is evident that SJ is detected in both SJ and non-SJ storms (not shown). Comparing the SJ to other features, the worst calibration is found against the CJ (Figure A.1b). The SJ label was set for less than 20 % of the samples.

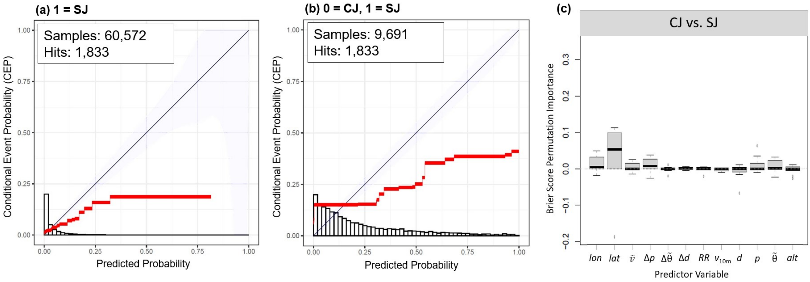


Figure A.1: Reliability diagrams for the SJ in first tests: (a) as Figure 5.7 for the SJ and (b) as Figure 5.8 for CJ vs. SJ and (c) predictor importance as Figure 5.12 but for CJ vs. SJ. Note that the training here was based on less case studies, but included a third SJ storm (*Diana II*, March 2020) and included location-specific parameters, i. e., longitude, latitude and altitude, and v_{10m}.

However, the SJ is overforecast in the CJ region, especially in non-SJ cases, suggesting that the RF actually does not learn distinct physical relations of an SJ.

It is evident from Figure A.1c that latitude is the most important predictor to distinguish CJ and SJ. This might be due to two of the SJ cases – *Egon* and *Diana II* – showing a more southern track compared to most other case studies with only the CJ occurring. Longitude acts as an indication that the SJ occurs before the CJ, i. e., in lower longitudes. Similarly, the pressure predictors can indicate the stage of the cyclone. As the distinction of the two features is not satisfactory and would be even worse without location-specific parameters, we decided to combine the SJ and CJ feature.

B Appendix to Chapter 6

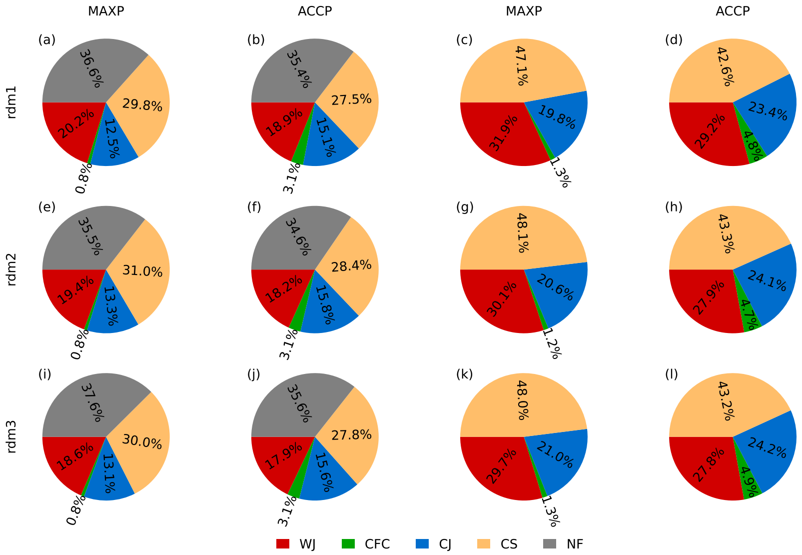


Figure B.1: As Figure 6.3m–p but for random subsets of nine winter seasons. rdm1: 2005/06, 2006/07, 2008/09, 2009/10, 2011/12, 2012/13, 2014/15, 2016/17, 2017/18; rdm2: 2001/02, 2003/04, 2007/08, 2008/09, 2011/12, 2012/13, 2013/14, 2014/15, 2016/17; rmd3: 2000/01, 2002/03, 2003/04, 2004/05, 2006/07, 2008/09, 2010/11, 2013/14, 2014/15

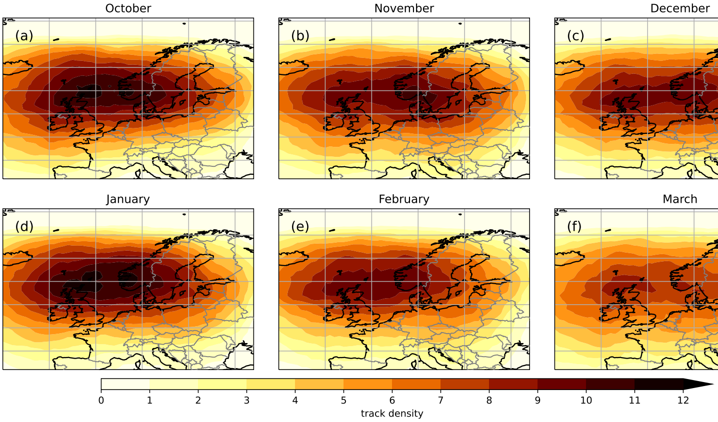


Figure B.2: As Figure 4.2 but for each month considered in the climatology.

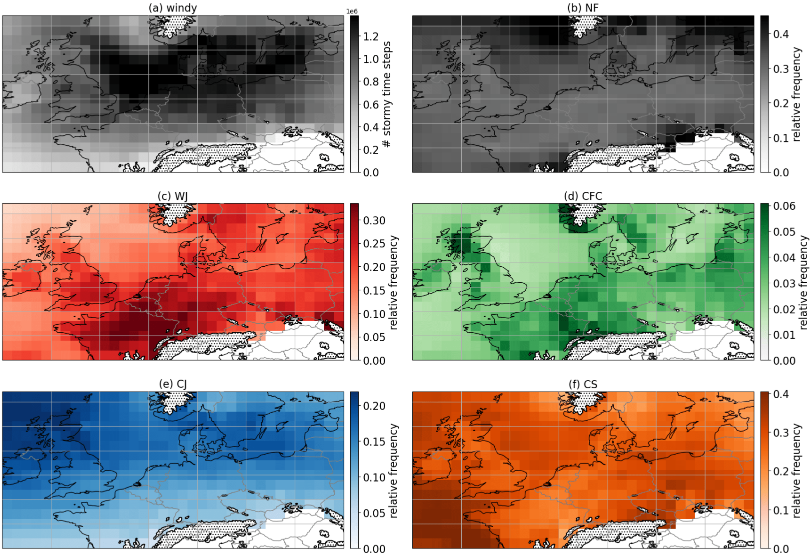


Figure B.3: As Figure 6.6 but ACCP.

C Appendix to Chapter 7

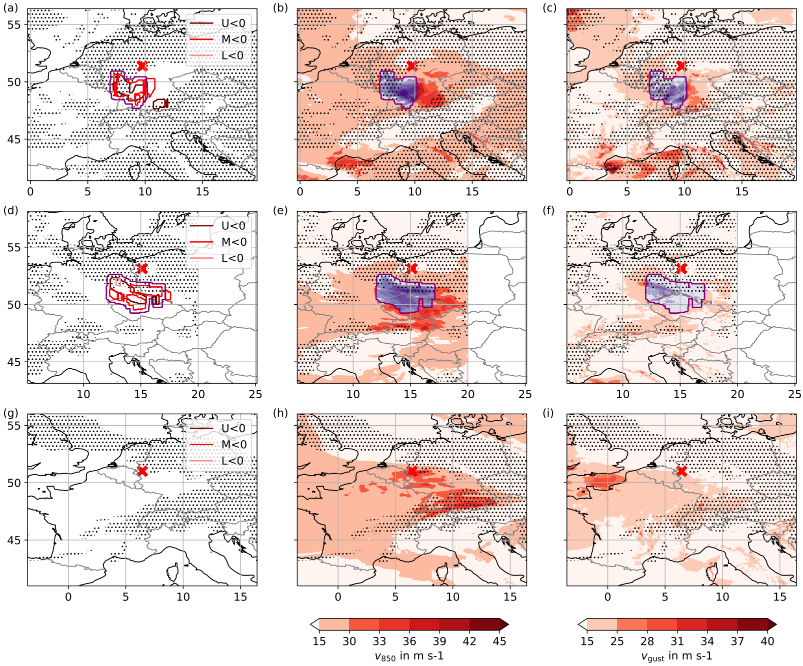


Figure C.1: As Figure 7.9 but for (a)-(c) SJ storm *Egon* (13 January 2017, 02 UTC), (d)-(f) SJ storm *Friederike* (18 January 2018, 15 UTC) and (g)-(i) non-SJ storm *Eberhard* (10 March 2019, 09 UTC).

Acronyms

ACCP	Accumulated probability
CCB	Cold conveyor belt
CFC	Cold-frontal convection
CJ	Cold jet
COSMO	COnsortium for Small-scale MOdelling
CS	Cold-sector winds
CSI	Conditional symmetric instability
DWD	Deutscher Wetterdienst (German weather service)
ECMWF	European Centre of Medium-range Weather Forecasting
EPS	Ensemble Prediction System
ERA	ECMWF Reanalysis
ICON	ICOsahedral Nonhydrostatic model
LAM	Limited Area Mode
MAXP	Most probable feature
MOGREPS-G	Met Office Global and Regional Ensemble Prediction System – Global
NAO	North Atlantic Oscillation

RAMEFI	RAndom-forest-based Mesoscale wind Feature Identification
RF	Probabilistic random forest
SJ	Sting jet
SSI	Storm Severity Index
WCB	Warm conveyor belt
WJ	Warm jet
WMO	World Meteorological Organization

p	Air pressure
Δp	Pressure difference along trajectories
p_{msl}	Mean sea level pressure
Δp_{msl}	Tendency of the mean sea level pressure
T	Temperature
v	Horizontal wind speed
$\vec{v}$	Three-dimensional wind vector
$\vec{v}_{\mathrm{h}}$	Wind velocity
$\Delta_{\mathrm{z}} v$	Vertical gradient of the horizontal wind speed
$\tilde{v}$	Normalised horizontal wind speed in 10 m
v_{gust}	Wind gust speed in 10 m
g_{v}	Gust factor
d	Wind direction in 10 m

Δd	Tendency of the wind direction 10 m
RR	Precipitation amount
θ	Potential temperature
$\tilde{\theta}$	Normalised potential temperature in 2 m
$\Delta\tilde{\theta}$	Tendency of the normalised potential temperature in 2 m
θ_{E}	Equivalent-potential temperature
θ_{E}^{*}	Saturated equivalent-potential temperature
θ_{W}	Wet-bulb potential temperature
RH	Relative humidity
RH_{ice}	Relative humidity with respect to ice
q	Specific humidity
ζ	Relative vorticity
ζ_{a}	Absolute vorticity
ζ_{z}	Vertical component of the absolute vorticity ζ_{a}
PV	Potential vorticity
MPV^{*}	Saturated moist potential vorticity
$DSCAPE$	Downdraught slantwise convective available potential energy
M_{g}	Geostrophic momentum
m	Saturation mixing ratio
ϕ	Latitude
R_d	Dry gas constant
v_{param}	Wind speed threshold for kinematic SJ detection
RH_{param}	RH_{ice} threshold for kinematic SJ detection

Bibliography

Bach, L., C. Schraff, J. D. Keller, and A. Hense, 2016: Towards a probabilistic regional reanalysis system for europe: evaluation of precipitation from experiments. *Tellus A: Dynamic Meteorology and Oceanography*, **68 (1)**, 32 209, doi: 10.3402/tellusa.v68.32209.

Baker, L. H., 2009: Sting jets in severe northern European wind storms. *Weather*, **64 (6)**, 143–148, doi: 10.1002/wea.397.

Baker, L. H., S. L. Gray, and P. A. Clark, 2014: Idealised simulations of sting-jet cyclones. *Quarterly Journal of the Royal Meteorological Society*, **140 (678)**, 96–110, doi: 10.1002/qj.2131.

Beckert, A. A., L. Eisenstein, A. Oertel, T. Hewson, G. C. Craig, and M. Rautenhaus, 2023: The three-dimensional structure of fronts in mid-latitude weather systems as represented by numerical weather prediction models. *Geoscientific Model Development Discussions [preprint]*, 1–40, doi: 10.5194/gmd-2022-278, in review.

Bivand, R., and N. Lewin-Koh, 2021: *maptools: Tools for Handling Spatial Objects*. URL `https://CRAN.R-project.org/package=maptools`, r package version 1.1-2.

Bjerknes, J., 1919: On the Structure of Moving Cyclones. *Monthly Weather Review*, **47 (2)**, 95–99, doi: 10.1175/1520-0493(1919)47<95:otsomc>2.0.co;2.

Bokeh Development Team, 2021: *Bokeh: Python library for interactive visualization*. URL `https://bokeh.pydata.org/en/latest/`, Python package version 2.3.2. Last accessed 17 May 2022.

Bollmeyer, C., J. D. Keller, C. Ohlwein, S. Wahl, S. Crewell, P. Friederichs, A. Hense, J. Keune et al., 2015: Towards a high-resolution regional reanalysis for the european cordex domain. *Quarterly Journal of the Royal Meteorological Society*, **141 (686)**, 1–15, doi: 10.1002/qj.2486.

Bolton, D., 1980: The Computation of Equivalent Potential Temperatu. *Monthly Weather Review*, **108**, 1046–1053, doi: 10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2.

Born, K., P. Ludwig, and J. G. Pinto, 2012: Wind gust estimation for Mid-European winter storms: Towards a probabilistic view. *Tellus, Series A: Dynamic Meteorology and Oceanography*, **64 (1)**, doi: 10.3402/tellusa.v64i0.17471.

Bowler, N. E., A. Arribas, K. R. Mylne, K. B. Robertson, and S. E. Beare, 2008: The mogreps short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **134 (632)**, 703–722, doi: 10.1002/qj.234.

Breiman, L., 1984: *Classification and regression trees*. Wadsworth Internat. Group.

Breiman, L., 2001: Random forests. *Machine Learning*, **45 (1)**, 5–32, doi: 10.1023/A:1010933404324.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, **78 (1)**, 1–3, doi: 10.1175/1520-0493(1950)078<0001:vofeit>2.0.co;2.

Browning, K. A., 1997: The dry intrusion perspective of extra-tropical cyclone development. *Meteorological Applications*, **4 (4)**, 317–324, doi: 10.1017/S1350482797000613.

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 (597)**, 375–399, doi: 10.1256/qj.02.143.

Browning, K. A., D. J. Smart, M. R. Clark, and A. J. Illingworth, 2015: The role of evaporating showers in the transfer of sting-jet momentum to the surface. *Quarterly Journal of the Royal Meteorological Society*, **141 (693)**, 2956–2971, doi: 10.1002/qj.2581.

Catto, J. L., 2016: Extratropical cyclone classification and its use. *Reviews of Geophysics*, **54**, 486–520, doi: 10.1002/2016RG000519.

Catto, J. L., and S. Pfahl, 2013: The importance of fronts for extreme precipitation. *Journal of Geophysical Research: Atmospheres*, **118 (19)**, 10,791–10,801, doi: 10.1002/jgrd.50852.

Catto, J. L., and S. Raveh-Rubin, 2019: Climatology and dynamics of the link between dry intrusions and cold fronts during winter. Part I: global climatology. *Climate Dynamics*, **53 (3-4)**, 1873–1892, doi: 10.1007/s00382-019-04745-w.

Catto, J. L., D. Ackerley, J. F. Booth, A. J. Champion, B. A. Colle, S. Pfahl, J. G. Pinto, J. F. Quinting, and C. Seiler, 2019: The future of midlatitude cyclones. *Current Climate Change Reports*, **5**, 407–420, doi: 10.1007/s40641-019-00149-4.

Clark, M. R., 2013: A provisional climatology of cool-season convective lines in the UK. *Atmospheric Research*, **123**, 180–196, doi: 10.1016/j.atmosres.2012.09.018.

Clark, P. A., K. A. Browning, and C. Wang, 2005: The sting at the end of the tail: Model diagnostics of fine-scale three-dimensional structure of the cloud head. *Quarterly Journal of the Royal Meteorological Society*, **131 (610 B)**, 2263–2292, doi: 10.1256/qj.04.36.

Clark, P. A., and S. L. Gray, 2018: Sting jets in extratropical cyclones: a review. *Quarterly Journal of the Royal Meteorological Society*, **144 (713)**, 943–969, doi: 10.1002/qj.3267.

Coronel, B., D. Ricard, G. Rivière, and P. Arbogast, 2016: Cold-conveyor-belt jet, sting jet and slantwise circulations in idealized simulations of extratropical cyclones. *Quarterly Journal of the Royal Meteorological Society*, **142 (697)**, 1781–1796, doi: 10.1002/qj.2775.

Dacre, H. F., M. K. Hawcroft, M. A. Stringer, and K. I. Hodges, 2012: An Extratropical Cyclone Atlas: A Tool for Illustrating Cyclone Structure and Evolution Characteristics. *Bulletin of the American Meteorological Society*, **93 (10)**, 1497–1502, doi: 10.1175/BAMS-D-11-00164.1, URL `https://journals.ametsoc.org/view/journals/bams/93/10/bams-d-11-00164.1.xml`.

Davies, H. C., C. Schär, and H. Wernli, 1991: The palette of fronts and cyclones within a baroclinic wave development. *Journal of the Atmospheric Sciences*, **48 (14)**, 1666–1689.

Davies-Jones, R., 2008: An Efficient and Accurate Method for Computing the Wet-Bulb Temperature along Pseudoadiabats. *Monthly Weather Review*, **136 (7)**, 2764–2785, doi: 10.1175/2007mwr2224.1.

Dimitriadis, T., T. Gneiting, and A. Jordan, 2021: Stable reliability diagrams for probabilistic classifiers. *Proceedings of the National Academy of Sciences*, **118 (8)**, e2016191 118, doi: 10.1073/pnas.2016191118.

Douglas Nychka, Reinhard Furrer, John Paige, and Stephan Sain, 2017: fields: Tools for spatial data. Boulder, CO, USA, URL `https://github.com/NCAR/Fields`, r package version 12.5, doi: 10.5065/D6W957CT.

Earl, N., S. Dorling, M. Starks, and R. Finch, 2017: Subsynoptic-scale features associated with extreme surface gusts in UK extratropical cyclone events. *Geophysical Research Letters*, **44 (8)**, 3932–3940, doi: 10.1002/2017GL073124.

Eckhardt, S., A. Stohl, H. Wernli, P. James, C. Forster, and N. Spichtinger, 2004: A 15-year climatology of warm conveyor belts. *Journal of Climate*, **17 (1)**, 218–237, doi: 10.1175/1520-0442(2004)017<0218:AYCOWC>2.0.CO;2.

Eisenstein, L., F. Pantillon, and P. Knippertz, 2020: Dynamics of sting-jet storm Egon over continental Europe: Impact of surface properties and model resolution. *Quarterly Journal of the Royal Meteorological Society*, **146 (726)**, 186–210, doi: 10.1002/qj.3666.

Eisenstein, L., B. Schulz, J. G. Pinto, and P. Knippertz, 2023a: Identification of high-wind features within extratropical cyclones using a probabilistic random forest – Part 2: Climatology. *Weather and Climate Dynamics Discussions*, 1–27, doi: 10.5194/wcd-2023-10, in review.

Eisenstein, L., B. Schulz, J. G. Pinto, and P. Knippertz, 2023b: Identification of high-wind features within extratropical cyclones using a probabilistic random forest – Part 2: Climatology – Video Supplement. Zenodo, doi: 10.5281/zenodo.7729357.

Eisenstein, L., B. Schulz, G. A. Qadir, J. G. Pinto, and P. Knippertz, 2022a: Identification of high-wind features within extratropical cyclones using a probabilistic random forest – Part 1: Method and case studies. *Weather and Climate Dynamics*, **3**, 1157–1182, doi: 10.5194/wcd-3-1157-2022.

Eisenstein, L., B. Schulz, G. A. Qadir, J. G. Pinto, and P. Knippertz, 2022b: Objective identification of high-wind features within extratropical cyclones using a probabilistic random forest. Part I: Method and illustrative case studies - Video Supplement. Zenodo, doi: 10.5281/zenodo.6541277.

Eisenstein, L., B. Schulz, G. A. Qadir, J. G. Pinto, and P. Knippertz, 2022c: RAMEFI (RAndom-forest based MEsoscale wind Feature Identification). Zenodo, doi: 10.5281/zenodo.6541303.

Feser, F., M. Barcikowska, O. Krueger, F. Schenk, R. Weisse, and L. Xia, 2015: Storminess over the North Atlantic and northwestern Europe—A review. *Quarterly Journal of the Royal Meteorological Society*, **141**, 350–382, doi: 10.1002/qj.2364.

Fink, A. H., T. Brücher, V. Ermert, A. Krüger, 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, doi: 10.5194/nhess-9-405-2009.

Gentile, E. S., and S. L. Gray, 2023: Attribution of observed extreme marine wind speeds and associated hazards to midlatitude cyclone conveyor belt jets near the british isles. *International Journal of Climatology*, **43 (6)**, 2735–2753, doi: 10.1002/joc.7999.

Genz, A., F. Bretz, T. Miwa, X. Mi, F. Leisch, F. Scheipl, and T. Hothorn, 2021: *mvtnorm: Multivariate Normal and t Distributions*. URL `https://CRAN.R-project.org/package=mvtnorm`, r package version 1.1-3.

Gilet, J.-B., M. Plu, and G. Rivière, 2009: Nonlinear Baroclinic Dynamics of Surface Cyclones Crossing a Zonal Jet. *Journal of the Atmospheric Sciences*, **66 (10)**, 3021–3041, doi: 10.1175/2009jas3086.1.

Gneiting, T., F. Balabdaoui, and A. E. Raftery, 2007: Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, **69 (2)**, 243–268, doi: 10.1111/j.1467-9868.2007.00587.x.

Gramlich, K., 2023: Objective identification of a sting jet by the example of storm Friederike. Karlsruhe Institute of Technology, Bachelor thesis.

Gray, S. L., O. Martínez-Alvarado, L. H. Baker, and P. A. Clark, 2011: Conditional symmetric instability in sting-jet storms. *Quarterly Journal of the Royal Meteorological Society*, **137 (659)**, 1482–1500, doi: 10.1002/qj.859.

Gray, S. L., O. Martínez-Alvarado, D. Ackerley, and D. Suri, 2021: Development of a prototype real-time sting-jet precursor tool for forecasters. *Weather*, **76 (11)**, 369–373, doi: 10.1002/wea.3889.

Greenwell, B. M., 2017: pdp: An R package for constructing partial dependence plots. *R Journal*, **9 (1)**, 421–436, doi: 10.32614/rj-2017-016.

Grønås, S., 1995: The seclusion intensification of the New Year's day storm 1992. *Tellus*, **47 A (5)**, 733–746.

Guttorp, P., and T. Gneiting, 2006: Studies in the history of probability and statistics XLIX: On the Matérn correlation family. *Biometrika*, **93**, 989–995, doi: 10.1093/biomet/93.4.989.

Hastie, T., R. Tibshirani, and J. Friedman, 2009: *The Elements of Statistical Learning*. Springer, doi: 10.1007/978-0-387-84858-7.

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey et al., 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 1999–2049, doi: 10.1002/qj.3803.

Hess, R., 2020: Statistical postprocessing of ensemble forecasts for severe weather at Deutscher Wetterdienst. *Nonlinear Processes in Geophysics*, **27 (4)**, 473–487, doi: 10.5194/npg-27-473-2020.

Hewson, T., I. Tsonevsky, L. Magnusson, and P. Bechtold, 2022: Wind gust predictions for storm Eunice. *ECMWF Newsletter No. 171 - Spring 2022*, G. Lentze, Ed., ECMWF, URL `https://www.ecmwf.int/en/elibrary/81299-newsletter-no-171-spring-2022`, last accessed 04 May 2023.

Hewson, T. D., 1998: Objective fronts. *Meteorological Applications*, **5 (1)**, 37–65, doi: 10.1017/S1350482798000553.

Hewson, T. D., and U. Neu, 2015: Cyclones, windstorms and the IMILAST project. *Tellus, Series A: Dynamic Meteorology and Oceanography*, **67 (1)**, 27 128, doi: 10.3402/tellusa.v67.27128.

Hill, A. J., G. R. Herman, and S. R. S., 2020: Forecasting severe weather with random forests. *Monthly Weather Review*, **148 (5)**, 2135–2161, doi: 10.1175/MWR-D-19-0344.

Hines, K. M., and C. R. Mechoso, 1993: Influence of surface drag on the evolution of fronts. *Monthly Weather Review*, **121**, 1152–1175, doi: 10.1175/1520-0493(1993)121<1152:IOSDOT>2.0.CO;2.

Hodges, K. I., 1994: A General Method for Tracking Analysis and Its Application to Meteorological Data. **122**, 2573–2586, doi: 10.1175/1520-0493(1994)122<2573:AGMFTA>2.0.CO;2.

Hoskins, B. J., and K. I. Hodges, 2002: New perspectives on the northern hemisphere winter storm tracks. *Journal of the Atmospheric Sciences*, **59 (6)**, 1041–1061, doi: 10.1175/1520-0469(2002)0592.0.CO;2.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quarterly Journal of the Royal Meteorological Society*, **111 (470)**, 877–946.

Hu, G., and C. L. E. Franzke, 2020: Evaluation of daily precipitation extremes in reanalysis and gridded observation-based data sets over germany. *Geophysical Research Letters*, **47 (18)**, e2020GL089 624, doi: 10.1029/2020GL089624.

Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, **269 (5224)**, 676–679, doi: 10.1126/science.269.5224.676.

Hussain, M., and I. Mahmud, 2019: pymannkendall: a python package for non parametric mann kendall family of trend tests. *Journal of Open Source Software*, **4 (39)**, 1556, doi: 10.21105/joss.01556.

Kern, M., T. D. Hewson, A. Schatler, R. Westermann, and M. Rautenhaus, 2018: Interactive 3D Visual Analysis of Atmospheric Fronts. *IEEE Transactions on Visualization and Computer Graphics*, **25 (1)**, 1080–1090, doi: 10.1109/tvcg.2018.2864806.

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 Science*, **3 (6)**, 725–732, doi: 10.5194/nhess-3-725-2003.

Lackmann, G., 2011: Chapter 6: Fronts. *Midlatitude Synoptic Meteorology: Dynamics, Analysis and Forecasting*, American Meteorological Society, Boston, Mass., 131–166.

Lagerquist, R., A. McGovern, and T. Smith, 2017: Machine learning for real-time prediction of damaging straight-line convective wind. *Weather and Forecasting*, **32 (6)**, 2175 – 2193, doi: 10.1175/WAF-D-17-0038.1.

Laurila, T. K., V. A. Sinclair, and H. Gregow, 2021: Climatology, variability, and trends in near-surface wind speeds over the North Atlantic and Europe during 1979–2018 based on ERA5. *International Journal of Climatology*, **41 (4)**, 2253–2278, doi: 10.1002/joc.6957.

Leckebusch, G. C., D. Renggli, and U. Ulbrich, 2008: Development and application of an objective storm severity measure for the Northeast Atlantic region. *Meteorologische Zeitschrift*, **17 (5)**, 575–587, doi: 10.1127/0941-2948/2008/0323.

Madonna, E., H. Wernli, H. Joos, and O. Martius, 2014: Warm conveyor belts in the era-interim dataset (1979–2010). part i: Climatology and potential vorticity evolution. *Journal of Climate*, **27 (1)**, 3–26, doi: 10.1175/JCLI-D-12-00720.1.

Manning, C., E. J. Kendon, H. J. Fowler, and N. M. Roberts, 2023: Projected increase in windstorm severity and contribution from sting jets over the uk and ireland. *Weather and Climate Extremes*, 100562, doi: 10.1016/j.wace.2023.100562.

Manning, C., E. J. Kendon, H. J. Fowler, N. M. Roberts, S. Berthou, D. Suri, and M. J. Roberts, 2022: Extreme windstorms and sting jets in convection-permitting climate simulations over europe. *Climate Dynamics*, **58 (9)**, 2387–2404, doi: 10.1007/s00382-021-06011-4.

Markowski, P., and Y. Richardson, 2010: Mesoscale instabilities. *Mesoscale meteorology in midlatitudes*, first impr ed., Wiley Blackwell, 41–70, doi: 10.1002/9780470682104.

Martínez-Alvarado, O., L. H. Baker, S. L. Gray, J. Methven, and R. S. Plant, 2014: Distinguishing the cold conveyor belt and sting jet airstreams in an intense extratropical cyclone. *Monthly Weather Review*, **142 (8)**, 2571–2595, doi: 10.1175/mwr-d-13-00348.1.

Martínez-Alvarado, O., S. L. Gray, J. L. Catto, and P. A. Clark, 2012: Sting jets in intense winter North-Atlantic windstorms. *Environmental Research Letters*, **7 (2)**, doi: 10.1088/1748-9326/7/2/024014.

Martínez-Alvarado, O., S. L. Gray, N. C. Hart, P. A. Clark, K. I. Hodges, and M. J. Roberts, 2018: Increased wind risk from sting-jet windstorms with climate change. *Environmental Research Letters*, **13 (4)**, doi: 10.1088/1748-9326/aaae3a.

Martínez-Alvarado, O., S. L. Gray, P. A. Clark, and L. H. Baker, 2013: Objective detection of sting jets in low-resolution datasets. *Meteorological Applications*, **20 (1)**, 41–55, doi: 10.1002/met.297.

Matérn, B., 1986: *Spatial Variation*. 2nd ed., Berlin:Springer-Verlag.

Matheron, G., 1963: Principles of geostatistics. *Economic geology*, **58 (8)**, 1246–1266.

McCarthy, M., S. Spillane, S. Walsh, and M. Kendon, 2016: The meteorology of the exceptional winter of 2015/2016 across the UK and Ireland. *Weather*, **71 (12)**, 305–313, doi: 10.1002/wea.2823.

McGovern, A., R. Lagerquist, D. J. Gagne, G. E. Jergensen, K. L. Elmore, C. R. Homeyer, and T. Smith, 2019: Making the black box more transparent: Understanding the physical implications of machine learning. *Bulletin of the American Meteorological Society*, **100 (11)**, 2175–2199, doi: 10.1175/BAMS-D-18-0195.1.

Merz, B., C. Kuhlicke, M. Kunz, M. Pittore, A. Babeyko, D. N. Bresch, D. I. V. Domeisen, F. Feser et al., 2020: Impact forecasting to support emergency management of natural hazards. *Reviews of Geophysics*, **58 (4)**, e2020RG000 704, doi: 10.1029/2020RG000704.

Murphy, D. M., and T. Koop, 2005: Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Quarterly Journal of the Royal Meteorological Society*, **131 (608)**, 1539–1565, doi: 10.1256/qj.04.94.

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.

Mühr, B., L. Eisenstein, J. G. Pinto, P. Knippertz, S. Mohr, and M. Kunz, 2022: CEDIM Forensic Disaster Analysis Group (FDA) ’Winter storm series: Ylenia, Zeynep, Antonia (int: Dudley, Eunice, Franklin)

February 2022 (NW & Central Europe)'. Karlsruhe Institute of Technology, URL `https://www.cedim.kit.edu/download/FDA_Storms_Feb2022_report1_all.pdf`, Last accessed 04 May 2023, doi: 10.5445/IR/1000143470.

Neu, U., M. G. Akperov, N. Bellenbaum, R. Benestad, R. Blender, R. Caballero, A. Cocozza, H. F. Dacre et al., 2013: IMILAST: A Community Effort to Intercompare Extratropical Cyclone Detection and Tracking Algorithms. *Bulletin of the American Meteorological Society*, **94 (4)**, 529–547, doi: 10.1175/BAMS-D-11-00154.1.

Original S code by Richard A. Becker and Allan R. Wilks. R version by Ray Brownrigg. Enhancements by Thomas P Minka and Alex Deckmyn., 2018: *maps: Draw Geographical Maps*. URL `https://CRAN.R-project.org/package=maps`, r package version 3.3.0.

Pantillon, F., S. Lerch, P. Knippertz, and U. Corsmeier, 2018: Forecasting wind gusts in winter storms using a calibrated convection-permitting ensemble. *Quarterly Journal of the Royal Meteorological Society*, **144 (715)**, 1864–1881, doi: 10.1002/qj.3380.

Parton, G., A. Dore, and G. Vaughan, 2010: A climatology of mid-tropospheric mesoscale strong wind events as observed by the MST radar, Aberystwyth. *Meteorological Applications*, **17 (3)**, 340–354, doi: 10.1002/met.203.

Parton, G. A., G. Vaughan, E. G. Norton, K. A. Browning, and P. A. Clark, 2009: Wind profiler observations of a sting jet. *Quarterly Journal of the Royal Meteorological Society*, **135 (March)**, 663–680, doi: 10.1002/qj.

Peterson, R. A., 2021: Finding Optimal Normalizing Transformations via bestNormalize. *The R Journal*, **13 (1)**, 310–329, doi: 10.32614/RJ-2021-041.

Pinto, J. G., E. L. Fröhlich, G. C. Leckebusch, and U. Ulbrich, 2007: Changing european storm loss potentials under modified climate conditions according to ensemble simulations of the ECHAM5/MPI-OM1 GCM. *Natural Hazards and Earth System Sciences*, **7 (1)**, 165–175, doi: 10.5194/nhess-7-165-2007.

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 (6)**, 823–838, doi: 10.1127/0941-2948/2005/0068.

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 (5)**, 711–737, doi: 10.1007/s00382-008-0396-4.

Pirret, J. S., P. Knippertz, and T. M. Trzeciak, 2017: Drivers for the deepening of severe European windstorms and their impacts on forecast quality. *Quarterly Journal of the Royal Meteorological Society*, **143 (702)**, 309–320, doi: 10.1002/qj.2923.

Priestley, M. D. K., and J. L. Catto, 2022: Future changes in the extratropical storm tracks and cyclone intensity, wind speed, and structure. *Weather and Climate Dynamics*, **3 (1)**, 337–360, doi: 10.5194/wcd-3-337-2022.

R Core Team, 2021: R: A language and environment for statistical computing. Vienna, Austria.

Rasmussen, C. E., and C. K. I. Williams, 2005: *Gaussian Processes for Machine Learning*. The MIT Press, doi: 10.7551/mitpress/3206.001.0001.

Rautenhaus, M., C. M. Grams, A. Schäfler, and R. Westermann, 2015: Three-dimensional visualization of ensemble weather forecasts – Part 2: Forecasting warm conveyor belt situations for aircraft-based field campaigns. *Geoscientific Model Development*, **8 (7)**, 2355–2377, doi: 10.5194/gmd-8-2355-2015.

Raveh-Rubin, S., 2017: Dry intrusions: Lagrangian climatology and dynamical impact on the planetary boundary layer. *Journal of Climate*, **30 (17)**, 6661–6682, doi: 10.1175/JCLI-D-16-0782.1.

Raveh-Rubin, S., and J. L. Catto, 2019: Climatology and dynamics of the link between dry intrusions and cold fronts during winter, Part II: Front-centred perspective. *Climate Dynamics*, **53 (3)**, 1893–1909, doi: 10.1007/s00382-019-04793-2.

Rivière, G., P. Arbogast, and A. Joly, 2015: Eddy kinetic energy redistribution within windstorms Klaus and Friedhelm. *Quarterly Journal of the Royal Meteorological Society*, **141 (688)**, 925–938, doi: 10.1002/qj.2412.

Rivière, G., D. Ricard, and P. Arbogast, 2020: The downward transport of momentum to the surface in idealized sting-jet cyclones. *Quarterly Journal of the Royal Meteorological Society*, doi: 10.1002/qj.3767.

Rotunno, R., W. C. Skamarock, and C. Snyder, 1998: Effects of surface drag on fronts within numerically simulated baroclinic waves. *Journal of the Atmospheric Sciences*, **55 (11)**, 2119–2129, doi: 10.1175/1520-0469(1998)0552.0.CO;2.

Sanders, F., 1963: On subjective probability forecasting. *Journal of Applied Meteorology*, **2 (2)**, 191–201, doi: 10.1175/1520-0450(1963)002<0191:OSPF>2.0.CO;2.

Sanders, F., and J. R. Gyakum, 1980: Synoptic-dynamic climatology of the "bomb". *Monthly Weather Review*, **108**, 1589–1606.

Santos, J. A., T. Woollings, and J. G. Pinto, 2013: Are the winters 2010 and 2012 archetypes exhibiting extreme opposite behavior of the north atlantic jet stream? *Monthly Weather Review*, **141 (10)**, 3626–3640, doi: 10.1175/MWR-D-13-00024.1.

Schemm, S., and H. Wernli, 2014: The linkage between the warm and the cold conveyor belts in an idealized extratropical cyclone. *Journal of the Atmospheric Sciences*, **71 (4)**, 1443–1459, doi: 10.1175/jas-d-13-0177.1.

Schultz, D. M., D. Keyser, and L. F. Bosart, 1998: The effect of large-scale flow on low-level frontal structure and evolution in midlatitude cyclones. *Monthly Weather Review*, **126 (7)**, 1767–1791, doi: 10.1175/1520-0493(1998)126<1767:teolsf>2.0.co;2.

Schultz, D. M., and J. M. Sienkiewicz, 2013: Using Frontogenesis to Identify Sting Jets in Extratropical Cyclones. *Weather and Forecasting*, **28 (3)**, 603–613, doi: 10.1175/waf-d-12-00126.1.

Schultz, D. M., and G. Vaughan, 2011: Occluded fronts and the occlusion process: A fresh look at conventional wisdom. *Bulletin of the American Meteorological Society*, **92 (4)**, 443–466, doi: 10.1175/2010BAMS3057.1.

Schultz, D. M., and F. Zhang, 2007: Baroclinic development within zonally-varying flows. *Quarterly Journal of the Royal Meteorological Society*, **133 (626)**, 1101–1112, doi: 10.1002/qj.87.

Schulz, B., and S. Lerch, 2022: Machine learning methods for postprocessing ensemble forecasts of wind gusts: A systematic comparison. *Monthly Weather Review*, **150 (1)**, 235–257, doi: 10.1175/MWR-D-21-0150.1.

Shapiro, M., H. Wernli, J.-W. Bao, J. Methven, X. Zou, J. Doyle, T. Holt, E. Donall-Grell, and P. Neiman, 1999: A Planetary-Scale to Mesoscale Perspective of the Life Cycles of Extratropical Cyclones: The Bridge between Theory and Observations. *The Life Cycles of Extratropical Cyclones*, M. A. Shapiro, and S. Grønås, Eds., American Meteorological Society, Boston, MA, 139–185, doi: 10.1007/978-1-935704-09-6_14.

Shapiro, M. A., and D. Keyser, 1990: Fronts, Jet Streams and the Tropopause. *Extratropical Cyclones: The Erik Palmen memorial volume*, C. W. Newton, and E. O. Holopainen, Eds., American Meteorological Society, 167–191.

Shutts, G. J., 1990: Dynamical aspects of the October storm, 1987: A study of a successful fine-mesh simulation. *Quarterly Journal of the Royal Meteorological Society*, **116 (496)**, 1315–1347, doi: 10.1002/qj.49711649604.

Slater, T. P., D. M. Schultz, and G. Vaughan, 2015: Acceleration of near-surface strong winds in a dry, idealised extratropical cyclone. *Quarterly Journal of the Royal Meteorological Society*, **141 (689)**, 1004–1016, doi: 10.1002/qj.2417.

Slater, T. P., D. M. Schultz, and G. Vaughan, 2017: Near-surface strong winds in a marine extratropical cyclone: Acceleration of the winds and the importance of surface fluxes. *Quarterly Journal of the Royal Meteorological Society*, **143 (702)**, 321–332, doi: 10.1002/qj.2924.

Smart, D. J., and K. A. Browning, 2014: Attribution of strong winds to a cold conveyor belt and sting jet. *Quarterly Journal of the Royal Meteorological Society*, **140 (679)**, 595–610, doi: 10.1002/qj.2162.

Sprenger, M., and H. Wernli, 2015: The LAGRANTO Lagrangian analysis tool - Version 2.0. *Geoscientific Model Development*, **8 (8)**, 2569–2586, doi: 10.5194/gmd-8-2569-2015.

Sprenger, M., G. Fragkoulidis, H. Binder, M. Croci-Maspoli, P. Graf, C. M. Grams, P. Knippertz, E. Madonna et al., 2017: Global climatologies of Eulerian and Lagrangian flow features based on ERA-Interim. *Bulletin of the American Meteorological Society*, **98 (8)**, 1739–1748, doi: 10.1175/BAMS-D-15-00299.1.

Taillardat, M., O. Mestre, M. Zamo, and P. Naveau, 2016: Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics. *Monthly Weather Review*, **144 (6)**, 2375–2393, doi: 10.1175/MWR-D-15-0260.1.

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 (509)**, 17–55, doi: 10.1002/qj.49711950903.

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*, Springer Wien, Vol. 96, 117–131, doi: 10.1007/s00704-008-0083-8.

Vannitsem, S., J. B. Bremnes, J. Demaeyer, G. R. Evans, J. Flowerdew, S. Hemri, S. Lerch, N. Roberts et al., 2021: Statistical postprocessing for weather forecasts: Review, challenges, and avenues in a big data world. *Bulletin of the American Meteorological Society*, **102 (3)**, E681–E699, doi: 10.1175/BAMS-D-19-0308.1.

Volonté, A., P. A. Clark, and S. L. Gray, 2018: The role of mesoscale instabilities in the sting-jet dynamics of windstorm Tini. *Quarterly Journal of the Royal Meteorological Society*, **144 (712)**, 877–899, doi: 10.1002/qj.3264.

Volonté, A., P. A. Clark, and S. L. Gray, 2020: Idealised simulations of cyclones with robust symmetrically unstable sting jets. *Weather and Climate Dynamics*, **1 (1)**, 63–91, doi: 10.5194/wcd-1-63-2020.

Volonté, A., S. L. Gray, P. A. Clark, O. Martinez-Alvarado, and D. Ackerley, 2023a: Strong surface winds in Storm Eunice. Part 1: storm overview and indications of sting-jet activity from observations and model data. *Weather*, doi: 10.1002/wea.4402.

Volonté, A., S. L. Gray, P. A. Clark, O. Martinez-Alvarado, and D. Ackerley, 2023b: Strong surface winds in Storm Eunice. Part 2: airstream analysis. *Weather*, doi: 10.1002/wea.4401.

Wang, C., H. Liu, and S.-K. Lee, 2010: The record-breaking cold temperatures during the winter of 2009/2010 in the northern hemisphere. *Atmospheric Science Letters*, **11 (3)**, 161–168, doi: 10.1002/asl.278.

Wanner, H., S. Brönnimann, C. Casty, D. Gyalistras, J. Luterbacher, C. Schmutz, D. B. Stephenson, and E. Xoplaki, 2001: North Atlantic oscillation - concepts and studies. *Surveys in Geophysics*, **22 (4)**, 321–381, doi: 10.1023/A:1014217317898/METRICS.

Wernli, H., and H. C. Davies, 1997: A Lagrangian-based analysis of extratropical cyclones. I: The method and some applications. *Quarterly Journal of the Royal Meteorological Society*, **123 (538)**, 467–489, doi: 10.1256/smsqj.53810.

Wernli, H., S. Dirren, M. A. Liniger, and M. Zillig, 2002: Dynamical aspects of the life cycle of the winter storm ‘lothar’ (24–26 december 1999). *Quarterly Journal of the Royal Meteorological Society*, **128 (580)**, 405–429, doi: 10.1256/003590002321042036.

Wieringa, J., 1973: Gust factors over open water and built-up country. *Boundary-Layer Meteorology*, **3**, 424–441, doi: 10.1007/BF01034986/METRICS.

Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences*, Vol. 100. Academic press, 704 pp.

Wright, M. N., and A. Ziegler, 2017: Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, **77 (1)**, 1–17, doi: 10.18637/jss.v077.i01, `1508.04409`.

Zadrozny, B., and C. Elkan, 2002: Transforming classifier scores into accurate multiclass probability estimates. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery (ACM), New York, New York, USA, 694–699, doi: 10.1145/775047.775151.

Zängl, G., D. Reinert, P. Rípodas, and M. Baldauf, 2015: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, **141 (687)**, 563–579, doi: 10.1002/qj.2378.

Acknowledgements

I want to take this opportunity to express my gratitude to several individuals who accompanied me during my PhD journey and helped me along the way.
First and foremost, I want to thank my first advisor Peter Knippertz, who enabled me to pursue my PhD in a topic I have been captivated by since my Bachelor thesis in his group. I appreciate all the time you spent helping me find solutions, sharing my enthusiasm and giving me new ideas on Friday mornings, 8am. I really appreciate all the input and support I experienced over the years. Further thanks go to my secondary supervisor Joaquim Pinto, who –albeit not being a morning person– joined our early meetings from time to time and brought new perspectives and insights to my work. I appreciate the opportunity to work more closely with you and your group in the coming years – rest assured, I won't suggest Friday, 8am meetings!
I am extremely grateful to have had Benedikt Schulz as my project partner, who always helped me if I had mathematical questions and brought a different view to my problems. Thanks for all the help, funny moments and bad jokes over the years – especially the bad jokes!
Further thanks go to Sebastian Lerch, who makes our C5 group complete, Marlon Maranan for not being annoyed by my many PAMORE requests – at least not showing it –, Andreas Beckert and Marc Rautenhaus for your visualisation support and Alex Beideck and the Waves to Weather colleagues, especially Audine Laurian and Robert Redl, for always helping me out.

Special thanks go to Sue Gray and the people in Reading who welcomed me at the University of Reading for an extended research stay and gave me many new ideas and invaluable input, especially going forward with the SJ identification.

I would also like to thank the Deutscher Wetterdienst, namely Sebastian Trepte, Christoph Gebhard and Reinhold Hess, as well as Meteo France for the data made available and interesting discussions. I am as well grateful to Colin Manning for sharing his code for the streamline-based SJ detection tool and welcoming me in Exeter for a very engaging SJ discussion.

I owe a big thank you to my family for never letting me down, supporting me at every step and making all this possible, as well as being so considerate over the last few months leading up to this submission. An individual thanks to my brother Fabian for proofreading this dissertation and giving me valuable comments.

My deepest gratitude goes to Thomas for always trying to cheer me up, encouraging me whenever I needed it, and just being you. The last few years were not the easiest, and I appreciate all you have done to support me and be there for me.

Last but certainly not least, I want to thank my colleagues, fellow sufferers and friends. Thanks to you guys, this was not just a work place but a place where some of my favourite people are day in day out. Thanks for enduring all my complaints and helping me along the way. Special thanks go to Gabi and Beth for always getting me, keeping me sane and making me feel better, Hyunju and my fellow office mates in The Corner Office for the fun atmosphere – unless we were complaining about life – and letting me grow my plant curtain – or rather try – and, finally, Marlon, Christoph and Raffi for seeing me through the lockdown times with fun game nights, always trying to help with my endless problems I needed to discuss and just being there.

Wissenschaftliche Berichte des Instituts für Meteorologie und Klimaforschung des Karlsruher Instituts für Technologie (0179-5619)

Bisher erschienen:

Nr. 1: *Fiedler, F. / Prenosil, T.*
Das MESOKLIP-Experiment. (Mesoskaliges Klimaprogramm im Oberrheintal)

Nr. 2: *Tangermann-Dlugi, G.*
Numerische Simulationen atmosphärischer Grenzschichtströmungen über langgestreckten mesoskaligen Hügelketten bei neutraler thermischer Schichtung

Nr. 3: *Witte, N.*
Ein numerisches Modell des Wärmehaushalts fließender Gewässer unter Berücksichtigung thermischer Eingriffe

Nr. 4: *Fiedler, F. / Höschele, K. (Hrsg.)*
Prof. Dr. Max Diem zum 70. Geburtstag

Nr. 5: *Adrian, G.*
Ein Initialisierungsverfahren für numerische mesoskalige Strömungs modelle

Nr. 6: *Dorwarth, G.*
Numerische Berechnung des Druckiderstandes typischer Geländeformen

Nr. 7: *Vogel, B.; Adrian, G. / Fiedler, F.*
MESOKLIP-Analysen der meteorologischen Beobachtungen von mesoskaligen Phänomenen im Oberrheingraben

Nr. 8: *Hugelmann, C.-P.*
Differenzenverfahren zur Behandlung der Advektion

Die Bände sind unter www.ksp.kit.edu als PDF frei verfügbar oder als Druckausgabe bestellbar.

Nr. 9: *Hafner, T.*
Experimentelle Untersuchung zum Druckwiderstand
der Alpen

Nr. 10: *Corsmeier, U.*
Analyse turbulenter Bewegungsvorgänge in der maritimen
atmosphärischen Grenzschicht

Nr. 11: *Walk, O. / Wieringa, J.(eds)*
Tsumeb Studies of the Tropical Boundary-Layer Climate

Nr. 12: *Degrazia, G. A.*
Anwendung von Ähnlichkeitsverfahren auf die turbulente
Diffusion in der konvektiven und stabilen Grenzschicht

Nr. 13: *Schädler, G.*
Numerische Simulationen zur Wechselwirkung zwischen
Landoberflächen und atmophärischer Grenzschicht

Nr. 14: *Heldt, K.*
Untersuchungen zur Überströmung eines mikroskaligen
Hindernisses in der Atmosphäre

Nr. 15: *Vogel, H.*
Verteilungen reaktiver Luftbeimengungen im Lee einer
Stadt – Numerische Untersuchungen der relevanten Prozesse

Nr. 16: *Höschele, K.(ed.)*
Planning Applications of Urban and Building Climatology –
Proceedings of the IFHP / CIB-Symposium Berlin,
October 14-15, 1991

Nr. 17: *Frank, H. P.*
Grenzschichtstruktur in Fronten

Nr. 18: *Müller, A.*
Parallelisierung numerischer Verfahren zur Beschreibung
von Ausbreitungs- und chemischen Umwandlungsprozessen
in der atmosphärischen Grenzschicht

Nr. 19: *Lenz, C.-J.*
Energieumsetzungen an der Erdoberfläche in gegliedertem
Gelände

Die Bände sind unter www.ksp.kit.edu als PDF frei verfügbar oder als Druckausgabe bestellbar.

Nr. 20: *Schwartz, A.*
Numerische Simulationen zur Massenbilanz chemisch reaktiver Substanzen im mesoskaligen Bereich

Nr. 21: *Beheng, K. D.*
Professor Dr. Franz Fiedler zum 60. Geburtstag

Nr. 22: *Niemann, V.*
Numerische Simulation turbulenter Scherströmungen mit einem Kaskadenmodell

Nr. 23: *Koßmann, M.*
Einfluß orographisch induzierter Transportprozesse auf die Struktur der atmosphärischen Grenzschicht und die Verteilung von Spurengasen

Nr. 24: *Baldauf, M.*
Die effektive Rauhigkeit über komplexem Gelände – Ein Störungstheoretischer Ansatz

Nr. 25: *Noppel, H.*
Untersuchung des vertikalen Wärmetransports durch die Hangwindzirkulation auf regionaler Skala

Nr. 26: *Kuntze, K.*
Vertikaler Austausch und chemische Umwandlung von Spurenstoffen über topographisch gegliedertem Gelände

Nr. 27: *Wilms-Grabe, W.*
Vierdimensionale Datenassimilation als Methode zur Kopplung zweier verschiedenskaliger meteorologischer Modellsysteme

Nr. 28: *Grabe, F.*
Simulation der Wechselwirkung zwischen Atmosphäre, Vegetation und Erdoberfläche bei Verwendung unterschiedlicher Parametrisierungsansätze

Nr. 29: *Riemer, N.*
Numerische Simulationen zur Wirkung des Aerosols auf die troposphärische Chemie und die Sichtweite

Nr. 30: *Braun, F. J.*
Mesoskalige Modellierung der Bodenhydrologie

Nr. 31: *Kunz, M.*
Simulation von Starkniederschlägen mit langer Andauer über Mittelgebirgen

Nr. 32: *Bäumer, D.*
Transport und chemische Umwandlung von Luftschadstoffen im Nahbereich von Autobahnen – numerische Simulationen

Nr. 33: *Barthlott, C.*
Kohärente Wirbelstrukturen in der atmosphärischen Grenzschicht

Nr. 34: *Wieser, A.*
Messung turbulenter Spurengasflüsse vom Flugzeug aus

Nr. 35: *Blahak, U.*
Analyse des Extinktionseffektes bei Niederschlagsmessungen mit einem C-Band Radar anhand von Simulation und Messung

Nr. 36: *Bertram, I.*
Bestimmung der Wasser- und Eismasse hochreichender konvektiver Wolken anhand von Radardaten, Modellergebnissen und konzeptioneller Betrachtungen

Nr. 37: *Schmoeckel, J.*
Orographischer Einfluss auf die Strömung abgeleitet aus Sturmschäden im Schwarzwald während des Orkans „Lothar"

Nr. 38: *Schmitt, C.*
Interannual Variability in Antarctic Sea Ice Motion: Interannuelle Variabilität antarktischer Meereis-Drift

Nr. 39: *Hasel, M.*
Strukturmerkmale und Modelldarstellung der Konvektion über Mittelgebirgen

---

Die Bände sind unter www.ksp.kit.edu als PDF frei verfügbar oder als Druckausgabe bestellbar.

Ab Band 40 erscheinen die Wissenschaftlichen Berichte des Instituts für Meteorologie und Klimaforschung bei KIT Scientific Publishing (ISSN 0179-5619). Die Bände sind unter www.ksp.kit.edu als PDF frei verfügbar oder als Druckausgabe bestellbar.

Nr. 40: *Lux, R.*
Modellsimulationen zur Strömungsverstärkung von orographischen Grundstrukturen bei Sturmsituationen
ISBN 978-3-86644-140-8

Nr. 41: *Straub, W.*
Der Einfluss von Gebirgswellen auf die Initiierung und Entwicklung konvektiver Wolken
ISBN 978-3-86644-226-9

Nr. 42: *Meißner, C.*
High-resolution sensitivity studies with the regional climate model COSMO-CLM
ISBN 978-3-86644-228-3

Nr. 43: *Höpfner, M.*
Charakterisierung polarer stratosphärischer Wolken mittels hochauflösender Infrarotspektroskopie
ISBN 978-3-86644-294-8

Nr. 44: *Rings, J.*
Monitoring the water content evolution of dikes
ISBN 978-3-86644-321-1

Nr. 45: *Riemer, M.*
Außertropische Umwandlung tropischer Wirbelstürme: Einfluss auf das Strömungsmuster in den mittleren Breiten
ISBN 978-3-86644-766-0

Nr. 46: Nicht erschienen

Nr. 47: Nicht erschienen

Nr. 48: Nicht erschienen

Nr. 49: *Peters, T.*
Ableitung einer Beziehung zwischen der Radarreflektivität, der Niederschlagsrate und weiteren aus Radardaten abgeleiteten Parametern unter Verwendung von Methoden der multivariaten Statistik
ISBN 978-3-86644-323-5

Nr. 50: *Khodayar Pardo, S.*
High-resolution analysis of the initiation of deep convection forced by boundary-layer processes
ISBN 978-3-86644-770-7

Nr. 51: *Träumner, K.*
Einmischprozesse am Oberrand der konvektiven atmosphärischen Grenzschicht
ISBN 978-3-86644-771-4

Nr. 52: Nicht erschienen

Nr. 53: *Lundgren, K.*
Direct Radiative Effects of Sea Salt on the Regional Scale
ISBN 978-3-86644-773-8

Nr. 54: *Sasse, R.*
Analyse des regionalen atmosphärischen Wasserhaushalts unter Verwendung von COSMO-Simulationen und GPS-Beobachtungen
ISBN 978-3-86644-774-5

Nr. 55: *Grenzhäuser, J.*
Entwicklung neuartiger Mess- und Auswertungsstrategien für ein scannendes Wolkenradar und deren Anwendungsbereiche
ISBN 978-3-86644-775-2

Nr. 56: *Grams, C.*
Quantification of the downstream impact of extratropical transition for Typhoon Jangmi and other case studies
ISBN 978-3-86644-776-9

Nr. 57: *Keller, J.*
Diagnosing the Downstream Impact of Extratropical Transition Using Multimodel Operational Ensemble Prediction Systems
ISBN 978-3-86644-984-8

Nr. 58: *Mohr, S.*
Änderung des Gewitter- und Hagelpotentials im Klimawandel
ISBN 978-3-86644-994-7

Nr. 59: *Puskeiler, M.*
Radarbasierte Analyse der Hagelgefährdung in Deutschland
ISBN 978-3-7315-0028-5

Nr. 60: *Zeng, Y.*
Efficient Radar Forward Operator for Operational
Data Assimilation within the COSMO-model
ISBN 978-3-7315-0128-2

Nr. 61: Nicht erschienen

Nr. 62: *Jerger, D.*
Radar Forward Operator for Verification of Cloud
Resolving Simulations within the COSMO Model
ISBN 978-3-7315-0172-5

Nr. 63: *Maurer, V.*
Vorhersagbarkeit konvektiver Niederschläge:
Hochauflösende Ensemblesimulationen für Westafrika
ISBN 978-3-7315-0189-3

Nr. 64: *Stawiarski, C.*
Optimizing Dual-Doppler Lidar Measurements of Surface
Layer Coherent Structures with Large-Eddy Simulations
ISBN 978-3-7315-0197-8

Nr. 65: *Mahlke, H.*
Mechanismen der Auslösung hochreichender Konvektion
im südwestdeutschen Mittelgebirgsraum
ISBN 978-3-7315-0203-6

Nr. 66: *Fosser, G.*
Precipitation statistics from regional climate model
at resolutions relevant for soil erosion
ISBN 978-3-7315-0227-2

Nr. 67: *Adler, B.*
Boundary-Layer Processes Producing Mesoscale
Water-Vapour Variability over a Mountainous Island
ISBN 978-3-7315-0247-0

Die Bände sind unter www.ksp.kit.edu als PDF frei verfügbar oder als Druckausgabe bestellbar.

Nr. 68: *Kraut, I.*
Separating the Aerosol Effect in Case of a „Medicane“
ISBN 978-3-7315-0405-4

Nr. 69: *Breil, M.*
Einfluss der Boden-Vegetation-Atmosphären Wechselwirkungen auf die dekadische Vorhersagbarkeit des Westafrikanischen Monsuns
ISBN 978-3-7315-0420-7

Nr. 70: *Lott, F. F.*
Wind Systems in the Dead Sea and Footprints in Seismic Records
ISBN 978-3-7315-0596-9

Nr. 71: *Rieger, D.*
Der Einfluss von natürlichem Aerosol auf Wolken über Mitteleuropa
ISBN 978-3-7315-0672-0

Nr. 72: *Loewe, K.*
Arctic mixed-phase clouds. Macro- and microphysical insights with a numerical model
ISBN 978-3-7315-0686-7

Nr. 73: *Piper, D. A.*
Untersuchung der Gewitteraktivität und der relevanten großräumigen Steuerungsmechanismen über Mittel- und Westeuropa
ISBN 978-3-7315-0701-7

Nr. 74: *Metzger, J.*
Wind Systems and Energy Balance in the Dead Sea Valley
ISBN 978-3-7315-0699-7

Nr. 75: *Deetz, K.*
Assessing the Aerosol Impact on Southern West African Clouds and Atmospheric Dynamics
ISBN 978-3-7315-0744-4

Nr. 76: *Ehmele, F. T.*
Stochastische Simulation großflächiger, hochwasserrelevanter Niederschlagsereignisse
ISBN 978-3-7315-0761-1

Nr. 77: *Hackenbruch, J.*
Anpassungsrelevante Klimaänderungen für städtische Baustrukturen und Wohnquartiere
ISBN 978-3-7315-0771-0

Nr. 78: *Schmidberger, M.*
Hagelgefährdung und Hagelrisiko in Deutschland basierend auf einer Kombination von Radardaten und Versicherungsdaten
ISBN 978-3-7315-0846-5

Nr. 79: *Gruber, S.*
Contrails and Climate Engineering - Process Studies on Natural and Artificial High-Level Clouds and Their Impact on the Radiative Fluxes
ISBN 978-3-7315-0896-0

Nr. 80: *Walter, C.*
Simulationen der Ausbreitung von Vulkanasche unter expliziter Berücksichtigung der optischen Eigenschaften der Aschepartikel
ISBN 978-3-7315-0939-4

Nr. 81: *Caldas-Alvarez, A.*
Atmospheric Moisture Effects on Deep Convection in the Western Mediterranean
ISBN 978-3-7315-0947-9

Nr. 82: *Sedlmeier, K.*
Near future changes of compound extreme events from an ensemble of regional climate simulations
ISBN 978-3-7315-0476-4

Nr. 83: *Brecht, B. M.*
Die urbane Wärmebelastung unter Einfluss lokaler Faktoren und zukünftiger Klimaänderungen
ISBN 978-3-7315-0990-5

Nr. 84: *Singh, S.*
Convective precipitation simulated with ICON over heterogeneous surfaces in dependence on model and land-surface resolution
ISBN 978-3-7315-1068-0

Nr. 85: *Wilhelm, J.*
Einfluss atmosphärischer Umgebungsbedingungen auf den Lebenszyklus konvektiver Zellen in der Echtzeit-Vorhersage
ISBN 978-3-7315-1182-3

Nr. 86: *Pickl, M.*
Perspectives on warm conveyor belts – sensitivities to ensemble configuration and the role for forecast error
ISBN 978-3-7315-1236-3

Nr. 87: *Wandel, J.*
Representation of warm conveyor belts in sub-seasonal forecast models and the link to Atlantic-European weather regimes
ISBN 978-3-7315-1249-3

Nr. 88: *Eisenstein, L. A.*
Objective identification and climatology of mesoscale high-wind features within extratropical cyclones
ISBN 978-3-7315-1374-2

88

LEA ANDREA EISENSTEIN

Objective identification and climatology of mesoscale high-wind features within extratropical cyclones

Strong winds accompanying extratropical cyclones are among Europe's most dangerous natural hazards and are commonly associated with mesoscale features like the warm jet, the cold jet (CJ), cold frontal convection, strong cold-sector winds, and the sting jet (SJ). We introduce RAMEFI (RAndom-forest-based MEsoscale wind Feature Identification), a tool to objectively identify these features based on key surface characteristics. RAMEFI works flexibly across various data sets independent of resolution, distinguishing features with high reliability, while combining CJs and SJs. We applied RAMEFI to compile a climatology over 19 extended winter seasons (2000–2019) using a reanalysis and an observational data set, offering a comprehensive analysis of feature frequency, distribution, and characteristics. Lastly, low-cost methods for SJ identification are tested and a kinematic approach to extend RAMEFI to identify SJs is presented.

ISSN 0179-5619
ISBN 978-3-7315-1374-2

Gedruckt auf FSC-zertifiziertem Papier