Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

DETECTION AND VISUAL ANALYSIS OF 3-D ATMOSPHERIC FRONTS – FROM INDIVIDUAL CASES TO ENSEMBLES

An der Universität Hamburg eingereichte Dissertation

zur Erlangung des akademischen Grades
Dr. rer. nat.
an der Fakultät für Mathematik, Informatik und
Naturwissenschaften im Fachbereich Informatik

vorgelegt von

ANDREAS ALEXANDER BECKERT

Hamburg – June 2024

Gutachter:	Dr. Marc Rautenhaus Prof. Dr. Felix Ament Prof. Dr. Thomas Ludwig
Vorsitzender der Prüfungskommission:	Prof. Dr. Chris Biemann
Tag der Disputation:	10.04.2025

Andreas Alexander Beckert: *Detection and Visual Analysis of 3-D Atmospheric Fronts – from Individual Cases to Ensembles,* an der Universität Hamburg eingereichte Dissertation zur Erlangung des akademischen Grades Doctor rerum naturalium an der Fakultät für Mathematik, Informatik und Naturwissenschaften im Fachbereich Informatik, June 2024.
Zur Veröffentlichung freigegebene Fassung vom 20. Oktober 2025.

ACKNOWLEDGMENTS

It is with deep gratitude that I would like to express my heartfelt appreciation to all the individuals and organisations involved in making this thesis a reality. I gratefully acknowledge the support of all those who have made this thesis possible. First and foremost, I would like to thank my supervisor, Dr. Marc Rautenhaus, for giving me the opportunity to write my thesis in the field of visualisation, for his unwavering guidance, endless support, and fruitful discussions. I am also immensely grateful to Prof. Dr. George Craig for sharing his expertise in meteorological matters and for his invaluable assessments. My sincere thanks also go to Prof. Dr. Stephan Olbrich, whose willingness to supervise me was crucial to the realisation of this work.

Special thanks to Dr. Lea Eisenstein, Dr. Annika Oertel, and Dr. Tim Hewson for their valuable collaboration and contribution to a joint publication. I am grateful to the Deutsche Forschungsgemeinschaft (DFG) and Waves to Weather (W2W) for their generous financial support, which made this research possible. I would like to thank all the members of the Visual Data Analysis Group, and in particular my office colleagues Thorwin Vogt and Kameswar Modali, for their engaging discussions and for creating a positive working environment that stimulated my productivity. I would also like to thank our former student assistant, Katharina Kaufmann, for her invaluable help with the ensemble visualisations. I am extremely grateful to Paul Keil and Christoph Sauter for their meticulous feedback on improving the linguistic quality of this thesis.

Last but not least, I would like to express my deep gratitude to my family, friends, and especially my partner Natalia for their unwavering love, encouragement and support. Their presence has been my rock throughout this academic journey.

ABSTRACT

Atmospheric fronts are a widely used conceptual model in meteorology, most encountered as two-dimensional (2-D) front lines on surface analysis charts. The three-dimensional (3-D) dynamical structure of fronts has been studied in the literature by means of "standard" 2-D maps and cross-sections and is commonly sketched in 3-D illustrations of idealised weather systems in atmospheric science textbooks. However, only recently has the feasibility of the objective detection and visual analysis of 3-D frontal structures and their dynamics within numerical weather prediction (NWP) data been proposed, and such approaches are not yet widely known in the atmospheric science community. In this thesis, I investigate the benefit of objective 3-D front detection for case studies of extra-tropical cyclones and for comparison of frontal structures between different NWP models. I build on a recent gradient-based detection approach, combined with modern 3-D interactive visual analysis techniques, and adapt it to handle data from state-of-the-art NWP models including those run at convection-permitting kilometre-scale resolution. The parameters of the detection method (including data smoothing and threshold parameters) are evaluated to yield physically meaningful structures. I illustrate the benefit of the method by presenting two case studies of frontal dynamics within mid-latitude cyclones. Examples include joint interactive visual analysis of 3-D fronts and Warm Conveyor Belt (WCB) trajectories, as well as identification of the 3-D frontal structures characterising the different stages of a Shapiro–Keyser cyclogenesis event. The 3-D frontal structures show agreement with 2-D fronts from surface analysis charts and augment the surface charts by providing additional pertinent information in the vertical dimension. A second application illustrates the relation between convection and 3-D cold-front structure by comparing data from simulations with parametrised and explicit convection. Investigation into "secondary fronts" that are commonly shown in UK Met Office surface analysis charts shows that for this event the secondary front is not a temperature-dominated but a humidity-dominated feature.

Building on this 3-D front detection approach, I extend the detection of individual 3-D fronts towards front-feature-based time series analysis and ensemble clustering. Ensemble simulations have become a standard in NWP. However, ensemble simulations generate large amounts of data and their analysis remains a challenge. In this thesis, I develop a manual and automated front-tracking algorithm based on geometric and physical characteristics to derive time series of frontal attributes from a selected cyclone system. These frontal attributes characterise the 3-D front by one-dimensional (1-D) physical properties, such as the average slope of the 3-D frontal structure. By tracking a selected front over successive time steps, time series of frontal attributes are derived to provide a compact view of the development of frontal attributes. To order and cluster ensemble simulations according to frontal attributes, a selected front automatically tracked across all ensemble members and front attribute time series are derived for each member. These feature time series are then ordered and clustered using time series distance measures in combination with k-means clustering, resulting in distinct clusters that represent different patterns of frontal evolution across ensemble members.

Integrated into the 3-D interactive visual analysis framework Met.3D, my approach allows a comprehensive analysis of the spatio-temporal evolution of 3-D atmospheric fronts and thus contributes to the challenge of rapid analysis of large ensemble weather forecasts, as well as having great potential for operational weather forecasting.

ZUSAMMENFASSUNG

Atmosphärische Fronten sind ein weit verbreitetes konzeptionelles Modell in der Meteorologie, das meist als zweidimensionale (2-D) Frontlinien auf Bodenanalysekarten dargestellt wird. Die dynamische, dreidimensionale (3-D) Struktur von Fronten wurde in der Literatur mit Hilfe von „standard“ 2-D Karten und Querschnitten erforscht und wird in der Meteorologie häufig in 3-D Illustrationen idealisierter Wettersysteme in Lehrbüchern dargestellt. Kürzlich wurde jedoch die Möglichkeit der objektiven Detektion und visuellen Analyse von 3-D Frontstrukturen und ihrer Dynamik in numerischen Wettervorhersagedaten (NWP) vorgeschlagen, allerdings sind solche Ansätze in der Meteorologie noch nicht weit verbreitet. In dieser Arbeit untersuche ich den Nutzen der objektiven 3-D Fronterkennung für Fallstudien extratropischer Zyklonen und für den Vergleich von Frontstrukturen zwischen verschiedenen NWP-Modellen. Ich baue auf einem kürzlich entwickelten gradientenbasierten Detektionsansatz auf, der mit modernen interaktiven 3-D Visualisierungstechniken kombiniert wird, und passe ihn an, um Daten von hochmodernen NWP-Modellen zu verarbeiten, einschließlich derer, die mit einer Konvektionsauflösung im Kilometerbereich arbeiten. Die Parameter der Detektionsmethode (einschließlich Datenglättung und Schwellwertparameter) werden evaluiert, um physikalisch sinnvolle Strukturen zu erhalten. Zwei Fallstudien zur Frontdynamik von Zyklonen in mittleren Breiten illustrieren die Vorteile dieser Methode. Beispiele hierfür sind die gemeinsame interaktive visuelle Analyse von 3-D Fronten und Trajektorien warmer Transportbänder sowie die Identifizierung von 3-D Frontstrukturen, die die verschiedenen Phasen eines Shapiro-Keyser-Zyklogeneseereignisses charakterisieren. Dabei entsprechen die 3-D Frontstrukturen den 2-D Fronten aus den Bodenanalysekarten und ergänzen die Bodenanalysekarten durch zusätzliche relevante Informationen in der vertikalen Dimension. Eine zweite Anwendung illustriert die Beziehung zwischen Konvektion und der 3-D Kaltfrontstruktur durch den Vergleich von Daten aus Simulationen mit parametrisierter und expliziter Konvektion. Die Untersuchung von „Sekundärfronten“, die üblicherweise in den Bodenanalysekarten des britischen Wetterdienstes (UK Met Office) dargestellt werden, zeigt, dass die Sekundärfront bei diesem Ereignis nicht von der Temperatur, sondern von der Feuchte dominiert wird.

Aufbauend auf diesem Ansatz zur 3-D Fronterkennung erweitere ich die Detektion einzelner 3-D Fronten in Richtung einer merkmalsbasierten Frontenzeitreihenanalyse und Ensemblegruppierung. Ensemble-Simulationen sind mittlerweile Standard in der NWP. Allerdings erzeugen Ensemble-Simulationen große Datenmengen, deren Analyse nach wie vor eine Herausforderung darstellt. In dieser Arbeit entwickle ich einen manuellen und automatischen Frontverfolgungsalgorithmus, basierend auf geometrischen und physikalischen Eigenschaften, um Zeitreihen von Frontmerkmalen aus einem ausgewählten Zyklonsystem

zu extrahieren. Diese Frontmerkmale charakterisieren die 3-D Front durch 1-D physikalische Eigenschaften, wie z. B. die mittlere Neigung der 3-D Frontstruktur. Um Ensemble-Simulationen nach Frontattributen zu ordnen und zu gruppieren, verfolge ich dieselbe Front durch alle Ensemble-Mitglieder und extrahiere Merkmalszeitreihen für jedes Mitglied. Diese Merkmalszeitreihen werden dann mit Hilfe von Zeitreihenabstandsmessungen in Verbindung mit k-Means-Clustering geordnet und gruppiert, wodurch verschiedene Gruppen entstehen, die unterschiedliche Muster der Frontenentwicklung über die Ensemblemitglieder repräsentieren.

Integriert in das 3-D interaktive Visualisierungs-Framework Met.3D ermöglicht mein Ansatz eine umfassende Analyse der räumlich-zeitlichen Entwicklung von 3-D atmosphärischen Fronten und trägt somit zur schnellen Analyse großer Ensemble-Wettervorhersagen bei.

CONTENTS

1 INTRODUCTION 1
  1.1 Motivation 2
  1.2 Objectives and research questions 4
  1.3 Outline 6
2 VISUALISATION AND COMPUTER SCIENCE BACKGROUND 7
  2.1 Visualisation in weather forecasting and atmospheric science 9
  2.2 Met.3D - Interactive 3-D visualisation of meteorological simulations 11
    2.2.1 Data processing pipeline 12
    2.2.2 Graphics pipeline 14
  2.3 Visualisation techniques for time series analysis 15
    2.3.1 Violin plot 16
    2.3.2 Horizon graph 17
  2.4 Foundations of Algorithms 18
    2.4.1 Data Representations 18
    2.4.2 Triangular Meshes 20
    2.4.3 Clustering techniques 24
    2.4.4 Distance measures 26
    2.4.5 The concept of normal curves 29
3 METEOROLOGICAL BACKGROUND 31
  3.1 Atmospheric structure and Motion 31
  3.2 Numerical weather prediction and ensemble simulations 35
  3.3 Extratropical cyclones and atmospheric fronts 36
  3.4 Front detection and feature tracking methods 41
4 DATA BASIS OF CASE STUDIES 47
  4.1 Vladiana 48
  4.2 Friederike 49
  4.3 Otto 51
5 THREE-DIMENSIONAL FRONT ANALYSIS OF MID-LATITUDE WEATHER SYSTEMS 53
  5.1 Method and implementation 54
    5.1.1 Conceptual and mathematical basis 54
    5.1.2 Filtering 55
    5.1.3 Supported data and methodological details 57
  5.2 Thermal quantity, smoothing length scale, and filter parameters 63
    5.2.1 Choice of thermal quantity 63
    5.2.2 Recommendations for filter thresholds and sensitivity of fronts to different smoothing length scales 64
    5.2.3 Impact on reduced vertical resolution 67

5.3 Case studies 69
  5.3.1 Meteorological theory 69
  5.3.2 The 3-D examination of conceptual model: fronts and warm conveyor belt 70
  5.3.3 Cold-front structure in the vicinity of convection 72
  5.3.4 The 3-D examination of conceptual model: Shapiro–Keyser cyclone 74
  5.3.5 Secondary fronts 78
6 FEATURE-BASED ENSEMBLE ANALYSIS AND CLUSTERING 81
6.1 Methodology: Feature tracking and feature time series analysis 82
  6.1.1 Manual front tracking 82
  6.1.2 Characterising 3-D fronts by frontal attributes 85
  6.1.3 Automated front tracking 88
  6.1.4 Architecture and implementation details in Met.3D 89
  6.1.5 Example of front-tracking workflow 90
  6.1.6 Postprocessing of frontal attribute time series 91
6.2 Case Study: Frontal attribute distributions and tracking 93
  6.2.1 Distribution of frontals attributes 93
  6.2.2 Time Resolution 94
  6.2.3 Vertical Resolution 97
  6.2.4 Ensemble Tracking 99
6.3 Case study: Feature-based time series similarities and clustering 99
  6.3.1 Feature-based time series similarities 102
  6.3.2 Feature-based time series clustering 104
7 SUMMARY AND CONCLUSION 113
7.1 Conclusions and answers to research questions 114
7.2 Outlook 118
8 PUBLICATION, AVAILABILITY OF CODE, DATA AND ADDITIONAL MATERIALS 121

LIST OF FIGURES 123
LIST OF TABLES 134
BIBLIOGRAPHY 135
VERSICHERUNG AN EIDES STATT 153

ACRONYMS

1-D	one-dimensional
2-D	two-dimensional
3-D	three-dimensional
COSMO	Consortium for Small-scale Modeling
CNN	Convolutional Neural Network
CT	computed tomography
DWD	Deutscher Wetterdienst
DTW	Dynamic Time Warping
ECMWF	European Centre for Medium-Range Weather Forecasts
EPS	Ensemble Prediction System
EMD	Earth Mover’s Distance
ENS	enseble
ERA5	Fifth Generation ECMWF Reanalysis for the Global Climate and Weather
GPU	Graphics Processing Unit
GRIB	General Regularly-distributed Information in Binary form
HRES	Atmospheric Model high resolution 10-day forecast
ICAO	International Civil Aviation Organization
ICON	ICOsahedral Nonhydrostatic
ISA	International Standard Atmosphere
ITCZ	Intertropical Convergence Zone
IVA	Interactive Visual Analysis
MC	Marching Cubes
ML	Machine Learning
MR	magnetic resonance
NAWDEX	North Atlantic Waveguide and Downstream Impact Experiment
NC	normal curve
NetCDF	Network Common Data Form
NWP	numerical weather prediction
SCAFET	Scalable Feature Extraction and Tracking
TFP	thermal front parameter

UK	United Kingdom
WCB	Warm Conveyor Belt

1

INTRODUCTION

Contents

1.1 Motivation 2
1.2 Objectives and research questions 4
1.3 Outline 6

Imagine going for a bike ride every weekend and recording metrics such as speed, time, or even position coordinates. You might wonder if your speed fluctuates along the route or if you tend to improve over time. While you could analyse the raw numbers from your tracked data, after a few rides you might lose the overview by just looking at the numbers. Would it not be more convincing to have some form of graphical representation, such as charts or maps, to turn those numbers into meaningful insights? With visual aids such as charts and time series or maps, patterns emerge, trends become apparent, and anomalies are easily spotted. Not only does it allow you to understand your progress more quickly, it also makes the tracking process more engaging and intuitive. Visualisation is the key to turning mere data points into actionable information. The importance of visualisation extends far beyond personal data tracking; it is a fundamental aspect of scientific data analysis. Scientific data visualisation transforms data sets into understandable forms, enabling researchers to identify complex patterns, trends, and anomalies that would otherwise be hidden in the raw data. It is multidisciplinary, combining mathematics and computer science to efficiently manage and process large amounts of data, cognitive and perceptual sciences to ensure that visualisations are aligned with the way the human brain interprets visual information, and engineering techniques to build the robust software and hardware tools needed to create and display visualisations (Telea, 2015).

Scientific visualisation is widely used in the analysis and interpretation of numerical weather prediction (NWP) data (Rautenhaus et al., 2018). Modern NWP models generate enormous amounts of data, including variables such as temperature, pressure, humidity, and wind speed at different elevations and locations, which are visualised, for example, through colour-coded maps or contour plots. However, scientific visualisation goes beyond the visualisation of colour-coded weather maps and contour plots; it also involves the development of methods to represent trends, clusters, or to detect and visualise features (Telea, 2015) such as atmospheric fronts.

The concept of atmospheric fronts, first introduced by Bjerknes (1919), plays a prominent role in meteorology. They are thought of as an interface separating two air masses of different densities, mostly caused by temperature differences (*Front - Glossary of Meteorology* 2012). Fronts are atmospheric features represented as imaginary surfaces in three-dimensional (3-D) space; however, most commonly they are encountered as two-dimensional (2-D) lines on surface analysis charts, where they still frequently originate from manual analysis of different atmospheric variables. Despite the prevalence of 2-D surface fronts in meteorological practice, several studies have highlighted the impact of the vertical structure of fronts on surface weather (Aemisegger et al., 2015; Bader et al., 1996; Browning and Monk, 1982; Locatelli et al., 1994, 2005).

The atmosphere is a chaotic system, and describing and predicting its behaviour is a major challenge for meteorologists and scientists. NWP models address this challenge by encapsulating the fundamental physical processes that govern the atmosphere. These models are based on the principles of conservation of mass, momentum, and energy, implemented through a specific set of mathematical formulations and numerical solutions, resulting in a system of equations. The objective of NWP models is to represent the physical processes within the atmosphere and, consequently, the time evolution of key atmospheric variables, including pressure, temperature, wind, humidity, cloud cover, and precipitation (*Wetter Und Klima* 2023). NWP models cannot provide an exact representation of the atmospheric state due to uncertainties in the initial conditions, model assumptions, and parametrisations. Given the chaotic nature of the atmosphere, even small deviations in the initial state of NWP predictions can lead to significantly different results (Wilks, 2019). To account for such uncertainties, NWP predictions are based on ensemble simulations, where the initial conditions are systematically perturbed for each member of the ensemble. As a result, the ensemble provides a spectrum of possible weather events. The objective of ensemble forecasting is to predict the true state of the atmosphere within the range of the ensembles and to represent forecast uncertainties. Ensemble simulations generate large datasets, which pose a major challenge for analysis, where visualisation plays an important role (Wang et al., 2019).

1.1 MOTIVATION

Algorithms for 2-D objective front detection have been developed since the 1960s (e.g. Huber-Pock and Kress, 1989; Jenkner et al., 2009; Renard and Clarke, 1965). A widely cited method based on the third derivative of a thermal variable was introduced by Hewson (1998) and recently extended from 2-D to 3-D by Kern et al. (2019). Kern et al. (2019) integrated the objective detection algorithm into the open-source meteorological interactive 3-D visualisation framework "Met.3D" (Met3d.documentation, 2024; Met3d.homepage, 2024; Rautenhaus, 2015; Rautenhaus et al., 2015a,b) and demonstrated the feasibility of interactive 3-D visualisation of frontal surfaces detected in NWP data from the European Centre for

Medium-Range Weather Forecasts (ECMWF) and Deutscher Wetterdienst (DWD). In this thesis, I address open issues such as parameter settings, its application to NWP datasets with different spatial resolutions, and demonstrate and evaluate its use for the analysis of atmospheric dynamics.

The methods based on Hewson (1998) and Kern et al. (2019) (as well as further detection methods proposed in the literature) build on extracting frontal feature candidates from fields of the third derivative of a thermal variable (cf. Thomas and Schultz, 2019b) that typically are smoothed to some extent to remove high-frequency fluctuations. The feature candidates are then filtered according to some filter criteria (most prominently, a so-called thermal front parameter (TFP) and the frontal strength) to yield the final frontal features. Two challenges arise when applying such an approach to modern NWP data. First, the current trend towards convection-permitting kilometre-scale resolution in NWP models leads to more small-scale fluctuations in the gradient fields. The question arises whether the existing approaches still extract meaningful structures that represent a frontal surface. A related issue is that smaller numerical differences between the values of neighbouring grid cells (caused by smaller grid-point spacing) require care to avoid numerical artefacts when computing higher-order derivatives (see Jenkner et al., 2009). Second, threshold values for the filtering of feature candidates need to be selected carefully to yield physically interpretable structures. In the literature addressing 2-D front detection, such thresholds have been set to "hard" thresholds, i.e. fixed values suitable for the data and elevation level used. Such thresholds may not be generalized across different model resolutions and vertical elevations (Hewson, 1998). Furthermore, hard thresholds can lead to undesired "holes" in the resulting frontal surfaces, e.g. where frontal strength or TFP is only slightly below the chosen threshold. Therefore, Kern et al. (2019) proposed a fuzzy filtering method with upper and lower filter thresholds, between which the frontal features are gradually faded. However, past literature focused little on the filtering process and how to select suitable thresholds.

Analysing the complete 3-D temporal evolution of atmospheric frontal surfaces presents a promising opportunity for improving weather forecasting and advancing atmospheric research. However, this potential is linked to the challenge of tracking these features over time. In tracking, we can think of a frontal surface as a distinctive atmospheric feature. Numerous techniques exist to track atmospheric phenomena, yet many concentrate solely on 2-D characteristics. Manual tracking, where forecasters follow the movement of the feature, is one option. However, this method is time-consuming and unsuitable for operational weather forecasting, where time is of the essence. An automated feature tracking method presents a more efficient solution. Such approaches utilise feature detection and sequentially filtering algorithms to preserve the best matching features as they evolve over time. Filtering is based on specific criteria such as size and feature properties, and is fine-tuned to accurately identify the feature in the next time step. Combining objective methods of identifying and tracking features with statistical analysis of these features can provide a valuable framework for studying climate patterns and seasonal variations (Hodges, 1999).

Feature analysis methods have evolved into feature-based ensemble analysis as ensemble simulations have become the standard in weather forecasting. Ensemble simulations generate vast amounts of data, and current research concentrates on overcoming the challenges tied to processing, utilising, and visualising this wealth of information. In operational weather forecasting, it is common practice to use ensemble analysis and clustering methods to assess forecast uncertainty. Traditionally, ensemble analyses and clustering methods have concentrated on clustering 2-D fields, specifically the 500 hPa geopotential height (Ferranti and Corti, 2011). Such approaches are well established in the literature, with the ECMWF and other institutions widely using these techniques. Wilks (2019) provides an overview of ensemble clustering. The primary technique utilised in 2-D field clustering is principal component analysis, which has demonstrated successful outcomes in identifying commonalities in phenomena such as tropical cyclones (Anwender et al., 2008; Harr et al., 2008; Keller et al., 2011) and heatwaves (Quandt et al., 2017). For these clustering results, researchers have developed visual representations of clusters and their uncertainties (Kumpf et al., 2018). As highlighted by (Schultz et al., 2019), the development of specialised tools and methods dedicated to the ensemble analysis of atmospheric features is needed to better exploit ensemble simulations for cyclones and related weather impacts.

1.2 OBJECTIVES AND RESEARCH QUESTIONS

For the analysis of 3-D features, recent advances in 3-D computer graphics and visualisation bear large potential for intuitive, rapid interpretation in the context of the underlying atmospheric situation. Such techniques are not yet widely used in weather forecasting and research, with reasons including a lack of suitable software tools and a lack of literature demonstrating the benefit of 3-D visual analysis (Rautenhaus et al., 2018). An overview of the current state of the art in visualisation in meteorology has been provided by Rautenhaus et al. (2018); recent examples of 3-D visual analysis being applied to meteorological research include the studies by Bader et al. (2020), Bösiger et al. (2022), Fischer et al. (2022), Kern et al. (2018, 2019), Meyer et al. (2021), Orf et al. (2017), and Rautenhaus et al. (2015a).

The first part of this thesis builds on the 3-D front detection method presented by Kern et al. (2019), focusing on its shortcomings and demonstrating the benefits of 3-D visual analysis for NWP data. The investigations include case studies and comparisons of frontal structures between simulations from different numerical models. In this analysis, I further contribute to the literature on the benefits of atmospheric feature detection and address the following objectives:

1a: Advance the Kern et al. (2019) approach to objectively detect 2-D and 3-D frontal structures independently of the grid-point spacing of the input

NWP data to be able to compare frontal structures between, for instance, different model resolutions (e.g. in convection-permitting vs. convection-parametrised simulations), different ensemble members, or different cases. My goal is to shed light on the smoothing and filtering processes in the detection method and to study the sensitivity of changing smoothing parameters on the resulting detected fronts: which smoothing parameters yield meaningful 3-D structures, and how do filtering thresholds need to be chosen accordingly?

1b: Evaluate the benefit of 3-D Interactive Visual Analysis (IVA) of the detected frontal structures for the analysis of mid-latitude cyclones. I focus on two case studies (Cyclone Vladiana, crossing the North Atlantic in September 2016, and Cyclone Friederike, hitting Germany in January 2018) and address the following questions: can I confirm known knowledge about the 3-D dynamical structure of fronts and related warm conveyor belts (WCBs) by means of 3-D IVA? How can the characteristic frontal development stages of a Shapiro–Keyser cyclone be distinguished in 3-D? How do 3-D frontal structures differ in (higher-resolution) convection-permitting vs. (lower-resolution) convection-parametrising simulations? How do the detected 3-D structures compare to official analyses by the UK Met Office, in particular with respect to “secondary warm fronts” often observed in UK Met Office charts?

The second part of this thesis focuses on a feature-based perspective of IVA in the context of NWP ensemble forecasts. The proposed approach develops a method to track 3-D fronts and to characterise the tracked fronts by feature attributes. This feature-based ensemble analysis approach is evaluated in two case studies. To ensure broad applicability to different ensemble NWP data, the developed methods are adaptable to common ensemble simulations and data formats. The main objective is to systematically organise and cluster ensemble simulations based on derived 3-D frontal attributes and to present the results in clear and comprehensible visualisations. For this analysis, I address the following research questions:

2a: Can 3-D atmospheric fronts be effectively isolated and characterised by assigning one-dimensional (1-D) frontal attributes? What are the key physical attributes of fronts that enhance meteorological analysis? How can the tracking of isolated 3-D fronts be automated using feature attributes?

2b: What are effective visualisation methods for displaying the evolution of frontal attributes over time? How can feature-based ensemble analysis be successfully performed on 3-D atmospheric fronts? What techniques can be used to perform feature-based ensemble similarity ordering and clustering, and how can these approaches be effectively visualised?

2c: How sensitive are clusters derived from 3-D fronts considering different frontal attributes and variations in filter thresholds applied to the detected

3-D frontal surfaces? What insights do these variations provide for understanding the behaviour and characteristics of 3-D atmospheric fronts within ensemble simulations?

1.3 OUTLINE

The structure of this thesis is as follows: Chapter 2 delves into the field of visualisation and computer science fundamentals relevant to this thesis. It focuses on the field of visualisation in meteorology, interactive 3-D visualisation software, data representations, and the essential algorithms that support this work.

Chapter 3 serves as an introduction to the meteorological background and basic atmospheric science concepts. This chapter provides a comprehensive overview of relevant meteorological background information, focusing on aspects such as NWP model simulations and the detection of atmospheric features, with particular attention to cyclones and fronts.

Chapter 4 introduces the cases and NWP data used throughout this thesis and provides background information on the meteorological situation in each case.

Chapter 5 introduces the front detection approach developed by Hewson (1998), its extension to 3-D by Kern et al. (2019), and develops improvements to enable front detection in high-resolution kilometre-scale data. This chapter also considers the selection of the appropriate thermal variable for the approach, and examines the sensitivity of detected fronts to different data resolutions and smoothing parameters. Furthermore, the benefits of 3-D front analysis for both weather forecasting and research are examined.

Chapter 6 introduces and describes the front tracking and feature-based ensemble analysis approach developed in this thesis. The methodology part of this chapter describes the front isolation and front-tracking algorithms and introduces the characterisation of fronts by frontal attributes. The results section evaluates the proposed feature-based NWP ensemble analysis approach in two case studies. The evaluation of this framework is carried out by examining the sensitivity of the clusters.

Chapter 7 concludes this thesis by addressing the research questions, summarising the key findings, and providing an outlook.

2

VISUALISATION AND COMPUTER SCIENCE BACKGROUND

Contents

2.1 Visualisation in weather forecasting and atmospheric science **9**
2.2 Met.3D - Interactive 3-D visualisation of meteorological simulations **11**
  2.2.1 Data processing pipeline 12
  2.2.2 Graphics pipeline 14
2.3 Visualisation techniques for time series analysis **15**
  2.3.1 Violin plot 16
  2.3.2 Horizon graph 17
2.4 Foundations of Algorithms **18**
  2.4.1 Data Representations 18
  2.4.2 Triangular Meshes 20
  2.4.3 Clustering techniques 24
  2.4.4 Distance measures 26
  2.4.5 The concept of normal curves 29

Data visualisation is a multidisciplinary field at the intersection of mathematics, computer science, cognitive and perceptual sciences, and engineering. Its primary purpose is to provide insight into various aspects of the processes of interest, such as scientific simulations or real-world phenomena (Telea, 2015). The definition of visualisation encompasses its essential role in understanding data and extracting knowledge from simulation results, computations, or measurements (McCormick et al., 1987). The visualisation process involves the generation of one or more images that provide valuable insight into the process under consideration. Furthermore, visualisation can be seen as a cognitive process performed by humans, in which mental images of a domain space are formed. In the context of computer and information science, it involves the visual representation of a domain space through graphics, images, animated sequences, and sound augmentation. This enables the representation of large, complex data sets representing systems, events, processes, objects, and concepts (Williams and Others, 1995).

By using visualisation techniques, researchers and analysts can ask specific questions, explore data in an interactive way, and potentially discover novel patterns, trends, or relationships that might otherwise remain hidden. Visualisation

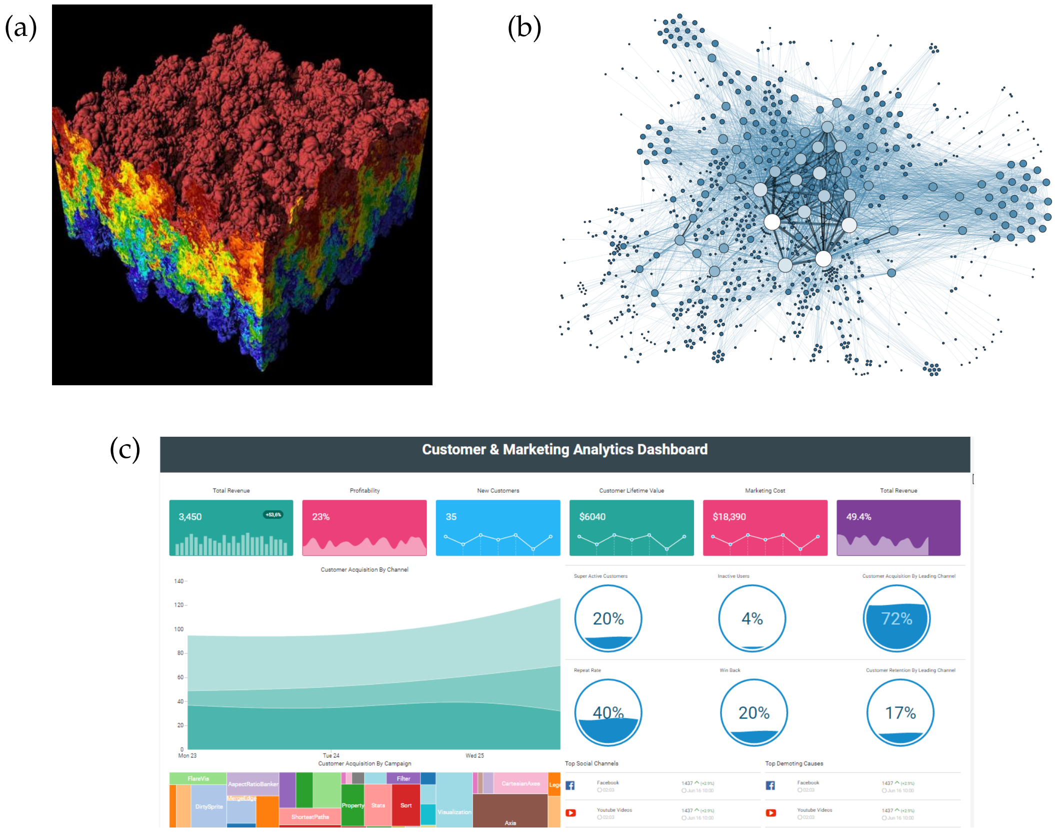


Figure 2.1: Example representations of the three visualisation disciplines. (a) Scientific visualisation of the Rayleigh-Taylor instability caused by the mixing of two fluids. (b) Information visualisation of a graph from a network analysis. (c) Visual analytics represented by a dashboard. Figures from (a) Lawrence Livermore National Laboratory (2009), (b) Grandjean (2014), used under CC BY-SA 3.0, and (c) HelicalInsight OpenSourceBI (2015), used under CC BY-SA 4.0.

enables data to be transformed into meaningful visual representations that can be understood and interpreted by humans, leading to improved decision-making, problem-solving and knowledge discovery. Following Telea (2015), the field of visualisation is divided into three distinct subfields: scientific visualisation (*scivis*), information visualisation (*infovis*), and visual analytics. Figure 2.1 shows example representations of each visualisation discipline.

Scientific visualisation focuses on the development of visualisation solutions to gain insight into scientific simulations. It specialises in the visualisation of 3-D phenomena and data sets that have a natural spatial embedding.

Information visualisation is dedicated to visualising relational datasets using trees, graphs, and networks. As the prevalence of big data continues to grow, the visualisation of data relations becomes increasingly important, considering the rising number, size, and variety of digital artefacts.

Visual analytics provides techniques and tools that support users in their analytical reasoning through interactive visual interfaces. It is characterised by a tight integration of data analysis, data mining, and visualisation technologies and tools. However, the boundary between visual analytics and the more traditional fields of information visualisation and scientific visualisation is not clearly defined, as they often overlap and share similar principles. As a result, the division into three distinct fields of visualisation has become blurred in recent years (*New VIS Conference* 2021).

The following sections discuss the role of visualisation in meteorology and introduce computer science and visualisation algorithms used in this thesis. First, Section 2.1 examines the role of visualisation in meteorology and for weather forecasting, as well as state-of-the-art visualisation tools. Second, Section 2.2 introduces the 3-D and interactive visualisation framework Met.3D, as well as the foundations of the visualisation and data pipelines. Most of the methods used in this thesis are implemented in Met.3D. Third, the foundations of visualisation and computer science algorithms are introduced in Section 2.4.

2.1 VISUALISATION IN WEATHER FORECASTING AND ATMOSPHERIC SCIENCE

The Earth's atmosphere is a complex and chaotic system, spanning three dimensions and governed by a multitude of interconnected variables. This complexity makes it a major challenge to gain a comprehensive understanding of atmospheric behaviour from raw data and numerical values alone. In meteorology, the use of visualisation techniques has become a powerful and indispensable tool for exploring this complex system.

Visualisation serves as a tool to interpret the vast amount of data derived from NWP models, measurements, and other sources. It performs the essential task of translating these numbers into intuitive, graphical representations that promote a deeper understanding of the dynamic nature of the atmosphere. By using visualisation, meteorologists aim to gain an insightful perspective within the numbers, uncover hidden relationships and correlations, and gain a comprehensive view of the data. The overall goal of visualisation of meteorological data is to facilitate the understanding, analysis, and communication of complex weather patterns, ensuring that meteorologists can uncover the underlying trends and patterns that govern atmospheric phenomena.

Weather forecasts play a crucial role in providing valuable information for various sectors and ensuring public safety. These forecasts are typically computed at national or international weather centres, such as the DWD, ECMWF, and Meteo France (the French national weather service). Weather forecasts generate large amounts of data, with approximately 150GB of data processed and simulated daily at the DWD (*Wetter und Klima - Deutscher Wetterdienst - NinJo - Meteorologischer Arbeitsplatz*). After computing a NWP simulation, it is the forecaster's task to analyse the data. Bosart (2003) outlined essential questions, leading the forecaster to understand the current and future weather situation:

- What happened and why? (diagnostic)
- What is happening and why? (diagnostic)
- What will happen and why? (prognostic)

To process the large amounts of data and support forecasters in understanding the weather situation and extracting the important information, weather services use a dedicated software. A detailed survey of visualisation in meteorology can be found in Rautenhaus et al. (2018), and a summary is given below.

The historical development of meteorological visualisation techniques has achieved several important milestones. It began with the creation of hand-drawn charts and diagrams, which represented the earliest methodology of visualising weather data. However, in the 1960s, a notable shift took place with the emergence of essential computer-based visualisations, exemplified by the work of Washington et al. (1968), which introduced 2-D line contours displayed on cathode ray tubes. Developments driven by the introduction of powerful workstations with dedicated Graphics Processing Units (GPUs) led to the development of 3-D visualisation tools. "Vis5D" (Hansen and Johnson, 2005) for instance, demonstrated the ability to render various meteorological variables and features through horizontal and vertical sections, 3-D isosurfaces, volumes, vector glyphs, streamlines, and trajectories, effectively expanding the dimensionality of weather visualisation. Continuous increase in computing power and the accessibility of high-end GPU hardware, have provided the foundation for powerful interactive 3-D visualisation. While general-purpose visualisation tools like Paraview (Ayachit, 2015) and VTK (Schroeder et al., 2018) have become integral to visualising NWP data in scientific research. However, conventional two-dimensional visualisation still remains the standard for operational weather forecasting, as exemplified by widely used software such as Metview (ECMWF, 2023b) and NinJo (*NinJo Workstation: NinJo Workstation*).

The DWD uses for example the NinJo Workstation Software (*NinJo Workstation: NinJo Workstation*), which is capable of processing a wide range of different data types such as NWP, radar, satellite, observations, and sounding data. The information is then visualised on 2-D maps and horizontal cross-sections, meteograms, or diagrams. Such software tools like Ninjo support the forecaster in making informed decisions and improving the forecasting ability. Once the forecast analysis is complete, weather centres communicate the current weather forecasts to the public through various channels, including apps, internet pages, videos, and other forms of communication.

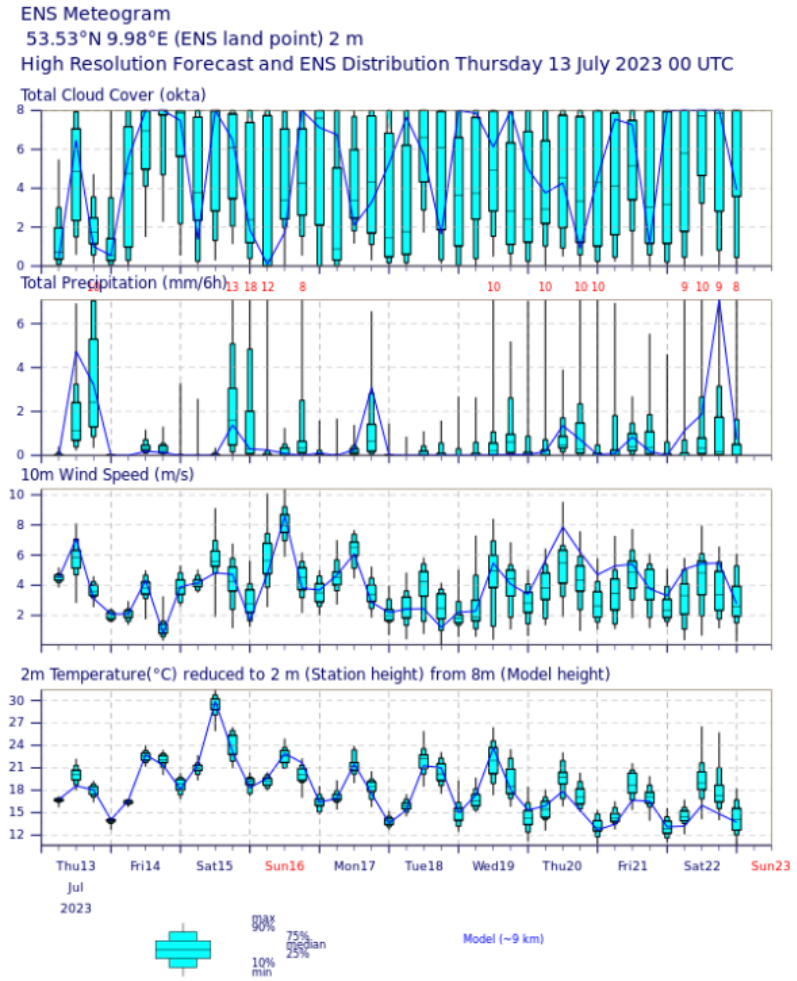


Figure 2.2: Example analysis plot of ECMWF ensemble forecast from 13. July 2023. The ensemble meteogram shows the high-resolution (9 km) and ensemble spread for the following 10 days. Figure from ECMWF (2023a), used under CC BY 4.0.

2.2 MET.3D - INTERACTIVE 3-D VISUALISATION OF METEOROLOGICAL SIMULATIONS

Figure 2.3 shows the open-source and interactive 3-D visualisation tool for NWP data, Met.3D (Rautenhaus, 2015). Originally developed for airborne field campaigns to improve detailed research flight planning using ensemble forecasts, it can be applied in various fields, such as weather forecasting, atmospheric research (see: Fischer et al., 2022; Meyer et al., 2021; Neuhauser et al., 2023), and educational purposes. Met.3D is designed as a "bridge" between common 2-D and new 3-D visualisation techniques in the field of weather forecasting and is intended to facilitate the introduction to 3-D visualisation.

The main objective of Met.3D is to provide a seamless and interactive analysis of different NWP data. Therefore, the software has been designed to directly read in a wide range of NWP data without the need for (vertical) regridding. For example, data from the ECMWF on a hybrid sigma-pressure grid or data from the

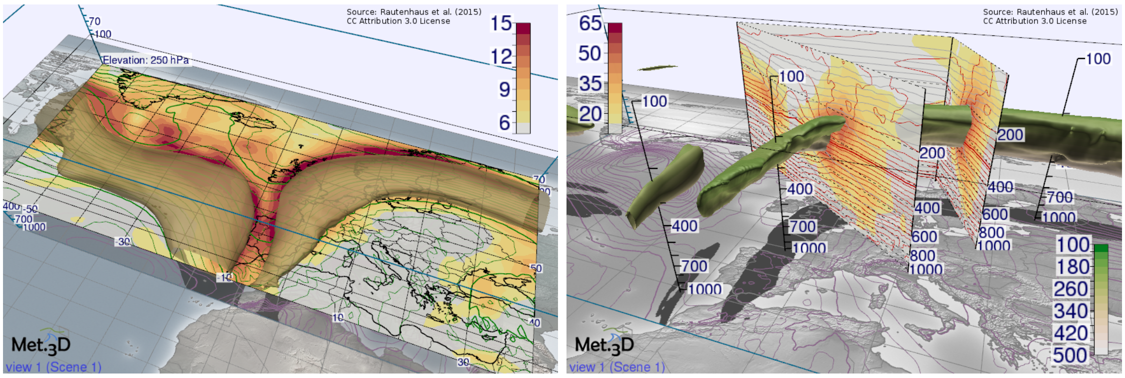


Figure 2.3: Screenshots of the interactive visualisation software Met.3D. Figure from Rautenhaus (2015), used under CC BY 3.0.

ICOsahedral Nonhydrostatic (ICON) model on an auxiliary pressure grid can be directly imported. Met.3D visualises data in a cylindrical projection by default, with vertical levels displayed on the commonly used logarithmic pressure scale.

The core visualisation components of the tool are a base map, graticules, horizontal and vertical cross-sections, 3-D isosurfaces, direct volume rendering, and trajectories. In Met.3D, these elements are called actors. A 3D field can be plotted on a horizontal or vertical slice, showing, for example, temperature as a texture or contour lines. Several 3D fields can be combined on an actor. 3-D isosurfaces can be used to represent areas of high wind speed, for example, and direct volume rendering can be used to represent clouds. The user can interact with the actors, e.g., move a horizontal or vertical cross-section through the scene to find an area of interest.

Met.3D is mainly developed in C++ and the OpenGL shader language. The architecture follows the data and graphics pipeline approach described above. A more detailed description of Met.3D can be found in Rautenhaus (2015) and Rautenhaus et al. (2015b). In this thesis, I use Met.3D for interactive visualisation and to develop and implement novel algorithms for feature extraction, feature characterisation, and feature tracking.

2.2.1 *Data processing pipeline*

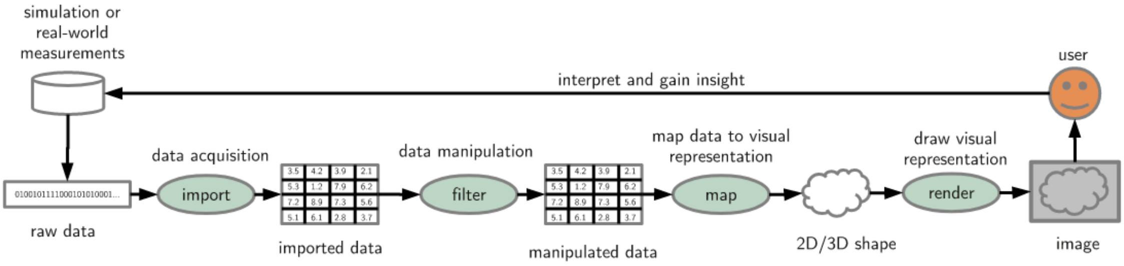


Figure 2.4: Illustration of the conceptual data pipeline in visualisation. Figure from Rautenhaus (2015), his figure 3.1., used with permission.

The visualisation process involves the concept of data flow and data representation through a data pipeline, which describes the sequence of steps from raw input data to the resulting images.

At the implementation level, this approach enables the construction of visualisations by assembling reusable and modular data processing operations. Each module in the pipeline is designed for a specific data transformation or manipulation task, with the aim of encoding the features of interest from the raw dataset required for analysis and visualisation.

The data pipeline typically consists of four stages: data import, data manipulation or filtering, data mapping, and data rendering. Figure 2.4 shows a visual representation of this data pipeline and how the different pipeline modules are connected. The data pipeline can be found in many textbooks, here I follow the descriptions and concept of Telea (2015), which is also illustrated in Figure 2.4.

Data import: The first stage involves obtaining raw data from simulations or measurements. This process involves reading the data from external storage and translating it from different data formats. For example, in the context of NWP data, this may involve translating data from Network Common Data Form (NetCDF) (Russ et al., 1989) or General Regularly-distributed Information in Binary form (GRIB) (Dey et al., 2007) formats into the internal data format of the visualisation program.

Data manipulation: The aim of the data manipulation pipeline module is to extract important aspects or features of interest. Since these aspects are usually not directly modelled in the raw data, data filtering or data enrichment techniques are applied. Filtering operations produce enriched datasets that represent the desired features for specific exploration tasks. The output of data manipulation consists of datasets that may have the same data format as the input dataset. An example of data manipulation is the computation of atmospheric fronts. Atmospheric fronts are not simulated directly, but the temperature field is simulated, which can be used to compute the desired atmospheric fronts.

Data mapping: In the data mapping stage, the manipulated dataset is mapped to a dataset of visual features. This process transforms invisible data into visible representations. For example, a gridded 2-D horizontal data field can be mapped to a 3-D elevation model. The 2-D data field contains elevation information for each grid point, which is then mapped to the 3-D representation of the elevation model. The output is a 3-D scene.

Rendering: The rendering stage combines the 3-D model from the mapping stage with additional viewing parameters such as viewpoint, lighting, and horizontal or vertical scaling. Viewing parameters are considered part of the rendering process. The rendered results can be interactively navigated and examined without recomputing the mapping stage operations. Section 2.2.2 describes the rendering stage and the rendering process in more detail.

In Met.3D, the data processing pipeline consists of various pipeline modules that are responsible for creating, reading, or manipulating data and storing it in internal data items. These data items are associated with a memory manager that caches all the intermediate results produced by the pipeline modules. The actor, responsible for visualising the final data item, is positioned at the end of the data processing pipeline. It communicates data requests to retrieve specific data items. These data requests then propagate backwards through the data modules from the final pipeline module to the input pipeline module.

Each pipeline module within the data processing pipeline interprets its data request. These data requests are combined into a task graph, which is then passed to a task scheduler. As computing resources become available, the scheduler initiates the execution of the task graph by propagating it from the first pipeline module to the last pipeline module. Each pipeline module computes its resulting data element and forwards it to the next pipeline module until the final data item is computed. If a data module's result is already cached in the data manager from prior computations, no further computation is needed and the cached result is returned. This setup enables the asynchronous execution of data requests and modulation of the data processing pipeline (Rautenhaus, 2015). For a detailed description of the data processing pipeline, see (Rautenhaus, 2015; Rautenhaus et al., 2015b).

2.2.2 *Graphics pipeline*

The graphics rendering pipeline is a fundamental component of real-time computer graphics rendering. This is the process of transforming 3-D geometric data into 2-D images that are displayed on the screen. Descriptions of the graphics pipeline can be found in many textbooks. Here, I follow the textbook of Akenine-Möller et al. (2018). The graphics pipeline is highly parallelised, allowing efficient processing of multiple vertices and fragments. Each stage of the pipeline builds on the output of the previous stage, leading to the final rendering of the scene. The core stages of the graphics pipeline include application, geometry processing, rasterisation, and pixel processing.

Application stage: The application stage is typically executed on the CPU and serves as the data pipeline for rendering. Its primary function is to generate the geometry for rendering. The geometry, known as rendering primitives, can be points, lines, or triangles. Once generated, the application sends these primitives to the geometry stage, which is implemented on the GPU, for further processing.

Geometry stage: The geometry stage operates either per vertex or per triangle and consists of several sub-stages, vertex shading, projection, clipping, and screen mapping. The *vertex shading* sub-stage is programmable and manipulates vertex positions while producing additional outputs such as normals and texture coordinates. Vertex shading transforms vertex positions from model space (e.g. latitude, longitude, and pressure coordinates) to clip space using the Model-View-Projection (MVP) matrix. The MVP matrix performs three transformations: first, from model space to world space, unifying primitives from different model spaces into a single world space; second, from world space to view space, positioning the camera at the origin of the coordinate system; and third, projection transformation, converting vertices from view space to the desired projection, usually orthographic or perspective projections.

The geometry stage provides optional sub-modules, including tessellation, geometry shading, and stream output. *Tessellation* allows the dynamic addition of scene detail, particularly useful when zooming in on objects. *Geometric shading* generates additional vertices from given primitives, allowing for more realistic representations, such as rendering trajectories as tubes. Material properties, light sources, and colours can be modelled during geometry shading if used, or during vertex shading otherwise. The *stream output* stage controls the output and can save the generated geometries for further processing.

The *clipping* sub-stage evaluates which geometries are fully or partially inside the view volume and discards those outside. Partially visible geometries are clipped to separate the part inside the viewing volume from the rest. The *screen mapping* sub-stage transforms the projected coordinates of the geometries into screen coordinates.

Rasterisation: Rasterisation is the transformation of geometry elements into screen pixels, which involves converting the screen space into the pixel space. The z-value of the geometry is used as a depth value to determine the visibility of overlapping elements.

Pixel Processing: The pixel processing stage consists of two sub-stages: pixel shading and merging. *Pixel shading* is performed by programmable GPU cores using the fragment shader. It operates on a per-pixel basis and involves texturing, i.e. 'gluing' one or more images onto objects, e.g., mapping a colour onto a 3D surface. The result is a colour value for each pixel. The following *merging* sub-stage is not programmable but configurable, and resolves pixel visibility using z-buffer algorithms. After merging, the final image is rendered on the screen.

2.3 VISUALISATION TECHNIQUES FOR TIME SERIES ANALYSIS

A time series is a common form of data presentation, characterised by its 1-D nature, that illustrates the evolution of a particular quantity over time. An example of a time series is the presentation of temperature trends at a particular location, showing how this parameter evolves over different time intervals. However, when looking at ensemble simulations, the analysis often extends beyond individual time series to include the evolution of a quantity across different ensemble members.

In scenarios where multiple ensemble members contribute to the overall trend, it can be challenging to plot the evolution of each member over time on a single plot, resulting in cluttered visualisations that are difficult to interpret. One approach to overcome this problem is to plot the ensemble spread using a box plot, as shown in Figure 2.2, to provide an overview of the variation between members over time. An alternative to the box plot is the violin plot (Hintze and Nelson, 1998), which combines the traditional box plot with smoothed histograms. The violin plot is used in this thesis in Chapter 6 to visualise the evolution of distributions over time. A detailed description of the violin plot is given in the following Section 2.3.1. However, such methods may sacrifice the individual insights provided by the time series of each ensemble member.

In order to effectively visualise the evolution of multiple ensemble members and to preserve detailed insights into individual ensemble members, the horizon graph of Saito et al. (2005) is used in this thesis in Chapter 6. The components and structure of the horizon graph are explained in Section 2.3.2.

2.3.1 *Violin plot*

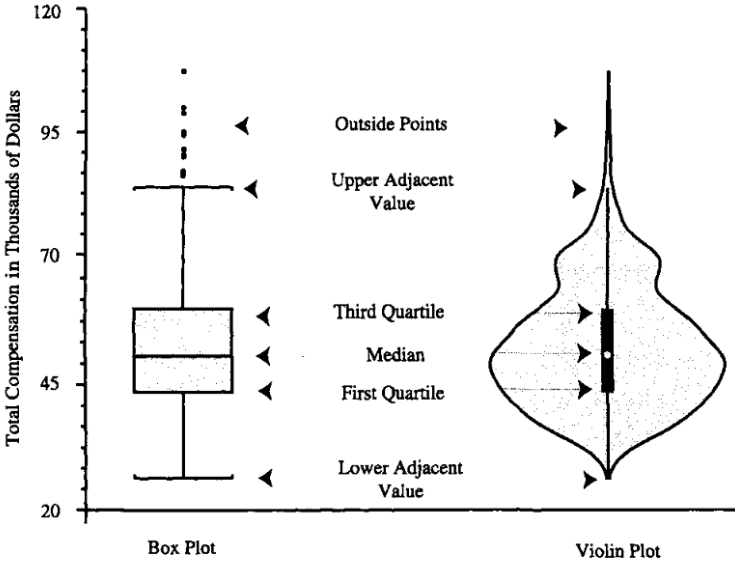


Figure 2.5: Example visualisations of Box Plot and Violin Plot show the total compensation for all academic ranks. Figure used with permission of Taylor & Francis Informa UK Ltd - Journals, from Hintze and Nelson (1998); permission conveyed through Copyright Clearance Center, Inc.

The violin plot, a statistical data visualisation tool introduced by Hintze and Nelson (1998), combines elements of box plots and smoothed histograms into a single plot. Figure 2.5 shows a comparison between the violin plot and the traditional box plot. This visualisation enhances the box plot by providing additional details about the data distribution without taking up more space. The violin plot is a powerful tool for comparing distributions over different variables or different time steps and for identifying clusters within the data (Hintze and Nelson, 1998). In this thesis, a basic version of the violin plot is used, which displays the distribution (violin) and marks only the first quantile, median, and third quantile.

2.3.2 *Horizon graph*

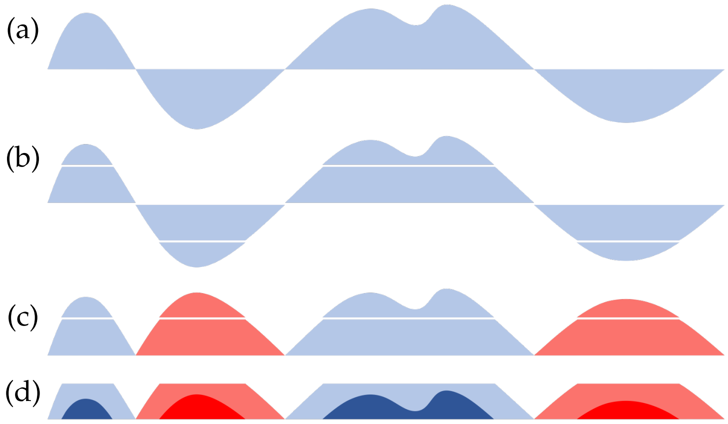


Figure 2.6: Construction of a horizon plot: (a) line plot, (b) segmented line plot, (c) layering, and (d) final horizon plot with visual differentiation. Figure adapted from Heer et al. (2009).

The horizon plot, initially referred to as "two-tone pseudo-colouring" by Saito et al. (2005), is a visualisation technique specifically designed for effectively representing large one-dimensional datasets, such as time series, with high detail and precision in a compact manner (Saito et al., 2005). A comprehensive example of a horizon plot showcasing multiple time series can be found in Few's (2008) work. Heer et al. (2009) conducted a thorough investigation and user study into the benefits of using a horizon plot, including improved processing speed, accuracy, visualisation quality, and data interpretation, in comparison to traditional line charts. The process of constructing a horizon plot involves the following steps:

Segmentation: The line chart is divided into non-overlapping bands of equal size along the y-axis (vertical axis).

Layering: These bands are stacked on top of each other, with negative values reflected around zero.

Visual Differentiation: To distinguish between negative and positive values as well as different band levels, various techniques such as colour, saturation, and intensity can be employed.

This technique significantly reduces the vertical space required to represent a classical line chart with both positive and negative values, achieving a vertical compression ratio of 2 times the number of bands (Heer et al., 2009). Consequently, the horizon plot proves to be highly suitable for visualising numerous time series, allowing for detailed and efficient representations of data while conserving space.

2.4 FOUNDATIONS OF ALGORITHMS

This section serves as an introduction to a collection of mathematical background and computer algorithms, laying the foundation for their subsequent application in this thesis. The aim is to provide the reader with the necessary background knowledge. Topics covered include data representation and coordinate transformations required for NWP simulation, data representation and further processing of such data (Section 2.4.1); triangular meshes for representing, e.g., 2-D or 3-D atmospheric features (Section 2.4.2); distance measurement algorithms (Section 2.4.4) for computing distances, e.g., in time series of atmospheric features, which are essential for the application of data clustering (Section 2.4.3).

2.4.1 *Data Representations*

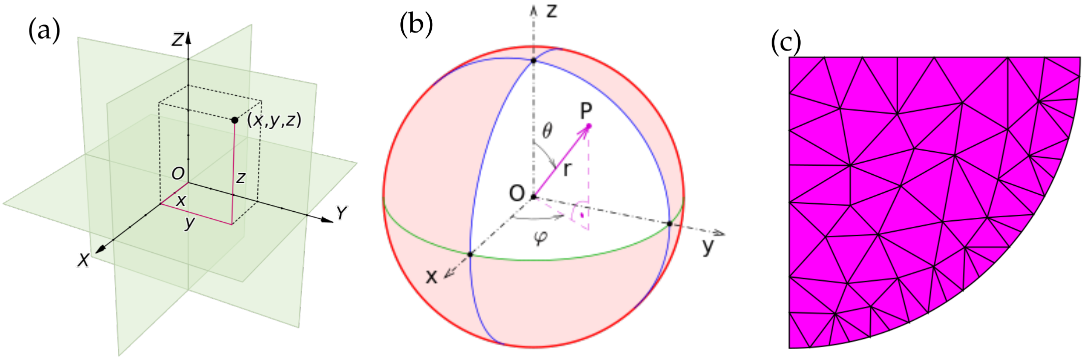


Figure 2.7: Representation of different coordinate systems and data grids. (a) Cartesian coordinate system, (b) spherical coordinate system, and (c) unstructured grid. Figures from (a) Stolfi (2009), (b) Ag2gaeh (2015), used under CC BY-SA 4.0, and (c) Slffea and Mysid (2006), used under CC BY-SA 3.0.

A coordinate system, or data grid, is a way of representing continuous data, such as data from numerical simulations. A coordinate system serves as a fundamental framework for specifying the position of a data point or object within a given space. It provides a systematic and standardised way to describe both the location and, when applicable, the orientation of an object relative to a reference point or axis. Coordinate systems play an important role in many fields, including NWP and computer graphics, where they are used to precisely define the position of objects and mathematical operations. The following list contains some common coordinate systems:

Cartesian coordinate system: Also known as the rectangular or Cartesian grid, uses a set of orthogonal axes (typically labelled x, y, and z in three-dimensional space) to define points by their distances along each axis. This system is widely used in mathematics and engineering for its simplicity and versatility. Figure 2.7a shows an example of a Cartesian coordinate system.

Spherical Coordinate System: In spherical coordinates, points are defined by their radial distance (r), polar angle (λ), and azimuthal angle (φ). This system is especially useful for describing positions in three-dimensional space when dealing with spherical objects or coordinates defined on the surface of a sphere. Figure 2.7b shows an example of a spherical coordinate system.

Geographic Coordinate System: This system is used to specify locations on the Earth's surface and is a special case of a spherical coordinate system. It uses latitude and longitude angles to define positions relative to the Earth's equator and prime meridian, respectively.

Unstructured grids: In addition to these structured coordinate systems, there are unstructured grids or meshes. These grids do not adhere to a regular pattern of points or elements and offer enhanced flexibility for discretising complex domains and geometries in computational simulations. Figure 2.7c shows an example of an unstructured grid.

When working with data on coordinate systems it might be necessary to transform coordinates into another coordinate system or to rotate the coordinate system. In the following these processes are described.

Coordinate transformation

The transformation from a geographical coordinate system to a Cartesian coordinate system is a useful process for various calculations, such as determining the area of a triangular mesh. The transformation equations can be found in many textbooks. The description provided by Hofmann-Wellenhof et al. (2001) is used below to describe the transformation from ellipsoidal to Cartesian coordinates. However, the transformation equations are slightly simplified and only valid for the case of a spherical coordinate system. NWP data are typically given in geographical coordinates, which is a spherical coordinate system.

The relation between the spherical (λ, φ, h) and Cartesian (x, y, z) coordinate systems is defined as follows:

$$\begin{aligned} x &= (r + h) * cos\,(\varphi) * cos\,(\lambda) \\ y &= (r + h) * cos\,(\varphi) * cos\,(\lambda) \\ z &= (r + h) * sin\,(\varphi) \end{aligned} \tag{2.1}$$

With r the radius of the spherical coordinate system, for geographical coordinates this is the Earth radius, h the height or elevation, φ the latitude in radians, and λ the longitude in radians.

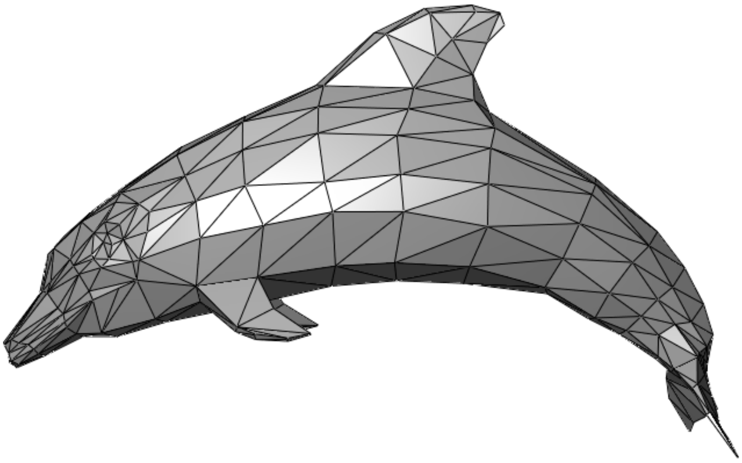


Figure 2.8: Example of a dolphin represented by triangular meshes. Figure from Chrschn (2007).

This coordinate transformation is used in this thesis to compute the geometric area of triangles as well as a preparation step to compute the geometric slope of a triangle.

Rotation of a Cartesian coordinate system

After the transformation of a triangle from the geographical coordinate system into a global Cartesian coordinate system, the position of the z-axis of the Cartesian coordinate system equals the Earth rotation axis (south pole - north pole orientation). In order to compute the slope of a triangle in the global Cartesian coordinate system, the global Cartesian coordinate system must be rotated, so that the z-axis of the coordinate system is perpendicular to the Earth surface at the location of the triangle. Neglecting the Earth curvature, it is sufficient if the z-axis of the rotated coordinate system crosses one vertex of the triangle.

This rotation is done in two steps. First, the x-axis is rotated around the z-axis, so that in the new rotated coordinate system, the y-axis goes through the point to which the coordinate system is rotated to (Hofmann-Wellenhof et al., 2001).

2.4.2 *Triangular Meshes*

A triangle mesh is a type of polygon mesh used in computer graphics, computational geometry and numerical simulation. It is a representation of a 2-D or 3-D surface or domain composed of connected triangles. In a triangular mesh, the surface or domain is divided into a collection of triangles, each triangle being defined by three vertices or corner points. These vertices are connected by edges to form the triangular elements. The connectivity information, which specifies which vertices are connected to form triangles, is essential in defining the overall geometry and topology of the mesh (see: Akenine-Möller et al., 2018; Telea, 2015).

Figure 2.8 shows an example of a dolphin represented by triangles. Triangular meshes have several benefits:

Simplicity: Triangles are the simplest polygonal shape, making triangular meshes easy to work with, both in terms of representation and computation.

Uniformity: Unlike irregular polygons, triangles have a consistent geometry, which simplifies many computational tasks.

Flexibility: Triangular meshes can accurately represent a wide range of shapes and surfaces, including complex and irregular geometries.

Numerical efficiency: Triangular meshes are computationally efficient and well suited for numerical simulations, allowing easy calculation of areas, normals, and other geometric properties.

Interpolation: Triangular meshes facilitate the interpolation of data values across the surface or within the domain, which is useful for various applications such as rendering, visualisation, and physical simulation.

Triangular meshes are commonly used in computer graphics for rendering 3-D models, in finite element analysis for structural and mechanical simulations, and in computational fluid dynamics for modelling fluid flow over surfaces.

Marching cubes

The Marching Cubes (MC) algorithm is a method for generating a triangular mesh of constant density from scalar volumetric data and was developed by Lorensen and Cline (1987). The algorithm was originally developed for medical data, such as 3-D computed tomography (CT) and magnetic resonance (MR) scans. Today, the MC is widely used in various visualisation applications, including computer graphics, medical imaging, and scientific visualisation. In short, the algorithm works as follows:

1. Derive scalar volumetric data into a grid of cubes. Each cube consists of eight vertices, each of which contains the corresponding scalar value of the scalar field.

2. The configuration of the cube is determined by comparing the scalar values of the vertices with the given constant density value. Each vertex can have two possible combinations, either above or below the constant density value. With eight vertices per cube, there are 2^8 (256) possible configurations.

3. Compare the resulting configuration with a pre-computed lookup table. The pre-computed table indicates which of the 14 possible triangulated cubes are present (see Figure 2.9).

4. When applicable, use the scalar values at the vertices of the cube to calculate the positions of the triangle vertices at the edge of the cube by linear interpolation.
5. Compute the normals at each vertex and interpolate the normal to the triangle vertices using linear interpolation.

The final surface mesh of constant density is constructed by rendering all the resulting triangles from the MC algorithm.

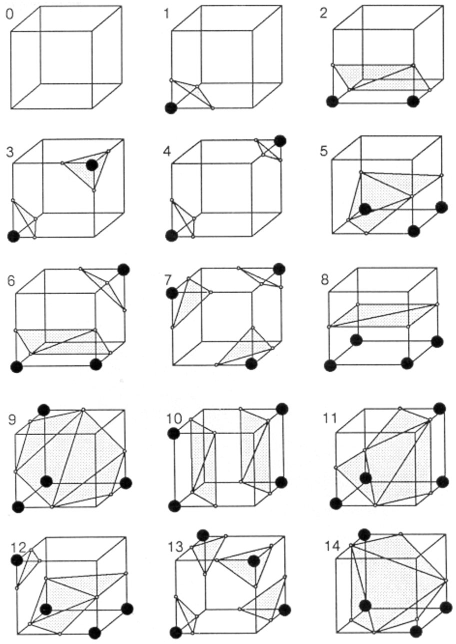


Figure 2.9: Triangulated cubes. Figure used with permission of ACM (Association for Computing Machinery), from Lorensen and Cline (1987); permission conveyed through Copyright Clearance Center, Inc.

Connected triangle sorting algorithm

After generating triangular meshes using the MC algorithm, the information about which triangles form a connected surface is not readily available. It might be possible to implement the MC algorithm in such a way as to obtain this information, but it will not be useful if triangles are subsequently filtered out. This would cause previously connected surfaces to be split, making this information useless. In this work, however, the information about connected triangles becomes essential. In the following, an algorithm is presented to obtain such batches of connected triangles. One triangle is considered connected to another triangle if they share one edge.

I would like to point out the following before diving into the algorithm. The available literature on this particular algorithm is limited and, to the best of my knowledge, I *could not find literature that explicitly describes the following algorithm. However, in general purpose 3-D computer graphics software, such as Blender (2024), similar algorithms are implemented. Therefore, I want to emphasize that I do not claim it as my original work.*

The algorithm is divided into three stages:

First stage: A list is created for each vertex, containing information about which triangles share this vertex.

Second stage: Each triangle is tested to determine if it shares an edge with another triangle. If this condition is met, the information is stored in a list. Subsequently, the algorithm checks if any other triangle connected to one of the vertices of the current triangle shares an additional vertex with the current triangle. If such a connection is found, that triangle is also added to the list.

Third stage: Batches of connected triangles are created by looping over each triangle and checking if it has a connection to another triangle. The result is a list of triangle batches, each batch describing a connected mesh of triangles.

Möller-Trumbore intersection algorithm

To interactively select certain features, such as frontal surfaces, it is necessary to determine the intersection points between a ray and geometries generated from a triangular mesh grid. The goal is to determine the exact location where a ray originating from the mouse cursor intersects a frontal surface in the scene. This information is crucial for selecting individual frontal surfaces with a simple mouse click. Once the intersected triangle is identified, it is trivial to find the corresponding triangle batch of the selected frontal surface.

The ray-triangle intersection algorithm introduced by Möller and Trumbore (1997) is a fast, efficient, and memory-sparing method for testing whether a ray intersects a triangle. The algorithm is designed to compute the intersection without precomputing of the plane equations as in earlier approaches.

Möller and Trumbore (1997) describe their ray triangle intersection algorithm as follows. A ray $R(t)$ is defined as:

$$R(t) = O + tD \tag{2.2}$$

Where D is the normalized direction with distance t from the ray origin O. A triangle is defined by three vortices: V_1, V_2, and V_3. The Möller-Trumbore intersection algorithm constructs a transformation to obtain if a ray intersects a triangle or not. It is even possible to compute the intersection point (u, v) as well as the distance t from the ray origin O to the intersection point (u, v). A point on a triangle $T(u, v)$ is defined in barycentric coordinates as:

$$T(u, v) = (1 - u - v) V_0 + uV_1 + vV_2 \tag{2.3}$$

where $u \geq 0$, $v \geq 0$, and $u = v \leq 1$. The intersection between the ray $R(t)$ and the triangle $T(u, v)$ can be computed by equating Equation 2.2 and Equation 2.3:

$$O + tD = (1 - u - v) V_1 + uV_2 + vV_3 \tag{2.4}$$

Rearranging the terms gives:

$$\begin{bmatrix} -D, & V_2 - V_1, & V_3 - V_1 \end{bmatrix} \begin{bmatrix} t \\ u \\ v \end{bmatrix} = O - V_1 \tag{2.5}$$

This system of linear equations can now be solved to obtain u, v, and t. A graphical solution can be found in Möller and Trumbore (1997), their Figure 1. Equation 2.5 has a solution if and only if the triangle vertices are not collinear and the ray is not parallel to the triangle's plane. This condition holds true if the matrix has three linearly independent column vectors in $\mathbb{R}$ and is invertible. To determine whether the ray intersects with the triangle, and to solve the linear equation system, Cramer's rule can be applied.

2.4.3 *Clustering techniques*

Clustering is a data analysis and machine learning technique used to identify patterns, group similar data points, and organise complex datasets (Steinley, 2006). Clustering is used to divide data into subsets, or clusters, where data points within the same cluster share similarities based on certain features or characteristics. By revealing such structures within data, clustering can help to gain valuable insights into the dataset. This section introduces a collection of algorithms and techniques for clustering time series. The objective is to provide the reader with the necessary background knowledge for the clustering of feature attribute time series. These include k-means clustering (Section 2.4.3), the silhouette score (Section 2.4.3), and the Rand index (Section 2.4.3). Ali et al. (2019) provide an extensive review focused on time series clustering, delving into the application of various distance measures for clustering purposes.

k-means clustering

K-means (MacQueen, 1967) is a clustering algorithm that aims to divide a dataset into k distinct clusters by iteratively assigning data points to the nearest centroid and updating the centroids based on the mean of the assigned points. Although the algorithm was originally developed by Lloyd in 1957, it was not published until 25 years later (Lloyd, 1982). It should also be noted that a similar algorithm was published independently by Forgy (1965). Later in this thesis (see Chapter 6), I employ k-means clustering combined with various distance metrics (Euclidean, Dynamic Time Warping (DTW), and soft-DTW) to cluster time series of feature attributes. For this purpose, I utilise the kernel from the *ts-learn* Python package

(Tavenard et al., 2020), which is based on a modified version of the kernel presented by Dhillon et al. (2004). In general, the k-means algorithm works as follows:

Define the number of clusters. The optimal number of clusters could be determined by the silhouette coefficient (Rousseeuw, 1987).

Initialize the cluster centroids. This is achieved by randomly assigning data points as the initial positions of the centroids.

Assign each data point to a cluster. Calculate for each data point the distance to the cluster centroids. Assign the data points to the cluster with the nearest centroid point. The distance can be calculated using a simple Euclidean distance measure.

Update the cluster centroid position based on the mean of each data point assigned to the cluster. This moves the cluster centroid towards the centre of the data assigned to this cluster.

Repeat the **assign** and **update** step until the convergence criteria are met. A convergence criterion is typically met after a predefined number of iterations or if the centroids of the clusters do not change significantly.

The k-means algorithm is designed to minimize the variance within clusters by selecting similar data points while maximizing the dissimilarity between different clusters.

Silhouette score

When applying k-means clustering, the number of clusters needs to be predefined. The silhouette score (Rousseeuw, 1987) can be a useful metric for determining the optimal number of clusters. In this thesis, I use the silhouette score not only to determine the optimal number of clusters, but also to evaluate the optimal distance metrics for time series clustering. The silhouette score measures the similarity of each data point within a cluster A to the cluster's centroid (coherence) and compares it to the data points of the nearest cluster B (separation). The silhouette score S_i for one data point i of cluster A is calculated as follows:

$$S_i = \frac{b_i - a_i}{max(b_i, a_i)} \tag{2.6}$$

Where a_i is the average dissimilarity of point i to all other data points in cluster A and b_i is the average dissimilarity of point i to all data points of the nearest cluster B. When computing the silhouette score, the difference between b_i and a_i is normalized with the largest difference of the two distances. This results always in a silhouette score S_i between $[-1, 1]$. Where 1 can be interpreted as well separated, 0 as overlapping, and -1 as arbitrary cluster (Shahapure and Nicholas, 2020). The overall silhouette of the clustering can be calculated by taking the arithmetic mean over all individual silhouette scores. This provides an aggregate measure of how well-separated the clusters are on average.

Rand index

The Rand index (Rand, 1971), also known as Rand measure or Rand score, is a statistical measure used to evaluate the similarity or agreement between two data clustering results or partitions. It quantifies the degree of similarity between the data points assigned to the same clusters in two different clustering solutions. Let X and Y be the two clusters to be compared. a is the number of data pairs that are correctly clustered together in the clustering solutions X and Y. b is the number of data pairs that are correctly not clustered together in the clustering solutions X and Y. c is the number of data pairs that are incorrectly clustered together in one clustering solution but not in the other. d is the number of data pairs that are incorrectly not clustered together in one clustering solution but are clustered together in the other. The Rand index is then calculated as follows:

$$R = \frac{a+b}{a+b+c+d} \tag{2.7}$$

A score of 1 indicates that the two clustering solutions are identical, and all data points are correctly clustered in both solutions. A score of 0 indicates that the two clustering solutions are completely dissimilar, and there is no agreement between them. The Rand index was employed by Kumpf et al. (e.g., 2018) to assess cluster robustness by comparing clusters generated with varying input parameters. In this thesis, I adopt the approach outlined by Kumpf et al. (2018) to test the sensitivity of my resulting time series clusters of feature attributes. This evaluation examines the sensitivity across various feature attributes and varying input parameters used to derive the features and, subsequently, the feature attributes.

2.4.4 *Distance measures*

Euclidean distance is a commonly used and effective measure for quantifying the dissimilarity or similarity between objects, data points, or distributions. In two dimensions, it represents the length of the line segment connecting the two points. In higher dimensions, it generalises to the square root of the sum of the squared differences in coordinates (Dokmanic et al., 2015). However, for time series similarity analysis, Euclidean distance is not well-suited because it cannot account for variations such as compression or stretching of elements (Ali et al., 2019). In the following sections, two distance measures – DTW and soft-DTW – are introduced, which address this limitation of Euclidean distance when applied to time series similarity analysis.

Dynamic Time Warping

DTW (Sakoe and Chiba, 1978) is a widely used technique employed to measure the similarity between two sequences that may differ in length or contain temporal distortions. Its applications span various fields, including music and signal processing (e.g., Müller, 2007; Müller et al., 2006) and time series analysis (e.g., Bahlmann and Burkhardt, 2004; Kahveci et al., 2002; Niennattrakul and Ratanamahatana, 2007). The foundation of DTW dates back to the 1970s when it was initially developed for speech and word recognition (Sakoe and Chiba, 1971; Velichko and Zagoruyko, 1970).

The fundamental concept underlying DTW involves determining an optimal alignment between the elements of two sequences, denoted as $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_m)$, with respective lengths of n and m. This alignment is achieved through the warping and stretching of the time axis, allowing for a non-linear alignment that can handle differences in the speed or timing of events in the sequences.

This work employs the DTW algorithm implemented in the Python package called *ts-learn* (Tavenard et al., 2020). The optimisation function is defined as follows:

$$DTW(x, y) = \min_{\pi} \sqrt{\sum_{(i,j)\in\pi} d\left(x_i, y_j\right)^2} \tag{2.8}$$

Let $\pi = [\pi_1, \ldots, \pi_K]$ denote the paths between the points of the two time sequences, each π subjects to the following conditions:

- $\pi_k = (i_k, j_k)$ with $1 \leq i_k \leq n$ and $1 \leq j_k \leq m$
- $\pi_1 = (1, j_k)$ or $\pi_1 = (i_k, 1)$ and $\pi_K = (n, j_k)$ or $\pi_K = (i_k, m)$
- for all $k > 1, \pi_k = (i_k, j_k)$ is related to the previous pair $\pi_{k-1} = (i_{k-1}, j_{k-1})$ with $i_{k-1} \leq i_k \leq i_{k-1} + 1$ and $j_{k-1} \leq j_k \leq j_{k-1} + 1$

π represents a temporal alignment of the two sequences, where the objective is to minimise the Euclidean distance between the aligned. Figure 2.10 shows a comparison between the paths π using Euclidean matching (matching same time steps in both sequences) and DTW.

Soft-DTW

Soft-DTW (Cuturi and Blondel, 2017) is a smoothed version of the original DTW algorithm. The key advantage of soft-DTW is its differentiability, allowing for the computation of both the loss function and its gradient. This differentiability property enables more efficient optimisation and analysis compared to the non-differentiable original DTW. Soft-DTW is particularly suitable for time series clustering tasks and has been shown to outperform the original DTW (Cuturi and

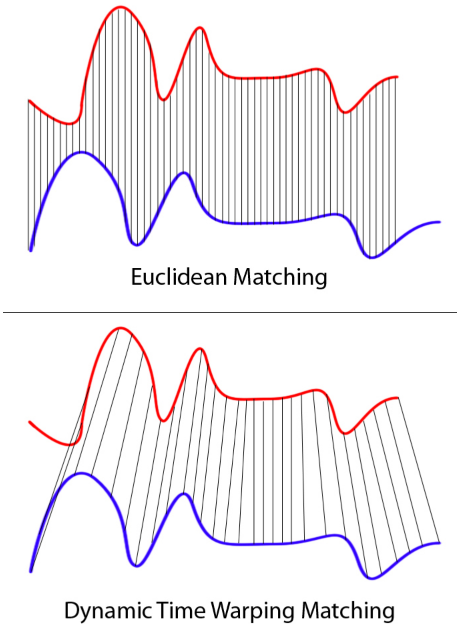


Figure 2.10: Difference in matching between Euclidean and DTW distance measures for time series analysis. Figure from XantaCross (2011), used under CC BY-SA 3.0.

Blondel, 2017). In soft-DTW, the non-differentiable *min* operator is replaced with a differentiable function, which facilitates gradient-based computations.

$$\text{soft-min}\,(a_1, \ldots, a_k) = -\gamma \log \sum_i e^{-a_i/\gamma} \tag{2.9}$$

Soft-DTW uses a hyper-parameter γ to control the smoothing of the resulting metric. As γ increases, the metric becomes smoother, approaching the limit case of squared DTW when γ approaches 0. γ allows for a trade-off between alignment flexibility and the desired level of smoothing.

Earth Mover's Distance

The Earth Mover's Distance (EMD) (Rubner et al., 2000) is a metric used to calculate the dissimilarity between two probability distributions. It quantifies the minimum "work" or effort required to transform one distribution into the other. Originally invented for image processing applications, EMD can be applied to measure the similarity between distributions. In image analysis, it allows images to be compared by considering the amount of 'mass' that needs to be moved from one image to match the distribution of another, providing a meaningful measure of image similarity. Beyond image analysis, EMD has applications in several fields, including transportation and logistics, where it can be used to measure the cost or effort required to transform one distribution of goods into another, aiding optimisation and planning processes (Ji and Shen, 2006). The EMD is more robust than other matching techniques because it tolerates a certain amount of feature deformation, resulting in greater accuracy (Rubner et al., 2000).

Ji and Shen (2006) apply the EMD to track time-varying features by comparing the distributions of geometric properties. Building on this approach, I have developed a methodology for tracking atmospheric features in NWP simulations. Unlike Ji and Shen (2006), my method does not use EMD to compare the geometric properties of features. Instead, I use the EMD as a cost function to evaluate feature similarities based on their physical properties. My feature tracking technique uses EMD to evaluate the minimum cost between distributions of physical quantities of atmospheric features, such as the distribution of thermal properties on feature surfaces.

2.4.5 *The concept of normal curves*

The concept of normal curves (NCs) is not a conventional distance measure and was introduced as a means of estimating distances within a scalar field that undergoes temporal displacement (Pfaffelmoser et al., 2011). These NCs traverse the surface orthogonally, following the direction of the surface's temporal gradient. Rautenhaus (2015) uses NCs to visualise the internal structure of isosurfaces, aiding in the identification of local extrema within the isosurface by observing the convergence of NCs towards these points.

A NC is defined as a path along the gradient direction of a scalar field. In the previously mentioned visualisation (Rautenhaus, 2015), where NCs are used to visualise the internal structures of isosurfaces, NCs start from predefined points on an isosurface. The end points of the NCs are defined by the isovalue of the isosurface. The NCs follow the gradient of the scalar field within the isosurface and end at the point where they would leave the isosurface. The definition of NCs from Kern et al. (2019) is adopted in this work.

The difference S of a scalar field Ψ along a NC is defined as follows:

$$S_{\Psi} \equiv \int_{NC} |\nabla_h \Psi| \, ds \tag{2.10}$$

3

METEOROLOGICAL BACKGROUND

Contents

3.1 Atmospheric structure and Motion 31
3.2 Numerical weather prediction and ensemble simulations 35
3.3 Extratropical cyclones and atmospheric fronts 36
3.4 Front detection and feature tracking methods 41

This chapter introduces the reader to the fundamentals of meteorology. It provides information on the current state of meteorological science, as well as some historical developments in the field, particularly with regard to atmospheric fronts and feature detection.

3.1 ATMOSPHERIC STRUCTURE AND MOTION

The vertical structure of the Earth's atmosphere is a fundamental aspect of atmospheric science. Starting at the surface and extending up to an altitude of around 11 kilometres, depending on the latitude, we encounter the troposphere. This layer is characterised by a steady decrease in temperature with increasing altitude. In particular, the troposphere is the primary domain in which the vast majority of weather phenomena and cloud formations occur. Beyond the troposphere is the stratosphere, ranging from approximately 11 to 47 kilometres of height. In this region, the temperature increases with height, which is caused by the ozone layer.

The height of the tropopause, a layer between the troposphere and stratosphere, and the troposphere varies considerably due to temperature differences from the equator to the poles. At the equator, the tropopause and troposphere extend to about 18 kilometres, while at the poles they are limited to about 8 kilometres. In the mid-latitudes, the tropopause typically occurs at an altitude of about 11 kilometres. The tropospheric structure is not static, it varies both temporally and spatially due to ever-changing weather conditions and solar activity, making it a dynamic and complex system that continually shapes the weather of our planet.

Figure 3.1 illustrates the global, large-scale circulation patterns within the Earth's troposphere. The tropospheric circulation is driven by the solar radiative heating imbalance caused by variable levels of solar radiation received at different

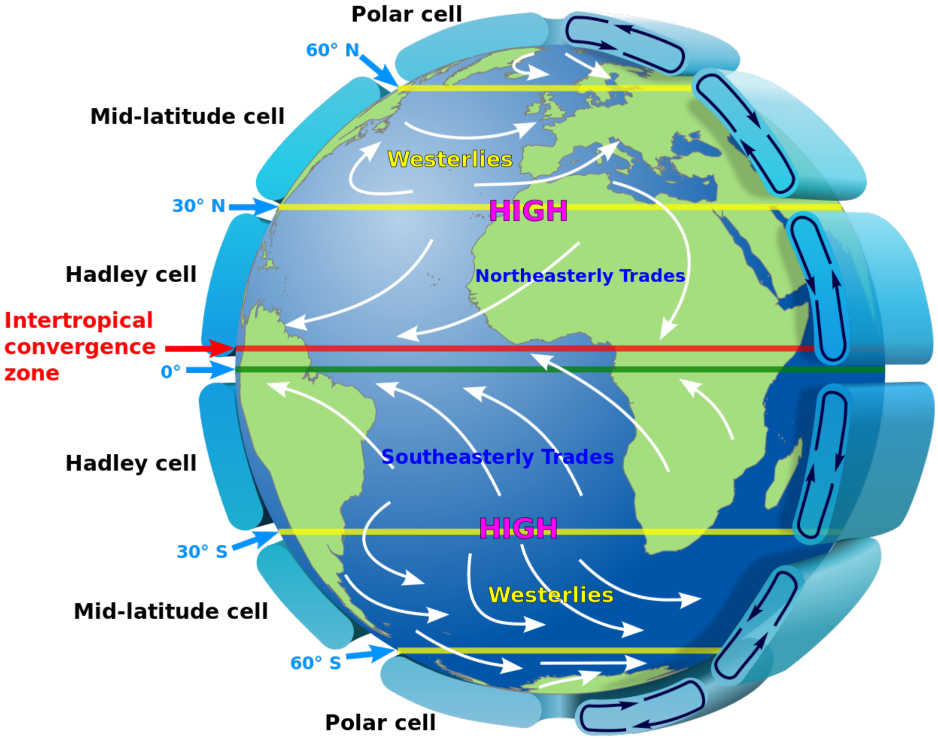


Figure 3.1: Illustration of idealized large-scale tropospheric circulation. Figure from Kaidor (2013), used under CC BY-SA 3.0.

latitudes. Strong solar radiation at the equator creates an energy surplus, while relatively weak solar radiation and constant radiative cooling at the poles create an energy deficit. The differential radiative energy is balanced by horizontal heat transport associated with the large-scale circulation. This circulation also has a strong zonal component due to the Earth's rotation. Focusing on the northern hemisphere for simplification, the general atmospheric circulation is defined by three primary bands of circulation.

The first of these bands, also known as the Hadley cell, extends from the equator to around 30 degrees north. In this cell, the predominant surface winds are the north-east trade winds. At the equator, the south-easterly and north-easterly trade winds converge, forming the Intertropical Convergence Zone (ITCZ), where the air rises. At around 30 degrees north, an area of high pressure known as the subtropical high prevails. These areas are characterised by subsidence, dry air, calm winds, clear skies, and the absence of precipitation.

The second band, the Ferrel or mid-latitude cell, develops between 30 and 60 degrees north. Within this cell, transient centres of low and high pressure develop. Low-pressure regions or cyclones typically bring severe weather conditions such as heavy rain, high humidity, and strong winds. In contrast, high-pressure regions, typified by fair weather conditions, are marked by clear skies, dry air, and the absence of precipitation. These pressure systems move from west to east, resulting in fluctuating but generally westerly winds and unstable weather. There is a region around 60 degrees north where a band of sub-polar lows forms near the surface. In this region, the weather is unstable with widespread cloud cover and precipitation.

The third band, the polar cell, forms between 60 degrees north and the pole. In the vicinity of the poles, the polar high, a region of high pressure in climatology, develops. Here, the sky stays clear, and the air is cold and dry.

In summary, the large scale circulation is driven by radiation imbalances and the Earth's rotation resulting in different characteristic regimes.

The description of atmospheric motion can be found in many textbooks. Here I follow the description of Etling (2008) and the summary by Rautenhaus (2015) and give a brief overview of the quantitative description of the atmosphere and its equations of motion. The state of the dry atmospheric is quantitatively described by the velocity vector $\vec{v}$, the pressure p, the density ρ, and the temperature T. The motion of the atmosphere can be described by the following six prognostic equations.

Conservation of momentum

Conservation of momentum is a fundamental principle that describes the motion of air particles under the influence of forces. This principle of conservation of momentum is derived from Newton's axiom, which states that the application of force to a body results in its acceleration.

When considering an air particle, it is subject to several forces that affect its motion. These forces are described by three prognostic equations and are summarised in Equation 3.1. First, the Coriolis force, which is caused by the rotation of the Earth; second, the gravitation force, which pulls particles towards the Earth's centre of mass, accelerating them downwards; third, the pressure gradient force, which results from pressure differences within the atmosphere.

$$\underbrace{\frac{\partial \vec{v}}{\partial t}}_{\text{Local wind speed tendencies}} + \underbrace{\vec{v} \cdot \nabla \vec{v}}_{\text{Advection}} + \underbrace{2\Omega \times \vec{v}}_{\text{Coriolis force}} = - \underbrace{-\nabla \Phi}_{\text{Gravitation force}} - \underbrace{-\frac{1}{\rho}\nabla p}_{\text{Pressure gradient force}} \tag{3.1}$$

Conservation of mass

The principle of conservation of mass states that, within a unit volume, the outflow or divergence of mass from the volume is equal to the loss of mass within the volume. In a unit volume, the change in the local mass content is the result of advection. According to the conservation of mass (Equation 3.2), the sum of advection and local mass changes must equal the negative divergence of mass within the volume.

$$\underbrace{\frac{\partial \rho}{\partial t}}_{\text{Local density tendencies}} + \underbrace{\vec{v} \cdot \nabla \rho}_{\text{Advection}} = - \underbrace{\rho \nabla \cdot \vec{v}}_{\text{Divergence}} \tag{3.2}$$

Conservation of energy

The first law of thermodynamics is a fundamental principle that describes the conservation of energy in a system. According to the law of conservation of energy, the change in internal energy of a closed system is equal to the sum of the heat supplied to the system and the thermodynamic work performed by the system. Equation 3.3 describes the conservation of energy, where c_p denotes the specific heat capacity at constant pressure and $\dot{Q}$ the energy input per unit mass.

$$\underbrace{\frac{\partial T}{\partial t}}_{\text{Local temperature tendencies}} + \underbrace{\vec{v} \cdot \nabla T}_{\text{Advection}} - \underbrace{\frac{1}{\rho c_p} \frac{dp}{dt}}_{\text{Adiabatic processes}} = \underbrace{\frac{\dot{Q}}{c_p}}_{\text{External heat exchange}} \tag{3.3}$$

Ideal gas equation

The equation for an ideal gas in a closed system can be derived from two fundamental laws of physics. First, for an ideal gas at constant pressure p, the ratio of volume V to temperature T is constant. Second, for an ideal gas at constant temperature T, the product of pressure p and volume V remains constant. Combining these two fundamental properties of gases, the ideal gas equation (Equation 3.4) can be derived, where R is the specific gas constant and may vary depending on the gas.

$$p = R \rho T \tag{3.4}$$

The six fundamental equations describing atmospheric motion have certain limitations and simplifications. They are generally valid only for dry air, which means that the equations do not take water vapour into account. However, water vapour plays a crucial role in many atmospheric processes, such as cloud formation and precipitation, and its omission limits the system's ability to fully capture these phenomena. Another simplification in the equations above is the absence of friction. Friction plays an important role in atmospheric dynamics, particularly at smaller scales and in interaction with the surface, where it affects the movement of air masses and the development of weather systems. Despite these limitations, the equation system is capable of describing a wide range of atmospheric phenomena. It can capture processes such as gravity waves, cyclones, and sound waves. For practical weather forecasting purposes, the equations are further simplified. Some processes, such as sound waves, are filtered out because their effect on the weather is considered negligible. This simplification allows for more efficient calculation and facilitates the use of the equations in real-time weather forecasting. For numerical weather prediction on the synoptic scale, the equations are transformed into the p-system and into a spherical coordinate system. The p-system uses pressure as the vertical coordinate instead of geometric height. This choice of coordinate is advantageous for the description of synoptic processes, as pressure is closely related to the large-scale dynamics of the atmosphere. The p-system provides a more convenient framework for the analysis and prediction of synoptic-scale weather patterns (Etling, 2008).

However, for some applications, it can be useful to convert pressure levels to geometric height. This is typically done using geopotential height. If geopotential height is not available then the height can be approximately calculated using the barometric formula and typical values of the International Standard Atmosphere (ISA) (DWD, 2014).

3.2 NUMERICAL WEATHER PREDICTION AND ENSEMBLE SIMULATIONS

To utilise Equations 3.1 - 3.4 for NWP prediction, they are approximated using finite difference methods. This approximation transforms the fundamental equations into an algebraic system of equations that can be solved numerically. The numerical solution is computed on a discrete grid.

In the horizontal plane, the grid spacing, denoted as Δx and Δy, determines the horizontal resolution of the numerical model. Smaller grid spacing leads to higher spatial resolution, allowing for more detailed representation of atmospheric features and processes. The vertical dimension can also be discretised with different methods. A common representation is vertical pressure coordinates, where the vertical levels are defined based on pressure levels in the atmosphere. The pressure levels are usually logarithmically spaced, with higher resolution in the lower atmosphere where most weather phenomena occur.

Modern NWP models often use hybrid vertical coordinates that follow the terrain in the lower troposphere. With increasing altitude, they become more and more similar to the geometric or pressure altitude isosurfaces. Above a certain altitude, they follow the iso-pressure or iso-geometric surfaces. The ECMWF models use so-called hybrid sigma-pressure coordinates, while the ICON model uses hybrid geometric coordinates.

NWP models run operationally on high-performance computers and are updated frequently with the latest weather observations, in a process called data assimilation, to produce the most accurate forecasts possible.

The accuracy of NWP diminishes substantially beyond a certain time range. These limitations arise from the non-linear nature of the equations of atmospheric motion and the imperfect knowledge of the initial state. Nearly similar initial conditions can produce different results. Furthermore, NWP simulations rely on approximations and parametrisations to represent complex atmospheric processes, such as cloud formation and dynamics. These parametrisations introduce

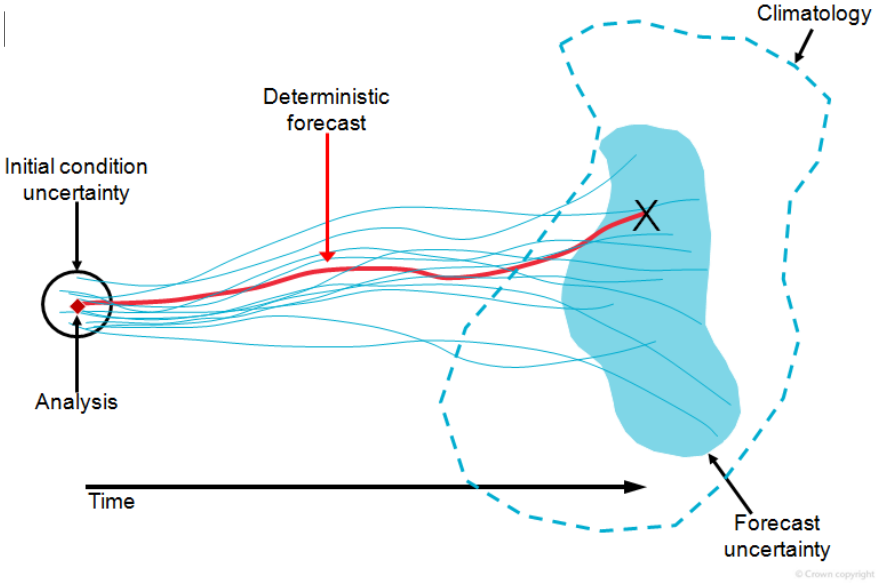


Figure 3.2: Schematic illustration of an ensemble forecast. Cropped figure from Met Office (2023), © British Crown copyright, Met Office, used with permission.

additional sources of uncertainty into the forecast. The predictability limit varies, but generally increases with the spatial and temporal scales of the meteorological phenomena. For synoptic scales (approx. 1000 km), the predictability limit is typically about four to five days (Coiffier, 2011).

To account for this uncertainty, ensemble simulations are used to estimate uncertainties in the forecast. The ensemble simulation is essentially a set of forecasts for the same area and time, but each forecast differs slightly in its initial conditions. The ensemble forecast can be thought of as a probability density function. Figure 3.2 illustrates an ensemble forecast schematically. When the forecast is initialised, the difference between the individual ensemble members is small, as it represents the perturbations of the initial forecast state. Over time, the divergence between the individual members of the ensemble gradually increases. An ideal ensemble forecast would represent the full range of possible weather events, and therefore the true weather state will always be somewhere within the ensemble spread.

3.3 EXTRATROPICAL CYCLONES AND ATMOSPHERIC FRONTS

Extratropical cyclones are a prominent and important feature of the atmosphere and are described in many meteorology textbooks. Here I follow the comments of Wallace and Hobbs (2006) and Stull (2017) and concentrate on extratropical cyclones in the Northern Hemisphere. In these cyclone systems, the primary force balance is maintained between the pressure gradient force and the Coriolis force, as expressed by Equation 3.1. Extratropical cyclones are large-scale weather systems characterised by low pressure near the surface, and are often associated

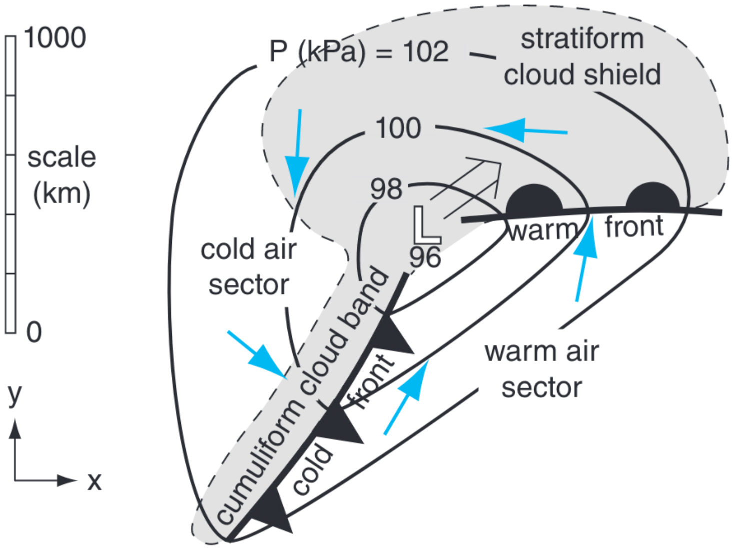


Figure 3.3: Illustration of an extratropical cyclone and accompanying cyclonic features: clouds (grey shaded areas), pressure isolines (black solid lines), near-surface wind direction (blue arrows), and fronts. Figure from Stull (2017), used under CC BY-NC-SA 4.0

with mesoscale fronts, strong winds, and heavy precipitation. The wind in extratropical cyclones rotates cyclonically around the cyclone centre, while the whole cyclone system typically moves from west to east and slightly poleward, driven by the jet stream. Figure 3.3 illustrates an extratropical cyclone, including features such as cloud formations, wind direction, pressure isobars, and fronts.

There is a wide variety of extratropical cyclones, and their development depends on factors such as the background flow, moisture availability, and surface characteristics, including surface temperature and roughness length (mainly differences between land and sea). Winter conditions are favourable for the development of strong extratropical cyclones.

Extratropical cyclones can last from a few days to about two weeks. The formation of a cyclone, called cyclogenesis, is characterised by an increase in vorticity (spinning motion), updraft, and a decrease in surface pressure. Cyclogenesis is controlled by three main factors. First, the updraft over the synoptic scale causes a decrease in surface pressure. Second, pressure differences between the cyclone centre and the surrounding regions generate strong horizontal winds that attempt to equalise the pressure gradients. These winds blow along the isobars, balancing the pressure gradient force and the Coriolis force. Third, surface friction counteracts the Coriolis force, resulting in spiral winds near the surface blowing towards the cyclone centre. These winds counteract the removal of air by the updrafts. The state of cyclolysis is reached when the horizontal convergence of spiral winds near the surface is larger than the divergence caused by the updraft. As a result, the centre of the cyclone fills, causing the horizontal pressure gradients to decrease and the cyclone to dissipate. Extratropical cyclones contribute up to 85-90% of annual precipitation (Hawcroft et al., 2012) and up to

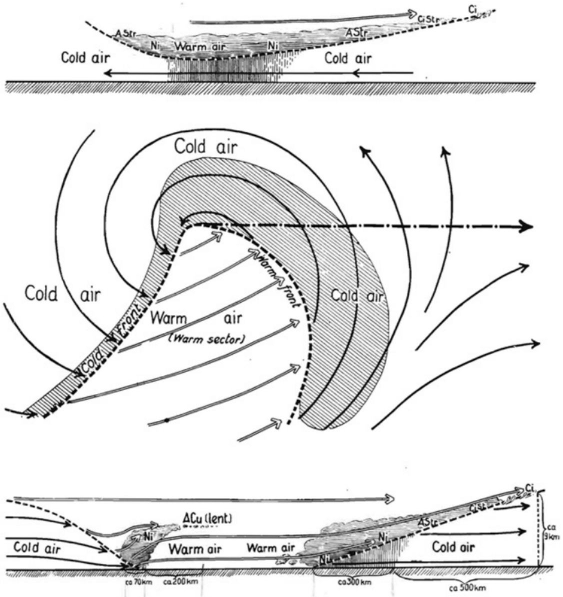


Figure 3.4: Idealised cyclone presented by the Bergen school (Bjerknes and Solber, 1922). Figure used with permission of the Norwegian Geophysical Society.

80% of intense precipitation events (Pfahl and Wernli, 2012) in some mid-latitude regions.

Extratropical cyclones are accompanied by atmospheric features called fronts. Fronts arise from baroclinic zones, which are long and narrow regions characterised by strong temperature gradients over a short horizontal distance. A strong baroclinic zone occurs in the mid-latitudes, driven by global circulations. This is where the warm air of the Ferrel cell meets the cold air of the Polar cell (see Section 3.1). Under certain conditions, the jet stream, strong westerly winds near the tropopause, can cause a perturbation in the baroclinic zone, leading to the formation of a small low-pressure system within the baroclinic zone. This phase is known as the *spin-up* phase of an atmospheric front.

During the *spin-up* phase, warm air from the south of the baroclinic zone is advected to the north, east of the centre of the low. Meanwhile, cold air from the north of the baroclinic zone is advected towards the equator, west of the centre of the low. This baroclinic zone takes on a wavy shape, called a *frontal wave*. The western part of the baroclinic zone, where warm air is advected towards colder regions, is called the *warm front*. The eastern part of the low centre of the baroclinic zone, where cold air is advected towards warmer regions, is called the *cold front*.

With favourable conditions in the jet stream, the low-pressure system continues to deepen. As the warm air is forced to rise above the colder air, cloud formation occurs. The pressure continues to drop as cyclogenesis progresses, leading to an intensification of winds and clouds around the centre of the cyclone, forming a vortex-shaped cloud layer. The cold front typically moves faster than the warm front. The most intense phase of the cyclone is reached when the pressure

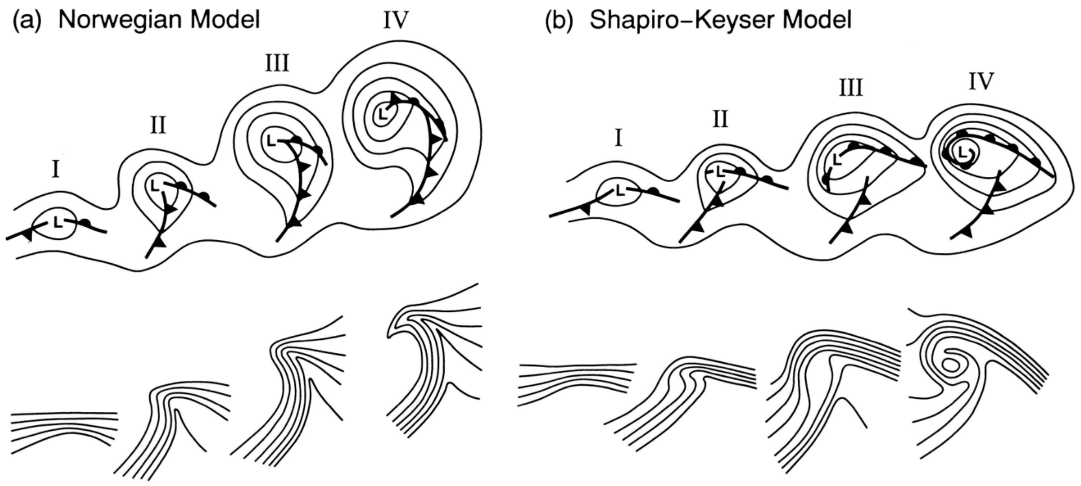


Figure 3.5: Conceptual models of cyclone evolution showing lower-tropospheric geopotential height and fronts (top), and lower-tropospheric potential temperature (bottom). (a) Norwegian cyclone model: (I) incipient frontal cyclone, (II) and (III) narrowing warm sector, (IV) occlusion; (b) Shapiro–Keyser cyclone model: (I) incipient frontal cyclone, (II) frontal fracture, (III) frontal T-bone and bent-back front, (IV) frontal T-bone and warm seclusion. Figure from Schultz et al. (1998), their Figure 15. © American Meteorological Society. Used with permission.

minimum at the centre of the cyclone is reached. At this point, the cold front catches up with the warm front and forms an occluding front near the centre of the cyclone. The dry, cold air begins to wrap around the centre of the cyclone and the occluded front, marking the start of cyclolysis. As cyclolysis progresses, the cyclone vorticity weakens and the central pressure begins to rise.

Frontal systems in mid-latitude cyclones were first comprehensively studied by Jacob Bjerknes (Bjerknes, 1919; Bjerknes and Solber, 1922), a Norwegian-American meteorologist and pioneer in weather analysis and forecasting. Bjerknes made significant contributions to the understanding of the dynamics of mid-latitude weather systems, including the development of the concept of atmospheric fronts. Figure 3.4 illustrates his work in describing the general structure of frontal systems, which is remarkably detailed given the limited availability of observational data and the lack of modern simulated numerical weather prediction methods at the time.

In the next decades, the study of extratropical cyclones and atmospheric fronts continued (e.g. Anderson et al., 1955; Berggren, 1952; Bjerknes, 1930, 1935; Bjerknes and Palmén, 1937; Nyberg et al., 1942). From Bjerknes (1919) and Bjerknes and Solber (1922), a general front model is derived, also known as the Norwegian cyclone model (Figure 3.5a). The Norwegian cyclone model describes the frontal development with four characteristic frontal stages:

I. Incipient frontal cyclone: A small perturbation along the polar front or baroclinic zone initiates the formation of a frontal wave as a depression forms. This perturbation causes a kink in the polar front, leading to the development of warm and cold fronts.

II. Intensification: The frontal wave continues to intensify and the surface pressure of the depression decreases. Temperature gradients become more pronounced and changes in wind direction can be observed at the warm front.

III. Mature stage: The frontal wave continues to intensify, accompanied by a further decrease in pressure. At this stage, the cold front begins to catch up with the warm front, resulting in the formation of an occluded front. The occluded front is where the cold front overtakes the warm front and marks the most mature stage of the cyclone.

IV. Occlusion: As the cold front continues to catch up with the warm front, the occlusion process intensifies. The low-pressure system begins to weaken as the cyclone and the frontal system starts to dissipate. This stage marks the beginning of the end of the cyclone's life.

The Norwegian cyclone model was further developed and supplemented by Shapiro and Keyser (1990), resulting in the Shapiro-Keyser cyclone model (Figure 3.5b). Unlike the original Norwegian cyclone model, the Shapiro-Keyser model introduces additional features as it develops. In Stage II of the Shapiro-Keyser model, the frontal waves break, leading to significant changes in the cyclone structure. Stage III is characterised by the formation of a T-bone structure, where the cold front is perpendicular to the warm front. Finally, in Stage IV, the occluded front wraps-up around the centre of the cyclone, resulting in the formation of a warm core seclusion.

Hewson and Titley (2010) added four stages to the cyclone model proposed by Shapiro and Keyser (1990), resulting in a total of eight stages. The first two stages, Stage 0 and Stage 1, come before the stages of the Shapiro-Keyser cyclone model. Stage 0 represents the initial state where the baroclinic zone is unperturbed. Stage 1, the diminutive frontal wave stage, represents the initial perturbation of the baroclinic zone. After the first two stages, stages 2 to 5 correspond to the cyclone stages I-IV of the Shapiro-Keyser cyclone model. The other two additional stages follow after the stages of the Shapiro and Keyser cyclone model. Stage 6, the mature cyclone state, is characterised by a pronounced wrap-up of the occluded front around the cyclone centre. Stage 7, the decay stage, represents the filling of the depression and frontolysis, leading to the dissipation of the cyclone and the weakening of the fronts.

The cyclone models can be accompanied by coherent circulation features called conveyor belts. The cold conveyor belt occurs ahead of the warm and occluded fronts, usually remaining below 850 hPa. It is often associated with high wind speeds in later stages, typically south-west of the cyclone centre. The Warm Conveyor Belt (WCB) (see Eckhardt et al., 2004; Madonna et al., 2014) occurs ahead of the cold front near the surface in the early stages and is also associated with high wind speeds. It typically ascends at least 600 hPa in the warm sector and over the warm front, and often splits into anti-cyclonically and cyclonically turning branches (Martínez-Alvarado et al., 2014).

The study of extratropical cyclones, their cyclonic features, and their weather impacts continues to be a major area of investigation in current scientific research (Dacre and Pinto, 2020; Kumler-Bonfanti et al., 2020; Raible et al., 2021; Schultz and Vaughan, 2011; Sinclair et al., 2020; Thomas and Schultz, 2019a,b). Schultz et al. (2019) provides a comprehensive and detailed review of a century of research on extratropical cyclones.

3.4 FRONT DETECTION AND FEATURE TRACKING METHODS

Atmospheric front detection methods aim to objectively identify and locate boundaries between air masses with distinct thermal properties. Since Renard and Clarke (1965) introduced the first front detection algorithm, different objective front detection methods have been developed and published. Based on Hewson (1998, their Table 1) and supplemented with recent approaches, Table 3.1 shows a selection of objective front detection methods. Most of the detection methods utilise gradient-based approaches, with the TFP being the most prominent and widely used. The TFP is employed both as a detection method and as a filter. A filter in the context of feature detection removes features that have been detected but are unwanted, such as a weak front. When the TFP is used as a detection method, it identifies narrow bands of high thermal gradients rather than detecting a frontal line directly. Manual drawing of frontal lines using the assistance of the TFP has been employed by Renard and Clarke (1965) and Zwatz-Meise and Mahringer (1988). Another approach, introduced by Jenkner et al. (2009), involves further filtering the TFP to narrow the band of strong thermal gradients. With T the temperature, θ the potential T, θ_w the wet-bulb θ, θ_e the equivalent θ, qv the specific humidity, u the zonal wind component, v the meridional wind component, w the vertical wind component.

Recent advancements in front detection include the work by Niebler et al. (2022), who utilises a Convolutional Neural Network (CNN) with a U-Net architecture (Ronneberger et al., 2015). The CNN is trained on Fifth Generation ECMWF Reanalysis for the Global Climate and Weather (ERA5) data and labelled polygons of frontal lines serving as ground truth. The polygon fronts are provided by weather services (National Weather Service of the United States) or extracted from surface analysis charts (DWD). The CNN is trained on two distinct regions: North America and Europe. A notable advantage of the CNN approach over traditional thermal variable methods is its ability to handle multiple input variables, such as temperature, specific humidity, zonal and meridional wind speed, and upward wind speed. This allows for a more comprehensive understanding of front dynamics. According to Niebler et al. (2022), the CNN demonstrates a critical success rate of approximately 66.9% and an object detection rate of about 77%. The authors claim that their approach outperforms traditional TFP front detection

Table 3.1: Overview and summary of selected objective front detection methods.

Scheme / Name	**Vertical level**	**Variable**	**Equation**	**Method**	**Filter / Masking**
Renard and Clarke (1965)	850 hPa 2-D	Potential temperature (θ)	$TFP(\tau) = -\nabla\lvert\nabla\tau\rvert \cdot \frac{\nabla\tau}{\lvert\nabla\tau\rvert}$	Maximum ridge of TFP	TFP $>$ small positive value
Clarke and Renard (1966)	850 hPa 2-D	Temperature (T)	$\nabla\tau \cdot \frac{\nabla(TFP(\tau))}{\lvert\nabla(TFP(\tau))\rvert}$	Zero contour	Where TFP $>$ 0
Cahir and Lottes (1982)	Surface 2-D	Wet-bulb potential temperature (θ_w)	$\lvert\nabla\lvert\nabla_\tau\rvert$	Discard gridboxes with x or y maximum	$\nabla^2\tau < 0$ $\nabla^2 P_{MSL} > 0$
Zwatz-Meise and Mahringer (1988)	500-850 hPa 2-D	Layer mean temperature	$TFP(\tau)$	Maximum ridge of TFP	Evaluation of cloud bands in satellite images
Hewson (1998)	∼1km 2-D	θ_w	$L(\tau) = \frac{\partial(\lvert\nabla_h\lvert\nabla_h\tau\rvert\rvert)_s}{\partial s}$	Zero contour of $L(\tau)$	TFP and ABZ small positive value
Jenkner et al. (2009)	700 hPa 2-D	Equivalent potential temperature (θ_e)	$TFP(\tau)$	Zero contour of TFP	$\nabla(\tau) > 4.5K$
Kern et al. (2019)	All 3-D	θ_w	$L(\tau)$	Zero contour	$TFP > 0$, Gradient within frontal zone
Niebler et al. (2022)	Surface 2-D	T, specific humidity, wind	CNN with U-Net architecture (Ronneberger et al., 2015).	Trained on surface analysis charts.	-
Niebler et al. (2023)	500 hPa to 1000 hPa 3-D	T, specific humidity, wind	CNN with U-Net architecture (Ronneberger et al., 2015).	Trained on surface analysis charts and extended to various height level.	-
Nellikkattil et al. (2024)	All 3-D	*Thermal variable*	SCAFET	Geometric shape descriptors	Small and weak object filter

methods. However, it should be noted that the network's transferability to regions other than those it was trained on is not straightforward, as highlighted in their study. Building upon their 2-D front detection method, Niebler et al. (2023) extend their approach to 3-D. The authors demonstrate that the CNN trained on surface fronts reliably detects fronts at pressure levels ranging from 500 hPa to 1000 hPa.

Scalable Feature Extraction and Tracking (SCAFET) (Nellikkattil et al., 2024) is a framework designed for generalised feature detection and feature tracking. This framework relies on feature shape descriptors, allowing it to be applied to a wide range of numerical datasets and grid structures. The authors of the framework demonstrate that SCAFET can detect various atmospheric features, such as jet streams, atmospheric rivers, tropical and extratropical cyclones, and sea surface temperature fronts. The desired feature is described by a feature shape index, which provides a quantitative description of the geometric shapes. Although SCAFET has not yet been specifically utilised for atmospheric fronts, the initial findings from Nellikkattil et al. (2024) are promising. Given its demonstrated capability to detect 2-D sea surface temperature fronts and 3-D atmospheric jet streams, it can be reasonably expected that the detection of 3-D atmospheric fronts is within its reach. With its scalable feature extraction and tracking capabilities, the authors claim that SCAFET can be effectively applied to a range of applications.

Feature tracking methods have been developed to understand the evolution of features and monitor their changes over time. Feature tracking is generally done by detecting features independently over consecutive time steps. Once the features of two time steps are detected, feature tracking is applied. Feature tracking is a key aspect of understanding time-varying datasets (Saikia and Weinkauf, 2017), particularly in the field of meteorology where such time-varying datasets are commonly generated and used. Feature tracking allows one to focus on specific regions of interest and to systematically monitor their transformations and interactions over time. In recent years, the field of feature tracking has developed several approaches tailored to its specific application in meteorology (Nellikkattil et al., 2024; Shields et al., 2018; Ullrich and Zarzycki, 2017). Feature tracking is usually a combination of a matching algorithm and a filter. The matching algorithm can be distance-based, or it can compare feature characteristics to match features from the previous time step that have very similar characteristics to features of the current time step. A filter can be some kind of region filter, because features can only move a certain distance between consecutive time steps. Features outside this region can then be excluded from the matching process. The following is a brief overview of a selection of these methods, highlighting their relevance and utility in a meteorological context.

Ji and Shen (2006) present a methodology for tracking 3-D time-varying ellipsoids in a force field dataset using the EMD. The EMD algorithm, discussed in detail in Section 2.4.4, facilitates the comparison of two distributions and the computation of the matching cost between them. In their study, Ji and Shen (2006) present a global optimisation algorithm specifically designed for tracking time-varying features. This algorithm is versatile, allowing the tracking of features represented by spatial cell distributions or other forms of distributional data. In particular, it uses the EMD as a metric to evaluate feature dissimilarity, ensuring precise computation of matching costs between source and target feature components.

Hewson and Titley (2010) present a framework that focuses on tracking cyclones and their associated features. Their approach is designed to track features at intervals of 12 hours or less, and introduces the concept of "half-time tracking" in combination with feature vectors. The algorithm works by linking root features at a given time to candidate features at the next time step. The process involves two checks: Check 1 discards certain candidate features based on a "half-time separation". Half-time separation interpolates feature positions between two consecutive time steps, and features that are too far away from the interpolated position are discarded. For the retained candidate features, Check 2 computes a "likelihood score" derived from a 3-D feature vector. This score quantifies the likelihood that candidate features match the root features, thus facilitating accurate feature pairing.

Neu et al. (2013) conduct a comprehensive study evaluating various cyclone detection and tracking algorithms. They find that most detection methods are based on mean sea level pressure minima. An intercomparison experiment with 15 commonly used detection and tracking algorithms highlights cyclone characteristics that remain robust between different schemes, as well as those that show marked discrepancies. Different tracking approaches, including nearest-neighbour and extrapolation-based methods, are investigated. They conclude that different methods of cyclone detection and tracking are most consistent for the intense phase of a cyclone's life. However, the methods diverge for the development and dissipation phases of a cyclone.

Ullrich and Zarzycki (2017) introduce an automated pointwise feature tracking technique that contributes to the objective identification and tracking of meteorological features, including extratropical and tropical cyclones and tropical easterly waves. The pointwise feature tracking is designed to work on both structured and unstructured grids. The proposed framework includes generalised kernels inspired by existing trackers. Ullrich et al. (2021) extend this framework to include areal feature tracking, increasing its versatility.

Shields et al. (2018) conduct a comprehensive review and comparison of atmospheric river identification and tracking methods. Their main focus is to examine the implications of different methods for climate studies, particularly in terms of the climatological, hydrological, and extreme impacts attributed to atmospheric rivers. This effort aims to elucidate the uncertainties associated with different identification methods, thus contributing to a more comprehensive understanding of atmospheric river behaviour.

Nellikkattil et al. (2024) develop a tracking approach similar to Hewson and Titley's (2010) method, but without "half-time tracking". Using positional information for the object centre, maximum and minimum values, objects are tracked

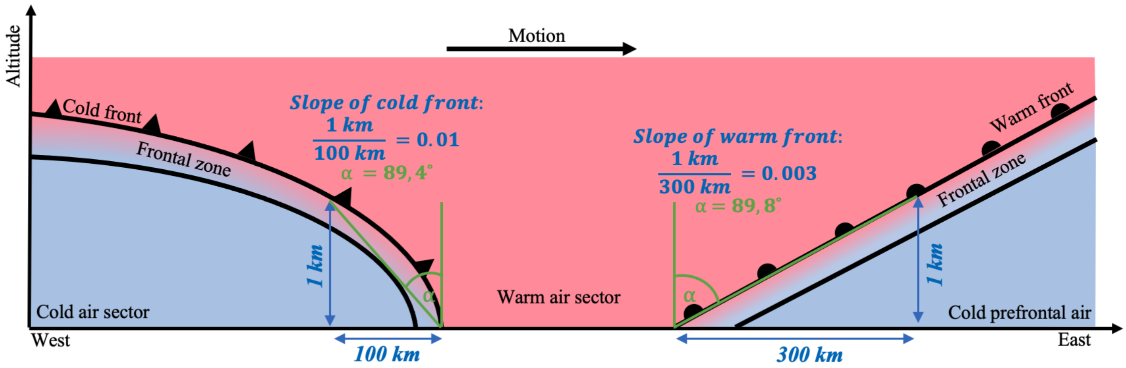


Figure 3.6: Illustration of a vertical cross-section from west to east through a frontal system to illustrate the computation of the frontal slope. Sketch not to scale.

over time. The algorithm clusters objects within a pre-defined (based on physical reasoning) radius and identifies them as the same moving object in subsequent time steps.

The frontal slope is often cited as an important parameter in frontal analysis (Hewson, 1998; Ward, 2020). The analysis of the frontal slope is usually carried out visually on plots of 2-D fronts at different heights (cf. Hewson (1998), superposition of frontal lines on a 2-D plot) or on vertical cross-sections of a thermal variable.

The frontal slope is quantitatively defined as the ratio of the vertical extent (lower edge, and upper edge of the front at a given location) to the horizontal component of the distance between the upper and lower edges of the front (Figure 3.6). The frontal slope is on average 0.01 for warm fronts and 0.03 for cold fronts, but can vary widely (Stull, 2017). The slope is calculated relative to the position of the frontal zone. The frontal zone is the area adjacent to the front where the temperature changes abruptly. The slope is positive when the front is tilted towards the frontal zone and negative when the front is tilted away from the frontal zone. However, this definition has a decisive disadvantage: the frontal inclination becomes infinitely large if the front is exactly perpendicular to the ground. Under this definition, calculating the average slope of a 3-D front surface is meaningless because there may be locations where the surface is exactly perpendicular to the ground, resulting in an infinitely large slope.

Therefore, we calculate the average of the slope angles to estimate the average inclination in Chapter 5. The slope of a front perpendicular to the ground is defined as 0°, for a front parallel to the ground and tilted towards the frontal zone as 90°, and for a front parallel to the ground and tilted away from the frontal zone as -90°.

4

DATA BASIS OF CASE STUDIES

Contents

4.1 Vladiana 48
4.2 Friederike 49
4.3 Otto 51

This chapter introduces the three storm cases used for the application, illustration, analysis, and evaluation of the proposed methods for front detection, front tracking, and feature-based ensemble analysis. The purpose of this chapter is to provide the reader with the essential context of these storm cases and a concise overview of their impact. Where available, references to other case studies related to these storms are included.

In addition to introducing the storm cases, this chapter also provides a detailed description of the NWP datasets used throughout this thesis. The following three severe storm cases are analysed throughout this thesis:

Vladiana: Cyclone Vladiana, which occurred over the Atlantic in September 2016 and is being studied extensively studied because of its connection to the North Atlantic Waveguide and Downstream Impact Experiment (NAWDEX) (Schäfler et al., 2018).

Friederike: Cyclone Friederike occurred in January 2018, had a significant impact on Europe and Germany, and stands out as one of the strongest winter storms in recent years.

Otto: Cyclone Otto is the most recent winter storm analysed in this thesis and occurred over Western Europe in February 2023.

In the following sections, each of these cyclones is discussed in detail, providing relevant background information and details of the data sources used in this thesis.

4.1 VLADIANA

Datasets

For the analysis of the Vladiana storm, ECMWF Atmospheric Model high resolution 10-day forecast (HRES) analysis data with parameterised convection are used. The analysis data cover the period from 23 September 2016, starting at 00:00 UTC and ending at 18:00 UTC on the same day. The output is available at 6-hourly intervals with a horizontal grid spacing of 0.15°. The study area is focused on the North Atlantic region (see Figure 4.1), with 137 vertical levels defined on hybrid sigma-pressure coordinates. To match the rotated grid of the following COSMO simulations, the data grid is rotated to centre the North Pole at 51°N and 160°E.

In addition, convection-permitting simulations of the COSMO NWP model are utilised. The data used in this thesis span from 23 September 2016, starting at 00:00 UTC to 18:00 UTC. The output is available at hourly intervals with a horizontal grid spacing of 0.02°. The study area remains consistent, covering the North Atlantic region. Vertical levels are defined on Gal-Chen coordinates (Gal-Chen and Somerville, 1975) with a vertical grid spacing of 60 levels. The COSMO model output is computed on a rotated grid with the North Pole centred at 51°N and 160°E. The initial and lateral boundary conditions are taken from the ECMWF HRES analysis (see Oertel et al., 2019, 2020, for a detailed description of the simulation setup). The COSMO simulation includes online trajectories (see Miltenberger et al., 2013), which were used to select strongly ascending trajectories with ascent rates of at least 600 hPa in 48 h, here referred to as WCB trajectories (Oertel et al., 2019, 2020). For the evaluation of the conceptual model

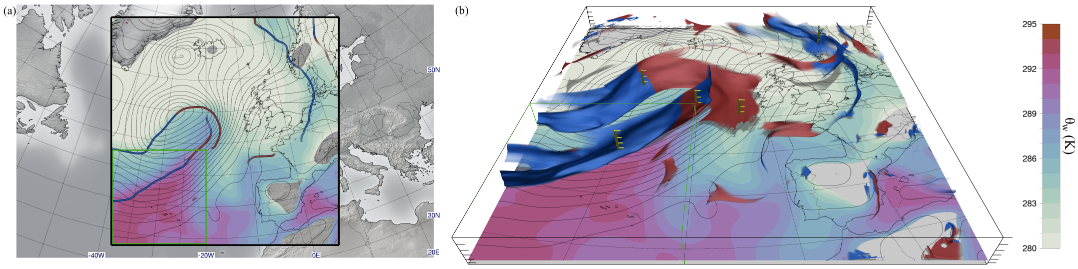


Figure 4.1: Cyclone Vladiana on 23 September 2016 at 06:00 UTC. **(a)** Detected 2-D warm (red line) and cold (blue line) fronts at 850 hPa, θ_w at 950 hPa (colours, in K), and mean sea level pressure (black contour lines, every 2 hPa) from a Consortium for Small-scale Modeling (COSMO) simulation (black frame shows domain boundaries; green frame shows the selected sub-region for studying convection in the vicinity of the cold front; see Section 5.3.3). **(b)** Detected 3-D warm (red) and cold (blue) fronts between 950 hPa and 500 hPa, on top of a horizontal map showing θ_w at 950 hPa and mean sea level pressure (black contour lines, every 2 hPa). Warm- and cold-front classification is computed according to warm- and cold-air advection at the front (following Hewson, 1998).

of 3-D fronts and WCBs, WCB trajectories ascending at least 25 hPa in 2 h at 06:00 UTC on 23 September 2016 were selected.

Atmospheric conditions

The extra-tropical cyclone Vladiana occurred during the NAWDEX (Schäfler et al., 2018). Vladiana formed near Newfoundland on 22 September 2016, and the frontal wave intensified as it moved eastward across the North Atlantic. As the cyclone continued to move north-eastward, it intensified until it reached its pressure minimum of 975 hPa at 18:00 UTC on 23 September. On 24 September, the cyclone reached Iceland and became stationary.

Figure 4.1 shows a horizontal section of θ_w with detected 2-D fronts at 850 hPa, as well as 3-D fronts on 23 September 2016 at 06:00 UTC. The frontal analysis of this case study builds on previous studies of Vladiana and its associated WCB ascent (Choudhary and Voigt, 2022; Kern et al., 2019; Oertel et al., 2019, 2020).

Based on the results of Oertel et al. (2019), I examine the conceptual model of WCB ascent in the vicinity of fronts (Section 5.3.2) and show how frontal surfaces from convection-permitting NWP simulations compare to those found in simulations in which convection is parametrised (Section 5.3.3).

4.2 FRIEDERIKE

Datasets

For the analysis of this case I use ERA5 reanalysis data from 16 January 2018, 12:00 UTC to 19 January 2018, 00:00 UTC. ERA5 data are available at hourly intervals with a horizontal grid spacing of 0.25° and 137 vertical levels defined on hybrid sigma-pressure coordinates. This analysis focuses on a region spanning from eastern North America across the North Atlantic towards Europe (see Figure 4.2).

In addition to the ERA5 data, ECMWF HRES forecast data initialized on 18 January 2018, 00:00 UTC and run for 36 hours are utilised. The forecast data are provided at hourly intervals with a horizontal grid spacing of 0.15° and 137 vertical levels defined on hybrid sigma-pressure coordinates. The forecast area covers eastern North America, the North Atlantic, and Europe. For the analysis of the frontal attributes within ensemble simulations, the ECMWF enseble (ENS) forecast initialised on 16 January 2018, 12:00 UTC, with 51 ensemble members, is used. All members have a horizontal grid spacing of 0.25°, include 12 vertical pressure levels, and are available in hourly resolution for the first 90 hours. In addition, the control run (ensemble member 0) is provided with a higher vertical grid spacing, consisting of 91 levels defined on hybrid sigma-pressure coordinates. The ECMWF ENS forecast covers an area that includes eastern North America, the North Atlantic, and Europe.

For the analysis of secondary frontal structures, fronts extracted from the UK Met Office surface analysis charts supplement the ECMWF HRES forecast.

Atmospheric conditions

The extra-tropical cyclone Friederike (named David in the United Kingdom (UK)) crossed Western Europe on 17 and 18 January 2018. The cyclone had formed east of Florida on 15 January 2018 and then moved north along the coast of Newfoundland before crossing the North Atlantic and making its first landfall in Europe on the west coast of Ireland on 17 January 2018. During its passage across the North Atlantic, the cyclone intensified, and its core pressure decreased to 985 hPa. The cyclone moved from Ireland across northern England and the North Sea, reaching the northern Netherlands on 18 January 2018 at 09:00 UTC with a core pressure of 976 hPa. From there, the cyclone moved east across northern Germany until it reached the Polish border at 18:00 UTC and dissipated in the following days. The cyclone caused high wind speeds with gusts of up to 203 km h^{-1} in the Harz Mountains, 144 km h^{-1} on the North Sea coast of the Netherlands, and 138 km h^{-1} in the lowlands of the Netherlands and central Germany (Wandel et al., 2018). Surface analysis charts from the UK Met Office indicate that this was a Shapiro–Keyser cyclone (Shapiro and Keyser, 1990). The 2-D front algorithm detects some of the characteristic frontal features of the Shapiro–Keyser cyclone, including the frontal wave stage, frontal fracture, and T-bone structure (Figure 4.2). This case allows for the first time (to my knowledge) the extraction and visualisation of 3-D frontal structure of a Shapiro–Keyser cyclone directly from NWP data, and enables the evaluation of time evolution in comparison with the conceptual model (Section 5.3.4). In Section 5.3.5, I analyse the occurrence of secondary warm frontal structures, as often seen in surface analysis charts from the UK Met Office. I also use this case to analyse the distribution and temporal evolution of derived frontal attributes in Section 6.2. This analysis considers different temporal and vertical grid spacings and how the front tracking responds to them.

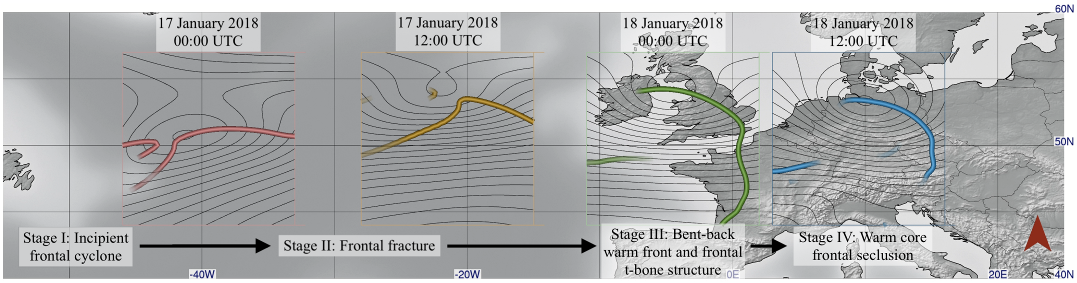


Figure 4.2: Successive time steps of objective 2-D frontal structures show the temporal development of Friederike (17 and 18 January 2018), as detected in ERA5 reanalysis data at 750 hPa and surface pressure (black lines). The displayed time steps are approximately assigned to the four ideal development stages of the Shapiro–Keyser cyclone model (Shapiro and Keyser, 1990). We find that not all characteristics of the individual stages can be observed in 2-D. As shown in Figure 5.13, 3-D front detection is required to observe all characteristics.

4.3 OTTO

Datasets

For the analysis of this case I use ICON-EU-Ensemble Prediction System (EPS) and ICON-EU NWP data provided by the DWD. The ICON-EU-EPS is the ensemble forecast with 40 members, providing probabilistic information for the European region. The 40 ensemble members are also referred to as ensemble members 1 to 40. In addition, the ICON-EU represents the deterministic forecast for the same region and is referred to as ensemble member 0. This forecast was initialised on 16 February 2023, 00:00 UTC, and runs for a period of 5 days. For the first 52 hours, all variables are available in hourly time steps. After that, only a 3-hourly output with reduced variables is provided. This forecast covers an area between 23.5°W to 23.5°E and between 29.5°N to 70.5°N. For my analysis, the forecast data was regridded from its original unstructured triangular grid to a regular grid with a grid spacing of 0.1°in both longitude and latitude.

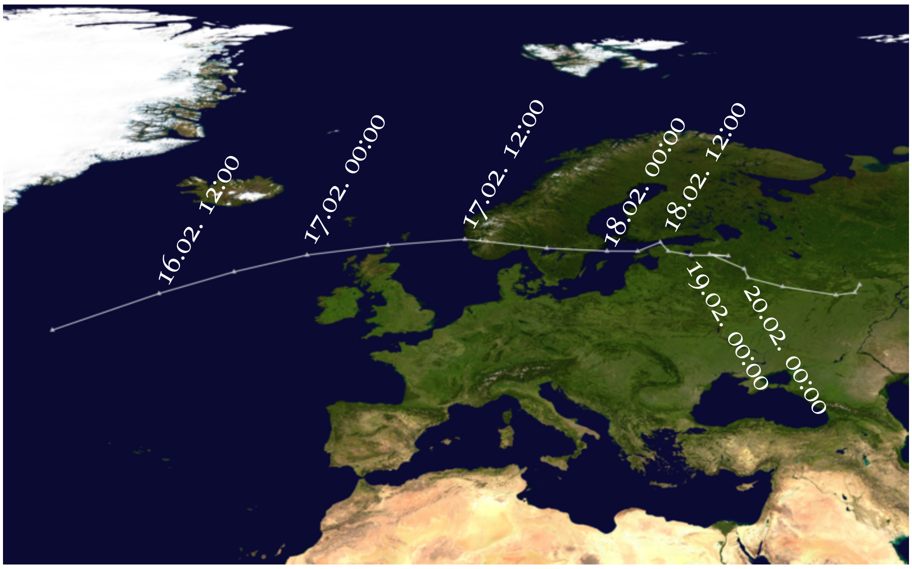


Figure 4.3: Track of the cyclone centre of winter storm *Otto* according to the University of Berlin. Figure from RandomInfinity17 (2023), used under CC BY-SA 4.0. The original image has been edited by adding date and timestamps to the storm track.

Atmospheric conditions

The low-pressure system *Otto* (*UK Met Office* 2023, named *Ulf* in Germany *Wetterpate* 2023) affected Europe between 16 February and 21 February 2023. Figure 4.3 shows the track of the cyclone centre. The storm originated in the North Atlantic and moved towards Europe, crossing Scotland in the northern region. It then moved east across the North Sea before making landfall on the Norwegian coast near the city of Bergen. Several weather warnings were issued by national weather services for Scotland and northern England (*UK Met Office* 2023), Denmark (Wenande, 2023), and Germany (*RBB* 2023). With a maximum wind speed of 193 km/h measured on 17 February 2023 in Cairngorm, Scotland (*Kachelmannwetter* 2023), it was one of the strongest winter storms of 2023. The winter storm *Otto* caused significant damage and disruption in Scotland and northern England. Around 30,000 homes lost power and many schools were forced to close. High winds brought down trees and damaged vehicles and buildings (*BBC News* 2023).

Chapter 6 uses this case to conduct a case study on feature-based ensemble analysis. In the case study, the cold front of Cyclone Otto is tracked and feature attributes are derived. Based on these feature attributes, ensemble similarity analysis in Section 6.3 and ensemble clustering in Section 6.3.2 are performed for a period of one day.

5

THREE-DIMENSIONAL FRONT ANALYSIS OF MID-LATITUDE WEATHER SYSTEMS

Contents

5.1 Method and implementation **54**
  5.1.1 Conceptual and mathematical basis 54
  5.1.2 Filtering 55
  5.1.3 Supported data and methodological details 57
5.2 Thermal quantity, smoothing length scale, and filter parameters **63**
  5.2.1 Choice of thermal quantity 63
  5.2.2 Recommendations for filter thresholds and sensitivity of fronts to different smoothing length scales 64
  5.2.3 Impact on reduced vertical resolution 67
5.3 Case studies **69**
  5.3.1 Meteorological theory 69
  5.3.2 The 3-D examination of conceptual model: fronts and warm conveyor belt 70
  5.3.3 Cold-front structure in the vicinity of convection 72
  5.3.4 The 3-D examination of conceptual model: Shapiro–Keyser cyclone 74
  5.3.5 Secondary fronts 78

This chapter focuses on atmospheric front detection and its 3-D visual analysis in weather forecasting and research. While a detailed description of the objectives and research questions are presented in Section 1.2, a brief overview of the primary objectives follows:

The first objective is to advance the methodology proposed by Kern et al. (2019) for detecting 2-D and 3-D frontal structures independently of the grid point spacing of the NWP data, facilitating cross-comparisons across different model resolutions and scenarios. In addition, the study investigates the impact of smoothing and filtering procedures and investigates the effects of varying parameters on the identification of 3-D frontal structures.

The second objective is to assess the benefits of 3-D IVA of the detected frontal structures for the analysis of mid-latitude cyclones. The main research objectives are to validate existing insights into 3-D dynamical structures, to identify 3-D frontal development stages in a Shapiro-Keyser cyclone, to contrast 3-D structures at different NWP simulation resolutions, and to compare detected 3-D frontal structures with official UK Met Office analyses.

The chapter is structured as follows: Section 5.1 describes the methodology and implementation of the 2-D and 3-D front detection method. Section 5.2 examines the choice of thermal quantity, filter thresholds, and the effects of different vertical resolutions on front detection. Finally, Section 5.3 applies the proposed front detection method in two severe mid-latitude cyclone case studies, focusing on 3-D frontal structures.

5.1 METHOD AND IMPLEMENTATION

Our algorithm follows the 2-D detection algorithm originally introduced by Hewson (1998) and extended to 3-D by Kern et al. (2019). We briefly explain the basics of the algorithm and focus on the parts that have been adapted for this study. For further details we refer to Hewson (1998) and Kern et al. (2019). In the following, we describe and illustrate the conceptual and mathematical basis (Section 5.1.1), the required filtering process for frontal candidates (Section 5.1.2), and implementation details we consider important (Section 5.1.3).

5.1.1 *Conceptual and mathematical basis*

Figure 5.1 illustrates the method. The goal is to detect the horizontal warm-air "boundaries" of frontal zones, i.e. regions with a strong horizontal gradient of a thermal variable τ (black line). In the simplified 1-D example shown in Figure 5.1, the first partial derivative of τ with respect to the spatial dimension $x(\partial\tau/\partial x)$ changes rapidly on both the warm- and cold-air boundaries of the frontal zone, with a maximum in between. Hence, the third derivative $\partial^3\tau/\partial x^3$ can be used to detect the locations of maximum gradient change; the locations where it is zero and the second derivative $\partial^2\tau/\partial x^2$ is negative coincide with the warm-air boundary of the frontal zone (Hewson, 1998). In the general 2-D case, points on a frontal line need to fulfil the "front location equation" (see Hewson, 1998) to account for curved fronts and corresponding along-front thermal gradients:

$$L_\tau \equiv \frac{\partial(|\nabla_\mathrm{h}|\nabla_\mathrm{h}\tau||)_s}{\partial\hat{s}} = 0, \tag{5.1}$$

with

$$\hat{s} = \pm\frac{\nabla_\mathrm{h}\,|\nabla_\mathrm{h}\tau|}{|\nabla_\mathrm{h}\,|\nabla_\mathrm{h}\tau||}.$$

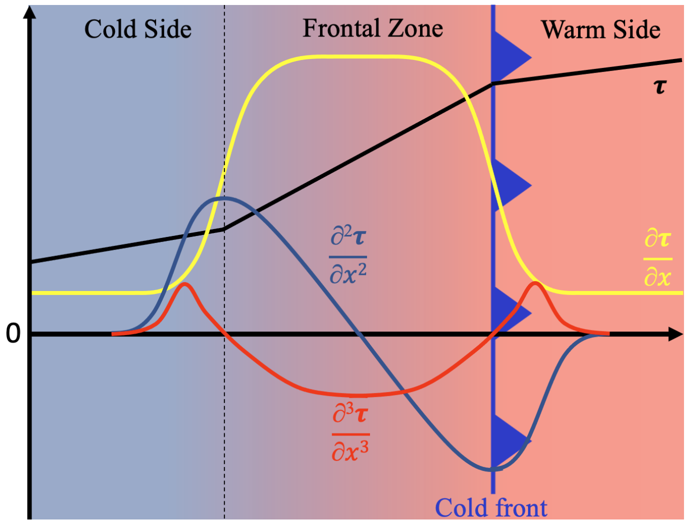


Figure 5.1: Illustration of the thermal-gradient-based detection method, using a simplified straight front and following Hewson (1998) and Kern et al. (2019). The goal is to determine the warm-air boundary of the frontal zone (i.e. the region of increased thermal gradient; see the yellow line). This boundary corresponds to the third derivative (red line) of a thermal variable τ (black line) being zero, under the condition that the second derivative of τ (blue line) is negative. The cold-front typing shown assumes air masses are moving from left to right across the figure.

Here, ∇_h denotes the horizontal derivative, and $\hat{s}$ is a unit axis (which possesses an orientation but no direction) oriented along $\nabla_h |\nabla_h \tau|$. To derive 3-D frontal surfaces the approach is extended to 3-D as proposed by Kern et al. (2019). In short, the front location equation Equation 5.1 is computed at every grid point of the gridded dataset; then “candidates” of frontal features are obtained by computing 3-D isosurfaces of $L_\tau = 0$ using a contouring algorithm such as marching cubes (Lorensen and Cline, 1987). This results in a large number of potential frontal surfaces; to obtain meaningful structures the feature candidates need to be filtered according to additional diagnostics including the strength of the thermal gradient within the frontal zone. For details, we refer the reader to Kern et al. (2019, their Sect. 4) . Note that only the horizontal gradient of the thermal variable is considered in this process; see Kern et al. (2019) for a discussion on the inclusion of vertical contributions.

5.1.2 *Filtering*

To obtain meaningful frontal surfaces (or frontal lines in the 2-D case), the feature candidates need to be filtered. Hewson (1998), following Renard and Clarke (1965), suggested to filter according to the thermal front parameter TFP , as well as to a frontal strength value estimated by the local thermal gradient at the frontal feature. The latter was improved by Kern et al. (2019) to estimate frontal strength by computing an average thermal gradient along “normal curves” traced through the frontal zone (basically streamlines computed on the gradient vector field). Here, we generalise these two filters to more generic types of filter mechanisms that can be interactively modified and combined during the analysis to investigate different aspects of the data.

a. *Masking*. The feature candidates are filtered according to an arbitrary 3-D scalar field that is sampled (i.e. interpolated) at all feature locations (e.g. if isosurfaces are extracted using marching cubes, at all vertices of the isosurface). User-defined thresholds of the scalar field are used to keep or discard features.

b. *Frontal zone traversal*. The frontal zone is traversed along "normal curves" started at feature candidate vertices and computed on the thermal gradient field (Kern et al., 2019); an arbitrary 3-D scalar field is sampled along the normal curves, and filtering thresholds are based on the obtained samples.

The generalization allows us, in addition to filtering with respect to TFP and frontal strength, to add filters that facilitate focus on the contribution of further quantities, including, for example, humidity and elevation. In this way, we can eliminate, for example, pure "humidity fronts" by tracing the changes in (dry) potential temperature (θ) along the normal curves. TFP and frontal strength, however, remain as the core filters.

TFP masking

TFP is a masking filter. Note that computing isosurfaces of $L_\tau = 0$ results in front feature candidates at both the cold and the warm sides of the frontal zone. Since we are interested in the warm side only (see Renard and Clarke, 1965), cold side feature candidates need to be discarded. We follow the approach of Hewson (1998) and use the TFP filter, first introduced by Renard and Clarke (1965). The TFP filter is defined as follows:

$$\mathrm{TFP}_\tau \equiv -\nabla_\mathrm{h} |\nabla_\mathrm{h}\tau| \cdot \frac{\nabla_\mathrm{h}\tau}{|\nabla_\mathrm{h}\tau|} > K_1, \tag{5.2}$$

where K_1 is a used-defined threshold. This equation can also be interpreted as the "negative curvature" of the thermal front parameter field (Kern et al., 2019), being positive at the warm side of the frontal zone and negative at the cold side. To obtain only frontal feature candidates at the warm side of the frontal zone, K_1 must be at least zero. Hewson (1998) suggested a slightly positive value for K_1 to eliminate spurious frontal pieces.

Frontal strength

Filters based on normal curves are evaluated for the remaining warm-air-side frontal candidates. We follow Kern et al. (2019) and estimate the frontal strength of the filter variable as "the average thermal gradient along a curved path through the frontal zone from the warm to the cold-air side". The frontal strength filter S_τ is defined as follows:

$$S_{\tau|\text{frontal zone}} \equiv \int_{\text{NC}} |\nabla_{\text{h}}\tau| \, \mathrm{d}s > K_2. \tag{5.3}$$

The integration through the frontal zone starts at the warm side of the frontal zone and stops once a "normal curve" reaches the cold side of the frontal zone (where L_τ again is zero). The threshold K_2 is used to eliminate weak fronts below a user-defined frontal strength.

Fuzzy filtering

The usage of distinct threshold values for K_1 and K_2 results in "hard" boundaries of the generated features. Such visualization can be misleading since a viewer can interpret distinct feature boundaries into the depiction (including, for example, "holes" in the front surfaces where, for example, frontal strength is just below the chosen threshold). For fronts, however, this is not the case, as thermal gradients are gradually decreasing in space. Kern et al. (2019) suggested a "soft" (or "fuzzy") filtering by providing two thresholds for each filter, between which opacity is faded from zero (completely transparent) to one (completely opaque). The feature candidates are subsequently rendered using the obtained opacity, resulting in "fuzzy" edges that visually indicate, , for example, a decreasing thermal gradient. The approach can also facilitate a visual distinction between weak fronts and strong fronts. When multiple filters are used in our implementation, every filter has individual threshold interval settings, and opacity information is accumulated accordingly.

5.1.3 *Supported data and methodological details*

The presented algorithm supports gridded data on horizontally regular and rotated latitude–longitude grids. In the vertical, the implementation can handle both pressure levels and model levels. For this study, we use data from the operational ECMWF high-resolution (HRES) forecast with 137 vertical model levels, horizontally interpolated to a regular grid with a grid-point spacing of 0.15° in both latitude and longitude; data from the global reanalysis ERA5 (Hersbach et al., 2020) (also 137 vertical model levels, interpolated to a horizontal grid spacing of 0.25°); and data from the COSMO model (Baldauf et al., 2011; Doms and Baldauf, 2018), available on a rotated latitude–longitude grid with 60 vertical model levels and a horizontal grid-point spacing of 0.02°, in both dimensions. The algorithm has been integrated into the interactive visualization framework Met.3D (Rautenhaus et al., 2015b) and is being made available as open-source.

In the following, we describe methodological details we deem important for understanding our approach. Figure 5.2 illustrates the main steps of the front detection process. For simplicity, the process is described for 2-D frontal lines (letters correspond to panels in Figure 5.2):

a. choice of a thermal input field τ (e.g. wet-bulb potential temperature; Figure 5.2a)

b. smoothing of τ (and further input fields used for filtering) to a user-defined length scale (Figure 5.2b)

c. computation of the magnitude of horizontal gradients $|\nabla_h \tau|$ (Figure 5.2c)

d. computation of the horizontal gradient of the magnitude of horizontal gradients $\nabla_h |\nabla_h \tau|$ (Figure 5.2d)

e. evaluation of the front location equation Equation 5.1 and computation of the zero isolines to obtain feature candidates (Figure 5.2e)

f. computation and application of the TFP masking filter (Figure 5.2f)

g. application of frontal strength and further "normal curve" filters (Figure 5.2g)

h. obtain final frontal structures (Figure 5.2h).

In the 2-D example in Figure 5.2, the 850 hPa pressure level is used. One important design decision for the 3-D variant of the algorithm is the choice of the vertical coordinate, as the numerical computations need to be implemented accordingly. For this study, we consistently use pressure as the vertical coordinate, i.e. all horizontal computations are evaluated on levels of constant pressure. This is also consistent with Met.3D's use of pressure as the vertical coordinate.

Smoothing

NWP data, especially at kilometre-scale resolution, include convective and thermal processes that are much smaller in scale than atmospheric fronts (Keyser and Shapiro, 1986). To obtain frontal features that meaningfully represent a scale of interest (e.g. synoptic-scale fronts), it is advisable to smooth small-scale thermal fluctuations in the thermal input field. Previous studies have used simple smoothing filters like a weighted moving average of neighbouring grid points (eg. Jenkner et al., 2009), well-known from image processing (Davies, 2017). Kern et al. (2019) point out that for data on a regular latitude–longitude grid, however, geometric distance between grid points varies with latitude, requiring the usage of a smoothing filter that considers all grid points based on a specified geometric smoothing distance. They propose the usage of a 2-D Gaussian smoothing kernel.

In our implementation, the smoothing distance is a user-defined method parameter that can be interactively changed in the analysis process. A disadvantage of a Gaussian smoothing filter, however, is its computational complexity that

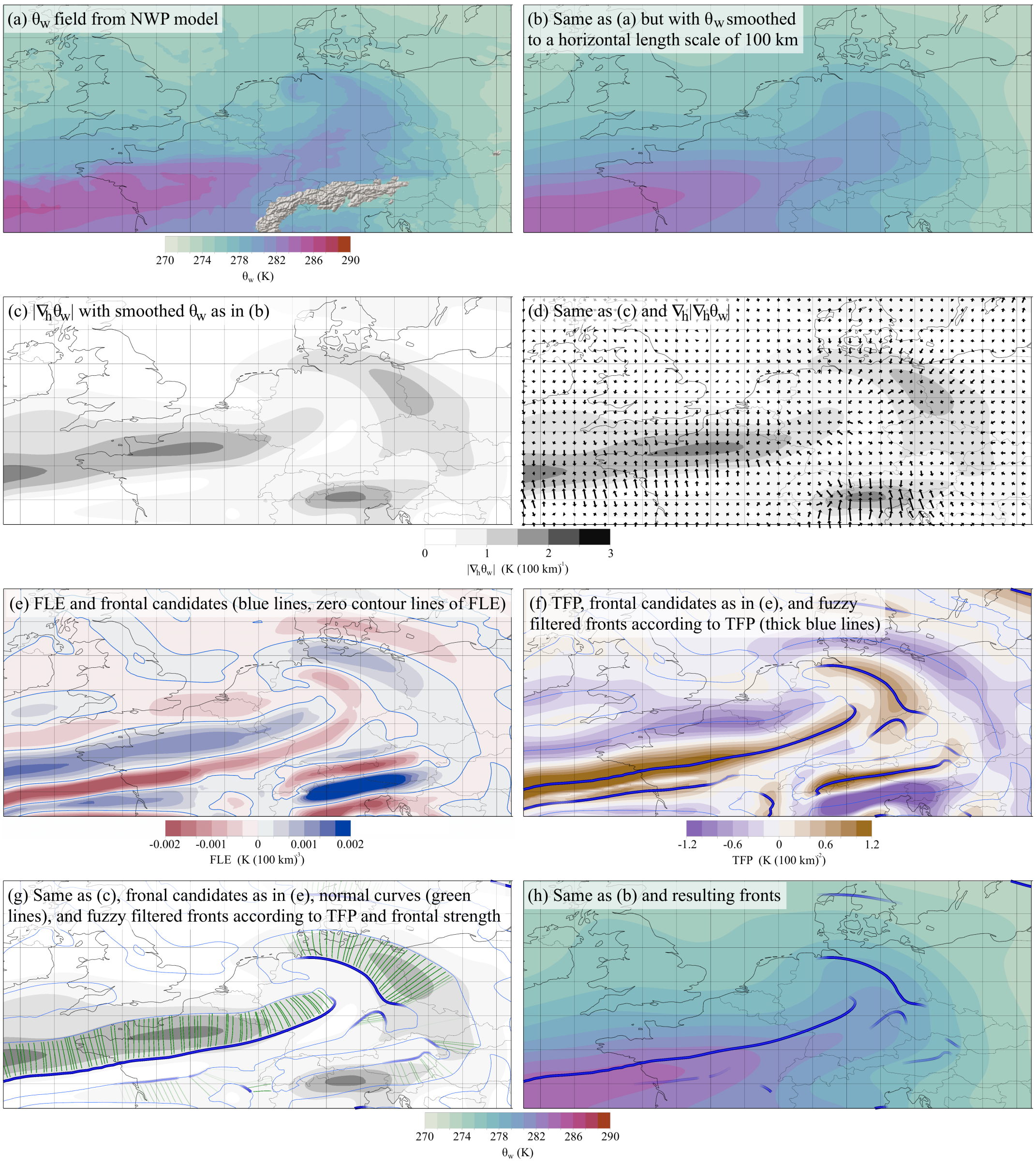


Figure 5.2: Step-by-step illustration of the 2-D front detection method. In the example, objective fronts are based on the 850 hPa wet-bulb potential temperature field (θ_w) from the ECMWF HRES forecast (horizontally regular grid-point spacing of 0.15° in both longitude and latitudes) initialised on 18 January 2018 at 00:00 UTC and valid on 18 January 2018 at 12:00 UTC. Fronts are "fuzzy filtered" using a fade-out range for TFP of 0.2–0.4 K $(100\,\mathrm{km})^{-2}$ and for frontal strength of 0.6–1 K $(100\,\mathrm{km})^{-1}$. See Sect. 2.3 for a description of panels **(a)–(h)**.

increases quadratically with smoothing distance – an important aspect for interactive use. We hence also provide an approximative smoothing method, the "fast almost-Gaussian filtering" presented by Kovesi (2010). The method uses a specified number of averaging passes. More averaging passes increase the accuracy of the approximative algorithm compared to Gaussian smoothing but at the cost of increasing computation time. Another important aspect to consider is that with an increased number of averaging passes the effect of "smoothing over the data field edges" propagates further into the data field centre (Kovesi, 2010). In our implementation, the smoothing computation complexity depends linearly on the averaging passes and the smoothing distance. We find that three averaging passes are a reasonable trade off between accuracy, computation time, and keeping the edged effect small. For illustration, we measured the performance of both smoothing algorithms on six cores of an AMD EPYC 7542 32-core processor at 2.9 GHz. In this set-up, it takes about 29.5 s to apply a horizontal Gaussian smoothing with a smoothing distance of 100 km to a 3-D data field of 1800×1800 horizontal grid points with a horizontal grid spacing of 0.02°, and 31 vertical level. For the same data field, the approximative algorithm requires 3.9 s. Both algorithms are optimised for OpenMP (OpenMP Architecture Review Board, 2015) and run in parallel.

Numerical implementation

For the computation of horizontal gradients, we use first-order finite central differences and at boundaries first-order finite right and left differences. As described above, we use pressure as the vertical coordinate and hence need to adapt the computations for data available on hybrid sigma pressure model levels or geometric altitude model levels. This leads to an additional coordinate transformation term (see Etling, 2008, pp. 129–131) in the derivatives. The horizontal gradient in pressure coordinates $|_p$ of the thermal variable τ is obtained from the partial derivative in the longitudinal direction on the original coordinate system $|_\sigma$ and an additional transformation term. The gradient component in the longitudinal direction hence becomes

$$\left.\frac{\partial\tau}{\partial\mathrm{long}}\right|_p = \left.\frac{\partial\tau}{\partial\mathrm{long}}\right|_\sigma + \left.\frac{\partial\tau}{\partial p}\right| \cdot \left.\frac{\partial \mathrm{p}}{\partial\mathrm{long}}\right|_\sigma \tag{5.4}$$

and the latitudinal component

$$\left.\frac{\partial\tau}{\partial\mathrm{lat}}\right|_p = \left.\frac{\partial\tau}{\partial\mathrm{lat}}\right|_\sigma + \left.\frac{\partial\tau}{\partial p}\right| \cdot \left.\frac{\partial p}{\partial\mathrm{lat}}\right|_\sigma . \tag{5.5}$$

Care needs to be taken for the numerical implementation of Equations 5.1–5.5. For numerical stability reasons, Hewson (1998) computed $\hat{s}$ as a "five-point-mean axis" – an average orientation axis derived from the gradient at the corresponding grid point and at the four surrounding grid points (for details see Hewson, 1998). We encountered challenges with this approach:

a. The studies by Hewson (1998) and Kern et al. (2019) used gridded data with a regular horizontal grid-point spacing on the order of 50 km (0.5°) to 100 km (1°). At the time of writing, current (e.g. limited-area) NWP models use finer grid spacings; e.g. the regional forecast model of the German Weather Service (DWD) runs with a horizontal grid spacing of 0.02°. At such resolutions and depending on the smoothing distance of previously applied smoothing, the differences between data values at neighbouring grid cells tend to be very small – in such cases, no numerically stable orientation of the five-point-mean axis can be obtained.

b. Analogous to the above reasons for the use of a distance-based Gaussian smoothing filter, the dependence of geometric distance between neighbouring grid points on latitude leads to inconsistent calculations of the five-point-mean axis.

c. The distance between neighbouring grid cells depends on the grid-point spacing of the specific dataset used. To compare fronts in different model simulations with a different grid-point spacing it is inconvenient to use a grid-point-based approach because the distance of the neighbouring grid cell changes with changing model resolutions.

Instead of taking the neighbouring grid points to calculate the five-point-mean axis, we propose using interpolated values at a specified distance to the considered central grid point. This improves numerical stability, makes the computation independent of geographic location, and facilitates objective comparison of frontal features obtained from NWP datasets with different grid-point spacings. From our experiments, we find that using a distance for the five-point-mean axis computation of half of the smoothing distance works well.

Front detection pipeline in Met.3D

The front detection algorithm is implemented in the data processing pipeline of Met.3D described in section 2.2.1. The extension of the data processing pipeline is shown in Figure 5.3, which illustrates the integration of additional pipeline modules and data structures to detect objective fronts in a thermal input field. The data pipeline modules are connected sequentially, with each module designed to take at least one input data representation and produce at least one output data representation that flows seamlessly into the next pipeline module. The final representation is then passed to the renderer, which further refines the 3-D geometries through additional filtering, e.g. fuzzy filtering, resulting in the visualisation of objective fronts. The front detection pipeline implementation in Met.3D includes the eight pipeline modules listed below, and the data flow between these pipeline modules is illustrated in 5.3. Filter pipeline modules are pipeline modules that modify the input field without changing the data structure. Source pipeline modules return data in a different structure to their input field.

1. **Smooth filter:** A horizontal low-pass filter that refines the input scalar field Ψ by applying a specified length scale for smoothing and returns the smoothed scalar field $\overline{\Psi}$.

2. **Partial derivative filter:** Computes a horizontal partial derivatives of the input field $\overline{\Psi}$ and returns its partial derivative e.g. $\partial\overline{\Psi}/\partial x$.

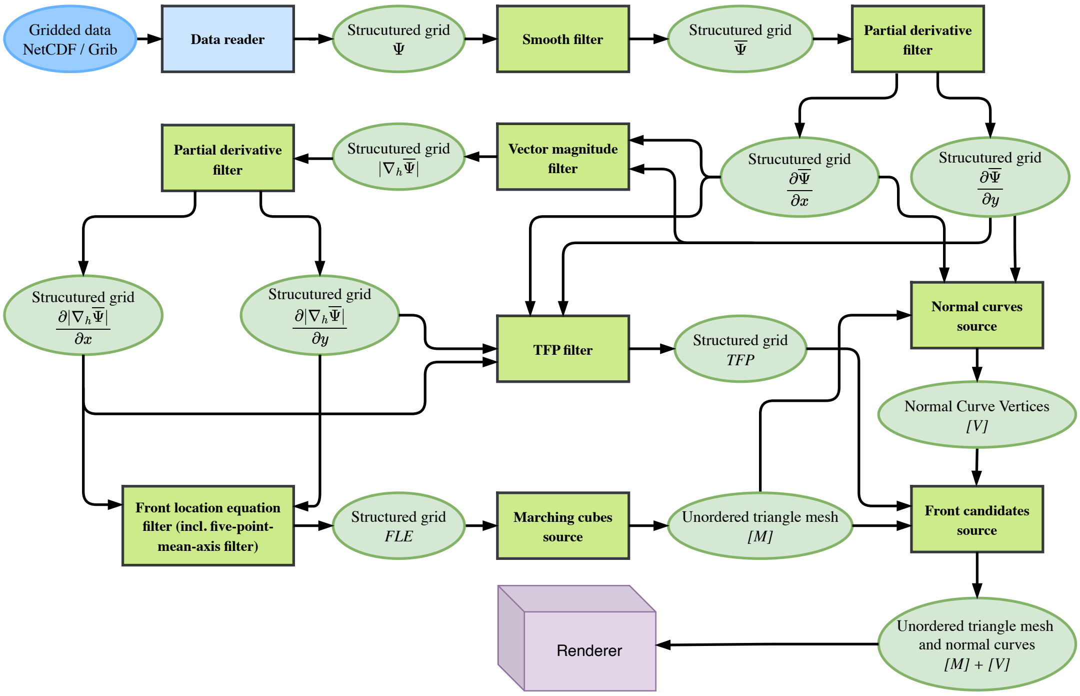


Figure 5.3: Illustration of the extended data pipeline in Met.3D for objective front detection. The existing pipeline module is shown in a blue box and new data pipeline modules are shown in green boxes. Data structures are represented by oval circles. Ψ represents the scalar input data field.

3. **Vector magnitude filter:** Computes the vector magnitude from two scalar fields e.g. from $\partial\overline{\Psi}/\partial x$ and $\partial\overline{\Psi}/\partial x$ and returns its magnitude $|\nabla_h\overline{\Psi}|$.
4. **TFP filter:** Computes the TFP from the following four specified scalar input fields: $\partial\overline{\Psi}/\partial x$, $\partial\overline{\Psi}/\partial y$, $|\nabla_h\overline{\Psi}|/\partial x$, and $|\nabla_h\overline{\Psi}|/\partial y$.
5. **Front location equation filter:** Derives and returns the front location equation scalar field FLE from the following two specified scalar input fields: $|\nabla_h\overline{\Psi}|/\partial x$ and $|\nabla_h\overline{\Psi}|/\partial y$.
6. **Marching cubes source:** Computes a triangular mesh $[M]$ representation of frontal surface candidates using FLE as data source.
7. **Normal curves source:** Computes normal curves $[V]$ from horizontal partial derivatives $\partial\overline{\Psi}/\partial x$ and $\partial\overline{\Psi}/\partial y$.
8. **Front candidates source:** Aggregates the triangular mesh $[M]$ representation of the frontal surface candidates, along with the normal curves $[V]$, before passing this combined data to the renderer for visualisation.

5.2 THERMAL QUANTITY, SMOOTHING LENGTH SCALE, AND FILTER PARAMETERS

To successfully apply front detection for case studies, three important aspects need to be considered: which thermal quantity should be used for detection, which smoothing distance should be applied to the data, and how do filter thresholds need to be adjusted (also with respect to the smoothing distance)?

5.2.1 *Choice of thermal quantity*

We first discuss the role of the chosen thermal quantity. Three candidates have frequently been used in the literature: (dry) potential temperature (θ), wet-bulb potential temperature (θ_w), and equivalent potential temperature (θ_e). There is an ongoing discussion in the scientific community regarding which thermal quantity is best suited to detect fronts (Berry et al., 2011; Hewson, 1998; Sanders and Doswell, 1995; Schemm et al., 2018; Thomas and Schultz, 2019a,b, e.g.). The following provides a brief overview of the potential thermal quantities and their advantages and disadvantages.

The dry potential temperature θ reflects the original, purely temperature-dominated definition of fronts and is most convenient from a rigorous dynamical point of view (Hewson, 1998). However, it is not conserved in moist processes, which often occur along fronts (Browning and Roberts, 1996). Alternative thermal quantities are θ_w or θ_e,which are both conserved in the reversible diabatic processes of evaporation and condensation (Thomas and Schultz, 2019b). Since both quantities have a one-to-one relationship (each θ_w value matches a unique θ_e value and vice versa; Bindon, 1940), they share the same advantages and disadvantages for front detection (Thomas and Schultz, 2019b). In the following, we consider only θ_w; the arguments are similar for θ_e (to detect similar structures, however, the filter thresholds need to be adjusted due to the non-linear relationship between θ_w and θ_e). The inclusion of humidity can help to better diagnose weak temperature gradients because humidity and temperature gradients are usually correlated, resulting in stronger θ_w gradients compared to θ gradients (Jenkner et al., 2009). However, if humidity and temperature are not correlated, gradients of θ_w could be weaker than gradients of θ. This may result in θ_w fronts being weaker than θ fronts, up to not being detected at all. Furthermore, in regions with humidity gradients but without temperature gradients, purely humidity-dominated fronts can be detected. Therefore, Thomas and Schultz (2019b) recommended examining the temperature and moisture fields separately when analysing frontal structures. On the other hand, Berry et al. (2011) found that in their study θ_w provided the closest match to manually prepared front analysis. In our experience, θ_w is best suited to detect continuous fronts and closely matches the frontal analysis provided by the UK Met Office (Figure 5.15). Note that some of the previously mentioned disadvantages of θ_w can be eliminated in our front algorithm. To facilitate the distinction between humidity- and temperature-dominated fronts, the implementation allows the mapping of different quantities on frontal surfaces, as well as the filtering of fronts according to multiple variables. Mapping the total change in θ or specific humidity within the frontal zone could help to distinguish between humidity- and temperature-dominated fronts. If desired, fronts can be filtered according to θ or humidity gradients within the frontal zone, which can help to eliminate purely temperature- or humidity-dominated fronts (Hewson and Titley, 2010).

5.2.2 *Recommendations for filter thresholds and sensitivity of fronts to different smoothing length scales*

The number of detected frontal features depends on filter thresholds and the smoothing length scale applied to the input fields. Depending on the scale of interest for the analysis, the horizontal smoothing length scale is chosen. The question arises of which filter thresholds for TFP and frontal strength filters should be recommended and how these values depend on the smoothing length scale. In this section, we explore these method parameters and provide recommendations. We first investigate how smoothing length scale affects the magnitude and distribution of TFP values, and then we consider the magnitude and distribution of frontal strength $|\nabla_h \theta_w|$. We present distributions of TFP and frontal strength obtained from 24 consecutive time steps of hourly ECMWF HRES forecast data on 18 January 2018 (initialised at 00:00 UTC) in a geographic region encompassing 30° N–70° N in latitude and 60° W–30° E in longitude (slightly larger than the region shown in Figure 5.2). The presented distributions provide guidance on the choice of suitable values for different smoothing length scales.

Dependence of filter thresholds K_1 and K_2 on smoothing length scale

Figure 5.4 shows the relative frequency of TFP values in the analysed area and for three different horizontal smoothing length scales of 100, 50, and 30 km. Large

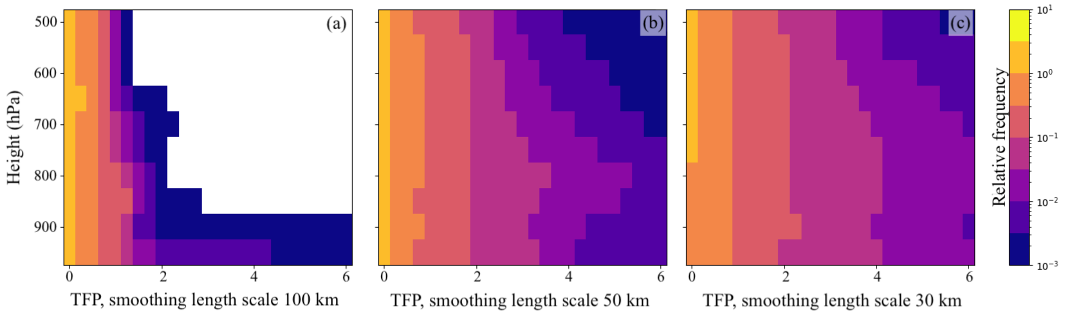


Figure 5.4: Distribution (relative frequencies) of thermal front parameter (TFP) values computed from hourly ECMWF HRES forecast data (horizontal grid-point spacing of 0.15°) from 18 January 2018, in the region 30° N–70° N, 60° W–30° E and between 950–500 hPa for different smoothing length scales: **(a)** 100 km, **(b)** 50 km, and **(c)** 30 km.

horizontal smoothing length scales result, in general, in lower TFP values and vice versa. With large smoothing applied, strong horizontal gradients are weakened, resulting in smaller horizontal gradients. The magnitude of the horizontal gradients is inversely proportional to the length scale of the horizontal smoothing, and the filter thresholds need to be adjusted accordingly. Table 1 provides our recommendations for fuzzy TFP filter thresholds for the discussed smoothing scales.

Figure 5.5 shows the relative frequency of $|\nabla_h \theta_w|$ for the same smoothing length scales as above, although this time only considering values at grid points within the frontal zone (i.e. where L_τ (Equation 5.1) > 0). The same effect encountered for TFP can be observed, and the horizontal smoothing length scale alters the relative frequency of $|\nabla_h \theta_w|$ as well. In general, $|\nabla_h \theta_w|$ decreases with increasing horizontal smoothing length scale. As for TFP , it is necessary to adapt frontal strength filter thresholds to the chosen horizontal smoothing length scale. Table 5.1 provides guidance.

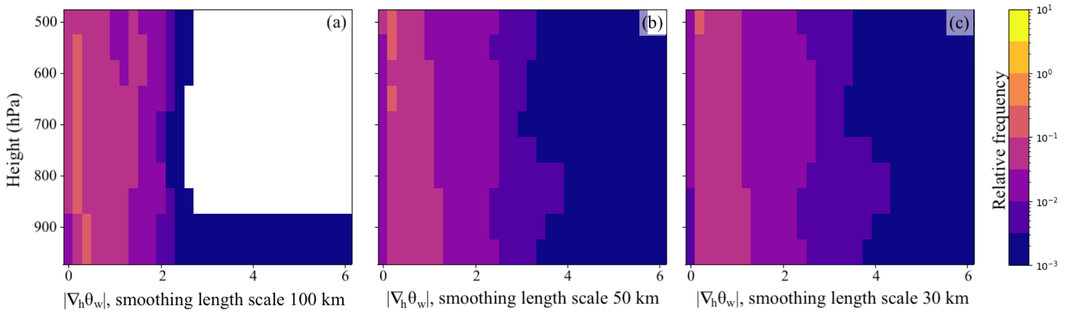
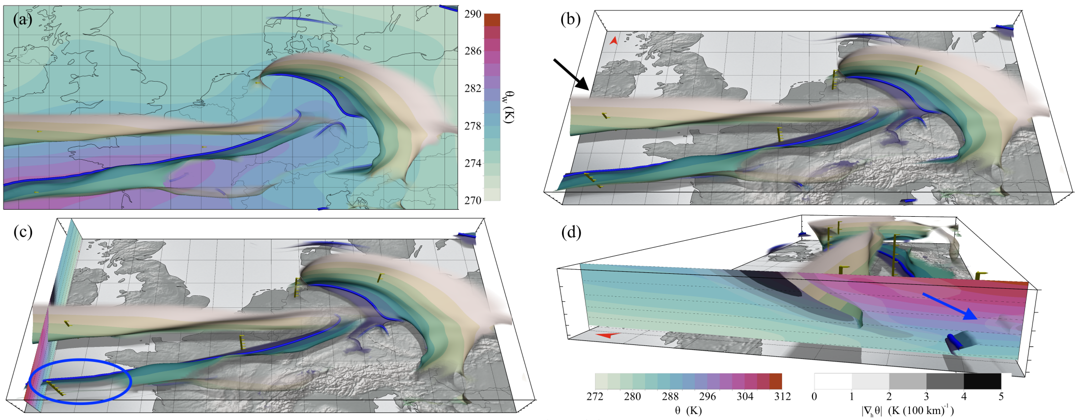


Figure 5.5: Distribution (relative frequencies) of $|\nabla_h \theta_w|$ within frontal zones between 950–500 hPa (same data, time, and region as in Figure 5.4) for different smoothing length scales: **(a)** 100 km, **(b)** 50 km, and **(c)** 30 km.

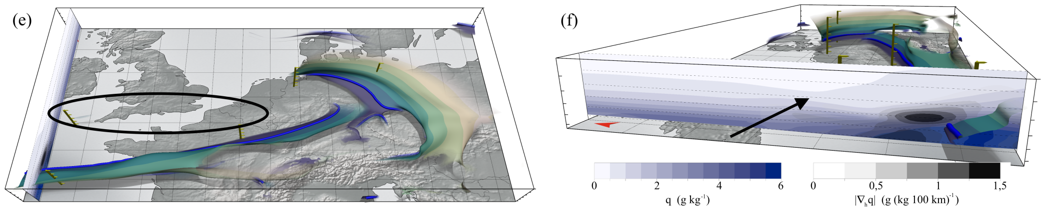
Example: impact of filtering and smoothing on detected frontal features

As mentioned above, NWP data at kilometre-scale resolution includes convective and thermal processes that are much smaller in scale than atmospheric fronts (Keyser and Shapiro, 1986). If the focus of an analysis is on large-scale frontal features, e.g. for large-scale weather analysis, the thermal variable can be smoothed with a distance between 50 and 100 km. If smaller-scale frontal surface phenomena, e.g. surface precipitation, are of interest, the smoothing distance can be reduced to a few kilometres. However, it should not be less than the grid spacing of the thermal input variable. In the following, we demonstrate how different smoothing length scales and filter thresholds impact the resulting frontal features. In particular, we show how different frontal strength filters can help distinguish between different front types (temperature- and humidity-dominated fronts).

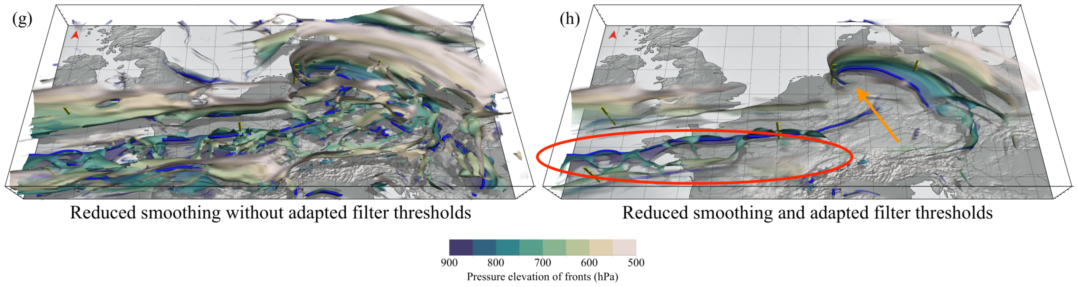
Figure 5.6a extends the 2-D visualization of Figure 5.2h to 3-D, depicting the full 3-D structure of the frontal surfaces. We would also like to point the reader to the Video supplement (Beckert et al., 2022c). We consider the interactive use of



Temperature-based fronts



Humidity-based fronts



Reduced smoothing without adapted filter thresholds

Reduced smoothing and adapted filter thresholds

Figure 5.6: From 2-D to 3-D objective fronts. Same data as in Figure 5.2 (18 January 2018, 12:00 UTC) but showing the full 3-D structure of frontal surfaces in the lower and middle atmosphere. All circles and arrows denote features discussed in text. **(a)** 850 hPa frontal lines from Figure 5.2h with 3-D frontal surfaces between surface and 500 hPa, viewed from the top. **(b)** Same as **(a)** but from a tilted viewpoint looking north. **(c)** Same as **(b)** but with additional fuzzy normal curve filter of θ between 0.6–1 K (100 km)$^{-1}$. **(d)** Same as **(c)** but viewed from west. Cross section shows θ and $|\nabla_h \theta|$. **(e)** Same as **(b)** but with additional fuzzy normal curve filter of specific humidity between 0.1–0.2 g (kg 100 km)$^{-1}$. **(f)** Same as **(e)** but viewed from west. Cross section shows q and $|\nabla_h q|$. **(g)** Input field smoothed to a horizontal length scale of 30 km with same filtering applied as in **(a)**. **(h)** Same as **(g)** but with adapted filter settings for TFP between 1.5–2.5 K (100 km)$^{-2}$ and frontal strength between 1.2–2.2 K (100 km)$^{-1}$.

the presented method as a key aspect of 3-D analysis, and the video provides an impression of the additional benefit gained through interaction.

The 3-D depiction in Figure 5.6a reveals further frontal structures such as the large-scale frontal surface in the north (marked with a black arrow in Figure 5.6b), which is located above the 850 hPa level and could easily be missed in a 2-D analysis. Not missing such potentially interesting structures is a key benefit of

3-D front detection compared to 2-D detection. Figure 5.1c–d show temperature-dominated fronts, obtained by applying an additional normal curve filter of θ with a fuzzy threshold interval of 0.6–1.0 K $(100\,\mathrm{km})^{-1}$, the same value range used for θ_w (see Figure 5.2). This filter discards all humidity-dominated fronts. Note that the interactive adjustment of the filter is also illustrated in the Video supplement (Beckert et al., 2022c). The blue circle in Figure 5.6c highlights an area of the cold front – note how upper-level parts (lighter green, towards the south) are discarded when the humidity contribution is filtered. The vertical cross-section in Figure 5.6d shows θ and $|\nabla_h \theta|$, with the black arrow pointing at the area of the filtered-out upper-level humidity-dominated front. The vertical cross-section also shows no temperature gradients, consistent with the interpretation that this is a humidity-dominated front. In Figure 5.6e–f a normal curve filter using a specific humidity filter is applied instead, shifting focus to humidity contribution and discarding temperature-dominated gradients in θ_w. In other words, temperature-dominated fronts are filtered out. The black circle in Figure 5.6e marks an area where a large-scale upper-level front is almost entirely discarded.

Finally, Figure 5.6g shows the impact of decreasing the smoothing length scale from 100 to 30 km. This reveals frontal features on a different length scale. However, without adjusting the filter thresholds, the resulting fronts become cluttered. Figure 5.6h shows the same fronts as in Figure 5.6g but with adapted filter thresholds to compensate for the reduced horizontal smoothing length scale. Due to reduced smoothing, the smoothness of the frontal surfaces is reduced. Especially at the cold front, fluctuations in θ_w cause less-continuous fronts (red circle). In addition, the reduced smoothing reveals other frontal features on smaller scales; for example, the wrap-up of the occluded front around the cyclone centre is more pronounced (orange arrow). Our recommendations for appropriate filter parameter intervals for different smoothing scales are summarised in Table 1 and are used throughout the paper, except where noted.

Table 5.1: Fuzzy frontal filter threshold recommendations for different smoothing length scales.

Smoothing length scale (km)	TFP $(\mathrm{K}\ (100\,\mathrm{km})^{-2})$	Frontal strength $\lvert\nabla_h \theta_w\rvert$ and $\lvert\nabla_h \theta\rvert$ $(\mathrm{K}\ (100\,\mathrm{km})^{-1})$	Scale of detected frontal features
100	0.2–0.4	0.6–1.0	~ larger than 500 km
50	0.4–0.8	1.0–1.6	~ 200–500 km
30	1.5–2.5	1.2–2.2	~ below 200 km

5.2.3 *Impact on reduced vertical resolution*

An estimation of the minimum vertical level in NWP data for front detection may be helpful, as in some cases only limited vertical levels are available. Therefore,

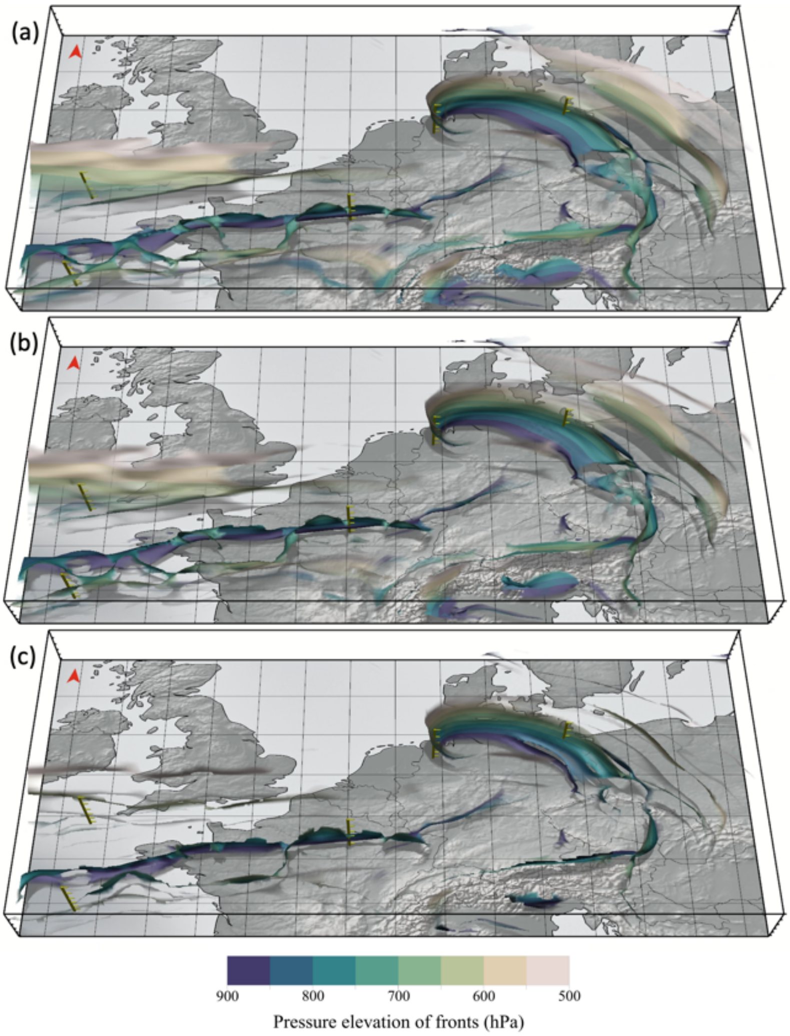


Figure 5.7: Impact of the vertical resolution on detected fronts. (a) original vertical resolution of ECMWF with 137 vertical level. (b) vertical level retained to 68 level. (c) vertical levels retained to 28 levels.

the effect of different vertical levels on the detected fronts is illustrated by an example. Figure 5.7 replicates Figure 5.6, showing the impact of decreasing vertical resolution in the ECMWF forecast data, which originally included 137 vertical levels. As expected, a reduction in vertical resolution results in reduced visibility of detail. The interconnected frontal surfaces visible at 137 levels (Figure 5.7a) become fragmented as the vertical resolution decreases. In particular, the blue encircled fronts in Figure 5.7c illustrate this effect: while halving the vertical resolution (from 137 to 68) has little effect, a further reduction to 28 levels (Figure 5.7c) leads to a breakdown of structures. Note that, similar to other figures in this thesis, only fronts in the lower troposphere (900 to 500 hPa) are shown, corresponding to about 24 vertical levels of the original 137. Halving the vertical resolution results in about 12 vertical levels between 900 hPa and 500 hPa, and at the lowest resolution shown in Figure 5.7c, only 5 vertical levels remain for front detection.

5.3 CASE STUDIES

We illustrate how meteorological analysis can be performed using 2-D and 3-D front detection by investigating two case studies of extra-tropical cyclones. The first case, Cyclone Vladiana, occurred in the North Atlantic in September 2016. In Section 4.1 of Chapter 4 describes the synoptic situation and the data used for our analysis. For Vladiana, we examine the conceptual model of WCB ascent in the vicinity of fronts (Section 5.3.2) and show how frontal surfaces from convection-permitting NWP simulations compare to those found in simulations in which convection is parametrised (Section 5.3.3). The second case, Cyclone Friederike, took place in western Europe in January 2018 (introduced in Section 4.2 of Chapter 4). For Friederike, we examine the development stages of a Shapiro–Keyser cyclone in 3-D (Section 5.3.4). Additionally, we compare our results to fronts analysed by the UK Met Office to discuss secondary fronts as often shown in surface analysis charts of the UK Met Office (Section 5.3.5). Before introducing our case studies, we briefly revisit the underlying meteorological theory in Section 5.3.1.

5.3.1 *Meteorological theory*

The frontal structure of extra-tropical cyclones is a key feature for the analysis of their development. Typically, extra-tropical cyclones are classified as either classical Norwegian cyclones (Bjerknes, 1919) or (the later proposed) Shapiro–Keyser cyclones (Shapiro and Keyser, 1990). The development of both cyclone types is classified into four characteristic stages. A cyclone first develops along a frontal wave as a small disturbance near the surface (stage I in both models). Meanwhile, this disturbance strengthens and extends to higher elevations, and the cyclone starts to rotate cyclonically and forms a warm sector (stage II). In stage II the warm sector has its maximum size and maximum energy conversion. For Norwegian cyclones the displacement speed of the cold front is faster than of the warm front, and the warm sector diminishes (stage III). The fronts occlude forcing the air to rise before the cyclone finally dissipates (stage IV). In contrast, a Shapiro–Keyser cyclone develops a frontal fracture in stage II separating the cold front from the warm front. While the cold front is usually weaker than in Norwegian cyclones (Schultz et al., 1998), the warm front is north of the cyclone centre and starts wrapping around it bending backwards and hence is also called bent-back front (stage III). This stage is also called "T-bone structure". With the warm front wrapping around the cyclone centre, a warm seclusion occurs (stage IV) before the cyclone decays. More recent literature proposes an extension of the four stages by three additional stages: the diminutive frontal wave stage and frontal wave stage which occur before stage I and a decay stage after stage IV (Hewson and Titley, 2010). However, in this publication we focus on the initially proposed four stages of the Shapiro–Keyser cyclone model.

Both cyclone models can be accompanied by coherent circulation features called conveyor belts. The cold conveyor belt occurs ahead of the warm and occlusion front, usually remaining below 850 hPa. It is often associated with high wind speeds in later stages, typically south-west of the cyclone centre. The WCB (see Eckhardt et al., 2004; Madonna et al., 2014) occurs ahead of the cold front near the surface in early stages and is also associated with high wind speeds. It typically ascends at least 600 hPa in the warm sector and over the warm front and often splits into anticyclonically and cyclonically turning branches (Martínez-Alvarado et al., 2014).

5.3.2 *The 3-D examination of conceptual model: fronts and warm conveyor belt*

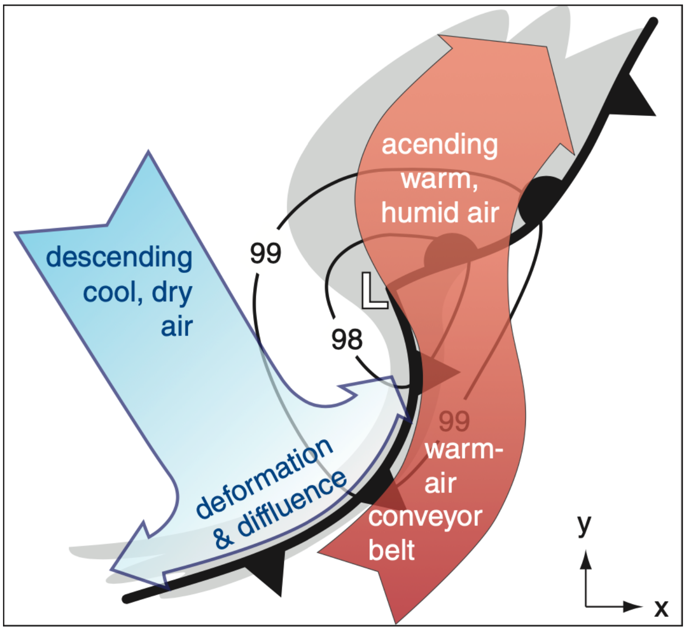


Figure 5.8: Conceptual model of fronts and WCB showing large-scale ascending and descending air in the vicinity of an extra-tropical cyclone. Figure adapted from Stull (2017), used under CC BY-NC-SA 4.0.

Conceptual models and simplified illustrations are frequently used to explain the relation and dynamics of fronts and the WCB. Figure 5.8 shows an example of such an illustration in 2-D, but a more sophisticated 3-D representation can be found, for example, in Martínez-Alvarado et al. (2014, their Figure 1). However, subsequent studies of these 3-D atmospheric features are usually conducted by means of horizontal or vertical 2-D slices through NWP data, and it is less common to use a 3-D representation of 3-D atmospheric features (Rautenhaus et al., 2018). In this section, we demonstrate the use of 3-D front detection to visualise such conceptual models against NWP data by directly representing these features in 3-D.

Figure 5.9a–c show the evolution of 3-D fronts from 03:00 to 09:00 UTC on 23 September 2016 of Vladiana, together with a selection of WCB trajectories that ascend at the selected times. During this period the frontal system moves eastwards. At 03:00 UTC the selected WCB trajectories are located in the lower troposphere near the surface in the warm sector and move along the cold front in a north-eastward direction (Figure 5.9a). At 06:00 UTC most of the WCB trajectories are in their ascent phase (Figure 5.9b), and at 09:00 UTC the majority

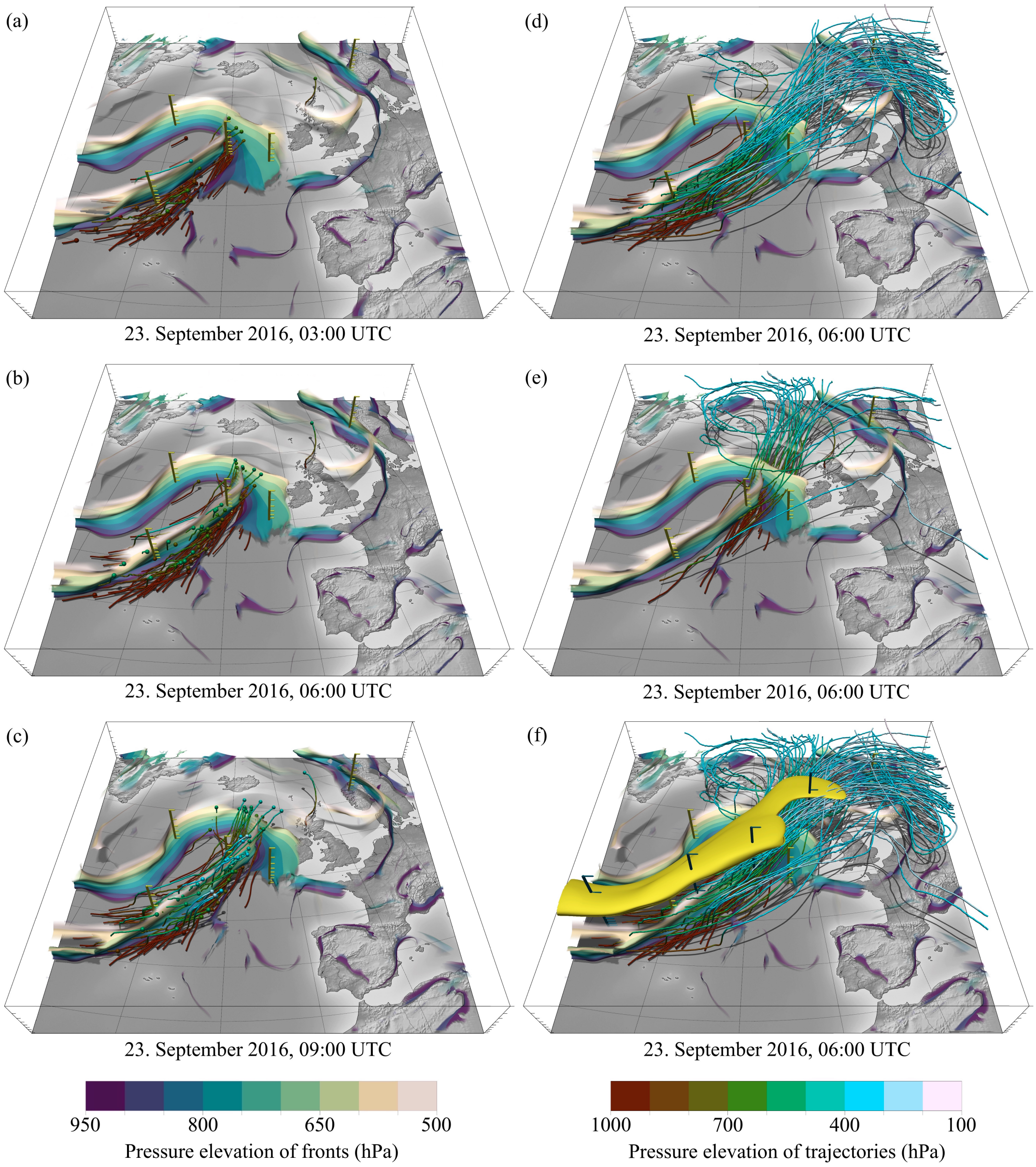


Figure 5.9: **(a–c)** Temporal evolution of 3-D frontal structures and WCB trajectories of Vladiana on 23 September 2016. **(d)** Same time as **(b)** but only fast-ascending WCB trajectories (minimum 200 hPa within 2 h) are displayed for a period of 48 h. **(e)** Same as **(d)** but only slow-ascending WCB trajectories (less than 200 hPa within 2 h) are displayed. **(f)** Same time as **(c)**, jet stream (yellow isosurface of 50 m s^{-1} wind speed) and WCB trajectories are displayed for a period of 48 h. For the full temporal development of this scene, see the Video supplement (Beckert et al., 2022a).

of the WCB trajectories have risen above 500 hPa (Figure 5.9c). The selected trajectories have different pathways for their ascent: some rise directly at or ahead of the cold front, and others rise above the warm front. While trajectories rapidly increase in altitude when lifted spontaneously at the cold front, trajectories at the warm front ascend more slowly and gradually. In Figure 5.9d–e the difference between cold-frontal and warm-frontal ascent is emphasised. Figure 5.9d shows frontal surfaces at 06:00 UTC, together with 48 h WCB trajectories with maximum ascent rates faster than 200 hPa within 2 h. Most of these fast-ascending WCB trajectories ascend at the cold front. In contrast, trajectories at the warm front ascend more slowly, with maximum ascent rates below 200 hPa in 2 h (Figure 5.9e). In the upper troposphere, the WCB splits into two outflow branches: a cyclonic branch which turns westward and an anticyclonic branch which turns eastwards. WCB trajectories ascending ahead of the cold front tend to take the anticyclonic outflow, while warm-frontal WCB trajectories tend to take the cyclonic outflow. We hypothesise that trajectories that rapidly ascend at the cold front experience jet wind speeds earlier following the anticyclonically turning jet stream and are thus deflected into the downstream ridge (see Figure 5.9f). The 3-D visualization corroborates the conceptual model of how WCB ascent relates to fronts and highlights the presence of smaller-scale convective ascent structures embedded in the WCB discussed in recent studies (see Blanchard et al., 2020; Oertel et al., 2019, 2020; Rasp et al., 2016). The 3-D visualization of rapidly and more slowly ascending high-resolution WCB trajectories further shows their similarity to the so-called “escalator–elevator” concept of WCB-embedded convection which was proposed by Neiman et al. (1993) to distinguish between fast ascent and more gradual frontal upglide. By looking at the 3-D structure of the trajectories, this concept appears suitable for this case study.

5.3.3 *Cold-front structure in the vicinity of convection*

Here we compare fronts of convection-permitting NWP simulations with fronts in simulations where convection is parametrised, using Vladiana as an example. We focus on the southern end of the cold front (green box in Figure 4.1) where mid- and small-scale convection occurs in this WCB. Oertel et al. (2019) highlight (embedded) convection with lightning near the trailing edge of the cold front on 23 September 2016 at 06:00 UTC. To detect mid-scale frontal features induced by convection the input field θ_w is smoothed to a horizontal length scale of 50 km and filtered according to TFP , θ_w, and θ (see Table 5.1). Figure 5.10 shows detected 2-D fronts at 850 hPa together with fronts of UK Met Office surface charts, at 700 hPa and at 500 hPa. The yellow dot at the southern end of the cold front marks the position of the observed embedded moist convection. The COSMO simulation shows strong ascending motion in this region at all plotted vertical levels (Figure 5.10d–f). In contrast, in the ECMWF data (Figure 5.10a–c) where convection is parametrised, the vertical velocity field shows no significant local maximum. The detected cold front of both simulations follows the cold front of

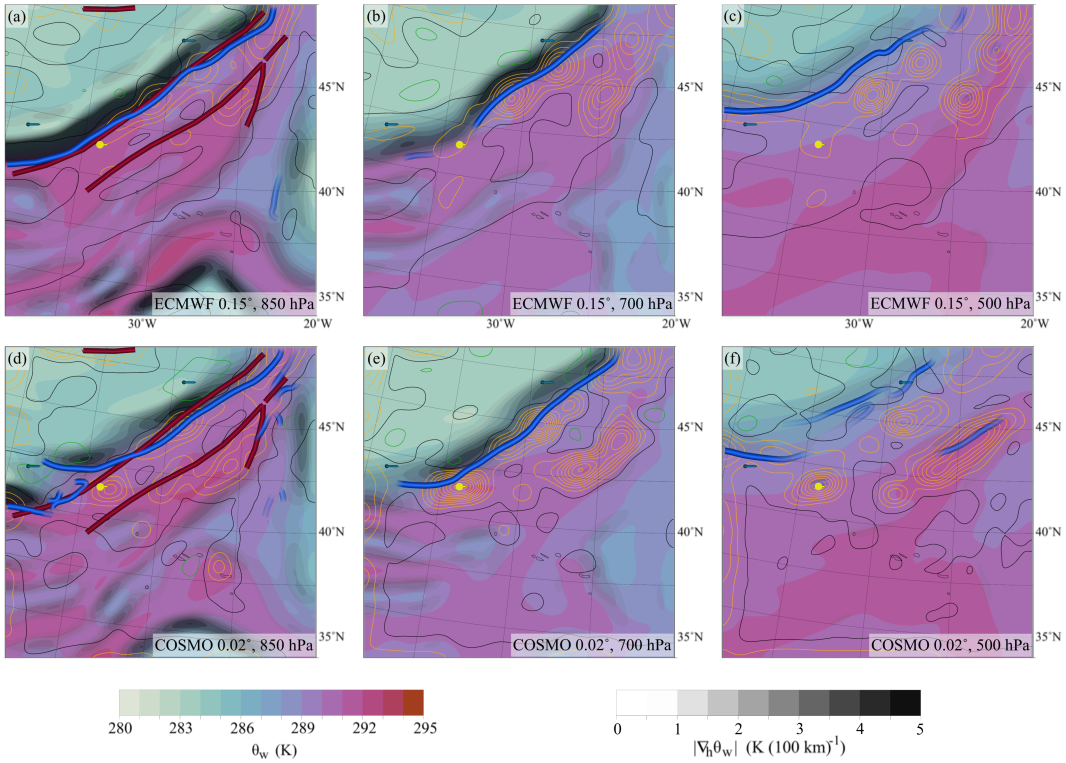


Figure 5.10: Convection and frontal structure on 23 September 2016 at 06:00 UTC. Region corresponds to green sub-area in Figure 4.1. ECMWF analysis **(a, b, c)** and COSMO analysis **(d, e, f)** at **(a, d)** 850 hPa, **(b, e)** 700 hPa, and **(c, f)** 500 hPa. Objective 2-D fronts (blue tubes) are shown along with UK Met Office fronts (red tubes), θ_w (colour), $|\nabla_h \theta_w|$ (grey shades), and upward air velocity (contour lines: orange is upwards, black is zero, and green is downwards; contour line spacing is 0.02 m s^{-1}).

the UK Met Office surface analysis chart. However, in the vicinity of convection and at 850 hPa the cold front of the COSMO simulation breaks apart, while the cold front detected in ECMWF is a continuous line. At 700 hPa the cold front detected in ECMWF data is weak and broken, while the cold front detected in COSMO data is a continuous line. At 500 hPa the cold front is shifted towards north and is less continuous in the COSMO data compared to ECMWF data.

Figure 5.11 shows the corresponding 3-D frontal structures. In the area where convective vertical motion differs between the two simulations, a gap can be observed in the frontal surface between 700–600 hPa in the ECMWF data, whereas the frontal surface is present in the COSMO simulation (red circle in Figure 5.11a–b). These kinds of gaps in the cold front have been observed in earlier studies (Geerts et al., 2006) and were associated with weaker temperature gradients at this elevation range. The time evolution of the COSMO 3-D front (Figure 5.12) suggests that the intensification of the mid-level cold front is a transient feature that occurs at the time of convection, which is associated with strong horizontal convergence (Figure 5.11c–d), and disappears as soon as the convection weakens again. In simulations where convection is parametrised, however, the convection

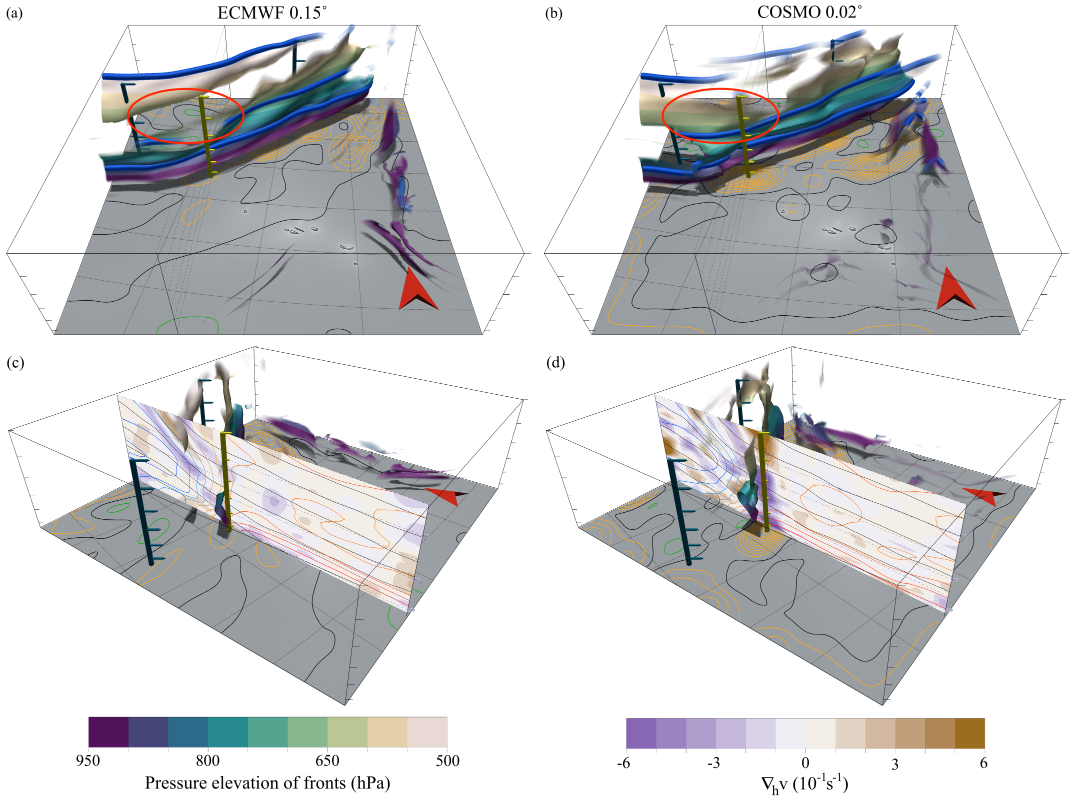


Figure 5.11: The 3-D view of the 2-D frontal structures from Figure 5.10. **(a)** 2-D objective fronts (blue tubes) at 850, 700, and 500 hPa (see Figure 5.10) in the context of full 3-D frontal structures, as found in ECMWF data. **(b)** Same as **(a)** but for COSMO data. Red circles in **(a)** and **(b)** mark the differences in the frontal surfaces. Contour lines on all surface maps represent upward air velocity at 700 hPa (orange is upwards, black is zero, and green is downwards; contour line spacing 0.02 m s^{-1}). **(c)** ECMWF 3-D fronts and vertical section of wind divergence (colour), θ_w (coloured contour lines, spacing 1 K), and θ (black contour lines, spacing 5 K). **(d)** Same as **(c)** but for COSMO data.

scheme may not activate at that time and location. Additionally, the feedback of the convection scheme on the grid-scale variables may differ from their explicit model representation (as shown in this example). We hypothesise that the model representation of convection and/or simulation grid spacing influences the feedback and interaction between convection, frontogenesis, and detailed frontal structures. The investigation of this relation between frontal structure, θ_w gradient, and convective ascent, however, will require more detailed and systematic analyses that are beyond the scope of this study.

5.3.4 *The 3-D examination of conceptual model: Shapiro–Keyser cyclone*

Figure 5.13 extends the 2-D frontal analysis of Friederike shown in Figure 4.1 and shows the temporal development of the 3-D structure. In 3-D, the typical characteristics of a Shapiro–Keyser cyclone (Shapiro and Keyser, 1990) with its

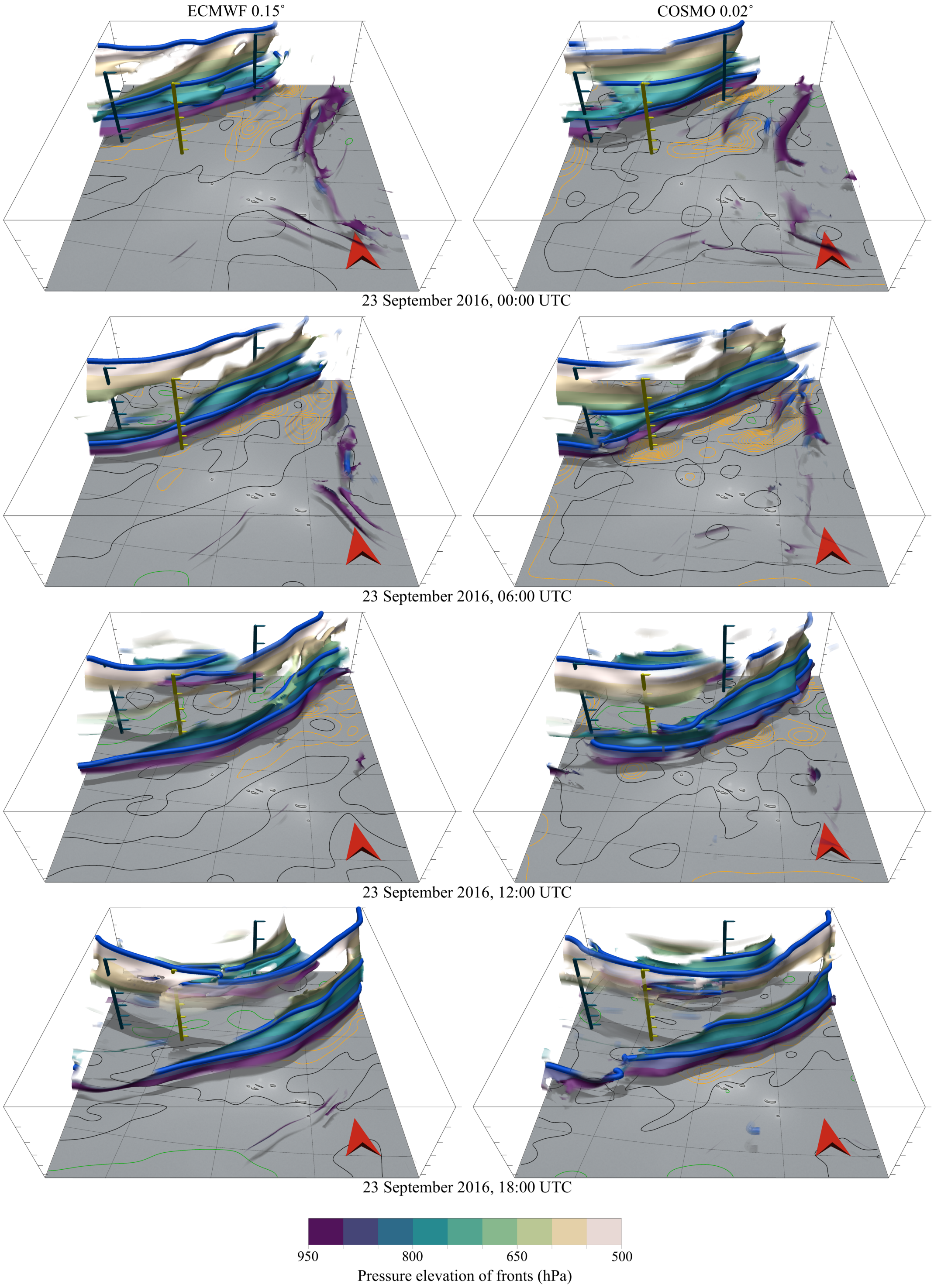


Figure 5.12: Temporal evolution of 3-D frontal structures in Figure 5.11, detected from (left) ECMWF analysis and (right) COSMO analysis. Contour lines projected onto the surface show upward air velocity at 700 hPa (orange is upwards, black is zero, and green is downwards; contour line spacing of 0.02 m s^{-1}). The yellow pole marks the centre of the convective updraft at 06:00 UTC, and the red arrow points northward.

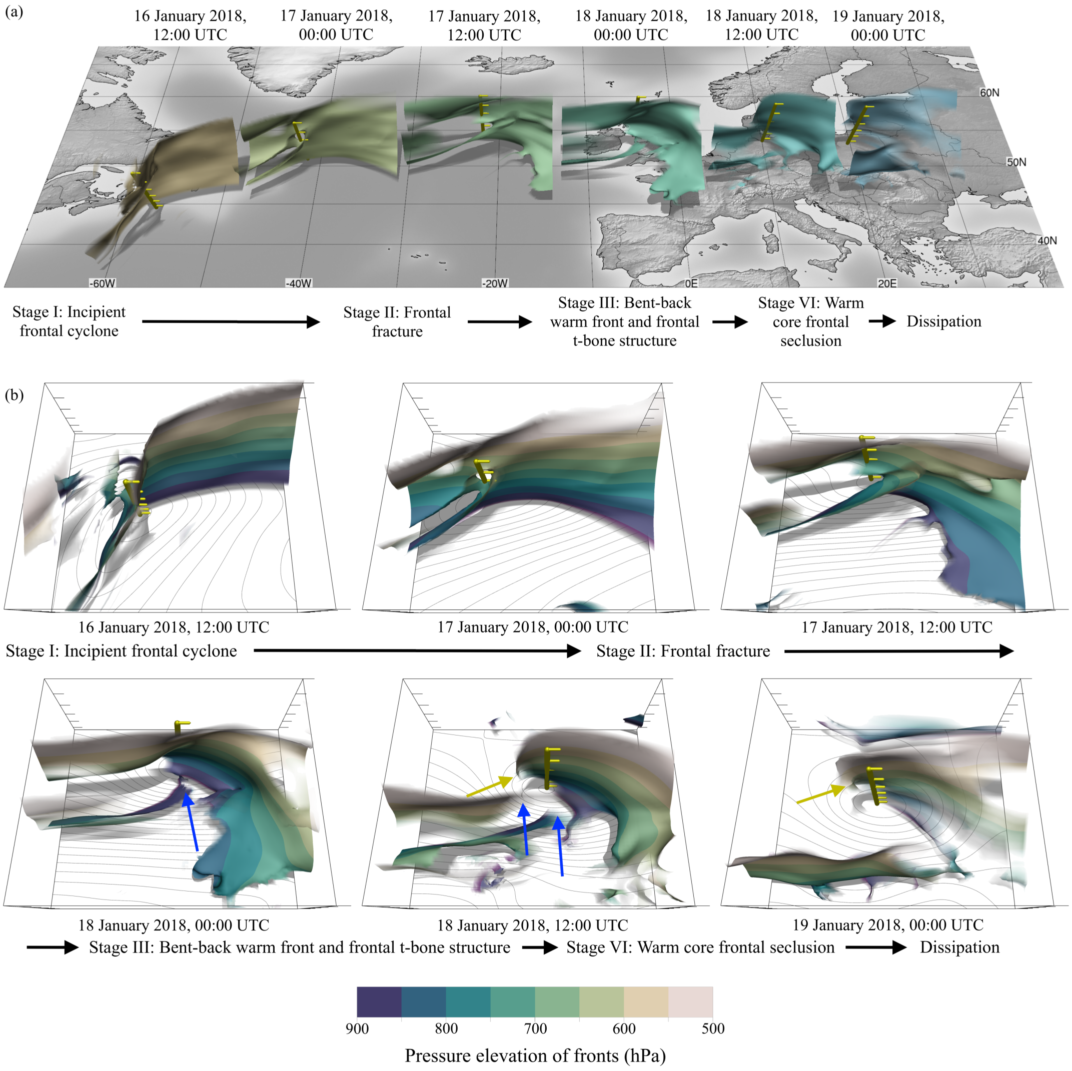


Figure 5.13: Temporal evolution of 3-D frontal structures of Friederike (16 to 19 January 2018), as detected in ERA-5 reanalysis data. **(a)** Different cyclone stages encountered along the cyclone track. Yellow poles mark centres of surface low, and front colours distinguish time steps. **(b)** The six stages from **(a)**, approximately centred around the cyclone centres for comparison of frontal structures. Blue arrows mark frontal fracture, yellow arrows mark warm-core frontal seclusion, and contour lines show surface pressure (spacing 2 hPa).

distinctive frontal T-bone structure and the four cyclone stages can be observed well. However, at different elevations the four stages, as described in Schultz and Vaughan (2011), occur at different times.

- *Red and orange front: stage I, incipient frontal cyclone*. A perturbation of the frontal structure is already present in the upper atmosphere. This disturbance will later develop into the frontal wave. However, the frontal surface in the lower atmosphere is unperturbed.
- *Orange, yellow, green front: stage II, frontal fracture*. The timing of frontal fracture strongly depends on the vertical level. In the lower troposphere the

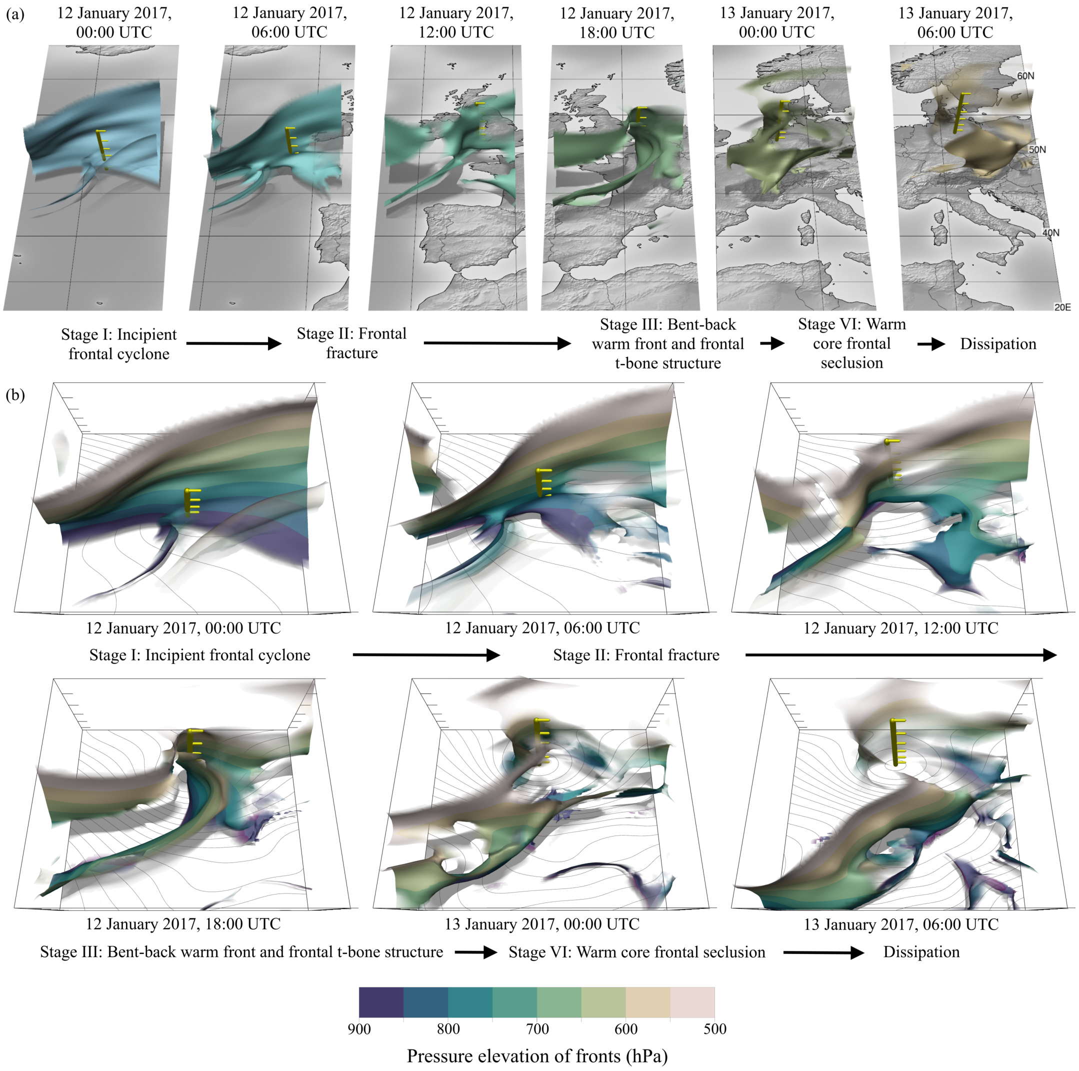


Figure 5.14: Temporal evolution of 3-D frontal structures of Egon (12 to 13 January 2017), as detected in ERA-5 reanalysis data. **(a)** Different cyclone stages encountered along the cyclone track. Yellow poles mark centres of surface low, and front colours distinguish time steps. **(b)** The six stages from **(a)**, approximately centred around the cyclone centres for comparison of frontal structures. Contour lines show surface pressure (spacing 2 hPa).

cold front is separating from the main front. In the upper troposphere, a connection between the cold front and the main part of the frontal surface still exists.

- *Green and blue front: stage III, bent-back warm front and frontal T-bone structure*. At lower levels, the cold front lies almost perpendicular to the warm front, showing the typical Shapiro–Keyser T-bone structure. Interestingly, the upper part of the cold front also bends slightly towards the south, following the lower part of the cold front, but a connection to the warm front remains.

- *Blue and purple front: stage IV, warm-core frontal seclusion.* The warm front wraps up around the warm air near the cyclone centre. The separated lower part of the cold front moves further south, and the upper cold front dissipates.

In this example, uniquely assigning the 3-D frontal structure at specific time steps to the Shapiro and Keyser stages is not possible. As described, frontal evolution does not occur synchronously at all elevations, creating a temporal offset of the stages at different elevations. We could also not find a height level where the 2-D fronts could be uniquely assigned (see Figure 4.2). It is important, however, that the 3-D front detection can detect all the characteristic structures of the Shapiro–Keyser model, even though a one-to-one assignment to the stages is not possible. Another example of the 3-D frontal development with typical characteristics of a Shapiro–Keyser cyclone, Cyclone Egon (11–13 January 2017; Eisenstein et al., 2020, is shown in Figure 5.14. Again, the visual analysis shows that frontal evolution does not occur synchronously at all elevations, creating a temporal offset. For example, frontal fracture does not occur at all elevations simultaneously. The time step on 13 January 2017 at 00:00 UTC shows the development of the bent-back warm front in upper levels, whereas the frontal fracture is not yet complete near the surface. These examples suggest a more nuanced view of the Shapiro–Keyser model, where there is a significant 3-D component to the evolution of a cyclone through the different stages of the conceptual model.

5.3.5 *Secondary fronts*

Secondary fronts are commonly analysed by the UK Met Office and seen in their surface analysis charts. Beside other variables, the UK Met Office uses the wet-bulb potential temperature as the primary thermal variable for their front

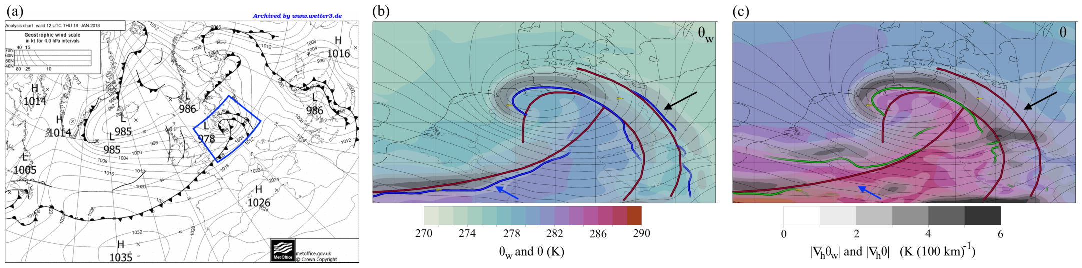


Figure 5.15: Comparison of UK Met Office fronts with objective fronts for case Friederike (18 January 2018, 12:00 UTC). **(a)** UK Met Office surface analysis chart. Blue box marks analysed area. **(b)** Objective 850 hPa 2-D fronts (blue lines) as detected from ECMWF HRES θ_w (colour; grey shading shows $|\nabla_h \theta_w|$), UK Met Office fronts (red lines), and mean sea level pressure (black contour lines, spacing 2 hPa). **(c)** Same as **(b)** but objective 2-D fronts (green lines) based on θ. The secondary front (black arrow) is only detected when using θ_w. When based on θ, the cold front (blue arrow) breaks up and is less continuous compared to the cold front based on θ_w.

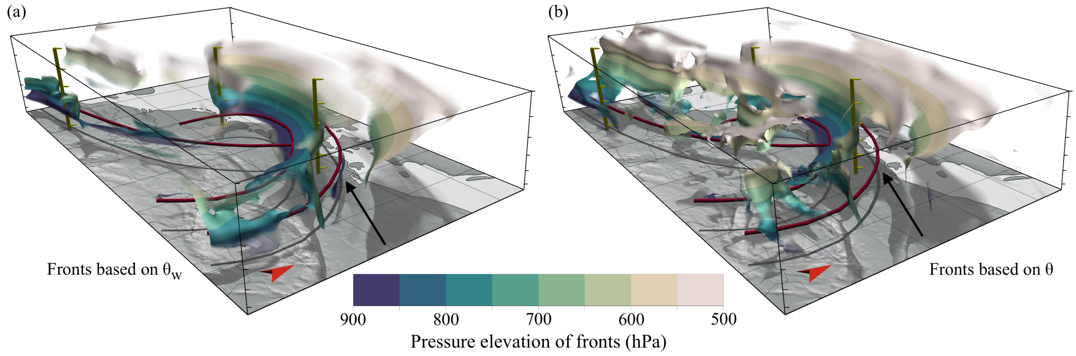


Figure 5.16: The 3-D view of Figure 5.15b–c. Red tubes show UK Met Office fronts, and 3-D objective fronts are coloured according to pressure elevation. Objective fronts based on **(a)** θ_w and **(b)** θ. The secondary front (black arrow) is a feature of θ_w and only occurs around 850 hPa. Yellow poles are to aid spatial perception. Compare the animated version in the Video supplement (Beckert et al., 2022b).

detection in surface analysis charts (Neil Armstrong, UK Met Office, personal communication, 2022). In this section, we consider a secondary front which occurs ahead of the warm front of Friederike. We investigate if the front detection algorithm can detect such secondary fronts and how secondary fronts depend on the detection variable. Red tubes in Figure 5.15 show the positions of fronts analysed by the UK Met Office for 18 January 2018 at 12:00,UTC. The most eastward front, extending from north-east Italy up to the southern border of Denmark, is a typical secondary warm front as often analysed by the UK Met Office. Figure 5.15b shows fronts detected in θ_w at 850 hPa (blue tubes). In general, the structure of fronts detected in θ_w agrees well with fronts of the UK Met Office, despite some smaller differences. In particular, the secondary front detected in θ_w is shorter in its horizontal extent, and the wrap-up of the occluded front around the cyclone centre is more pronounced.

Figure 5.15c shows fronts detected in θ at 850 hPa (green tubes). There is no indication for secondary fronts in this analysis, as no strong horizontal gradients of θ are present in this area. Hence, the presence of the secondary front detected by θ_w results from moisture gradients. Furthermore, the structure of the primary fronts is less continuous and deviates more from the UK Met Office analysis. Figure 5.16 shows the 3-D frontal surfaces of θ_w (Figure 5.16a) and θ (Figure 5.16b). The 3-D frontal structure illustrates that the secondary front detected in θ_w is a shallow atmospheric feature and is only present in the lower troposphere at around 850 hPa. For this case study we conclude that the lower-atmospheric secondary front is a moisture feature and thus can only be detected in a variable that includes humidity formation. Furthermore, θ_w as the detection variable results in more-continuous fronts compared to θ. We again would like to point the reader to the Video supplement (Beckert et al., 2022c), which illustrates the benefit of interactive exploration and analysis of the detected fronts within Met.3D.

Impact on surface weather from secondary front

In the ECMWF HRES forecast data, a total column rainwater content signature is observed near the secondary cold front, as indicated in Figure 5.17. Additionally, a review of a UK radar image, though not displayed here due to copyright constraints, reveals a fragmented convective rain band in the vicinity of the secondary cold front over the UK. Although significant surface weather is observable, there is no conclusive evidence to establish a direct link between this convective rain band and the secondary cold front. Therefore, further investigation is considered necessary to confirm any possible link between the two phenomena.

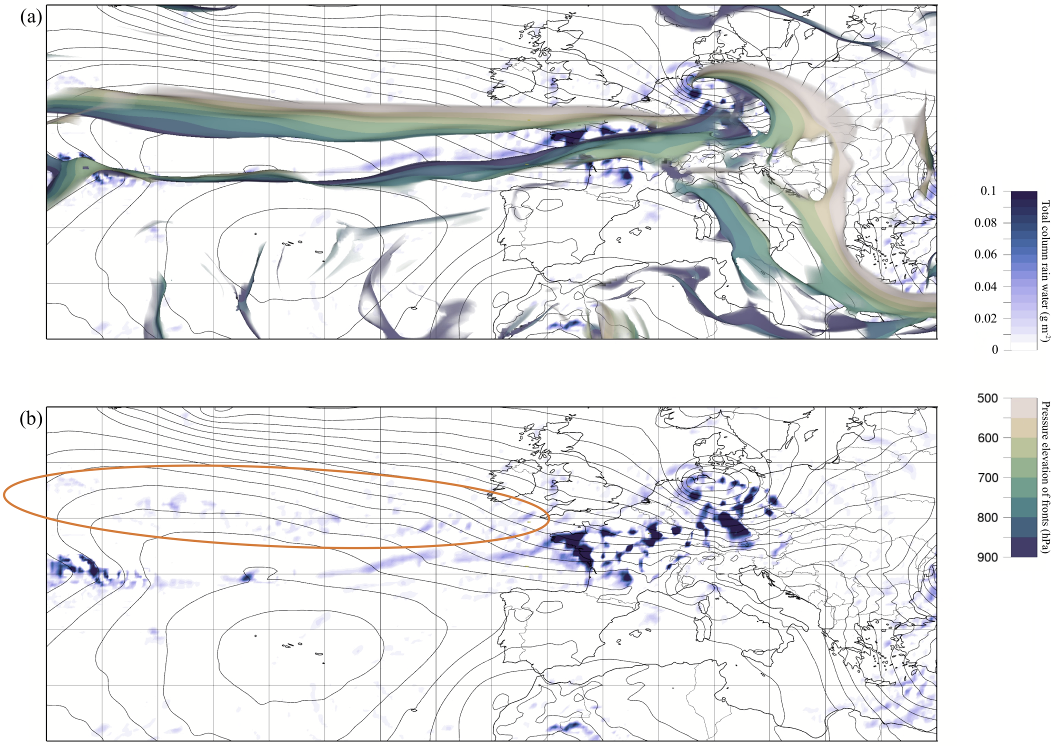


Figure 5.17: Total column rain water and 3-D fronts of ECMWF HRES simulation on 18 January 2018, 12:00 UTC, initialised on 18 January 2018, 00:00 UTC. (a) 3-D fronts and total column rain water. (b) Total column rain water. The orange circle highlights the position of the secondary cold front shown in (a).

6

FEATURE-BASED ENSEMBLE ANALYSIS AND CLUSTERING

Contents

6.1 Methodology: Feature tracking and feature time series analysis 82
  6.1.1 Manual front tracking 82
  6.1.2 Characterising 3-D fronts by frontal attributes 85
  6.1.3 Automated front tracking 88
  6.1.4 Architecture and implementation details in Met.3D 89
  6.1.5 Example of front-tracking workflow 90
  6.1.6 Postprocessing of frontal attribute time series 91
6.2 Case Study: Frontal attribute distributions and tracking 93
  6.2.1 Distribution of frontals attributes 93
  6.2.2 Time Resolution 94
  6.2.3 Vertical Resolution 97
  6.2.4 Ensemble Tracking 99
6.3 Case study: Feature-based time series similarities and clustering 99
  6.3.1 Feature-based time series similarities 102
  6.3.2 Feature-based time series clustering 104

Clustering of NWP ensemble simulations is a widely used tool in weather forecasting and atmospheric research to derive representative weather scenarios and to analyse forecast uncertainties based on scalar fields (e.g. Ferranti and Corti, 2011; Kumpf et al., 2018). A review of common ensemble analysis and clustering techniques in atmospheric science can be found in Wilks (2019). In this chapter, I design, develop and apply an approach to order and cluster ensemble simulations based on similarities of 3-D atmospheric fronts. 3-D atmospheric fronts are represented by 3-D surfaces. To my knowledge, an approach of using 3-D atmospheric features, in particular 3-D fronts represented by 3-D surfaces, to cluster ensemble simulations has not been proposed yet. The objectives of this method are described in detail in Section 1.2. However, a brief summary of the main objectives is given here. The objective is to design and develop a method for isolating and tracking 3-D atmospheric fronts. From the tracked 3-D fronts, time series of 1-D feature attributes are generated for each ensemble member of NWP ensemble simulations. These feature attribute time series are then ordered and clustered based on their similarities. Can such an approach lead to meaningful clusters and enhance the meteorological analysis of ensemble simulations?

This chapter is structured as follows: Section 6.1 describes methods developed to isolate, track, and characterise 3-D fronts based on feature attributes. Section 6.2 analyses the derived frontal attributes and examines their suitability for automatic front tracking. Based on two case studies of NWP ensemble simulations, a feature-based ensemble time series similarity analysis and a feature-based clustering of ensemble simulations are presented in Section 6.3 and Section 6.3.2, respectively.

6.1 METHODOLOGY: FEATURE TRACKING AND FEATURE TIME SERIES ANALYSIS

To generate time series of frontal attributes and subsequently sort and cluster these time series based on their similarities, it is necessary to track the front of interest over time. To track objectively detected 3-D fronts, the previously presented objective front detection method (see Section 5.1) is slightly adapted. Subsequently, 1-D feature attributes are derived that characterise the physical and geometrical properties of 3-D fronts. These 1-D frontal attributes are then computed for each time step of a manually tracked 3-D front and represent the temporal evolution of the frontal characteristics.

This section is structured as follows: a manual front tracking method, including adaptations of the objective front detection method, is presented in Section 6.1.1. The computation of frontal attributes is described in Section 6.1.2 and an automated front tracking method is proposed in Section 6.1.3. A detailed workflow example is presented in Section 6.1.5.

6.1.1 *Manual front tracking*

To implement front tracking, the objective 3-D front detection method introduced in Section 5.1 is adapted and extended by three additional data pipeline modules. First, the fuzzy front filtering (see Section 5.1.2) is replaced by hard filter thresholds. Second, the computation of contiguous 3-D front surfaces. Third, the implementation of an interactive front selection method, including the handling of front splitting and merging events. The implementation details for all four adaptations and extensions of front detection are given below.

3-D front filter

The initial implementation for filtering front candidates includes fuzzy filtering of the 3-D front. Unwanted front geometries are not filtered out, but are made transparent or semi-transparent based on fuzzy filter thresholds applied during

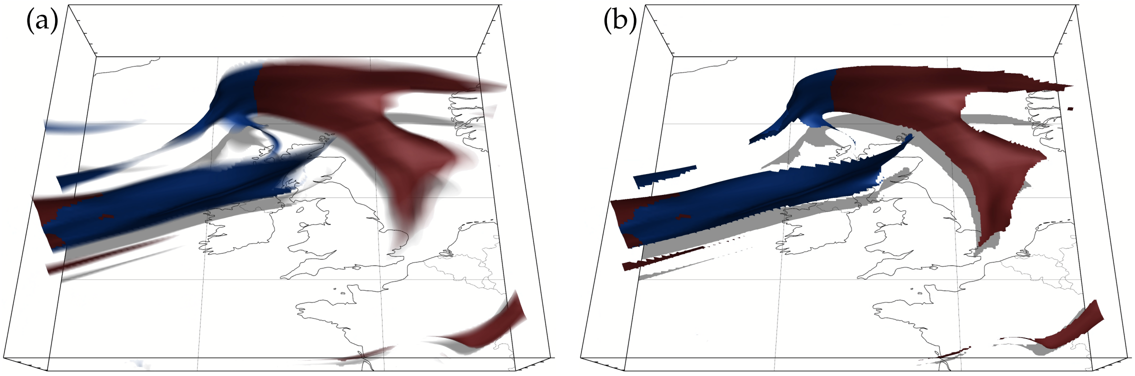


Figure 6.1: Comparison of 3-D fronts with (a) fuzzy-filter thresholds and (b) hard filter thresholds.

the rendering process. In addition, the filtering process for the most prominent front filters (TFP and frontal strength filter) is implemented on the GPU. The GPU implementation significantly improves filter efficiency while maintaining interactivity.

However, this filtering technique has two disadvantages when implementing the selection of individual fronts and the definition of frontal attributes. Firstly, a frontal surface is not defined with a clear boundary and its geometry is not separated from other frontal surfaces. The transparency mapping during the rendering process causes fronts to appear only as separate structures, but fronts removed by filters are not removed from the underlying triangle mesh, only rendered fully transparent. Secondly, it is unclear how to handle fuzzy frontal areas (areas between the lower and upper fuzzy filter thresholds) where fronts are rendered semi-transparent. Imagine calculating the frontal area of a cold front. How can such semi-transparent areas be taken into account?

For the above reasons, an additional 3-D front filter (*front hard borders filter*) implements hard filter thresholds that are applied prior to rendering. These hard filter thresholds ensure that the boundaries of frontal surfaces are precisely defined and that the geometries of frontal surfaces are separated from different fronts by deleting all previously transparent and semi-transparent surfaces. The algorithm is implemented in a new pipeline module in Met.3D, which checks for each vertex of the triangle whether the individual threshold criteria are met or not; if any of the three vertices do not meet the threshold criteria, the entire triangle is discarded. All triangles that pass this filter are passed to the next pipeline module. Figure 6.1 compares fuzzy-filtered fronts and fronts filtered with hard filter thresholds.

3-D front orderer source

The *front orderer source* is a new data pipeline module that determines which triangles belong to a contiguous 3-D front surface. The input of arbitrarily arranged triangles originates from the MC algorithm and is filtered in the previous pipeline module. In the initial front detection pipeline, each triangle is rendered separately, and the information about which triangles form a contiguous 3-D front surface is not required. The 3-D front orderer source pipeline module structures the triangle mesh by computing batches of contiguous frontal surfaces. The filter uses an algorithm described in Section 2.4.2. A triangle is considered connected to another triangle if they share at least one edge. The algorithm is briefly described below. A detailed description can be found in Section 2.4.2. The algorithm consists of the following three steps:

1. Compute for each vertex a list of surrounding triangles.
2. Compute for each triangle a list of surrounding triangles.
3. Create batches of contiguous triangles.

On completion, the filter returns a list of contiguous triangle batches, where each batch represents a contiguous frontal surface. The frontal batches are rendered as frontal surfaces, similar to the original front detection pipeline. However, the ordered representation allows further analysis and processing of the frontal surfaces in subsequent stages of the pipeline.

3-D front selector source

The objective of the *front selector source* pipeline module is to interactively select individual contiguous 3-D front surfaces. It provides a direct interaction between the user and the interactive visualisation software Met.3D. Its purpose is to allow the user to select specific front surfaces using the computer mouse. By positioning the mouse pointer over a desired frontal surface and clicking the left mouse button, a ray is cast from the mouse pointer position perpendicular to the plane of the monitor into the scene. The intersection of this ray with the frontal surfaces is determined using the Möller-Trumbore (Möller and Trumbore, 1997) intersection algorithm, which is described in detail in Section 2.4.2. In addition, the centroid of the selected 3-D front is rendered as a point in the scene. Fronts can now be manually tracked by selecting a front in successive time steps. To illustrate the temporal displacement of a front during tracking, the centroids of selected fronts are connected by lines. To enhance the visual distinction between different ensemble members, the colour of the line connecting the centroids varies between ensemble members, allowing them to be distinguished and compared. Figure 6.2 shows an example of interactive front selection.

SPLIT AND MERGE EVENT: During the selection process, it is possible to encounter front split events, which involve the splitting of a front that was once a continuous surface at the previous time step. These events can occur in two different ways. Firstly, a front split event may occur temporarily, where the separated fronts reconnect to form a continuous surface after a few time steps. Secondly, the front separation event can persist for all subsequent time steps, resulting in permanently separated fronts.

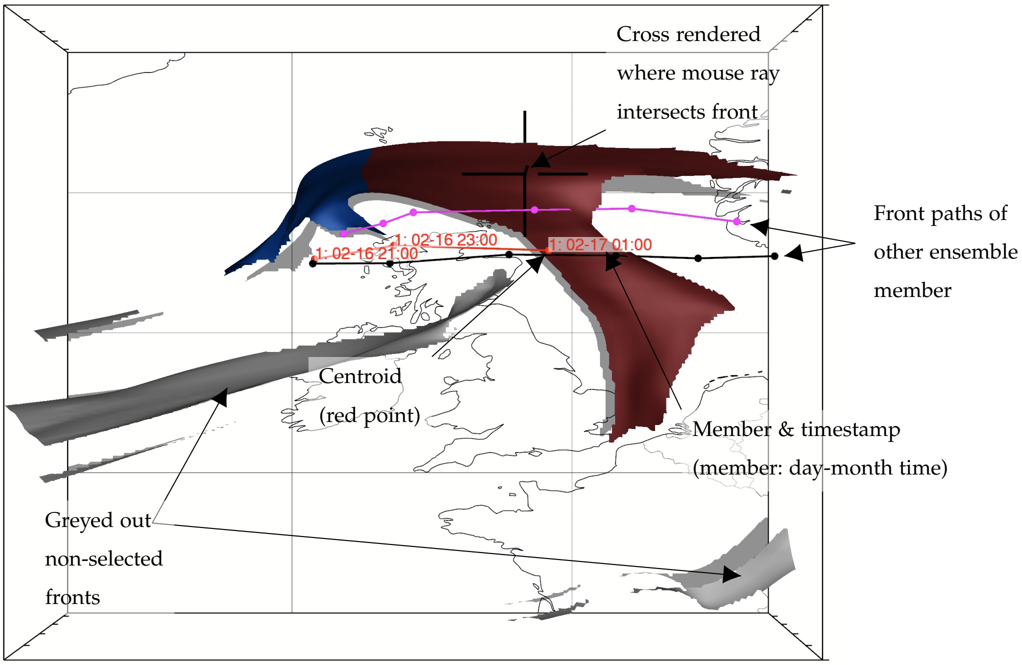


Figure 6.2: Example of the 3-D front selection process. Selected fronts remain in colour, unselected fronts are shaded light grey. The path lines show the evolution of the frontal centroids of the current and other ensemble members. The path lines of the current ensemble member are annotated with a timestamp.

To address this issue, the user has the option to select multiple fronts from the same time step. In this case, the split fronts are still treated as a single front when assigning frontal attributes. This allows the user to consider several divided fronts as a cohesive structure. Figure 6.3 shows an example of a split followed by a merge event.

6.1.2 *Characterising 3-D fronts by frontal attributes*

To characterise fronts based on their physical and geometric properties, the following frontal attributes are computed for each selected and tracked 3-D front:

Centroid x, y, z: The area-weighted centre of the selected front. The centroids are also rendered as dots during the selection process. The centroids of successive time steps from the same ensemble member are connected by lines, giving the user a visual feedback regarding the displacement of selected fronts.

Front area: The area of the selected front in square metres or square kilometres. To compute the area of a front, the original geographic coordinates of the triangles, with pressure as the height coordinate, are transformed to a Cartesian coordinate system using the coordinate transformation described in Section 2.4.1. The conversion from pressure altitude to metric altitude uses the geopotential as the geometric altitude, if available. If the geopotential

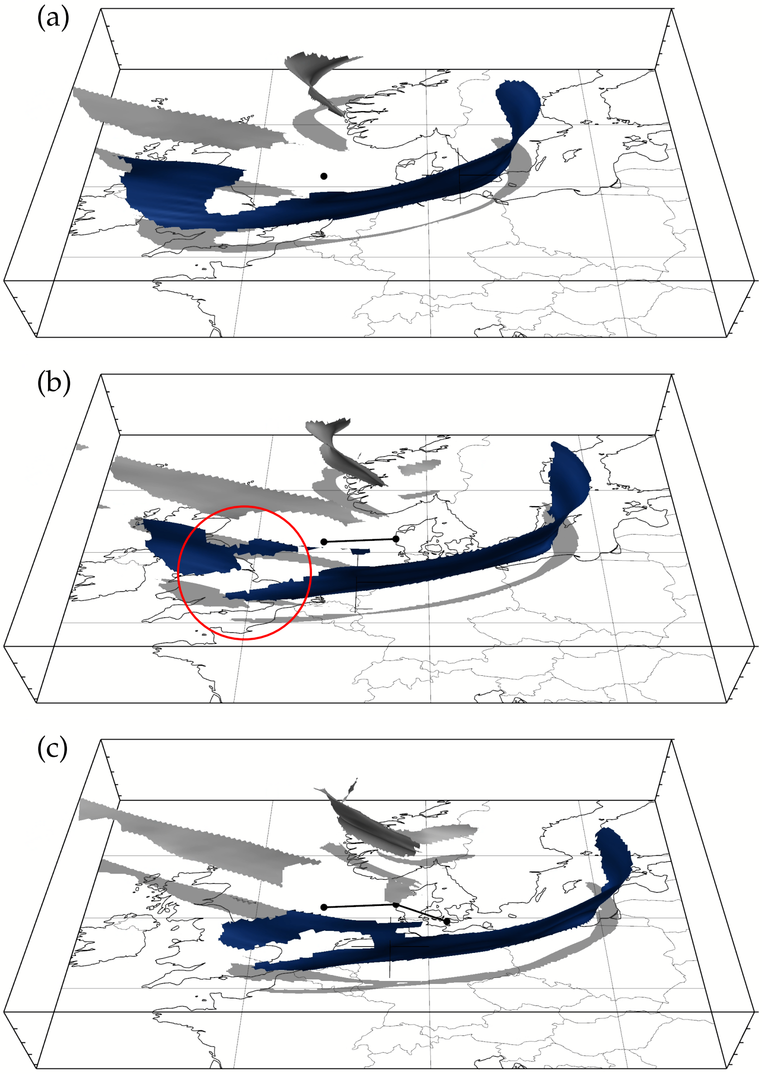


Figure 6.3: Example of a temporarily split cold front event during frontal tracking. (a) The cold front is compact and forms a contiguous surface. (b) The red circle marks the location of a cold front split. (c) The cold front is compact and contiguous again.

height is not available, the International Civil Aviation Organization (ICAO) Standard Atmosphere is used as an approximation for this conversion. The following equations are used to calculate the area of a triangle given by the points a_c, b_c, and c_c, where c indicates the Cartesian coordinate system:

$$\begin{aligned}\overrightarrow{AB} &= b_c - a_c \\ \overrightarrow{AC} &= c_c - a_c \\ Area &= 0.5 \cdot \|\overrightarrow{AB} \times \overrightarrow{AC}\|\end{aligned} \tag{6.1}$$

The total area of a front is the sum of the area of each triangle of the selected surface.

Mean and maximum TFP: The maximum and area-weighted mean TFP. The mean TFP is calculated as the area-weighted average over each triangle of a frontal surface.

Mean and maximum frontal strength: The maximum and area-weighted mean frontal strength. The mean frontal strength is calculated as the area-weighted average over each triangle of a frontal surface.

Mean, minimum negative and positive slope: The slope is defined in Section 3.4 as the ratio of the vertical extent, lower edge and upper edge of the front at a given location, to its horizontal extent, which is the horizontal component of the distance between the upper and lower edges of the front. To compute the slope of a front, the original geographical coordinates of the triangles, with pressure as the height coordinate, are transformed to a local Cartesian coordinate system using the coordinate transformation described in Section 2.4.1. Subsequently, and for each triangle of the frontal mesh, the Cartesian coordinate system is rotated so that the z-axis of the Cartesian coordinate system intersects one of the triangle's vertices. In the local Cartesian coordinate system, the horizontal s_h and the vertical s_v of a triangle are calculated. The slope is calculated as follows:

$$Slope = \left(\arccos\left(\frac{s_h}{\sqrt{s_h^2 + s_v^2}} \right) \times \frac{180}{\pi} - 90 \right) * -1 \tag{6.2}$$

The mean slope is calculated as an area-weighted average over each triangle of the frontal surface.

For each selected frontal surface, the frontal attributes are computed accordingly. The frontal attributes are stored in a dictionary-like structure and are computed for each selected time step and ensemble member. Due to some front splitting events, it may be necessary to select multiple frontal surfaces from the same ensemble member and time step. The frontal attributes of such multi-selected fronts are computed using area-weighted averages for mean values, and the minimum or maximum for minimum or maximum values.

Once all the fronts required for analysis have been selected and the frontal attributes have been calculated, the frontal attributes can be transferred to Python for further analysis. The data structure of the frontal attributes is based on the structure of the trajectories according to the NetCDF Climate and Forecast (CF) Metadata Conventions (short NetCDF conventions, Hassell et al. (2017)).

Data interfaces for frontal attributes

The frontal attributes are stored in a trajectory-like structure. Each trajectory corresponds to an ensemble member and each time step corresponds to a trajectory time step. The longitude, latitude and vertical coordinate represent the computed centroid of a selected frontal surface. Frontal attributes are represented as auxiliary trajectory variables.

Storing frontal attributes in a trajectory structure has the following advantages: The file format and structure of the file are known and defined by NetCDF conventions. It can be imported into other programmes as a trajectory file. This means that stored frontal attributes can be re-imported into Met.3D and other meteorological applications without further modification. This allows, for example, frontal attributes to be displayed along the frontal displacement paths (example shown in Figure 6.5e) using the multi-parameter trajectory display of Neuhauser et al. (2023). In addition, with the Python interface built into Met.3D, Python packages can be used directly to perform further analysis on frontal attributes and send the results back to Met.3D.

6.1.3 *Automated front tracking*

Manual tracking can be time-consuming, especially when analysing ensemble simulations. For example, tracking fronts for a 24-hour period in a ICON-EU-EPS forecast that includes 40 ensemble members would require the selection of 40 x 24 (=960) fronts. As this approach is impractical for ensemble analysis, I have developed a front-tracking system that automatically tracks a front of interest over a given period of time.

The automated front tracking process starts with a manual selection of a front at time step 0 (t), which serves as a reference front to be tracked over successive time steps. The subsequent automatic front tracking consists of two steps.

First, all fronts in the next time step ($t + 1$) of the NWP simulation are detected, and then fronts whose centroids (see Section 6.1.2) are located too far away from the reference front (t) are filtered out. Second, from the remaining fronts, the TFP distributions on each front surface are calculated and compared with the TFP distribution from the reference front at time step t. The front for which the distribution at time step $t + 1$ is most similar to the distribution at time step t is then identified as the target and serves as the reference front for the next time step ($t + 2$) of the NWP simulation. These two steps are repeated until front tracking is complete for the entire period.

The two steps of the front tracking are implemented as follows:

Distance-based filter: A distance-based filter is used to exclude fronts in the next time step whose centroid is unreasonably far from the centroid of the reference front. Based on the case study in Section 6.2, a centroid can vary by about 500 km between two successive time steps. This is caused by segments of the tracked front surface being disconnected or connected within a time step. As a result, the centroid can shift over a range of less than 500 km. In addition, the displacement speed of the front is also considered with 100 km to 150 km per hour. The true displacement of cyclones and thus also of fronts is notably slower than 150 km (cf. Nellikkattil et al., 2024).

Attribute distribution matching: From the remaining potential front matches at the next time step ($t+1$), the distributions of TFP of each remaining front candidate are compared with the distribution of the reference front. The distance between these distributions is measured using EMD, an algorithm that finds the minimum cost to transform one distribution into the other. A detailed description of the EMD algorithm can be found in Section 2.4.4. The front at time $t+1$ with the smallest EMD is then identified as the target front and used as the reference front for the next time step ($t+2$).

In Section 6.2, the feasibility of this approach is evaluated through a case study that examines different vertical grid spacings and time resolutions of NWP.

6.1.4 *Architecture and implementation details in Met.3D*

The above-proposed manual and automated front tracking is implemented into the interactive visualisation framework Met.3D. The implementation details in Met.3D, including changes to the front detection method introduced in Section 5.1, are outlined below. Four additional data pipeline modules are introduced to enable manual and automated front tracking and feature attribution computation.

1. **3-D front hard borders filter:** Implements hard filter thresholds to isolate 3-D fronts.
2. **3-D front orderer source:** Orders triangular frontal meshes into individual contiguous frontal surfaces.
3. **3-D front selector source:** Allows the manual selection of individual 3-D fronts.
4. **3-D front attribute source:** Derives frontal characteristics by computing frontal attributes.

Figure 6.4 shows the adapted front detection data pipeline to enable front selection and front characterisation by frontal attributes. Three additional data pipeline modules have been added to the processing data pipeline. The fourth module, the *front attribute source*, is not a pipeline module in the context of the data processing pipeline (see Section 2.2.1). This module is executed after the rendering process and does not feed back into the visualisation during the rendering process. However, as it is an essential component for feature-based attribute analysis and has the same structure as a normal pipeline module, it will be treated as part of the data processing pipeline in the following.

In the front detection pipeline of Section 5.1, the MC algorithm computes a triangular mesh of frontal candidates. The resulting triangular mesh consists of two datasets: First, a list of all vertex positions, where each vertex position is a 3-D vector containing its x, y, z coordinates. Second, a list of vertex indices, where each triangle contains a list of three vertex indices (i, j, k), identifying the three vertices that form the corresponding triangle. In this data structure, it is not known which triangles form a contiguous frontal surface. In order to select a single contiguous frontal surface, this information becomes relevant.

As mentioned above, four additional data pipeline modules are introduced, extending the data pipeline to allow feature-based ensemble analysis. The new pipeline modules are placed after the *front candidates source* module. The first module is the *front hard borders filter*, which filters 3-D front candidates according to TFP, frontal strength, and pressure thresholds. The second module is the *front orderer source*, which orders the 3-D triangular meshes into contiguous mesh batches. This pipeline module implements an algorithm to compute batches of triangles following the algorithm introduces in Section 2.4.2. A triangle batch is a data structure that describes a mesh of adjacent triangles, where each triangle shares at least one edge with another triangle. The third module is the *front selector source*, which allows the interactive selection of contiguous frontal surfaces (e.g. a warm front from a particular cyclone system). This module implements the Möller-Trumbore intersection algorithm introduced in Section 2.4.2 to select a single frontal surface. The fourth module is the *front attribute source*, which computes the attributes of the previously selected front.

For automatic front detection (introduced in Section 6.1.3), the distance-based filtering and TFP distribution matching are computed in Python, utilising the Met.3D Python interface.

6.1.5 *Example of front-tracking workflow*

First, fronts are computed regularly, but with hard filter thresholds to obtain sharp-edged frontal surfaces. Figure 6.5a shows the resulting fronts coloured according to warm and cold air advection. The fronts can then be further filtered according to warm air (warm front) and cold air (cold front) advection. In this example, warm fronts are filtered out, leaving only cold fronts. Once the warm fronts have been filtered out, the front of interest is selected and the front-tracking process begins. The selected front is highlighted, unselected fronts are grey (Figure 6.5b). Figure 6.5c shows the front tracking over a period of one day. The black annotated dots indicate the position of the centroid of the front at the selected time steps. The dots are connected by black lines, indicating the movement of the front over multiple time steps. In the case of front tracking over multiple ensemble members, multiple front paths are shown, each path representing the front evolution for one ensemble member (Figure 6.5d). Figure 6.5e shows an example visualisation of frontal attribute evolution using the multi-parameter trajectory display of Neuhauser et al. (2023). This can facilitate the analysis of frontal attributes and the identification of ensemble members with, for example, rapid increases or decreases in mean TFP, mean frontal strength, or mean frontal slope.

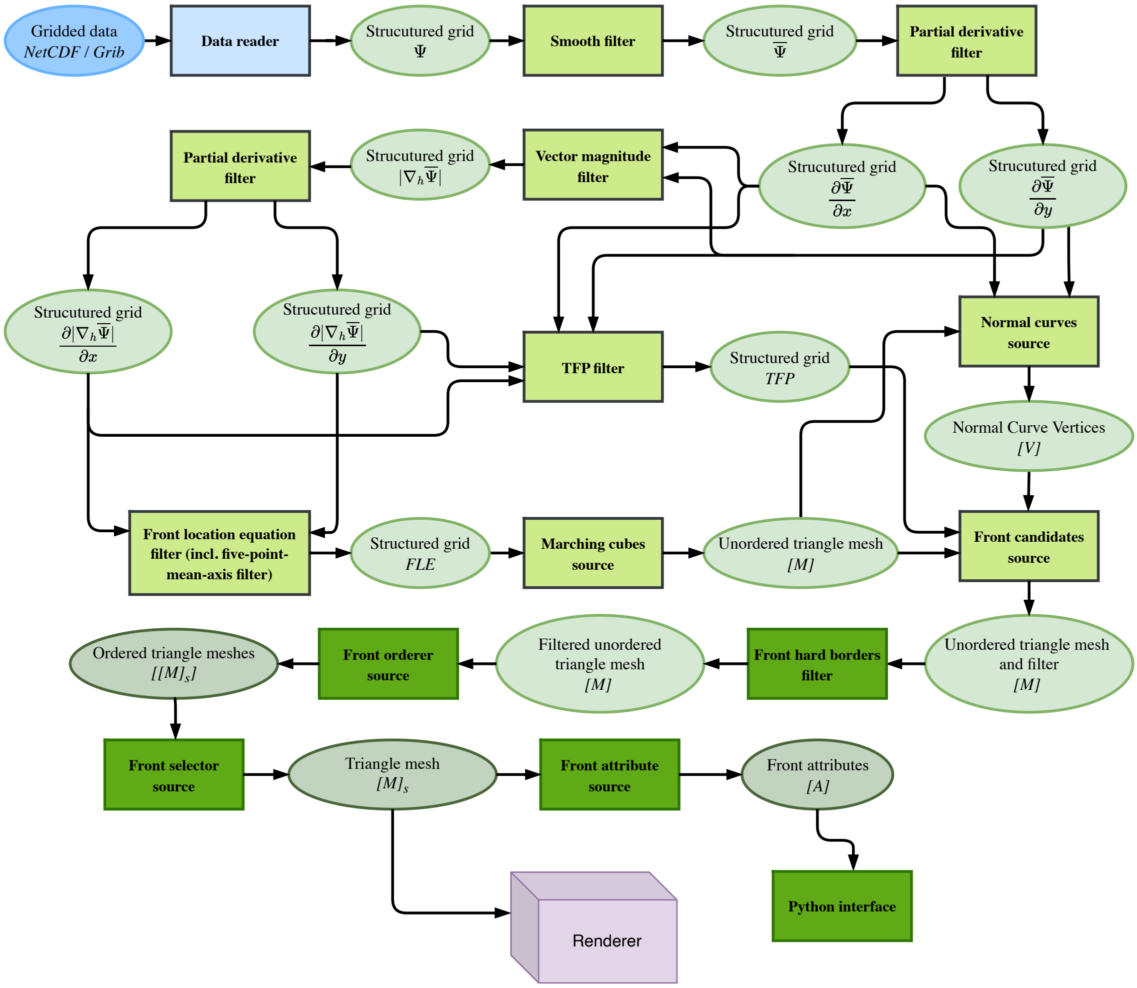


Figure 6.4: Adapted data pipeline for the feature selection framework. Additional pipeline modules and data input/output formats are illustrated in darker colours.

6.1.6 *Postprocessing of frontal attribute time series*

After successfully tracking frontal surfaces across different ensemble members and computing the frontal attribute time series, the aim is to provide analysis techniques for these time series of frontal attributes. For this, I developed a time series similarity analysis, which provides resources for exploring ensemble time series in Python.

The first objective of time series similarity analysis is to present ensemble time series data in a clear and comprehensive format. To achieve this goal, I chose to graphically represent time series in a horizon plot, which is specifically designed to effectively represent large one-dimensional data sets, such as time series, with high detail and precision in a compact manner (Saito et al., 2005). For a detailed description of the horizon plot, see Section 2.3.2.

The second objective of time series similarity analysis is to order and cluster ensemble members based on time series of frontal attributes. To achieve this,

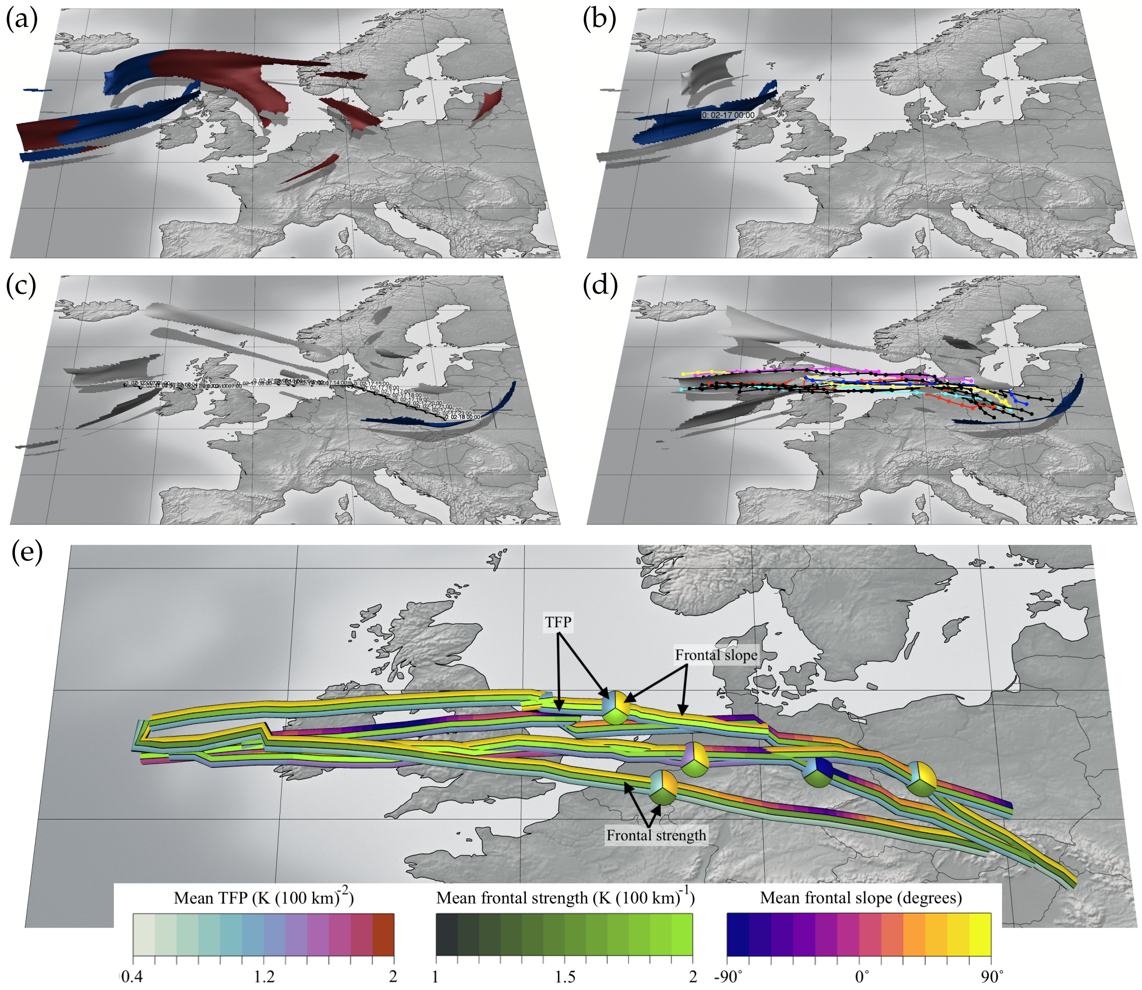


Figure 6.5: Example of the interactive front selection workflow. (a) Frontal surfaces filtered with hard thresholds and coloured according to front type (blue = cold front, red = warm front). (b) Warm fronts are filtered out, the selected cold front is highlighted in blue, and non-selected cold fronts are grey. The text indicates the ensemble member, date and the time of the selected front. (c) 24-hour front track together with the cold front of the last tracking time step. (d) Front tracks of multiple ensemble members and the cold front of the currently selected member. (e) Example of front tracks of five selected ensemble members. Front tracks are presented using the multi-parameter trajectory display of Neuhauser et al. (2023). The band of each front track shows the evolution of the mean TFP, mean frontal strength and mean frontal slope attributes, including a highlighted time step annotated as a pie chart.

similarities are assessed by measuring the distances between the time series data. Three different metrics are used to calculate these distances: Euclidean distance, DTW and soft-DTW. A detailed description of these distance metrics is given in Section 2.4.4. Based on these calculated distances, the ensemble members can be sorted according to a reference member. Additionally, these distance metrics are used to cluster the ensemble members using k-means clustering (MacQueen, 1967). The different distance metrics for sorting and clustering ensemble attribute time series are evaluated in a case study in Section 6.3 and Section 6.3.2.

To test the sensitivity of the clusters between different frontal attributes as well as different frontal filter settings, I follow the idea of (Kumpf et al., 2018) and use the Rand index (Rand, 1971). The Rand index is a statistical measure used to evaluate the similarity or agreement between two data clustering results or partitions. A detailed description of the Rand index can be found in Section 2.4.3. The full source code of this time series analysis can be found in Beckert et al. (2024).

6.2 CASE STUDY: FRONTAL ATTRIBUTE DISTRIBUTIONS AND TRACKING

Using Storm Friederike as a case study, I analyse the capabilities and limitations of the automated front-tracking method presented in Section 6.1. For this evaluation, I use NWP data from the ECMWF ENS forecast initialised on 16 January 2018, 12:00 UTC, with 51 ensemble members. The front tracking is evaluated from 16 January 2018, 12:00 UTC until 19 January 2018, 12:00 UTC. Afterwards, the cyclone and the fronts are almost completely dissipated (cf. Figure 5.13). All objectively detected 3-D fronts for this analysis are defined between 900 hPa and 600 hPa, with a hard TFP filter threshold of 0.4 $K(100km)^2$ and a hard frontal strength filter threshold of 1.0 $K(100km)^1$.
I concentrate my analysis on four key aspects:

1. How do the distributions of frontal attributes compare at different times for a tracked front at?
2. What temporal resolution is required for successful automated front tracking?
3. Can automated front tracking be used in NWP simulations with limited vertical grid spacing?
4. How similar are the frontal attributes from the same front but from different ensemble members, at the initial time steps of a NWP simulation?

6.2.1 *Distribution of frontals attributes*

The distribution of three frontal attributes from the cold front of Cyclone Friederike is visualised as violin plots (for more details on violin plots, see Section2.3.1) in Figure 6.6 for a tracking period of 73 hours. This figure provides an overview of how these attribute distributions develop over time for a tracked frontal surface.

The thermal-based frontal attribute distributions (TFP and frontal strength) follow a similar pattern, showing a slight increase from the initial time step on 16 January at 12:00 UTC for 24 hours, followed by a decrease over the next 48 hours. The evolution of these distributions is characterised by gradual changes with no abrupt differences between successive time steps, indicating a smooth transition.

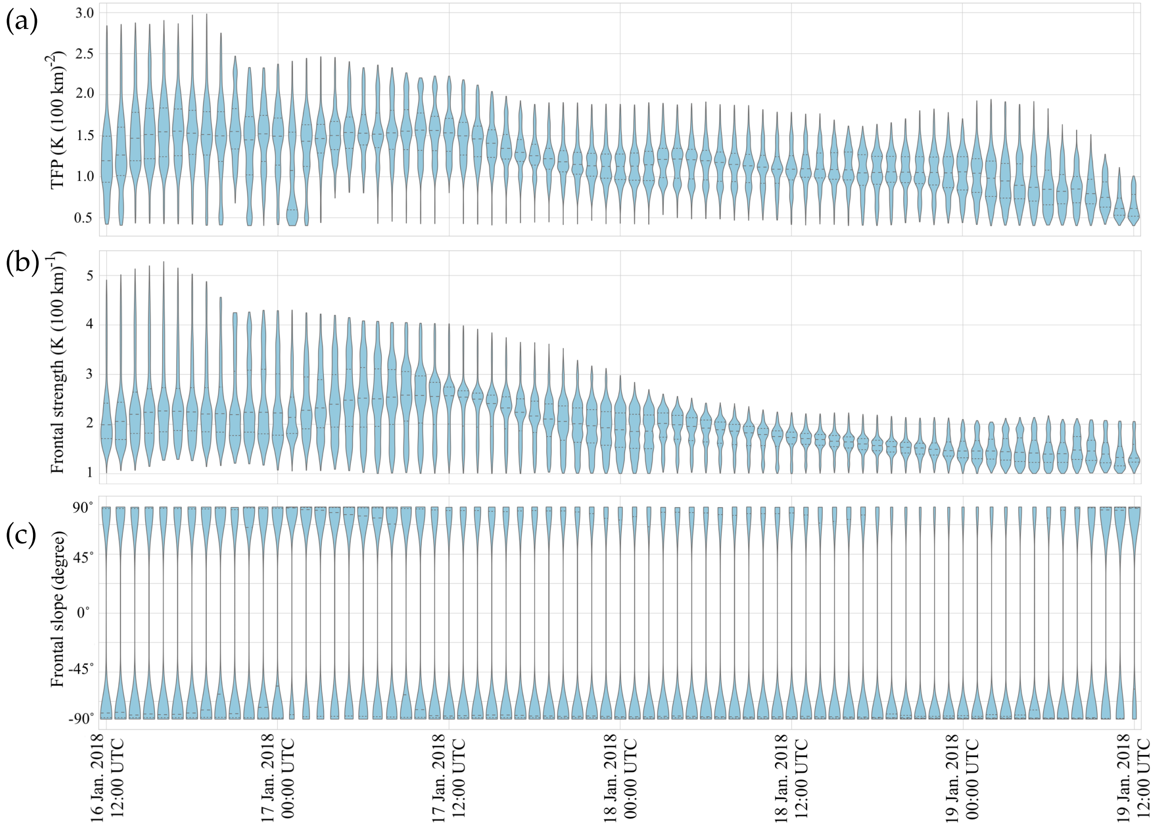


Figure 6.6: Distribution of (a) TFP, (b) frontal strength, and (c) frontal slope of the tracked 3-D cold front of Storm Friederike, shown as a violin plot. The front is tracked in ECMWF ENS forecast, initialised on 16 January 2018, 12:00 UTC, with 91 vertical levels defined on hybrid sigma-pressure coordinates. The distributions are shown for 73 time steps of the tracked front and for ensemble member 0. While the distributions of TFP and frontal strength have similar characteristics, the frontal slope has a dipole structure. Individual parts of the front are either steeply sloped forwards (close to 90°) or backwards (close to -90°).

In contrast, the slope attribute distributions show a dipole pattern, with parts of the front either positively or negatively tilted near +90° and -90°. While the interpretation of these attributes may not be immediately obvious when looking at 3-D representations of fronts, it is important to note that the vertical axis is significantly stretched in all 3-D visualisations, causing the fronts to appear less tilted and more vertical. As shown in Figure 3.6, the typical angles for a positive frontal slope are in the range of 89.4° to 89.8°.

6.2.2 *Time Resolution*

The automatic front-tracking algorithm developed in section 6.1.3 compares the EMD of the TFP distributions between a reference front and all potential front matches. The front from the potential matches with the lowest EMD is selected as the target front. Therefore, successful automatic front tracking requires, at most,

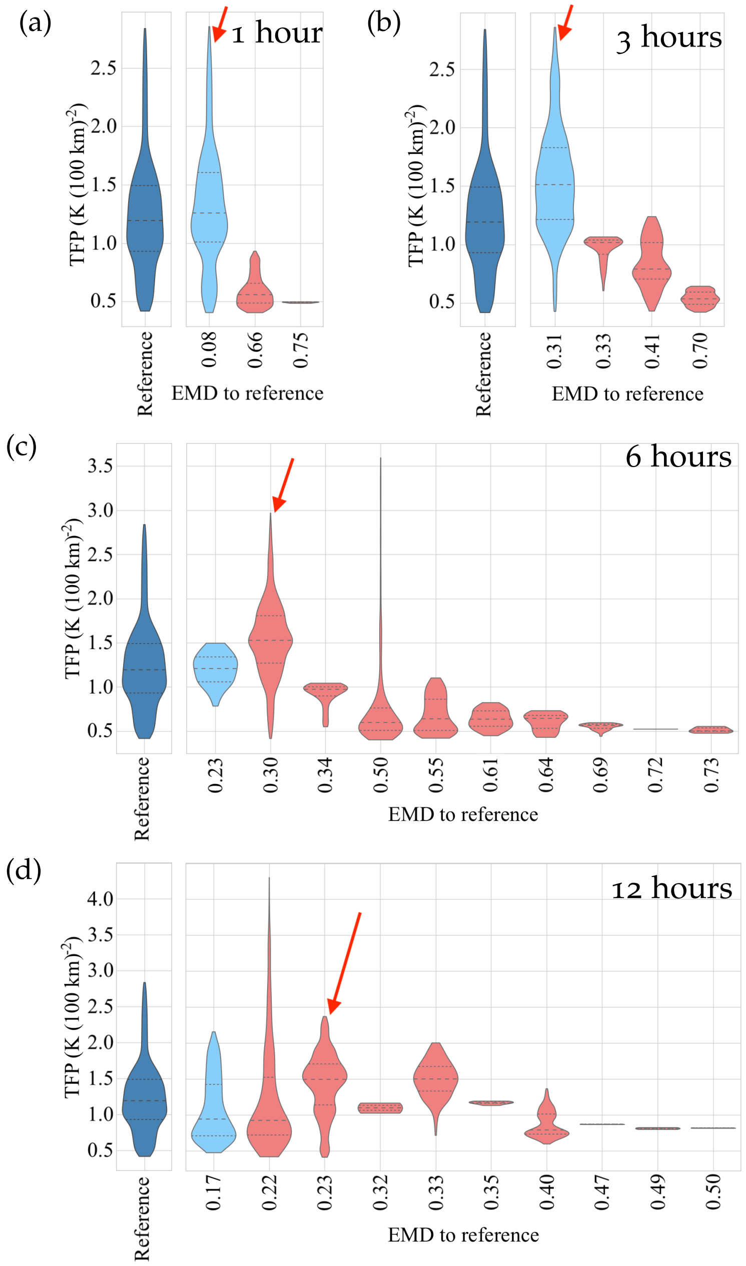


Figure 6.7: Distribution of TFP of the reference front and potential front matches for different time intervals. Manually selected target TFP distributions are marked with a red arrow. For (a) 1 hour and (b) 3 hour intervals, the smallest EMD distance between the TFP distributions of the reference and target fronts is correct and automatic front tracking is successful. For (c) 6 hour and (d) 12 hour intervals, the minimum EMD distance between the TFP distribution of the reference does not represent the correct target front and automatic front tracking fails. Figure 6.8 shows the corresponding 3-D fronts. Fronts detected in ensemble member 0 of ECMWF ENS forecast, initialised on 16 January 2018, 12:00 UTC, with 91 vertical levels defined on hybrid sigma-pressure coordinates.

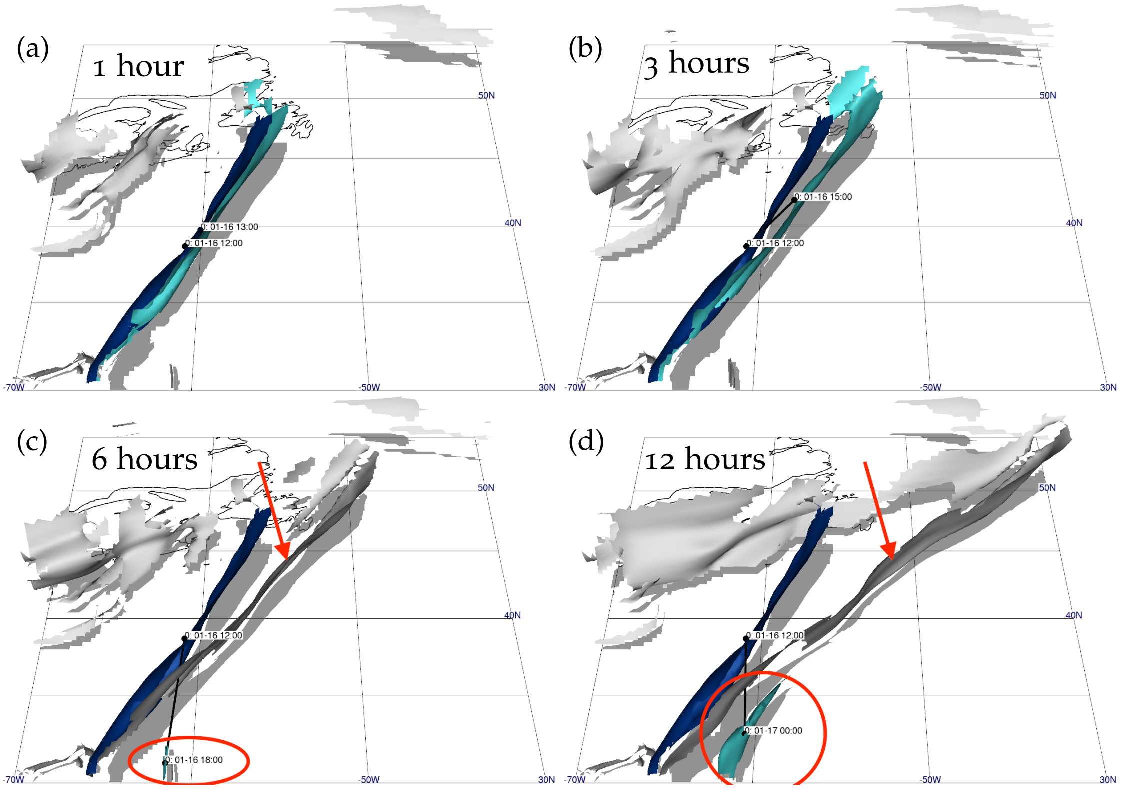


Figure 6.8: Illustration of automated front tracking for different time intervals. The reference front is shown in dark blue, the target front obtained by automated tracking is shown in light blue. As shown in Figure 6.7, automated tracking is successful for time intervals of (a) 1 and (b) 3 hours. Tracking fails for time intervals of (c) 6 and (d) 12 hours. The manually selected target front is marked with a red arrow and the front selected by automated front tracking is marked with a red circle in (c) and (d), where automated front tracking fails. All fronts are detected in ensemble member 0 of ECMWF ENS forecast, initialised on 16 January 2018, 12:00 UTC, with 91 vertical levels defined on hybrid sigma-pressure coordinates.

gradual changes in the distribution of TFP between successive time steps, so that the EMD is low. As shown in Figure 6.6, the development of TFP distributions is characterised by gradual changes between successive time steps. Figure 6.7 quantifies these changes in EMD between possible front matches and the reference front for different time intervals. For a time interval of 1 hour and 3 hours (Figure 6.7a and b), the front with the lowest EMD corresponds to the manually selected target front (marked with a red arrow). All other fronts have a higher EMD. However, for the higher time intervals of 6 hours and 12 hours (Figure 6.7c and d), the front with the lowest EMD does not correspond to the manually selected target front (marked with a red arrow). In these cases, the automatic front-tracking algorithm does not agree with the manually selected target front.

Figure 6.8 shows the 3-D structure of the reference front, the front with the smallest EMD to the reference front, and the manually selected target front. For a time interval of 1 hour (Figure 6.8a) and 3 hours (Figure 6.8b), the manually selected target and the target front selected by automatic front tracking are identical. For 6 hours (Figure 6.8c) and 12 hours (Figure 6.8d), the front selected by the automatic front-tracking algorithm (marked with a red circle) differs from the manually selected front (marked with a red arrow).

In the case presented, the automatic front-tracking algorithm successfully tracks fronts when the time interval between consecutive time steps is 3 hours or less. For larger time intervals (6 and 12 hours), the front-tracking algorithm selects unreasonably small fronts. A possible factor for this behaviour is the normalisation of distributions by the EMD algorithm before determining the minimum cost of transforming one distribution into another. This means that the size of the 3-D frontal surface is not taken into account. For tracking fronts over successive time steps greater than 3 hours, it may be beneficial to implement an additional pre-filter based on the size of the 3-D frontal surface to eliminate unreasonably small frontal surfaces.

6.2.3 *Vertical Resolution*

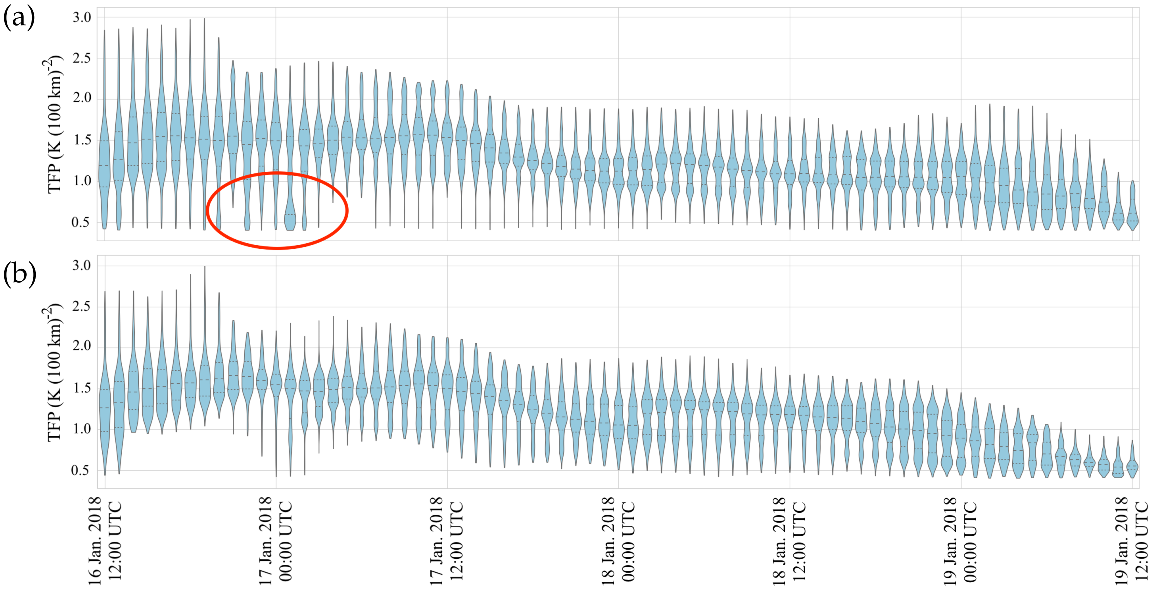


Figure 6.9: Distribution of (a) TFP on model level and (b) on pressure level on a frontal surface, shown as violin plots. The front is tracked in the ECMWF ENS forecast, initialised on 16 January 2018, 12:00 UTC. Model levels consist of 91 vertical levels defined on hybrid sigma-pressure coordinates. Pressure levels consist of 12 vertical levels defined on pressure coordinates. Distributions are shown for 73 time steps of the tracked front and for ensemble member 0. The TFP distributions in both model levels (a) and pressure levels (b) have a similar structure. However, the distributions in the model level data show an anomaly around 17 January 2018, 00:00 UTC, with an accumulation of low TFP values (red circle).

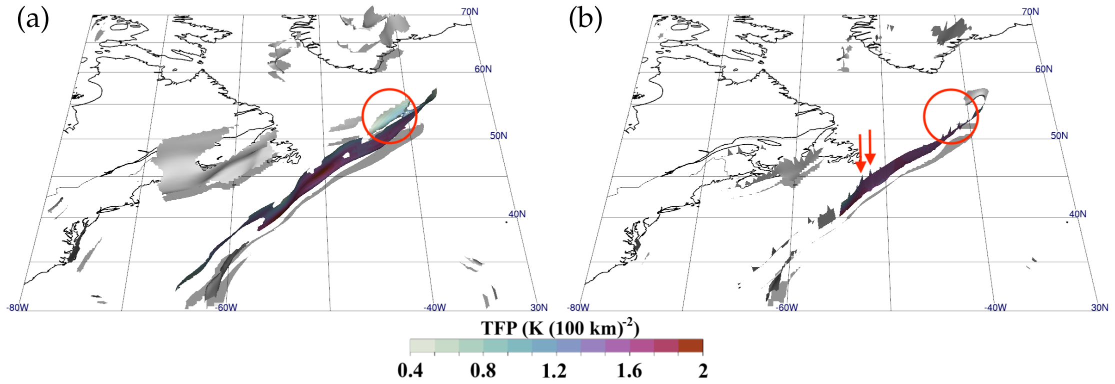


Figure 6.10: Comparison of fronts detected in (a) model level and (b) pressure level data on 17 January 2018, using the same data as used in Figure 6.9. The red circle indicates a part of the frontal surface with low TFP values in the model level data, which is not present in the pressure level data. This leads to the anomalies of low TFP in the model level data shown in Figure 6.9. Red arrows in (b) point to "spikes" in the frontal surface detected in the pressure level data due to the low vertical grid spacing.

To evaluate the effect of different vertical grid spacing on frontal attributes, Figure 6.9 shows the evolution of the TFP distribution over a tracked front for two different vertical resolutions. As discussed in Section 6.1.3, the TFP distribution plays a central role in front tracking by identifying the front in the next time step. However, in this case study, the full ensemble data of the ECMWF ENS simulation is only available for 12 vertical pressure levels. Can the automatic front detection algorithm successfully track fronts in data with low vertical resolution, and how comparable are the derived frontal attributes from the tracked front in data with low and higher vertical grid spacing?

Figure 6.9 shows that for both scenarios, full vertical grid spacing (91 vertical levels on hybrid sigma-pressure coordinates) and reduced vertical grid spacing (12 levels on pressure coordinates), the pattern and characteristics of the TFP distributions are analogous, although the two distributions show differences when examined in detail. For example, in Figure 6.9a, TFP distributions of fronts detected in model level data show an anomaly around 17 January 2018, 00:00 UTC, characterised by an aggregation of low TFP values (marked with a red circle). This anomaly is not present in TFP distributions of fronts detected in the pressure level data. Comparing the structure of the 3-D fronts between model and pressure level data, shown in Figure 6.10, the front in the model level data includes a subset of the frontal surface with low TFP values (marked with a red circle) that is not present in the pressure level data. This results in the low TFP anomalies shown in Figure 6.9a in the model level data. Due to the reduced vertical resolution in the pressure level data, the fronts are smaller in vertical extent and are characterised by "spikes" at the top and bottom of the front (marked with red arrows in Figure 6.10b). Although a lower vertical resolution preserves the basic characteristics of the TFP distributions on the 3-D front, details may be lost.

6.2.4 *Ensemble Tracking*

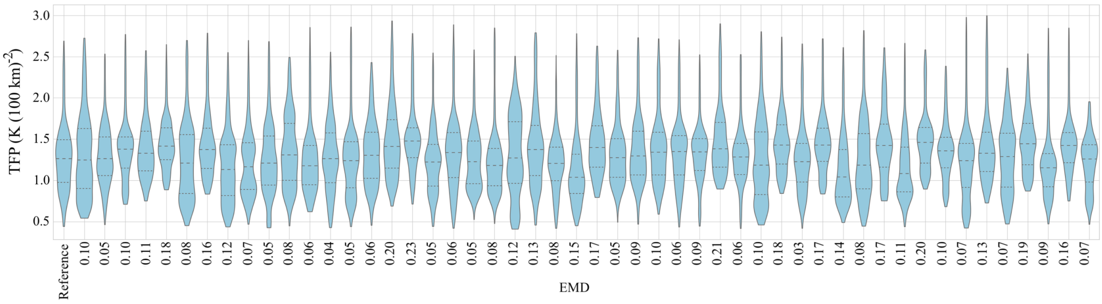


Figure 6.11: Distribution of TFP detected in the same front across all 51 ensemble members at the simulation initialisation time step. Annotated on the x-axis is the EMD relative to the reference ensemble member (member 0) for each TFP distribution. 3-D fronts detected in ECMWF ENS forecast data, initialised on 16 January 2018, 12:00 UTC, with 12 vertical levels defined on pressure coordinates.

For front tracking in ensemble simulations, it may be useful to select only one front from the first time step of an ensemble member and track the front across all ensemble members. To do this, I evaluate the differences in the EMD of the TFP distributions of the first time step for a selected front across all ensemble members. If the EMD is small, it is expected that selecting a single front from one ensemble member would be sufficient to identify the same front in all other ensemble members. Figure 6.11 shows the TFP distributions over the same front but different ensemble members and the EMD to a reference front. The distributions and EMD relative to the reference front show only small differences, with the highest EMD being 0.23 (cf. Figure 6.7). This suggests that it is reasonable that frontal tracking could be initialised based on only one specific front at the initialisation time step of the simulation. This front is then identified in all other ensemble members and can be tracked over the desired time steps.

6.3 CASE STUDY: FEATURE-BASED TIME SERIES SIMILARITIES AND CLUSTERING

Computer simulations are used to study complex real-world phenomena. The output of simulations is influenced by a number of parameters, such as initial conditions and the parametrisation of physical processes. A single run is not sufficient to account for such uncertainties in the configuration. Therefore, scientists often employ ensemble simulations with slightly different model configurations (Wang et al., 2019), as seen in ensemble weather prediction (see Section 3.2).

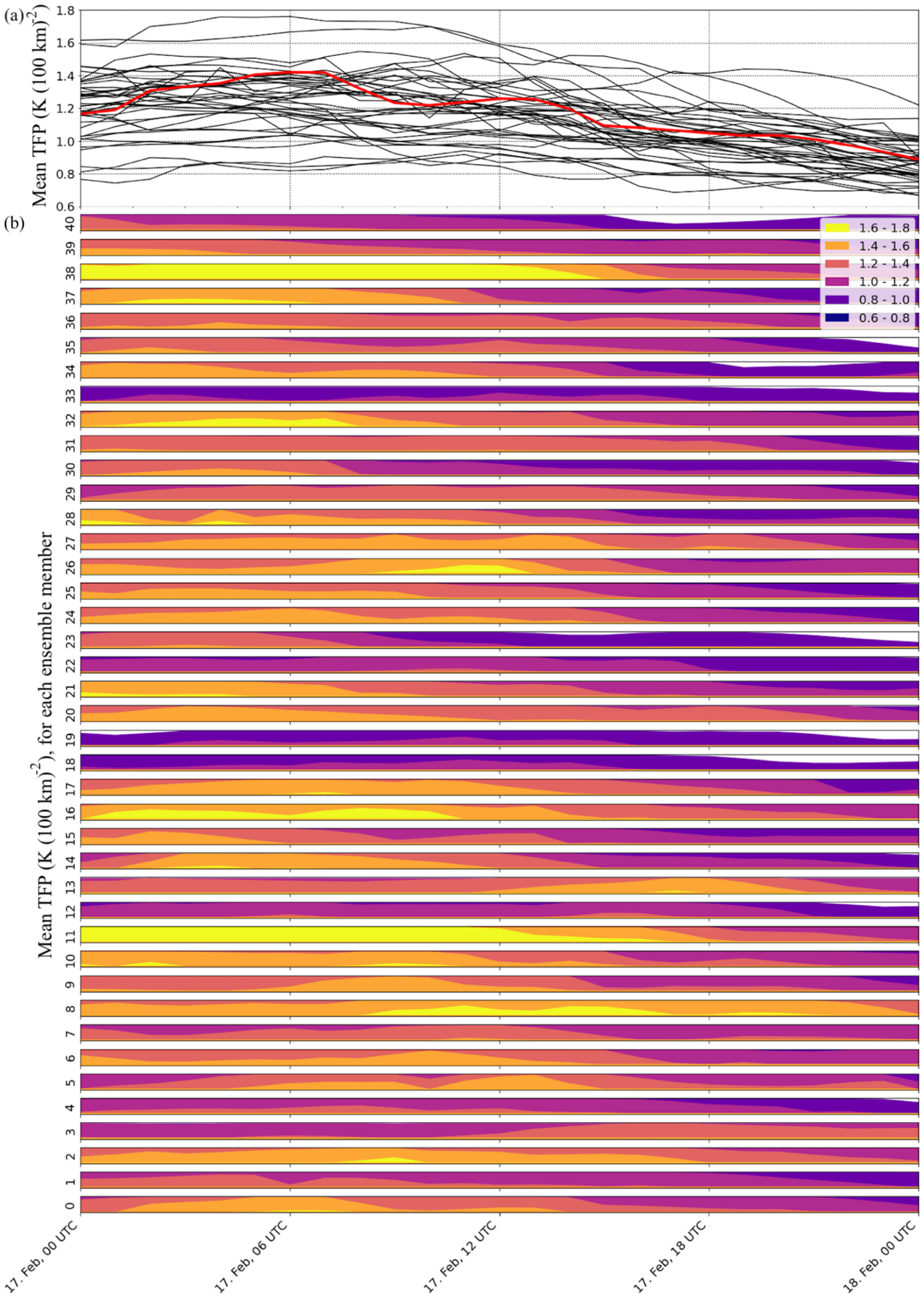


Figure 6.12: (a) Spaghetti plot of mean TFP frontal attribute development over a 25-hour period. Each black line represents an ensemble member, the red line represents the reference member (member 0). (b) Horizon plot of *TFP mean* frontal attribute time series. Each row represents an ensemble member. The frontal attributes are computed from tracked 3-D cold fronts of Cyclone Otto using the DWD ICON-EU-ENS forecast, initialised on 16 February 2023 at 00:00 UTC.

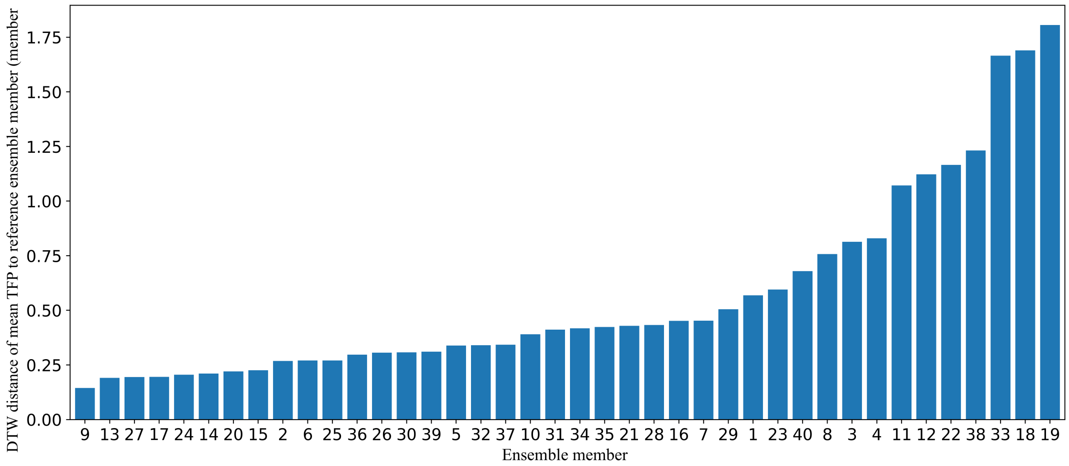


Figure 6.13: DTW distances of each ensemble member to the reference member (member 0). Ensemble members are sorted by distance in ascending order.

Ensemble simulations generate large amounts of data, making the analysis of such complex data challenging, and visualisation plays a crucial role in the analysis (Wang et al., 2019). Visualisation of uncertainties and clustering in ensemble simulations facilitates the analysis of these large datasets to identify trends, clusters, and outliers. Various visualisation approaches have been proposed in the literature to characterise uncertainties of ensemble simulations in 2-D and 3-D, such as contour box plots (Whitaker et al., 2013) and streamline visualisation to represent uncertainties in vector fields (Ferstl et al., 2016b). For the representation of uncertainty in 3-D surfaces, (Pfaffelmoser et al., 2011) presents a direct volume rendering approach that visualises the variability of isosurfaces in colour-coded form. Further advances in the visualisation of uncertainties for NWP ensemble simulations have been made by (Kumpf et al., 2018). Their approach focuses on the analysis of cluster robustness and representativeness. The recently published work by Chaves-de-Plaza et al. (2024) focuses on the visualisation of outliers in contour ensemble data. However, to my knowledge, feature-based time series similarity ordering and clustering of 3-D atmospheric features, such as objectively detected 3-D fronts, has not yet been proposed in the literature.

In the following, I present a case study for my ensemble ordering and clustering approach presented in 6.1. This approach is based on objectively identified 3-D fronts and derived frontal attributes. The objective of the feature-based time series similarity analysis is to compare time series of feature attributes and derive feature-based ensemble clusters.

For the following case study, I use the ICON-EU-EPS and ICON-EU NWP forecasts, initialised on 16 February 2023, 00:00 UTC, where ensemble member 0 refers to the deterministic ICON-EU forecast and ensemble members 1 to 40 refer to the ensemble members of the ICON-EU-EPS forecast. For a detailed description of the datasets and the weather situation, see section 4.3. To derive feature attributes of 3-D fronts, the cold front of cyclone *Otto* is tracked for 25 hours starting on 17 February 2023, 00:00 UTC. This period is chosen because the ICON-EU forecast is a limited-area simulation and the cold front does not enter the simulation area of the ICON-EU forecast until 17 February 2023, 00:00 UTC. After 18 February 2023, 00:00 UTC, the front dissipates. Unless explicitly stated otherwise, all objectively detected 3-D fronts for this analysis are defined between 850 hPa and 500 hPa with a hard TFP filter threshold of 0.4 $K(100km)^2$ and a hard frontal strength filter threshold of 1.0 $K(100km)^1$.

6.3.1 *Feature-based time series similarities*

Figure 6.12a shows the ensemble time series of the mean TFP frontal attribute. The time series spaghetti plot provides an overview of the overall development and the ensemble spread. However, extracting specific information about the evolution of a particular ensemble member relative to others can be challenging, as this plot is cluttered with the superimposed lines of individual ensemble members. In addition, there is no assignment of lines to ensemble members. To provide additional information about the evolution of individual time series, Figure 6.12b shows the frontal attribute time series in a horizon plot, with each ensemble member represented in a row. The composition of the horizon plot is described in detail in Section 2.3.2. The horizon plot facilitates the identification of differences between ensemble members. This allows a precise assessment of how the frontal attribute of one member evolves relative to others. For example, compared to members 18, 19 and 33, members 11 and 38 have high mean TFP values throughout the period.

As shown, the horizon plot can be used to visually identify differences between ensemble members. To quantify the differences between ensemble members and to quantitatively compare the time series differences, DTW is used to compute the similarity between a reference member time series and all other ensemble member time series of a selected frontal attribute. The DTW algorithm is described in section 2.4.4. The advantage of using DTW instead of e.g. Euclidean distance is that it takes into account temporal distortions in time series. For example, if the mean TFP in the reference member at time step t and in the compared ensemble member at time step $t+2$ change in a similar way, the distance between these time series calculated using DTW instead of Euclidean distance would be significantly smaller.

A user-defined reference member is used as the "ground truth". For each ensemble member, the distance to the reference member is calculated using DTW. Here the reference member is ensemble 0. Figure 6.13 shows the distances between the ensemble members and the reference member. As already mentioned after the visual interpretation of Figure 6.12b, ensemble members 11, 18, 19, 33 and 38 are among the most different ensemble members compared to the reference member 0. Figure 6.20 shows a visual representation of the tracked cold front of reference member 0, member 11 and member 18.

Using the calculated distance between the ensemble members and the reference member, the horizon plot is rearranged, and the ensemble members are ordered

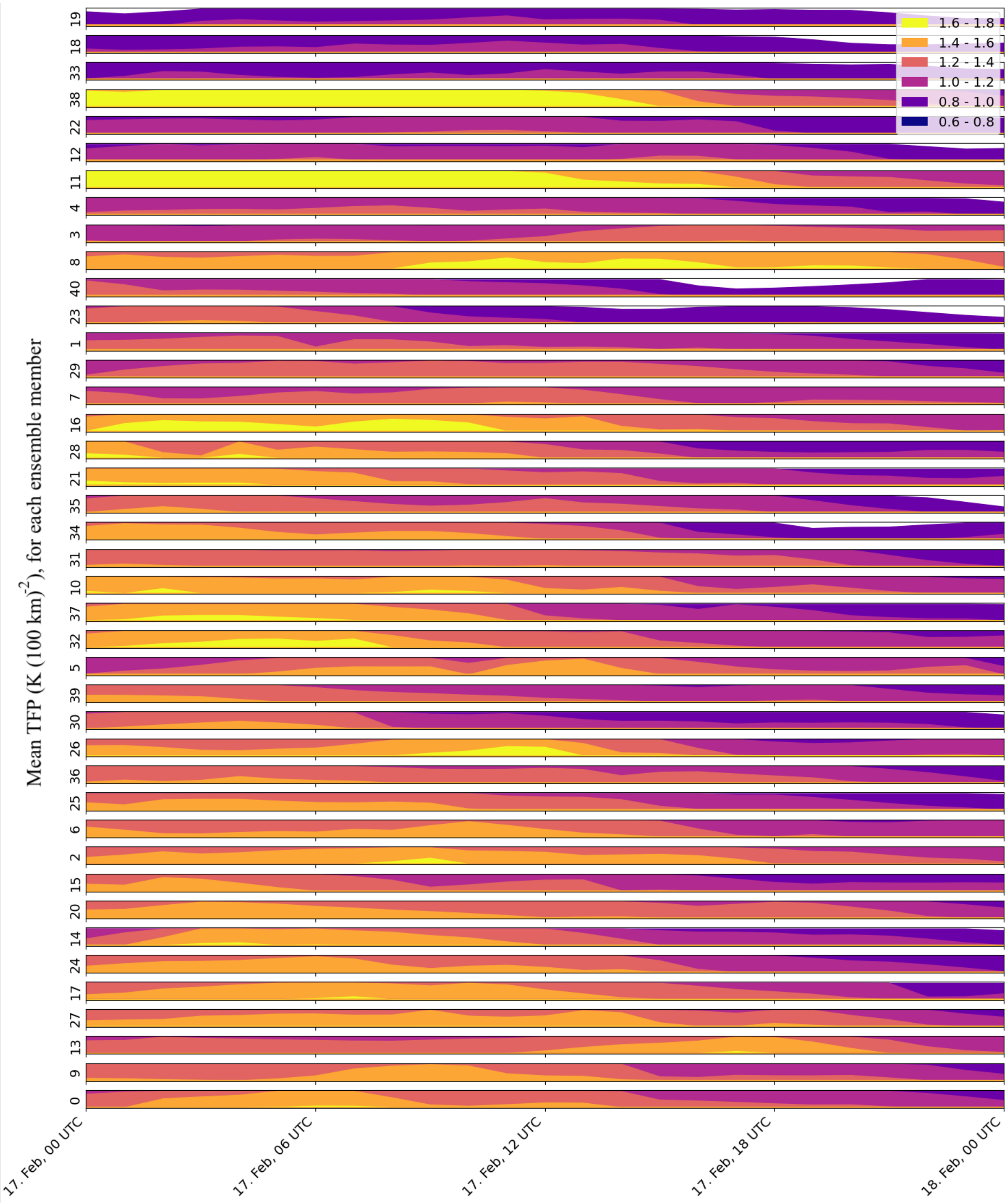


Figure 6.14: Horizon plot of the mean TFP frontal attribute. Ensemble members are sorted by similarity to the reference member (member 0) using DTW, from most similar member (bottom) to most different member (top).

according to their similarity to the reference member. Figure 6.14 shows the horizon plot with rearranged ensemble members. Ensemble members with similar characteristics to the reference member are at the bottom of the plot. The further up the plot the ensemble member is, the more it differs from the reference member. For ensemble members with similar characteristics to the reference member, the mean TFP first increases slightly and then decreases until the end of the analysed period.

6.3.2 *Feature-based time series clustering*

The presented time series clustering builds on 3-D frontal attributes and the time series distance measures previously described and used in Section 6.3.1. The aim is to categorise ensemble members based on the time series feature attributes and to explore potential similarity patterns and clusters in a case study.

Following (Kumpf et al., 2018), the clustering itself is done with k-means, which is described in detail in section 2.4.3. To apply k-means to time series, the distance between time series must be calculated using a distance metric. In addition to the previously applied distance metric (DTW), two other distance metrics are evaluated: Euclidean distance and soft-DTW (for more details on these distance metrics, see Section 2.4.4).

Before applying the k-means algorithm, the number of desired clusters must be specified. Here, I follow Shahapure and Nicholas (2020) and use the silhouette score to determine the optimal number of clusters. The silhouette score measures how well clusters are separated relative to their internal coherence. A brief description of the silhouette score and the algorithm is given in Section 2.4.3. Based on the silhouette score, I evaluate the optimal number of clusters and which distance metric performs best.

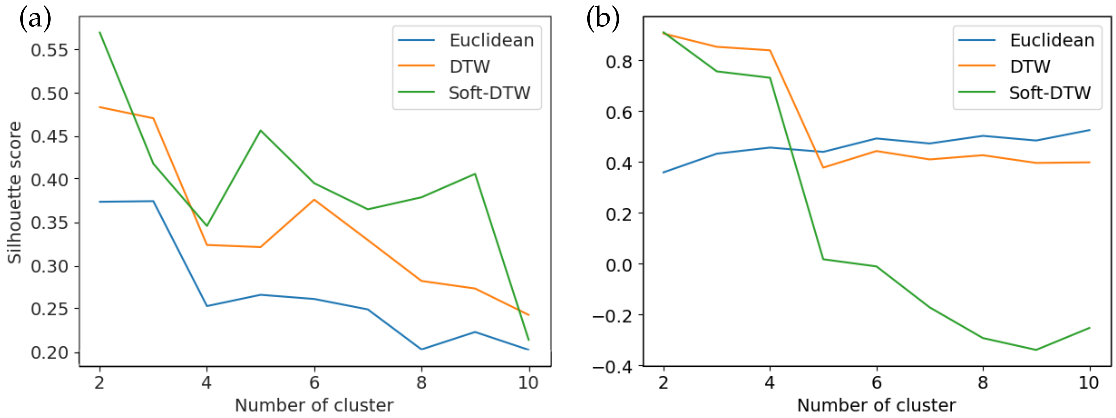


Figure 6.15: Silhouette score for k-means clustering of (a) mean TFP and (b) mean frontal slope for different numbers of clusters. The silhouette score shows the highest intra-cluster coherence and inter-cluster separation for the DTW and soft-DTW distance function and 2 to 4 clusters. Overall, the silhouette score is significantly lower for mean TFP compared to mean frontal slope.

Figure 6.15 shows the silhouette scores for k-means clustering for 2 to 10 cluster centres of the frontal attributes mean TFP (Figure 6.15a) and mean frontal slope (Figure 6.15b). For both frontal attributes, the silhouette score is highest for a small number of cluster centres (2 to 4) and for DTW and soft-DTW. This means that intra-cluster coherence and inter-cluster separation are highest for a small number of clusters and when using DTW or soft-DTW. At least, for a small number of cluster centres, the Euclidean distance is outperformed by DTW or soft-DTW.

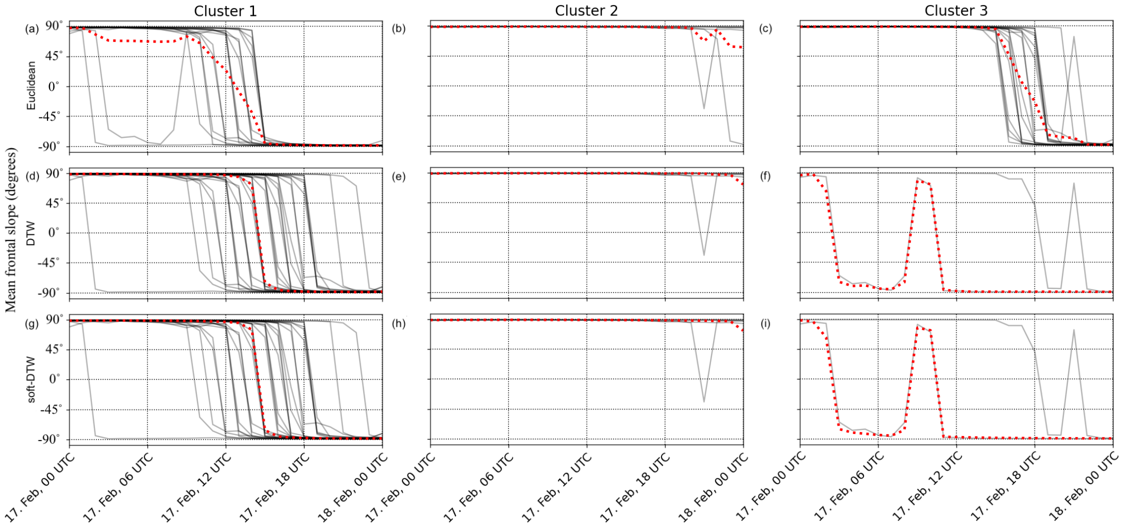


Figure 6.16: Spaghetti plots of clustered frontal attribute mean frontal slope using the k-means algorithm for 3 cluster centres. The figure shows cluster results for different distance metrics: (a) - (c) Euclidean, (d) - (f) DTW, and (g) - (i) soft-DTW. Red lines highlight the cluster centres.

Figure 6.16 shows a spaghetti plot of an example of k-means clustering for the mean frontal slope attribute using 3 cluster centres and the three different distance metrics. For the Euclidean distance, cluster 1 (17 members) and cluster 3 (18 members) represent the dominant time series structure. For DTW and soft-DTW, only cluster 1 (34 members) represents the dominant time series structure. For cluster 1 and all distance metrics, the mean frontal slope attribute is positive (90°) in the early stages and transitions sharply to a negative slope (-90°) around 14 UTC. Note that the mean frontal slope is either strongly positive or strongly negative, i.e. the front is predominantly positively or negatively sloped. Cluster 2 also represents the same structure for all three distance metrics, where all ensemble members are characterised by a positive frontal slope over the whole time. Cluster 3 represents a cluster similar to cluster 1 for the Euclidean distance metric, but with a later sharp transition from positive to negative slope. While the different distance metrics did not lead to significantly different clusters for clusters 1 and 2, for cluster 3 the cluster centre using soft-DTW and DTW differs significantly from that using Euclidean distance. For DTW and soft-DTW, the cluster centre has three sharp transitions, from positive to negative, from negative to positive, and from positive to negative again. This pattern represents the pattern of the frontal slope attribute of the two ensemble members in this cluster. Both ensemble members follow this distinct pattern, but at different times. Unlike using the Euclidean distance, DTW and soft-DTW detect these patterns and group them into a separate cluster. This also reflects the characteristics of DTW and soft-DTW because, as mentioned above, these distance metrics consider temporal shifts differently than the Euclidean distance metric. For DTW and soft-DTW, time series with the same characteristics, e.g. an abrupt transition from positive to negative values, but shifted in time, are still considered similar (small distance).

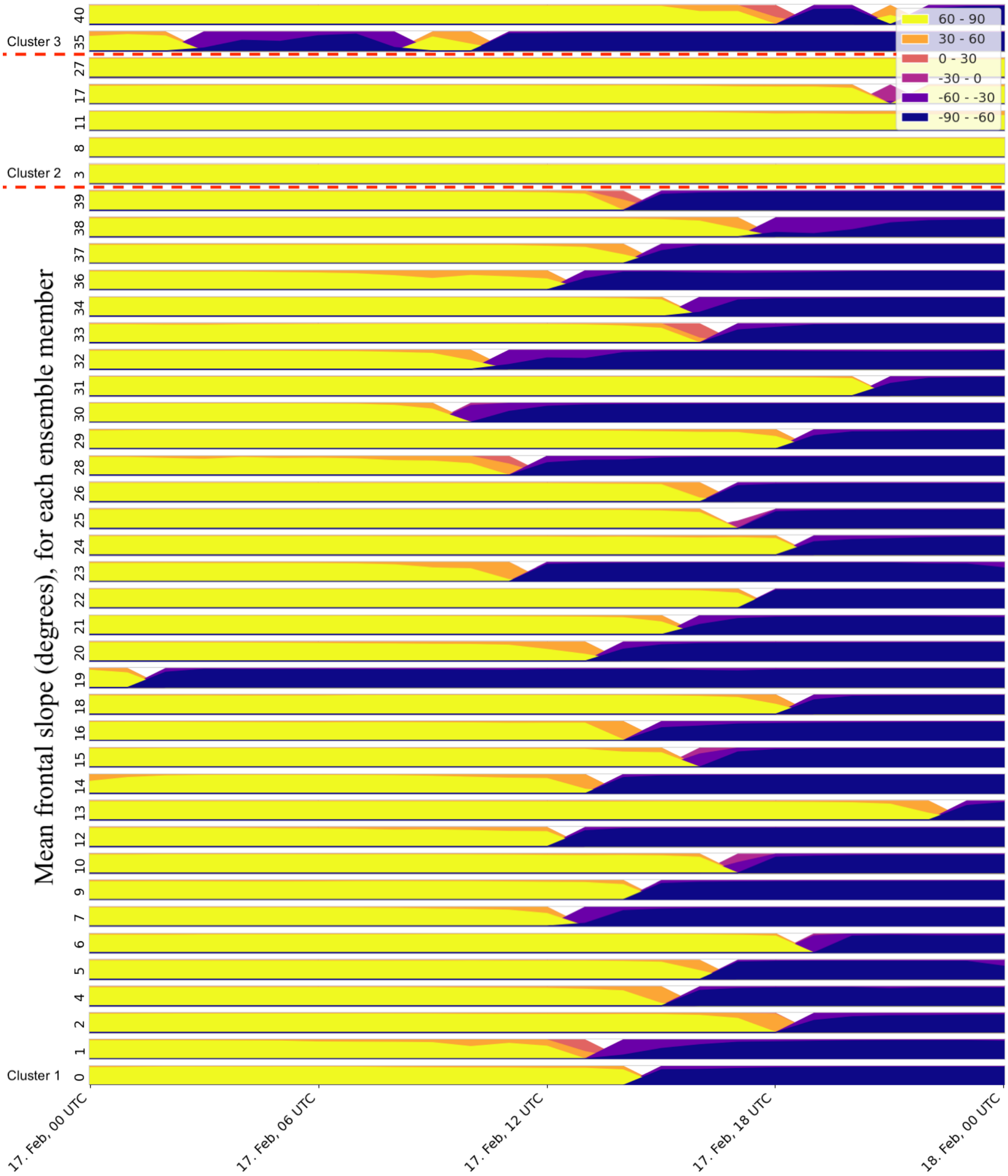


Figure 6.17: Horizon plot visualisation of mean frontal slope attribute time series clustered by k-means. K-means clustering was performed using DTW as the distance metric. Mean frontal slope attribute time series are derived from tracked 3-D cold fronts of Cyclone Otto using the DWD ICON-EU-ENS forecast, initialised on 16 February 2023 at 00:00 UTC.

This results in clusters of time series with similar patterns but with temporal distortion.

In addition to the previously discussed spaghetti plot of the clustered time series, an alternative representation of the clusters can be obtained using the horizon plot. As shown in Figure 6.17, the horizon plot provides a visual representation of the clusters obtained by using DTW as the distance metric.

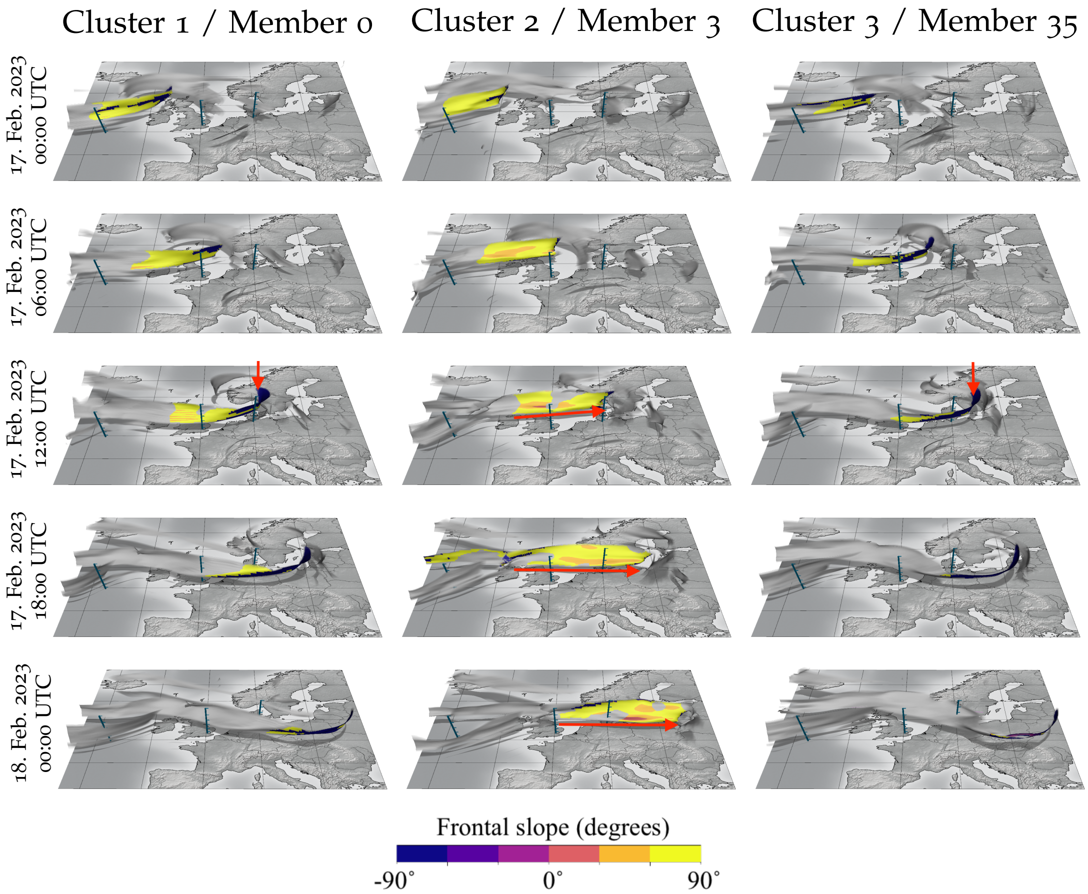


Figure 6.18: 3-D fronts of a sample member for each cluster. The tracked cold front is colour-coded according to the frontal slope. All other fronts are shown in grey. Fronts detected in ensemble member 0 (right row) and ensemble member 35 (left row) follow the Norwegian cyclone model, and the cold front wraps up around the cyclone centre (see red arrows on 17 February 2023, 12:00 UTC). Fronts detected in ensemble member 3 (middle row) are more similar to the Shapiro-Keyser cyclone model. The front is structured in a straight line from west to east (horizontal red arrow on 17 February, 12:00 UTC, 18:00 UTC and 18 February 2023, 00:00) and perpendicular is to the warm front. All fronts are detected in the ICON-EU-ENS forecast, initialised on 16. February 2023 at 00:00 UTC.

The variability of 3-D fronts between clusters is represented by selected ensemble members in Figure 6.18. Comparing the frontal structure of representative ensemble members, ensemble member 0 (representing cluster 1) and ensemble member 35 (representing cluster 3), the cold front in both begins to wrap up around the cyclone centre during the course of 17 February 2023 (see vertical red arrow on 17 February 2023, 12:00 UTC, in Figure 6.18). The structure of the cold front follows the structure of the Norwegian cyclone model. In contrast, the cold front in ensemble member 3 (representing cluster 2) is structured in a straight line from west to east (see horizontal red arrow on 17 February, 12:00 UTC, 18:00 UTC, and 18 February 2023, 00:00 UTC in Figure 6.18), perpendicular to the warm front, and does not show the wrap-up process as in the other representative ensemble members. The structure of the cold front follows the structure of the Shapiro–Keyser cyclone model (see Figure 3.5).

In addition to the analysis of the frontal slope attribute, Figure 6.19 shows the clustered horizon plot for the development of the mean TFP. In analogy to the previous clustering, the number of clusters is also set to 3, but here only the DTW distance metric is analysed. Cluster 1 is characterised by a transition

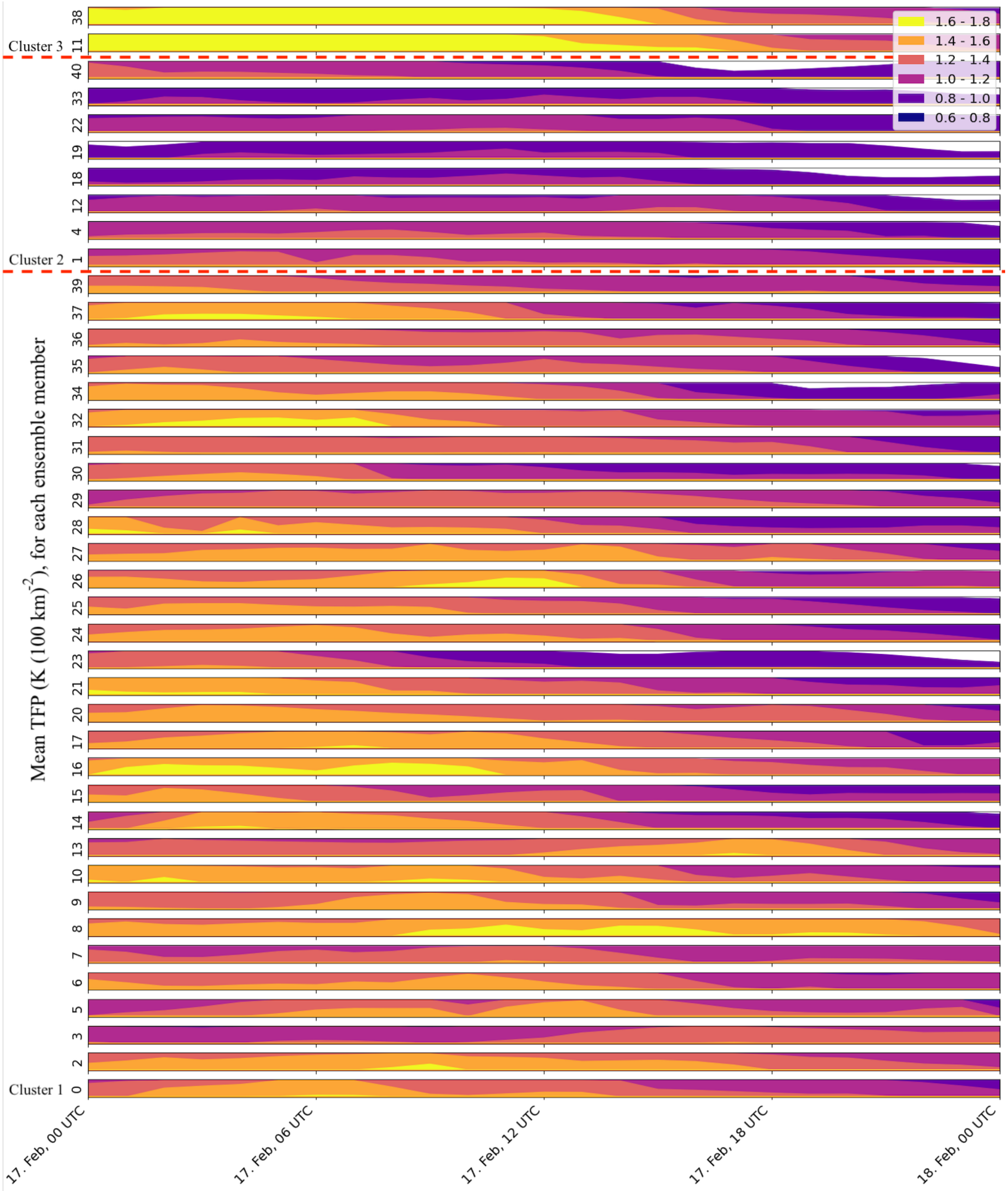


Figure 6.19: Horizon plot visualisation of mean TFP frontal attribute time series clustered by k-means. K-means clustering was performed using DTW as the distance metric. Mean TFP frontal attribute time series are derived from tracked 3-D cold fronts of Cyclone Otto using the DWD ICON-EU-ENS forecast, initialised on 16 February 2023 at 00:00 UTC.

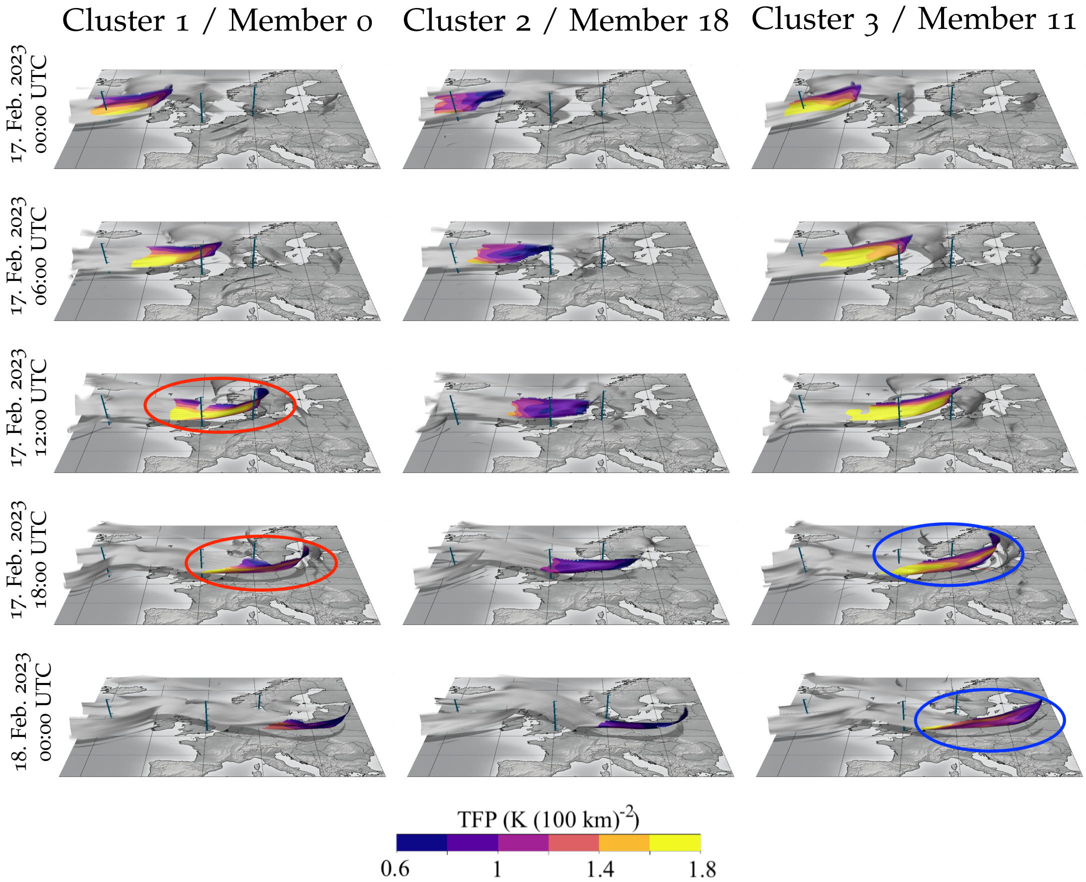


Figure 6.20: 3-D fronts of a sample member for each cluster. The tracked cold front is colour coded according to TFP. All other fronts are shown in grey. Fronts detected in ensemble member 0 (left row) are characterised by a transition from high TFP values between 17 February 2023, 12:00 UTC, to low TFP values at 18:00 UTC (marked with red circles in both time steps). Fronts detected in ensemble member 2 (middle row) are characterised by low TFP values throughout the period shown. Fronts detected in ensemble member 11 (right row) are characterised by high TFP values throughout the period shown, with a slight decrease after 17 February 2023, 12:00 UTC (marked with blue circles in both time steps). All fronts are detected in the ICON-EU-ENS forecast, initialised on 16. February 2023 at 00:00 UTC.

from high to low mean TFP frontal attribute values, cluster 2 is characterised by consistently low mean TFP frontal attribute values and cluster 3 is characterised by consistently high mean TFP frontal attribute values. Analogous to the analysis of the mean frontal slope attribute, the variability of 3-D fronts between clusters is illustrated by selected ensemble members in Figure 6.20. For ensemble member 0 (representing cluster 1), TFP transitions from high values between 17 February 2023, 12:00 UTC, to low TFP values at 18:00 UTC (marked with a red circle for both time steps in Figure 6.20). This coincides with the transition of the frontal slope from positive to negative values in the same periods (cf. Figure 6.17 and Figure 6.18). For ensemble member 18 (representing cluster 2), the TFP

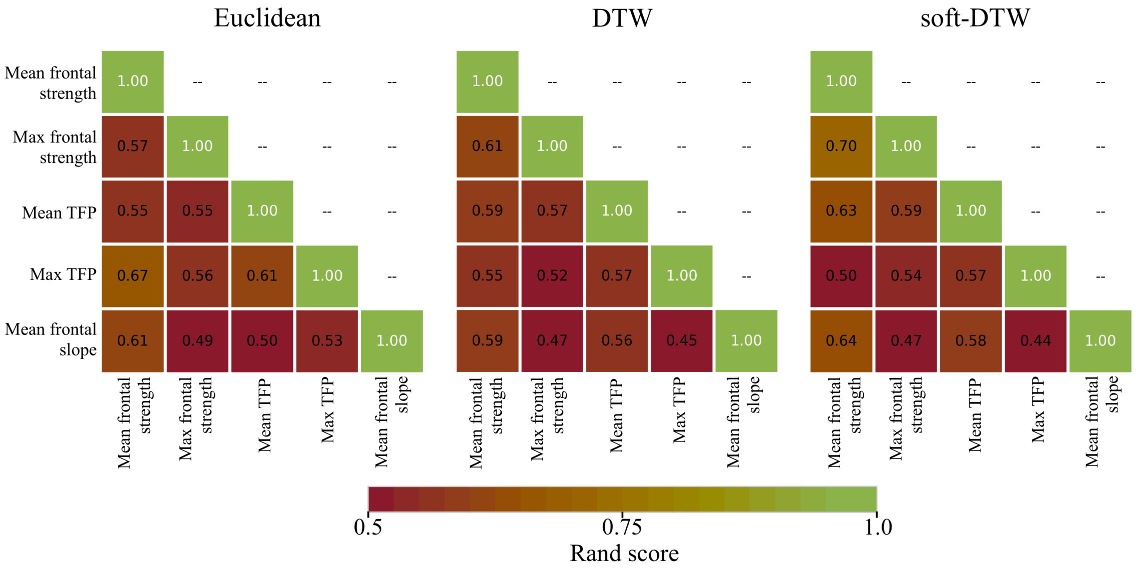


Figure 6.21: Sensitivities of clusters derived from 3-D frontal attribute time series using different distance metrics.

values are below 1 $K(100km)^{-2}$ throughout the period. For ensemble member 11 (representing cluster 3), the TFP are higher compared to ensemble members 0 and 18, and the 3-D cold front appears larger. The TFP decreases slightly after 17 February 2023, 12:00 UTC (marked with blue circles in Figure 6.20), but more moderately than in ensemble member 0. Are high TFP values associated with a positive frontal slope? This question arises because the frontal slope for ensemble member 11 remained positive throughout the period analysed (cf. Figure 6.17). In the following analysis, I will examine the sensitivity between these clusters and clusters of other frontal attributes in order to derive possible relationships between clusters.

Attribute sensitivity analysis

To assess the sensitivity of the clusters to different frontal attributes, I follow Kumpf et al. (2018) and use the Rand index. The Rand index is a statistical measure used to assess the similarity or agreement between two data clustering results or partitions, where a score of 1 represents identical clusters and a score of 0 indicates no relationship between clusters. See Section 2.4.3 for a detailed description of the Rand index. The cluster sensitivity analysis is performed for the three previously used distance metrics: Euclidean, DTW and soft-DTW, as well as the previously used filter thresholds for TFP 0.4 $K(100km)^{-2}$ and frontal strength of 0.4 $K(100km)^{-2}$ to detect the 3-D fronts.

The results shown in Figure 6.21 indicate that using the DTW or soft-DTW distance metrics results in an increased similarity between the mean frontal strength and maximum frontal strength attributes compared to using the Euclidean distance metric. DTW shows a Rand index of 0.61, while soft-DTW shows a Rand index of 0.7. The Euclidean distance metric identifies the most similarities in clusters between the TFP maximum and mean frontal strength attributes, with a Rand index of 0.67. On the other hand, the clusters of the mean frontal slope attribute consistently show a lower level of similarity with clusters derived from the thermal variable (TFP and frontal strength). Overall, and compared to the Rand indices in Kumpf et al. (2018), the clusters in this case study tend to be independent of each other. However, these results cannot be generalised, as this would require statistics from several case studies.

Cluster sensitivities across filter thresholds

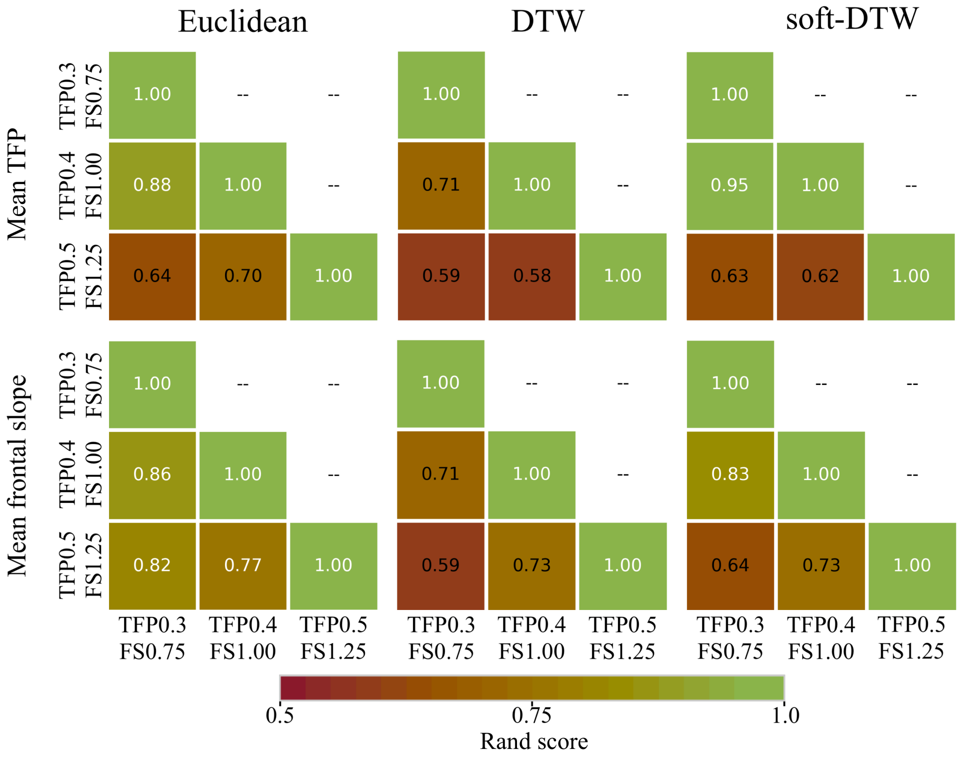


Figure 6.22: Cluster sensitivities for different filter thresholds for mean TFP (top) and mean frontal slope (bottom).

To test how sensitive the clusters are when altering the filter thresholds of detected 3-D fronts, clusters are derived for different front filter thresholds. For the previous analysis, the front filter TFP was set to 0.4 K $(100\,\text{km})^{-2}$ and the frontal strength filter to 1 K $(100\,\text{km})^{-2}$ (abbreviated as TFP0.4 FS1.00). For this sensitivity analysis, the frontal attributes are again derived from 3-D fronts, but the 3-D fronts are computed using different filter criteria: once with less strict filters (TFP: 0.3 K $(100\,\text{km})^{-2}$, front strength: 0.75 K $(100\,\text{km})^{-2}$, abbreviated as TFP0.3 FS0.75), and once with stricter filters (TFP: 0.5 K $(100\,\text{km})^{-2}$, frontal strength: 1.25 K $(100\,\text{km})^{-2}$, abbreviated as TFP0.5 FS1.25).

Figure 6.21 shows the sensitivity of the clusters, quantified by the Rand index. This sensitivity analysis covers all three different distance metrics: Euclidean, DTW and soft-DTW, and considers two different frontal attributes, the mean TFP and the mean frontal slope. Overall, clusters with the lowest sensitivity to changes in filter thresholds are computed using Euclidean distance, with an average Rand score of 77.8. The average Rand score of clusters computed with DTW is 65.2 and with soft-DTW is 73.3. Furthermore, clusters based on mean frontal slope are less sensitive to changes in filter thresholds than those based on mean TFP. In addition, clusters tend to be less sensitive to less strict filter thresholds, but more sensitive to stricter filter thresholds. This may be due to the fact that stricter filter thresholds typically result in fewer and smaller detected 3-D fronts. Therefore, it is possible that stricter filtering may exclude certain frontal features, resulting in significantly different frontal attributes and consequently different clusters.

7

SUMMARY AND CONCLUSION

Contents

7.1 Conclusions and answers to research questions **114**
7.2 Outlook **118**

This thesis explored and investigated how objective 2-D and 3-D front detection and visualisation, integrated into an interactive 3-D visual analysis environment for atmospheric data, can be used to study frontal dynamics within mid-latitude cyclones and thus be useful for weather forecasting and research. The method presented builds on approaches previously introduced by Hewson (1998) and Kern et al. (2019) and is applicable to gridded data from state-of-the-art NWP models. It facilitates rapid analysis of 3-D frontal dynamics, including objective comparison of detected frontal structures between data sets from different numerical models or ensemble members, even at different model resolutions. The integration of 3-D front detection with 3-D IVA (in this case the open-source visual analysis framework Met.3D) facilitates rapid analysis of complex weather situations, as the detected fronts can be visualised jointly with interactively placed depictions of other meteorological quantities.

Furthermore, a feature-based ensemble analysis approach based on objectively detected 3-D atmospheric fronts has been developed and applied in case studies. This methodology consists of the following three key elements: First, an atmospheric front of interest is isolated from a group of fronts, e.g. a cold front of a particular cyclone system from all other detected fronts. This isolated front is characterised by physical frontal attributes such as frontal slope or mean frontal strength and can be tracked using two different approaches: manual and automated tracking. Automated tracking uses a two-pass filtering mechanism. A distance-based filter is applied to narrow down potential matches. Next, a filter that compares frontal attribute distributions is applied to identify the target front. Second, time series of frontal attributes are computed to characterise the tracked front. These data provide a dynamic representation of the physical and geometric characteristics of the tracked front and their evolution over time. Third, the development and application of feature-based ensemble analysis. This analysis compares time series of frontal attributes across different ensemble members. It provides a means of ordering and clustering ensemble simulations based on derived frontal attribute time series.

In summary, this methodology provides a core framework for conducting feature-based ensemble analysis in the context of atmospheric fronts. By systematically detecting, tracking, and analysing these features across ensemble members, this approach provides a novel method for ensemble simulation analysis. Details of obtaining the code and datasets necessary to replicate the analysis and results presented in this thesis are provided in Section 8.

7.1 CONCLUSIONS AND ANSWERS TO RESEARCH QUESTIONS

Here, I revisit and answer the research questions and objectives presented in the introduction to conclude the thesis.

1a: Identifying appropriate detection parameters, including data smoothing and filtering thresholds, to ensure objective comparability between different model resolutions, different ensemble members, or different cases.

The choice of the thermal variable is essential for the presented approach. For the cases presented in this thesis, I show that θ_w is most suitable, since, in contrast to θ, it considers reversible moist processes in the atmosphere. The resulting fronts are longer and more continuous. A disadvantage of θ_w is that it also detects purely humidity-dominated fronts. Separately filtering frontal feature candidates according to humidity and θ gradients, however, allows one to distinguish humidity-dominated from temperature-dominated fronts. The choice of filter parameters and filter thresholds to obtain meaningful frontal structures is challenging. These settings depend on the thermal input variable's horizontal smoothing length scale, which determines the "spatial scales" of detected frontal features (large-scale smoothing of the thermal input field results in the detection of large-scale frontal features and vice versa). The distribution of gradient magnitudes shows that different smoothing length scales require different filter thresholds to obtain meaningful fronts. Large-scale smoothing requires less restrictive filter thresholds compared to small-scale smoothing. We present recommendations to future users on how to tune filter thresholds according to the previously applied smoothing length scale (Table 5.1).

1b: Evaluation of the benefit of 3-D IVA of frontal surfaces through case study investigations, including interpretations based on conceptual models, and comparison of frontal structures between different numerical models and with manually produced surface analysis charts.

The application of the proposed approach to case studies of mid-latitude cyclones provides detailed information about the temporal evolution of 3-D front characteristics. I demonstrate the use of 3-D front detection to visualise dynamic relations of features in the context of fronts in NWP data by directly representing these features in 3-D. In a case study of Cyclone Vladiana (September 2016), I examine the conceptual model of the WCB as represented by NWP data. At the cold front, WCB trajectories ascend quickly, experience jet wind speeds early, and follow the anticyclonically turning jet stream. In contrast, WCB trajectories ascending at the warm front show a slower ascent rate and tend to take the cyclonic outflow branch in the upper troposphere. These observations agree well with conceptual models of fronts and WCB as proposed in the literature. The next example considers the relation between convection and cold-front structure. For Vladiana, the cold front at mid-tropospheric levels is temporarily strengthened in the vicinity of resolved convection; I hypothesize that the model representation of convection and/or simulation grid spacing influences the feedback and interaction between convection, frontogenesis, and detailed frontal structures. In a second case study of Cyclone Friederike (January 2018), I visually analyse the 3-D temporal evolution of fronts in a Shapiro–Keyser cyclone and compare these results to the conceptual model proposed in the literature. I observe that the different Shapiro–Keyser cyclone stages do not occur simultaneously at all elevations. However, all characteristic stages of the conceptual model of the Shapiro–Keyser cyclone could be observed in NWP data. Finally, I compare the objective 3-D frontal structures with 2-D fronts in UK Met Office surface analysis charts and investigate the occurrence of secondary fronts often present in UK Met Office surface analyses. The objective 3-D fronts are consistent with the UK Met Office fronts if θ_w is used for front detection. This is no coincidence, as θ_w is the primary thermal variable used for the manual front detection by the UK Met Office. For the storm case Friederike, I show that the secondary front corresponds to a humidity-dominated rather than a temperature-dominated front. An, in parts, similar front detection approach – only two-dimensional but also applicable to kilometre-scale resolution data – was proposed by Jenkner et al. (2009). Because of high sensitivity to local noise in higher derivatives, their approach uses the zero lines of the TFP (second derivative) as frontal candidates, which correspond to the steepest gradient within the frontal zone. However, this does not match the most common definition of a front as the boundary of the frontal zone located on the warm-air side (see Renard and Clarke, 1965). I argue that an advantage of my approach, in particular for case studies, is that also in kilometre-scale data fronts are detected at this warm-air side, albeit at the cost of potential smoothing artefacts.

2a: Can 3-D atmospheric fronts be effectively isolated and characterised by assigning 1-D frontal attributes? What are the key physical attributes of fronts that enhance meteorological analysis? How can the tracking of isolated 3-D fronts be automated using feature attributes?

While fuzzy filtering allows for a visual representation of frontal strength and other filter criteria, the challenge lies in the visual fading of the 3-D front at its boundaries, causing difficulties in assigning well-defined frontal attributes. To address this issue, a shift from fuzzy filtering to the implementation of hard filter thresholds is proposed, resulting in the generation of well-defined 3-D frontal surfaces characterised by sharp boundaries. The analysis shows that meaningful frontal attributes, including thermal properties such as mean TFP and mean frontal slope, can be computed on these distinct surfaces. These attributes serve to characterise the 3-D frontal surface, effectively reducing its dimensionality and facilitating comprehensive analysis. The study shows that the unique signature of the frontal attribute distributions for individual fronts enables effective frontal tracking. In particular, the application of the EMD to compute a distance between individual TFP distributions of frontal surfaces in combination with a distance filter is shown to be an effective approach for automatic 3-D frontal tracking. The automated tracking of a front over time facilitates the generation of time series of frontal attributes, providing a robust basis for a comprehensive analysis of the evolution of frontal characteristics.

2b: What are effective visualisation methods for displaying the evolution of frontal attributes over time? How can feature-based ensemble analysis be successfully performed on 3-D atmospheric fronts? What techniques can be used to perform feature-based ensemble similarity ordering and clustering, and how can these approaches be effectively visualised?

To comprehensively analyse time series of derived frontal attributes, a clear visualisation method is crucial. In the context of ensemble analysis, the horizon plot proves to be an effective tool for displaying similarities and differences in time series of feature attributes. However, ordering these time series based on similar characteristics is challenging, as conventional Euclidean distance measures may inaccurately emphasise time offsets of characteristic events between two series. This limitation is particularly evident when events with identical characteristics occur at different times, leading to an overestimation of the distance between time series. Such a characteristic event could be, for example, a drop in the mean TFP attribute occurring in two time series but at different time steps. To address this issue, the study proposes the use of DTW and soft-DTW as suitable distance measures between attribute time series. The application of DTW and soft-DTW allows time series to be ordered according to their similarity, taking into account characteristic developments while allowing for time shifts. This methodology allows the production of similarity-based ordered horizon plots, supports a comprehensive analysis of feature-based ensemble clusters, and provides precise and condensed insights into the temporal evolution of frontal attributes.

2c: How sensitive are clusters derived from 3-D fronts, considering different frontal attributes and variations in filter thresholds applied to the detected 3-D frontal surfaces? What insights do these variations provide

for understanding the behaviour and characteristics of 3-D atmospheric fronts within ensemble simulations?

The sensitivity of the proposed clustering method is systematically evaluated across different frontal attributes and by varying the frontal detection filter thresholds, including variations in both directions, more and less strict filter thresholds. Following Kumpf et al. (2018), the sensitivity analysis uses the Rand index, a metric that ranges from 1 (indicating no sensitivity) to 0 (indicating high sensitivity).

The results show that the derived clusters are generally sensitive between different frontal attributes. The differences in sensitivity values derived from the different distance metrics for k-means clustering show almost similar results, with the Euclidean distance metric having the highest overall Rand index. For this case study, I conclude that clusters derived from different frontal attributes have a degree of independence. This means that clustering different frontal attributes leads to different clusters. However, with only one case study, the sample size is too small to derive generally valid sensitivity statistics.

To test the sensitivity of clusters to front filter thresholds, the filter thresholds for the detected 3-D fronts were adjusted, resulting in slightly larger fronts for less strict filter thresholds and smaller fronts for more strict filter thresholds. The resulting sensitivity analysis shows that clusters generated from frontal attributes are generally not very sensitive to small changes in the filter thresholds of the underlying 3-D fronts. However, clusters are slightly less sensitive when less strict filter thresholds are used, whereas they show more sensitivity when stricter filter thresholds are used. This may be due to the fact that more restrictive filter thresholds typically result in fewer and smaller detected 3-D fronts, and therefore it is possible that such restrictive filtering may exclude certain frontal features, creating different clusters and, consequently, higher sensitivities due to cluster dissimilarity.

As a final remark, the feature-based ensemble analysis and visualisation approach presented here has the potential to be used operationally. Being integrated in Met.3D, other meteorological variables can be analysed in conjunction with the 3-D frontal structures. This facilitates the rapid analysis of complex weather situations, as required in operational settings (see Rautenhaus et al., 2018). The results suggest that the use of a feature-based ensemble analysis technique, which automates 3-D feature tracking and characterisation by feature attributes, has potential in both research and operational weather forecasting. The method's stability and adaptability make it a valuable tool for weather forecasting and weather analysis, enhancing the ability to analyse and comprehend the variability of fronts and frontal dynamics in ensemble simulations. Further possible fields of application include climatological studies of frontal characteristics derived from the 3-D features, and investigation of the relation of frontal structures to other physically meaningful features in the 3-D atmosphere, including the jet stream –

this will be beneficial for studies that contribute to the understanding of complex dynamical processes in the atmosphere.

7.2 OUTLOOK

The present study introduces a novel approach to ensemble analysis through the examination of feature attributes, and provides valuable insights into selected cases. However, the scope of my analysis is currently limited, suggesting the need for further research to explore broader applications and improve the methodology. This outlook highlights potential paths for future investigation.

Machine Learning (ML) for 3-D feature detection: Opportunities for future novel feature detection methods may be facilitated by recent advances in ML. For example, an approach using artificial neural networks to detect 2-D fronts (Niebler et al., 2022) and its 3-D extension (Niebler et al., 2023) has recently been proposed. Their approach learns from fronts depicted on analysis charts issued by national weather services and hence mimics the approaches of human forecasters. One advantage over the objective 2-D and 3-D front detection presented here is that there is no need for parameter configuration and adjustments, as the fronts are detected directly in NWP model data output. However, their ML approach learns from fronts that are, to some extent, subjectively depicted by human forecasters. Can such a ML approach identify objective fronts based on statistics of subjectively depicted fronts? Their 3-D approach to front detection is represented by 2-D horizontal and vertical slices, lacking the geometric 3-D representation of fronts. However, exploring the feasibility of obtaining a comprehensive 3-D representation and comparing the results with my proposed 3-D objective front detection method could provide insights into the accuracy and capabilities of such an ML approach.

3-D visualisation of cluster uncertainties: Visualisation of uncertainties and clustering in ensemble simulations facilitates the analysis of large datasets to identify trends, clusters, and outliers. Various visualisation approaches have been proposed in the literature to characterise uncertainties of ensemble simulations in 2-D, such as contour box plots (Whitaker et al., 2013), variability plots of iso-contour lines (Ferstl et al., 2016a) and visualisation of outliers in contour ensemble data (Chaves-de-Plaza et al., 2024) and in 3-D for iso-surfaces (Pfaffelmoser et al., 2011) and clusters of streamlines (Ferstl et al., 2016b). Combining such approaches to represent clusters and their uncertainties of 3-D fronts would open new possibilities for visualising, analysing, and interpreting clusters of 3-D fronts derived from frontal attributes.

Climatologies of feature attributes: The proposed feature-based ensemble analysis approach could be extended to produce climatologies of feature attributes, providing a comprehensive understanding of the temporal (seasonal) and spatial (location) variations in these attributes. The use of extensive datasets, such as ERA5 data, holds promise for evaluating the effectiveness of the approach under different climatic conditions. My results indicating a successful application to ERA5 data suggest that this extension is both feasible and promising.

Collections of ensembles and intercomparison: Extending the analysis to collections of ensembles provides a novel opportunity to study specific meteorological phenomena, such as severe winter storms over Europe. By comparing the characteristics of similar storms, researchers can explore the existence of common patterns, clusters, and footprints. This comparative analysis can reveal hidden relationships and contribute to a deeper understanding of the dynamics underlying these high-impact weather events.

Characterisation of additional atmospheric features and measures of feature relationships: Beyond storm-related investigations, future research should aim to characterise other key atmospheric features, including the jet core, tropopause, and polar vortex. Applying the feature-attribute methodology to these features opens opportunities to derive geometric and physical measurements, providing valuable insights into their behaviour and potential interactions. Such an extension could contribute significantly to our knowledge of atmospheric dynamics. Investigating the relationships between feature attributes and the geometric and physical properties of atmospheric features is a critical next step. By establishing these relationships, researchers can gain a deeper understanding of the underlying processes that drive atmospheric phenomena. This approach has the potential to contribute to improved predictive modelling and forecasting capabilities.

In conclusion, the presented front detection and ensemble analysis approach provides a promising foundation for future research. By extending the methodology to climatologies, collections of ensembles, and additional atmospheric features, researchers can gain new insights into the complexity of the Earth's atmosphere. This perspective encourages interdisciplinary collaboration to advance the field of ensemble analysis towards more comprehensive applications.

8

PUBLICATION, AVAILABILITY OF CODE, DATA AND ADDITIONAL MATERIALS

PUBLICATION

Major parts of Chapter 5 and parts of Chapter 1 and 7 are based on the following publication:

> Andreas A. Beckert et al. (2023a). "The Three-Dimensional Structure of Fronts in Mid-Latitude Weather Systems in Numerical Weather Prediction Models." In: *Geoscientific Model Development* 16.15, pp. 4427–4450. ISSN: 1991-959X. DOI: `10.5194/gmd-16-4427-2023`. URL: `https://gmd.copernicus.org/articles/16/4427/2023/` (visited on 08/18/2023)

Author contributions

I designed and implemented the algorithm, performed the analysis, created the visualisations, and wrote the manuscript. Lea Eisenstein contributed to the meteorological-related results, especially to the characterisation and discussion of the Shapiro–Keyser cyclone. Annika Oertel contributed to the meteorological related results, especially to the characterisation and discussion of the WCB. Tim Hewson contributed to the design of the algorithm and secondary front discussion. Marc Rautenhaus and George C. Craig proposed, supervised, and administered the study. All authors contributed to writing and revising the manuscript.

AVAILABILITY OF CODE, DATA AND ADDITIONAL MATERIALS

Code and data

The code of the specific version of the open-source visualisation framework Met.3D, including the code for front detection and example configuration files to reproduce figures of Chapter 5, is available at Beckert et al., 2023b.

The ECMWF ERA5 and ECMWF HRES forecast and analysis datasets used in this thesis are available at Beckert, 2023.

The code of the specific version of the open-source visualisation framework Met.3D, including the code and data for front tracking and feature-based ensemble analysis to reproduce figures and results of Chapter 6, is available at Beckert et al., 2024.

User and developer documentation of Met.3D is available at Met3d.documentation, 2024; Met3d.homepage, 2024.

Video supplement

The following movies illustrate interactive visual data analysis using Met.3D and provide supplementary insights into the 3-D dynamics of frontal structures, jet stream, and WCB trajectories. They also illustrate the benefit gained from interactive use of 3-D visual analysis:

- Comparison of objectively detected 3-D fronts in wet-bulb potential temperature and potential temperature of Friederike on 18 January 2018 at 12:00 Beckert et al. (2022b).
- Development of 3-D frontal structures, jet stream, and WCB trajectories of Vladiana Beckert et al. (2022a).
- Interactive front analysis of storm Friederike using the open-source meteorological 3-D visualization framework "Met.3D" Beckert et al. (2022c).

LIST OF FIGURES

Figure 2.1 Example representations of the three visualisation disciplines. (a) Scientific visualisation of the Rayleigh-Taylor instability caused by the mixing of two fluids. (b) Information visualisation of a graph from a network analysis. (c) Visual analytics represented by a dashboard. Figures from (a) Lawrence Livermore National Laboratory (2009), (b) Grandjean (2014), used under CC BY-SA 3.0, and (c) HelicalInsight OpenSourceBI (2015), used under CC BY-SA 4.0. 8

Figure 2.2 Example analysis plot of ECMWF ensemble forecast from 13. July 2023. The ensemble meteogram shows the high-resolution (9 km) and ensemble spread for the following 10 days. Figure from ECMWF (2023a), used under CC BY 4.0. 11

Figure 2.3 Screenshots of the interactive visualisation software Met.3D. Figure from Rautenhaus (2015), used under CC BY 3.0. 12

Figure 2.4 Illustration of the conceptual data pipeline in visualisation. Figure from Rautenhaus (2015), his figure 3.1., used with permission. 12

Figure 2.5 Example visualisations of Box Plot and Violin Plot show the total compensation for all academic ranks. Figure used with permission of Taylor & Francis Informa UK Ltd - Journals, from Hintze and Nelson (1998); permission conveyed through Copyright Clearance Center, Inc. 16

Figure 2.6 Construction of a horizon plot: (a) line plot, (b) segmented line plot, (c) layering, and (d) final horizon plot with visual differentiation. Figure adapted from Heer et al. (2009). 17

Figure 2.7 Representation of different coordinate systems and data grids. (a) Cartesian coordinate system, (b) spherical coordinate system, and (c) unstructured grid. Figures from (a) Stolfi (2009), (b) Ag2gaeh (2015), used under CC BY-SA 4.0, and (c) Slffea and Mysid (2006), used under CC BY-SA 3.0. 18

Figure 2.8 Example of a dolphin represented by triangular meshes. Figure from Chrschn (2007). 20

Figure 2.9 Triangulated cubes. Figure used with permission of ACM (Association for Computing Machinery), from Lorensen and Cline (1987); permission conveyed through Copyright Clearance Center, Inc. 22

Figure 2.10 Difference in matching between Euclidean and DTW distance measures for time series analysis. Figure from XantaCross (2011), used under CC BY-SA 3.0. 28

Figure 3.1 Illustration of idealized large-scale tropospheric circulation. Figure from Kaidor (2013), used under CC BY-SA 3.0. 32

Figure 3.2 Schematic illustration of an ensemble forecast. Cropped figure from Met Office (2023), © British Crown copyright, Met Office, used with permission. 36

Figure 3.3 Illustration of an extratropical cyclone and accompanying cyclonic features: clouds (grey shaded areas), pressure isolines (black solid lines), near-surface wind direction (blue arrows), and fronts. Figure from Stull (2017), used under CC BY-NC-SA 4.0 37

Figure 3.4 Idealised cyclone presented by the Bergen school (Bjerknes and Solber, 1922). Figure used with permission of the Norwegian Geophysical Society. 38

Figure 3.5 Conceptual models of cyclone evolution showing lower-tropospheric geopotential height and fronts (top), and lower-tropospheric potential temperature (bottom). (a) Norwegian cyclone model: (I) incipient frontal cyclone, (II) and (III) narrowing warm sector, (IV) occlusion; (b) Shapiro–Keyser cyclone model: (I) incipient frontal cyclone, (II) frontal fracture, (III) frontal T-bone and bent-back front, (IV) frontal T-bone and warm seclusion. Figure from Schultz et al. (1998), their Figure 15. © American Meteorological Society. Used with permission. 39

Figure 3.6 Illustration of a vertical cross-section from west to east through a frontal system to illustrate the computation of the frontal slope. Sketch not to scale. 45

Figure 4.1 Cyclone Vladiana on 23 September 2016 at 06:00 UTC. **(a)** Detected 2-D warm (red line) and cold (blue line) fronts at 850 hPa, θ_w at 950 hPa (colours, in K), and mean sea level pressure (black contour lines, every 2 hPa) from a COSMO simulation (black frame shows domain boundaries; green frame shows the selected sub-region for studying convection in the vicinity of the cold front; see Section 5.3.3). **(b)** Detected 3-D warm (red) and cold (blue) fronts between 950 hPa and 500 hPa, on top of a horizontal map showing θ_w at 950 hPa and mean sea level pressure (black contour lines, every 2 hPa). Warm- and cold-front classification is computed according to warm- and cold-air advection at the front (following Hewson, 1998). 48

Figure 4.2 Successive time steps of objective 2-D frontal structures show the temporal development of Friederike (17 and 18 January 2018), as detected in ERA5 reanalysis data at 750 hPa and surface pressure (black lines). The displayed time steps are approximately assigned to the four ideal development stages of the Shapiro–Keyser cyclone model (Shapiro and Keyser, 1990). We find that not all characteristics of the individual stages can be observed in 2-D. As shown in Figure 5.13, 3-D front detection is required to observe all characteristics. 50

Figure 4.3 Track of the cyclone centre of winter storm *Otto* according to the University of Berlin. Figure from RandomInfinity17 (2023), used under CC BY-SA 4.0. The original image has been edited by adding date and timestamps to the storm track. 51

Figure 5.1 Illustration of the thermal-gradient-based detection method, using a simplified straight front and following Hewson (1998) and Kern et al. (2019). The goal is to determine the warm-air boundary of the frontal zone (i.e. the region of increased thermal gradient; see the yellow line). This boundary corresponds to the third derivative (red line) of a thermal variable τ (black line) being zero, under the condition that the second derivative of τ (blue line) is negative. The cold-front typing shown assumes air masses are moving from left to right across the figure. 55

Figure 5.2 Step-by-step illustration of the 2-D front detection method. In the example, objective fronts are based on the 850 hPa wet-bulb potential temperature field (θ_w) from the ECMWF HRES forecast (horizontally regular grid-point spacing of 0.15° in both longitude and latitudes) initialised on 18 January 2018 at 00:00 UTC and valid on 18 January 2018 at 12:00 UTC. Fronts are "fuzzy filtered" using a fade-out range for TFP of 0.2–0.4 K $(100\,\mathrm{km})^{-2}$ and for frontal strength of 0.6–1 K $(100\,\mathrm{km})^{-1}$. See Sect. 2.3 for a description of panels **(a)**–**(h)**. 59

Figure 5.3 Illustration of the extended data pipeline in Met.3D for objective front detection. The existing pipeline module is shown in a blue box and new data pipeline modules are shown in green boxes. Data structures are represented by oval circles. Ψ represents the scalar input data field. 62

Figure 5.4 Distribution (relative frequencies) of thermal front parameter (TFP) values computed from hourly ECMWF HRES forecast data (horizontal grid-point spacing of 0.15°) from 18 January 2018, in the region 30° N–70° N, 60° W–30° E and between 950–500 hPa for different smoothing length scales: **(a)** 100 km, **(b)** 50 km, and **(c)** 30 km. 64

Figure 5.5 Distribution (relative frequencies) of $|\nabla_h \theta_w|$ within frontal zones between 950–500 hPa (same data, time, and region as in Figure 5.4) for different smoothing length scales: **(a)** 100 km, **(b)** 50 km, and **(c)** 30 km. 65

Figure 5.6 From 2-D to 3-D objective fronts. Same data as in Figure 5.2 (18 January 2018, 12:00 UTC) but showing the full 3-D structure of frontal surfaces in the lower and middle atmosphere. All circles and arrows denote features discussed in text. **(a)** 850 hPa frontal lines from Figure 5.2h with 3-D frontal surfaces between surface and 500 hPa, viewed from the top. **(b)** Same as **(a)** but from a tilted viewpoint looking north. **(c)** Same as **(b)** but with additional fuzzy normal curve filter of θ between 0.6–1 K $(100\,\mathrm{km})^{-1}$. **(d)** Same as **(c)** but viewed from west. Cross section shows θ and $|\nabla_h \theta|$. **(e)** Same as **(b)** but with additional fuzzy normal curve filter of specific humidity between 0.1–0.2 g $(\mathrm{kg}\ 100\,\mathrm{km})^{-1}$. **(f)** Same as **(e)** but viewed from west. Cross section shows q and $|\nabla_h q|$. **(g)** Input field smoothed to a horizontal length scale of 30 km with same filtering applied as in **(a)**. **(h)** Same as **(g)** but with adapted filter settings for TFP between 1.5–2.5 K $(100\,\mathrm{km})^{-2}$ and frontal strength between 1.2–2.2 K $(100\,\mathrm{km})^{-1}$. 66

Figure 5.7 Impact of the vertical resolution on detected fronts. (a) original vertical resolution of ECMWF with 137 vertical level. (b) vertical level retained to 68 level. (c) vertical levels retained to 28 levels. 68

Figure 5.8 Conceptual model of fronts and WCB showing large-scale ascending and descending air in the vicinity of an extra-tropical cyclone. Figure adapted from Stull (2017), used under CC BY-NC-SA 4.0. 70

Figure 5.9 **(a–c)** Temporal evolution of 3-D frontal structures and WCB trajectories of Vladiana on 23 September 2016. **(d)** Same time as **(b)** but only fast-ascending WCB trajectories (minimum 200 hPa within 2 h) are displayed for a period of 48 h. **(e)** Same as **(d)** but only slow-ascending WCB trajectories (less than 200 hPa within 2 h) are displayed. **(f)** Same time as **(c)**, jet stream (yellow isosurface of 50 $\mathrm{m\,s^{-1}}$ wind speed) and WCB trajectories are displayed for a period of 48 h. For the full temporal development of this scene, see the Video supplement (Beckert et al., 2022a). 71

Figure 5.10 Convection and frontal structure on 23 September 2016 at 06:00 UTC. Region corresponds to green sub-area in Figure 4.1. ECMWF analysis **(a, b, c)** and COSMO analysis **(d, e, f)** at **(a, d)** 850 hPa, **(b, e)** 700 hPa, and **(c, f)** 500 hPa. Objective 2-D fronts (blue tubes) are shown along with UK Met Office fronts (red tubes), θ_w (colour), $|\nabla_h \theta_w|$ (grey shades), and upward air velocity (contour lines: orange is upwards, black is zero, and green is downwards; contour line spacing is 0.02 $\mathrm{m\,s^{-1}}$). 73

Figure 5.11 The 3-D view of the 2-D frontal structures from Figure 5.10. **(a)** 2-D objective fronts (blue tubes) at 850, 700, and 500 hPa (see Figure 5.10) in the context of full 3-D frontal structures, as found in ECMWF data. **(b)** Same as **(a)** but for COSMO data. Red circles in **(a)** and **(b)** mark the differences in the frontal surfaces. Contour lines on all surface maps represent upward air velocity at 700 hPa (orange is upwards, black is zero, and green is downwards; contour line spacing 0.02 $\mathrm{m\,s^{-1}}$). **(c)** ECMWF 3-D fronts and vertical section of wind divergence (colour), θ_w (coloured contour lines, spacing 1 K), and θ (black contour lines, spacing 5 K). **(d)** Same as **(c)** but for COSMO data. 74

Figure 5.12 Temporal evolution of 3-D frontal structures in Figure 5.11, detected from (left) ECMWF analysis and (right) COSMO analysis. Contour lines projected onto the surface show upward air velocity at 700 hPa (orange is upwards, black is zero, and green is downwards; contour line spacing of 0.02 $\mathrm{m\,s^{-1}}$). The yellow pole marks the centre of the convective updraft at 06:00 UTC, and the red arrow points northward. 75

Figure 5.13 Temporal evolution of 3-D frontal structures of Friederike (16 to 19 January 2018), as detected in ERA-5 reanalysis data. **(a)** Different cyclone stages encountered along the cyclone track. Yellow poles mark centres of surface low, and front colours distinguish time steps. **(b)** The six stages from **(a)**, approximately centred around the cyclone centres for comparison of frontal structures. Blue arrows mark frontal fracture, yellow arrows mark warm-core frontal seclusion, and contour lines show surface pressure (spacing 2 hPa). 76

Figure 5.14 Temporal evolution of 3-D frontal structures of Egon (12 to 13 January 2017), as detected in ERA-5 reanalysis data. **(a)** Different cyclone stages encountered along the cyclone track. Yellow poles mark centres of surface low, and front colours distinguish time steps. **(b)** The six stages from **(a)**, approximately centred around the cyclone centres for comparison of frontal structures. Contour lines show surface pressure (spacing 2 hPa). 77

Figure 5.15 Comparison of UK Met Office fronts with objective fronts for case Friederike (18 January 2018, 12:00 UTC). **(a)** UK Met Office surface analysis chart. Blue box marks analysed area. **(b)** Objective 850 hPa 2-D fronts (blue lines) as detected from ECMWF HRES θ_w (colour; grey shading shows $|\nabla_\mathrm{h}\theta_\mathrm{w}|$), UK Met Office fronts (red lines), and mean sea level pressure (black contour lines, spacing 2 hPa). **(c)** Same as **(b)** but objective 2-D fronts (green lines) based on θ. The secondary front (black arrow) is only detected when using θ_w. When based on θ, the cold front (blue arrow) breaks up and is less continuous compared to the cold front based on θ_w. 78

Figure 5.16 The 3-D view of Figure 5.15b–c. Red tubes show UK Met Office fronts, and 3-D objective fronts are coloured according to pressure elevation. Objective fronts based on **(a)** θ_w and **(b)** θ. The secondary front (black arrow) is a feature of θ_w and only occurs around 850 hPa. Yellow poles are to aid spatial perception. Compare the animated version in the Video supplement (Beckert et al., 2022b). 79
Figure 5.17 Total column rain water and 3-D fronts of ECMWF HRES simulation on 18 January 2018, 12:00 UTC, initialised on 18 January 2018, 00:00 UTC. (a) 3-D fronts and total column rain water. (b) Total column rain water. The orange circle highlights the position of the secondary cold front shown in (a). 80
Figure 6.1 Comparison of 3-D fronts with (a) fuzzy-filter thresholds and (b) hard filter thresholds. 83
Figure 6.2 Example of the 3-D front selection process. Selected fronts remain in colour, unselected fronts are shaded light grey. The path lines show the evolution of the frontal centroids of the current and other ensemble members. The path lines of the current ensemble member are annotated with a timestamp. 85
Figure 6.3 Example of a temporarily split cold front event during frontal tracking. (a) The cold front is compact and forms a contiguous surface. (b) The red circle marks the location of a cold front split. (c) The cold front is compact and contiguous again. 86
Figure 6.4 Adapted data pipeline for the feature selection framework. Additional pipeline modules and data input/output formats are illustrated in darker colours. 91

Figure 6.5 Example of the interactive front selection workflow. (a) Frontal surfaces filtered with hard thresholds and coloured according to front type (blue = cold front, red = warm front). (b) Warm fronts are filtered out, the selected cold front is highlighted in blue, and non-selected cold fronts are grey. The text indicates the ensemble member, date and the time of the selected front. (c) 24-hour front track together with the cold front of the last tracking time step. (d) Front tracks of multiple ensemble members and the cold front of the currently selected member. (e) Example of front tracks of five selected ensemble members. Front tracks are presented using the multi-parameter trajectory display of Neuhauser et al. (2023). The band of each front track shows the evolution of the mean TFP, mean frontal strength and mean frontal slope attributes, including a highlighted time step annotated as a pie chart. 92

Figure 6.6 Distribution of (a) TFP, (b) frontal strength, and (c) frontal slope of the tracked 3-D cold front of Storm Friederike, shown as a violin plot. The front is tracked in ECMWF ENS forecast, initialised on 16 January 2018, 12:00 UTC, with 91 vertical levels defined on hybrid sigma-pressure coordinates. The distributions are shown for 73 time steps of the tracked front and for ensemble member 0. While the distributions of TFP and frontal strength have similar characteristics, the frontal slope has a dipole structure. Individual parts of the front are either steeply sloped forwards (close to 90°) or backwards (close to -90°). 94

Figure 6.7 Distribution of TFP of the reference front and potential front matches for different time intervals. Manually selected target TFP distributions are marked with a red arrow. For (a) 1 hour and (b) 3 hour intervals, the smallest EMD distance between the TFP distributions of the reference and target fronts is correct and automatic front tracking is successful. For (c) 6 hour and (d) 12 hour intervals, the minimum EMD distance between the TFP distribution of the reference does not represent the correct target front and automatic front tracking fails. Figure 6.8 shows the corresponding 3-D fronts. Fronts detected in ensemble member 0 of ECMWF ENS forecast, initialised on 16 January 2018, 12:00 UTC, with 91 vertical levels defined on hybrid sigma-pressure coordinates. 95

Figure 6.8 Illustration of automated front tracking for different time intervals. The reference front is shown in dark blue, the target front obtained by automated tracking is shown in light blue. As shown in Figure 6.7, automated tracking is successful for time intervals of (a) 1 and (b) 3 hours. Tracking fails for time intervals of (c) 6 and (d) 12 hours. The manually selected target front is marked with a red arrow and the front selected by automated front tracking is marked with a red circle in (c) and (d), where automated front tracking fails. All fronts are detected in ensemble member 0 of ECMWF ENS forecast, initialised on 16 January 2018, 12:00 UTC, with 91 vertical levels defined on hybrid sigma-pressure coordinates. 96

Figure 6.9 Distribution of (a) TFP on model level and (b) on pressure level on a frontal surface, shown as violin plots. The front is tracked in the ECMWF ENS forecast, initialised on 16 January 2018, 12:00 UTC. Model levels consist of 91 vertical levels defined on hybrid sigma-pressure coordinates. Pressure levels consist of 12 vertical levels defined on pressure coordinates. Distributions are shown for 73 time steps of the tracked front and for ensemble member 0. The TFP distributions in both model levels (a) and pressure levels (b) have a similar structure. However, the distributions in the model level data show an anomaly around 17 January 2018, 00:00 UTC, with an accumulation of low TFP values (red circle). 97

Figure 6.10 Comparison of fronts detected in (a) model level and (b) pressure level data on 17 January 2018, using the same data as used in Figure 6.9. The red circle indicates a part of the frontal surface with low TFP values in the model level data, which is not present in the pressure level data. This leads to the anomalies of low TFP in the model level data shown in Figure 6.9. Red arrows in (b) point to "spikes" in the frontal surface detected in the pressure level data due to the low vertical grid spacing. 98

Figure 6.11 Distribution of TFP detected in the same front across all 51 ensemble members at the simulation initialisation time step. Annotated on the x-axis is the EMD relative to the reference ensemble member (member 0) for each TFP distribution. 3-D fronts detected in ECMWF ENS forecast data, initialised on 16 January 2018, 12:00 UTC, with 12 vertical levels defined on pressure coordinates. 99

Figure 6.12 (a) Spaghetti plot of mean TFP frontal attribute development over a 25-hour period. Each black line represents an ensemble member, the red line represents the reference member (member 0). (b) Horizon plot of *TFP mean* frontal attribute time series. Each row represents an ensemble member. The frontal attributes are computed from tracked 3-D cold fronts of Cyclone Otto using the DWD ICON-EU-ENS forecast, initialised on 16 February 2023 at 00:00 UTC. 100
Figure 6.13 DTW distances of each ensemble member to the reference member (member 0). Ensemble members are sorted by distance in ascending order. 101
Figure 6.14 Horizon plot of the mean TFP frontal attribute. Ensemble members are sorted by similarity to the reference member (member 0) using DTW, from most similar member (bottom) to most different member (top). 103
Figure 6.15 Silhouette score for k-means clustering of (a) mean TFP and (b) mean frontal slope for different numbers of clusters. The silhouette score shows the highest intra-cluster coherence and inter-cluster separation for the DTW and soft-DTW distance function and 2 to 4 clusters. Overall, the silhouette score is significantly lower for mean TFP compared to mean frontal slope. 104
Figure 6.16 Spaghetti plots of clustered frontal attribute mean frontal slope using the k-means algorithm for 3 cluster centres. The figure shows cluster results for different distance metrics: (a) - (c) Euclidean, (d) - (f) DTW, and (g) - (i) soft-DTW. Red lines highlight the cluster centres. 105
Figure 6.17 Horizon plot visualisation of mean frontal slope attribute time series clustered by k-means. K-means clustering was performed using DTW as the distance metric. Mean frontal slope attribute time series are derived from tracked 3-D cold fronts of Cyclone Otto using the DWD ICON-EU-ENS forecast, initialised on 16 February 2023 at 00:00 UTC. 106

Figure 6.18 3-D fronts of a sample member for each cluster. The tracked cold front is colour-coded according to the frontal slope. All other fronts are shown in grey. Fronts detected in ensemble member 0 (right row) and ensemble member 35 (left row) follow the Norwegian cyclone model, and the cold front wraps up around the cyclone centre (see red arrows on 17 February 2023, 12:00 UTC). Fronts detected in ensemble member 3 (middle row) are more similar to the Shapiro-Keyser cyclone model. The front is structured in a straight line from west to east (horizontal red arrow on 17 February, 12:00 UTC, 18:00 UTC and 18 February 2023, 00:00) and perpendicular is to the warm front. All fronts are detected in the ICON-EU-ENS forecast, initialised on 16. February 2023 at 00:00 UTC. 107

Figure 6.19 Horizon plot visualisation of mean TFP frontal attribute time series clustered by k-means. K-means clustering was performed using DTW as the distance metric. Mean TFP frontal attribute time series are derived from tracked 3-D cold fronts of Cyclone Otto using the DWD ICON-EU-ENS forecast, initialised on 16 February 2023 at 00:00 UTC. 108

Figure 6.20 3-D fronts of a sample member for each cluster. The tracked cold front is colour coded according to TFP. All other fronts are shown in grey. Fronts detected in ensemble member 0 (left row) are characterised by a transition from high TFP values between 17 February 2023, 12:00 UTC, to low TFP values at 18:00 UTC (marked with red circles in both time steps). Fronts detected in ensemble member 2 (middle row) are characterised by low TFP values throughout the period shown. Fronts detected in ensemble member 11 (right row) are characterised by high TFP values throughout the period shown, with a slight decrease after 17 February 2023, 12:00 UTC (marked with blue circles in both time steps). All fronts are detected in the ICON-EU-ENS forecast, initialised on 16. February 2023 at 00:00 UTC. 109

Figure 6.21 Sensitivities of clusters derived from 3-D frontal attribute time series using different distance metrics. 110

Figure 6.22 Cluster sensitivities for different filter thresholds for mean TFP (top) and mean frontal slope (bottom). 111

LIST OF TABLES

Table 3.1 Overview and summary of selected objective front detection methods. 42
Table 5.1 Fuzzy frontal filter threshold recommendations for different smoothing length scales. 67

BIBLIOGRAPHY

Aemisegger, F., J. K. Spiegel, S. Pfahl, H. Sodemann, W. Eugster, and H. Wernli (2015). "Isotope Meteorology of Cold Front Passages: A Case Study Combining Observations and Modeling." In: *Geophysical Research Letters* 42.13, pp. 5652–5660. ISSN: 00948276. DOI: `10.1002/2015GL063988`. URL: `http://doi.wiley.com/10.1002/2015GL063988`.

Ag2gaeh (2015). *Kugelkoordinaten: Definition*. URL: `https://commons.wikimedia.org/wiki/File:Kugelkoord-def.svg` (visited on 10/09/2023).

Akenine-Möller, Tomas, Eric Haines, Naty Hoffman, Angelo Pesce, Michal Iwanicki, and Sebastien Hillaire (2018). *Real-Time Rendering*. Fourth edition. Boca Raton: CRC Press, Taylor and Francis Group. 1 p. ISBN: 978-1-138-62700-0.

Ali, Mohammed, Ali Alqahtani, Mark W. Jones, and Xianghua Xie (2019). "Clustering and Classification for Time Series Data in Visual Analytics: A Survey." In: *IEEE Access* 7, pp. 181314–181338. ISSN: 2169-3536. DOI: `10.1109/ACCESS.2019.2958551`. URL: `https://ieeexplore.ieee.org/document/8930535` (visited on 12/20/2023).

Anderson, R., B. W. Boville, and D. E. McClellan (1955). "An Operational Frontal Contour-Analysis Model." In: *Quarterly Journal of the Royal Meteorological Society* 81.350, pp. 588–599. ISSN: 1477-870X. DOI: `10.1002/qj.49708135008`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.49708135008` (visited on 06/16/2023).

Anwender, Doris, Patrick A. Harr, and Sarah C. Jones (2008). "Predictability Associated with the Downstream Impacts of the Extratropical Transition of Tropical Cyclones: Case Studies." In: *Monthly Weather Review* 136.9, pp. 3226–3247. ISSN: 1520-0493, 0027-0644. DOI: `10.1175/2008MWR2249.1`. URL: `https://journals.ametsoc.org/view/journals/mwre/136/9/2008mwr2249.1.xml` (visited on 08/02/2023).

Ayachit, Utkarsh (2015). *The ParaView Guide: A Parallel Visualization Application*. Clifton Park, NY, USA: Kitware, Inc. 276 pp. ISBN: 978-1-930934-30-6.

Bader, M.J., G. S. Forbes, J. R. Grant, R. B. E. Lilley, and A. J. Waters (1996). *Images in Weather Forecasting : A Practical Guide for Interpreting Satellite and Radar Imagery*. 523 pp. ISBN: 978-0-521-45111-6.

Bader, Robin, Michael Sprenger, Nikolina Ban, Stefan Radisuhli, Christoph Schar, and Tobias Ganther (2020). "Extraction and Visual Analysis of Potential Vorticity Banners around the Alps." In: *IEEE Transactions on Visualization and Computer Graphics* 26.1, pp. 1–1. ISSN: 1077-2626. DOI: `10.1109/TVCG.2019.2934310`. pmid: `31425096`. URL: `https://ieeexplore.ieee.org/document/8805433/`.

Bahlmann, C. and H. Burkhardt (2004). "The Writer Independent Online Handwriting Recognition System Frog on Hand and Cluster Generative Statistical Dynamic Time Warping." In: *IEEE Transactions on Pattern Analysis and Ma-*

chine Intelligence 26.3, pp. 299–310. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2004.1262308. URL: http://ieeexplore.ieee.org/document/1262308/ (visited on 05/26/2024).

Baldauf, Michael, Axel Seifert, Jochen Förstner, Detlev Majewski, Matthias Raschendorfer, and Thorsten Reinhardt (2011). "Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities." In: *Monthly Weather Review* 139.12, pp. 3887–3905. ISSN: 00270644. DOI: 10.1175/MWR-D-10-05013.1.

Beckert, Andreas A. (2023). *Datasets Associated with the Publication: "The Three-Dimensional Structure of Fronts in Mid-Latitude Weather Systems in Numerical Weather Prediction Models"*. Version 1.0. DOI: 10.5281/zenodo.7875629. URL: https://zenodo.org/records/7875629 (visited on 06/19/2024).

Beckert, Andreas A., Lea Eisenstein, Annika Oertel, Tim Hewson, George C. Craig, and Marc Rautenhaus (2022a). "Comparison of Objectively Detected 3-D Fronts in Wet-Bulb Potential Temperature and Potential Temperature." In: *TIB AV Portal [video suppl.]* DOI: 10.5446/57600.

– (2022b). "Development of 3-D Frontal Structures, Jet Stream and WCB Trajectories of Vladiana." In: *TIB AV Portal [video suppl.]* DOI: 10.5446/57570.

– (2022c). "Interactive Front Analysis of Storm Friederike Using the Open-Source Meteorological 3-D Visualization Framework "Met. 3D"." In: *TIB AV Portal [video suppl.]* DOI: 10.5446/57944.

– (2023a). "The Three-Dimensional Structure of Fronts in Mid-Latitude Weather Systems in Numerical Weather Prediction Models." In: *Geoscientific Model Development* 16.15, pp. 4427–4450. ISSN: 1991-959X. DOI: 10.5194/gmd-16-4427-2023. URL: https://gmd.copernicus.org/articles/16/4427/2023/ (visited on 08/18/2023).

Beckert, Andreas A., Marc Rautenhaus, Michael Kern, and Met 3D Contributors (2023b). *Met.3d-1.8.0_3DFronts_v1.0*. Version 1.8.0-3DFronts_v1.0. Zenodo. DOI: 10.5281/zenodo.7870254. URL: https://zenodo.org/records/7870254 (visited on 06/19/2024).

Beckert, Andreas A., Marc Rautenhaus, Michael Kern, Katharina Kaufmann, and Met 3D Contributors (2024). *Met.3D front tracking and clustering*. Gitlab. URL: https://gitlab.com/ki1roy/met.3d-front-tracking-and-clustering (visited on 06/19/2024).

Berggren, Roy (1952). "The Distribution of Temperature and Wind Connected with Active Tropical Air in the Higher Troposphere and Some Remarks Concerning Clear Air Turbulence at High Altitude." In: *Tellus* 4.1, pp. 43–53. ISSN: 0040-2826. DOI: 10.3402/tellusa.v4i1.8668. URL: https://doi.org/10.3402/tellusa.v4i1.8668 (visited on 02/27/2023).

Berry, Gareth, Michael J. Reeder, and Christian Jakob (2011). "A Global Climatology of Atmospheric Fronts." In: *Geophysical Research Letters* 38.4, pp. 1–5. ISSN: 00948276. DOI: 10.1029/2010GL046451.

Bindon, H. H. (1940). "RELATION BETWEEN EQUIVALENT POTENTIAL TEMPERATURE AND WET-BULB POTENTIAL TEMPERATURE." In: *Monthly*

Weather Review 68.9, pp. 243–245. ISSN: 0027-0644. DOI: 10.1175/1520-0493(1940)068<0243:RBEPTA>2.0.CO;2. URL: http://journals.ametsoc.org/doi/10.1175/1520-0493(1940)068%3C0243:RBEPTA%3E2.0.CO;2.

Bjerknes, J (1919). "On the Structure of Moving Cyclones." In: pp. 95–99. ISSN: 03044076.

– (1930). "Practical Examples of Polar-Front Analysis over the British Isles in 1925–6." In: *Geophys. Mem.* 5.10.

– (1935). "Investigations of Selected European Cyclones by Means of Serial Ascents, Case 3: December 30-31, 1930." In: *Geof. Publ.* 11.4, pp. 3–18. URL: http://www.ngfweb.no/docs/NGF_GP_Vol11_no4.pdf (visited on 02/27/2023).

Bjerknes, J and E Palmén (1937). "Investigations of Selected European Cyclones by Means of Serial Ascents, Case 4: February 15-17, 1935." In: *Geof. Publ.* 12.2, pp. 5–62. URL: http://www.ngfweb.no/docs/NGF_GP_Vol12_no2.pdf (visited on 02/27/2023).

Bjerknes, J and H Solber (1922). "Life of the Cyclones and the Polar Front Theory of Atmopheric Circulation." In: pp. 1–18.

Blanchard, Nicolas, Florian Pantillon, Jean-Pierre Chaboureau, and Julien Delanoë (2020). "Organization of Convective Ascents in a Warm Conveyor Belt." In: *Weather and Climate Dynamics* 1.2, pp. 617–634. DOI: 10.5194/wcd-1-617-2020.

Blender (2024). *A 3D Modelling and Rendering Package*. Version 4.1. Blender Online Community, Stichting Blender Foundation, Amsterdam: Blender Foundation. URL: https://www.blender.org/ (visited on 06/16/2024).

Bosart, Lance F. (2003). "Whither the Weather Analysis and Forecasting Process?" In: *Weather and Forecasting* 18.3, pp. 520–529. ISSN: 1520-0434, 0882-8156. DOI: 10.1175/1520-0434(2003)18<520:WTWAAF>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/wefo/18/3/1520-0434_2003_18_520_wtwaaf_2_0_co_2.xml (visited on 07/14/2023).

Browning, K. A. and G. A. Monk (1982). "A Simple Model for the Synoptic Analysis of Cold Fronts." In: *Quarterly Journal of the Royal Meteorological Society* 108.456, pp. 435–452. ISSN: 00359009. DOI: 10.1002/qj.49710845609. URL: https://onlinelibrary.wiley.com/doi/10.1002/qj.49710845609.

Browning, K. A. and N. M. Roberts (1996). "Variation of Frontal and Precipitation Structure along a Cold Front." In: *Quarterly Journal of the Royal Meteorological Society* 122.536, pp. 1845–1872. ISSN: 00359009. DOI: 10.1002/qj.49712253606. URL: https://onlinelibrary.wiley.com/doi/10.1002/qj.49712253606.

Bösiger, Lukas, Michael Sprenger, Maxi Boettcher, Hanna Joos, and Tobias Günther (2022). "Integration-Based Extraction and Visualization of Jet Stream Cores." In: *Geoscientific Model Development* 15.3, pp. 1079–1096. ISSN: 1991-9603. DOI: 10.5194/gmd-15-1079-2022. URL: https://gmd.copernicus.org/articles/15/1079/2022/.

Cahir, John J and Walter D Lottes (1982). "An Objective Diagnostic Aid in Locating Meteorologically Significant Boundaries." In: Conference on Weather Forecasting and Analysis, 9 Th, Seattle, WA, pp. 296–299.

Chaves-de-Plaza, N.F., M. Molenaar, P. Mody, M. Staring, R. Van Egmond, E. Eisemann, A. Vilanova, and K. Hildebrandt (2024). "Depth for Multi-Modal Contour Ensembles." In: *Computer Graphics Forum*, e15083. ISSN: 0167-7055, 1467-8659. DOI: 10.1111/cgf.15083. URL: https://onlinelibrary.wiley.com/doi/10.1111/cgf.15083 (visited on 06/11/2024).

Choudhary, Anubhav and Aiko Voigt (2022). "Impact of Grid Spacing, Convective Parameterization and Cloud Microphysics in ICON Simulations of a Warm Conveyor Belt." In: *Weather and Climate Dynamics* 3.4, pp. 1199–1214. DOI: 10.5194/wcd-3-1199-2022. URL: https://wcd.copernicus.org/articles/3/1199/2022/ (visited on 07/25/2023).

Chrschn (2007). *An Example of a Polygon Mesh. Illustration of a Dolphin, Represented with Triangles.* URL: https://commons.wikimedia.org/wiki/File:Dolphin_triangle_mesh.png (visited on 10/07/2023).

Clarke, Leo C. and Robert J. Renard (1966). "The U. S. Navy Numerical Frontal Analysis Scheme: Further Development and a Limited Evaluation." In: *Journal of Applied Meteorology* 5.6, pp. 764–777. ISSN: 0021-8952. DOI: 10.1175/1520-0450(1966)005<0764:TUSNNF>2.0.CO;2. URL: http://journals.ametsoc.org/doi/10.1175/1520-0450(1966)005%3C0764:TUSNNF%3E2.0.CO;2.

Coiffier, Jean (2011). *Fundamentals of Numerical Weather Prediction*. 1st ed. Cambridge University Press. ISBN: 978-1-107-00103-9 978-0-511-73445-8. DOI: 10.1017/CBO9780511734458. URL: https://www.cambridge.org/core/product/identifier/9780511734458/type/book (visited on 07/18/2023).

Cuturi, Marco and Mathieu Blondel (2017). "Soft-DTW: A Differentiable Loss Function for Time-Series." In: *Proceedings of the 34th International Conference on Machine Learning - Volume 70*. ICML'17. Sydney, NSW, Australia: JMLR.org, pp. 894–903.

DWD (2014). *ICAO-Standardatmosphäre*. Deutscher Wetterdienst. URL: https://www.dwd.de/DE/service/lexikon/begriffe/S/Standardatmosphaere_pdf.pdf? (visited on 06/09/2023).

Dacre, Helen F. and Joaquim G. Pinto (2020). "Serial Clustering of Extratropical Cyclones: A Review of Where, When and Why It Occurs." In: *npj Climate and Atmospheric Science* 3.1 (1), pp. 1–10. ISSN: 2397-3722. DOI: 10.1038/s41612-020-00152-9. URL: https://www.nature.com/articles/s41612-020-00152-9 (visited on 08/17/2023).

Davies, E. Roy (2017). *Computer Vision, Principles, Algorithms, Applications, Learning*. Academic Press. 900 pp. ISBN: 978-0-12-809284-2.

RBB (2023). *Deutscher Wetterdienst warnt vor Sturmböen am Montag*. URL: https://www.rbb24.de/panorama/beitrag/2023/02/amtliche-warnung-sturmboeen-berlin-deutscher-wetterdienst.html (visited on 06/05/2023).

Dey, C, C Sanders, J Clochard, and J Hennessy (2007). "Guide to the WMO Table Driven Code Form Used for the Representation and Exchange of Regularly Spaced Data in Binary Form: FM 92 GRIB." In: *Tech. Rep. No. 98*.

Dhillon, Inderjit S., Yuqiang Guan, and Brian Kulis (2004). "Kernel K-Means: Spectral Clustering and Normalized Cuts." In: *Proceedings of the Tenth ACM*

SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '04. New York, NY, USA: Association for Computing Machinery, pp. 551–556. ISBN: 978-1-58113-888-7. DOI: 10.1145/1014052.1014118. URL: https://dl.acm.org/doi/10.1145/1014052.1014118 (visited on 06/02/2023).

Dokmanic, Ivan, Reza Parhizkar, Juri Ranieri, and Martin Vetterli (2015). "Euclidean Distance Matrices: Essential Theory, Algorithms, and Applications." In: *IEEE Signal Processing Magazine* 32.6, pp. 12–30. ISSN: 1053-5888. DOI: 10.1109/MSP.2015.2398954. URL: http://ieeexplore.ieee.org/document/7298562/ (visited on 05/26/2024).

Doms, Günther and Michael Baldauf (2018). *A Description of the Nonhydrostatic Regional COSMO-Model. Part I: Dynamics and Numerics*. Deutscher Wetterdienst. DOI: 10.5676/DWD_pub/nwv/cosmo-doc_5.05_I.

ECMWF (2023a). *ENS Meteogram*. URL: https://charts.ecmwf.int/products/opencharts_meteogram?base_time=202307130000&epsgram=classical_10d&lat=53.53&lon=9.98 (visited on 07/13/2023).

– (2023b). *Metview - Meteorological Workstation Application*. Metview. URL: https://confluence.ecmwf.int/display/METV/Metview (visited on 10/14/2023).

Eckhardt, Sabine, Andreas Stohl, Heini Wernli, Paul James, Caroline Forster, and Nicole Spichtinger (2004). "A 15-Year Climatology of Warm Conveyor Belts." In: *Journal of Climate* 17.1, pp. 218–237. ISSN: 0894-8755. DOI: 10.1175/1520-0442(2004)017<0218:AYCOWC>2.0.CO;2. URL: http://journals.ametsoc.org/doi/10.1175/1520-0442(2004)017%3C0218:AYCOWC%3E2.0.CO;2.

Eisenstein, Lea, Florian Pantillon, and Peter Knippertz (2020). "Dynamics of Sting-Jet Storm Egon over Continental Europe: Impact of Surface Properties and Model Resolution." In: *Quarterly Journal of the Royal Meteorological Society* 146.726, pp. 186–210. ISSN: 1477870X. DOI: 10.1002/qj.3666.

Etling, Dieter (2008). *Theoretische Meteorologie Eine Einführung*. Springer-Verlag Berlin Heidelberg. 376 pp. ISBN: 978-3-540-75978-2.

Ferranti, Laura and Susanna Corti (2011). "New Clustering Products." In: *ECMWF*. URL: https://www.ecmwf.int/en/elibrary/80187-new-clustering-products (visited on 08/02/2023).

Ferstl, F., M. Kanzler, M. Rautenhaus, and R. Westermann (2016a). "Visual Analysis of Spatial Variability and Global Correlations in Ensembles of Iso-Contours." In: *Computer Graphics Forum* 35.3, pp. 221–230. ISSN: 1467-8659. DOI: 10.1111/cgf.12898. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.12898 (visited on 06/21/2024).

Ferstl, Florian, Kai Bürger, and Rüdiger Westermann (2016b). "Streamline Variability Plots for Characterizing the Uncertainty in Vector Field Ensembles." In: *IEEE Transactions on Visualization and Computer Graphics* 22.1, pp. 767–776. ISSN: 1941-0506. DOI: 10.1109/TVCG.2015.2467204. URL: https://ieeexplore.ieee.org/document/7192675 (visited on 06/14/2024).

Few, Stephen (2008). "Time on the Horizon." In: *Visual Business Intelligence Newsletter*. URL: http://www.perceptualedge.com/articles/visual_business_intelligence/time_on_the_horizon.pdf.

Fischer, Christoph, Andreas H. Fink, Elmar Schömer, Roderick Van der Linden, Michael Maier-Gerber, Marc Rautenhaus, and Michael Riemer (2022). "A Novel Method for Objective Identification of 3-D Potential Vorticity Anomalies." In: *Geoscientific Model Development* 15.11, pp. 4447–4468. ISSN: 1991-9603. DOI: 10.5194/gmd-15-4447-2022. URL: https://gmd.copernicus.org/articles/15/4447/2022/ (visited on 12/20/2022).

Forgy, Edward W (1965). "Cluster Analysis of Multivariate Data: Efficiency versus Interpretability of Classifications." In: *Biometrics* 21.3, pp. 761–777. ISSN: 0006-341X. JSTOR: 2528559. URL: https://www.jstor.org/stable/2528559 (visited on 06/01/2023).

Front - Glossary of Meteorology (2012). URL: https://glossary.ametsoc.org/wiki/Front (visited on 02/28/2022).

Gal-Chen, Tzvi and Richard C. J Somerville (1975). "On the Use of a Coordinate Transformation for the Solution of the Navier-Stokes Equations." In: *Journal of Computational Physics* 17.2, pp. 209–228. ISSN: 0021-9991. DOI: 10.1016/0021-9991(75)90037-6. URL: https://www.sciencedirect.com/science/article/pii/0021999175900376 (visited on 05/31/2024).

Geerts, Bart, Rick Damiani, and Samuel Haimov (2006). "Finescale Vertical Structure of a Cold Front as Revealed by an Airborne Doppler Radar." In: *Monthly Weather Review* 134.1, pp. 251–271. ISSN: 1520-0493. DOI: 10.1175/MWR3056.1. URL: http://journals.ametsoc.org/doi/10.1175/MWR3056.1.

Grandjean, Martin (2014). "La connaissance est un réseau: Perspective sur l'organisation archivistique et encyclopédique." In: *Les Cahiers du numérique* 10.3, pp. 37–54. ISSN: 1622-1494. DOI: 10.3166/LCN.10.3.37-54. URL: https://shs.cairn.info/revue-les-cahiers-du-numerique-2014-3-page-37 (visited on 10/12/2023).

Hansen, Charles D. and Chris R. Johnson, eds. (2005). *The Visualization Handbook*. Amsterdam ; Boston: Elsevier-Butterworth Heinemann. 962 pp. ISBN: 978-0-12-387582-2.

Harr, Patrick A., Doris Anwender, and Sarah C. Jones (2008). "Predictability Associated with the Downstream Impacts of the Extratropical Transition of Tropical Cyclones: Methodology and a Case Study of Typhoon Nabi (2005)." In: *Monthly Weather Review* 136.9, pp. 3205–3225. ISSN: 1520-0493, 0027-0644. DOI: 10.1175/2008MWR2248.1. URL: https://journals.ametsoc.org/view/journals/mwre/136/9/2008mwr2248.1.xml (visited on 08/02/2023).

Hassell, David, Jonathan Gregory, Jon Blower, Bryan N. Lawrence, and Karl E. Taylor (2017). "A Data Model of the Climate and Forecast Metadata Conventions (CF-1.6) with a Software Implementation (Cf-Python v2.1)." In: *Geoscientific Model Development* 10.12, pp. 4619–4646. ISSN: 1991-9603. DOI: 10.5194/gmd-10-4619-2017. URL: https://gmd.copernicus.org/articles/10/4619/2017/ (visited on 06/19/2023).

Hawcroft, M. K., L. C. Shaffrey, K. I. Hodges, and H. F. Dacre (2012). "How Much Northern Hemisphere Precipitation Is Associated with Extratropical Cyclones?" In: *Geophysical Research Letters* 39.24. ISSN: 1944-8007. DOI: 10.1029/

2012GL053866. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2012GL053866 (visited on 02/27/2023).

Heer, Jeffrey, Nicholas Kong, and Maneesh Agrawala (2009). "Sizing the Horizon: The Effects of Chart Size and Layering on the Graphical Perception of Time Series Visualizations." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09: CHI Conference on Human Factors in Computing Systems. Boston MA USA: ACM, pp. 1303–1312. ISBN: 978-1-60558-246-7. DOI: 10.1145/1518701.1518897. URL: https://dl.acm.org/doi/10.1145/1518701.1518897 (visited on 05/24/2023).

HelicalInsight OpenSourceBI (2015). *Marketing Dashboard*. URL: https://commons.wikimedia.org/wiki/File:Marketing_dashboard.png (visited on 10/12/2023).

Hersbach, Hans, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, et al. (2020). "The ERA5 Global Reanalysis." In: *Quarterly Journal of the Royal Meteorological Society* 146.730, pp. 1999–2049. ISSN: 0035-9009. DOI: 10.1002/qj.3803. URL: https://onlinelibrary.wiley.com/doi/10.1002/qj.3803.

Hewson, Tim D. (1998). "Objective Fronts." In: *Meteorological Applications* 5.1, pp. 37–65. ISSN: 13504827. DOI: 10.1017/S1350482798000553. URL: http://doi.wiley.com/10.1017/S1350482798000553.

Hewson, Tim D. and Helen A. Titley (2010). "Objective Identification, Typing and Tracking of the Complete Life-Cycles of Cyclonic Features at High Spatial Resolution." In: *Meteorological Applications* 17.3, pp. 355–381. ISSN: 13504827. DOI: 10.1002/met.204. URL: https://onlinelibrary.wiley.com/doi/10.1002/met.204.

Hintze, Jerry L. and Ray D. Nelson (1998). "Violin Plots: A Box Plot-Density Trace Synergism." In: *The American Statistician* 52.2, pp. 181–184. ISSN: 0003-1305. DOI: 10.1080/00031305.1998.10480559. URL: https://www.tandfonline.com/doi/abs/10.1080/00031305.1998.10480559 (visited on 01/13/2024).

Hodges, K. I. (1999). "Adaptive Constraints for Feature Tracking." In: *Monthly Weather Review* 127.6, pp. 1362–1373. ISSN: 1520-0493, 0027-0644. DOI: 10.1175/1520-0493(1999)127<1362:ACFFT>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/mwre/127/6/1520-0493_1999_127_1362_acfft_2.0.co_2.xml (visited on 07/19/2023).

Hofmann-Wellenhof, Bernhard, Herbert Lichtenegger, and James Collins (2001). *Global Positioning System*. Vienna: Springer. ISBN: 978-3-211-83534-0 978-3-7091-6199-9. DOI: 10.1007/978-3-7091-6199-9. URL: http://link.springer.com/10.1007/978-3-7091-6199-9 (visited on 06/09/2023).

Huber-Pock, F. and Ch Kress (1989). "An Operational Model of Objective Frontal Analysis Based on ECMWF Products." In: *Meteorology and Atmospheric Physics* 40.4, pp. 170–180. ISSN: 0177-7971. DOI: 10.1007/BF01032457. URL: http://link.springer.com/10.1007/BF01032457.

Jenkner, J., Michael Sprenger, I. Schwenk, C. Schwierz, S. Dierer, and D. Leuenberger (2009). "Detection and Climatology of Fronts in a High-Resolution

Model Reanalysis over the Alps." In: *Meteorological Applications* 17.1, pp. 1–18. ISSN: 14698080. DOI: 10.1002/met.142.

Ji, Guangfeng and Han-Wei Shen (2006). "Feature Tracking Using Earth Mover's Distance and Global Optimization." In: *Pacific Graphics*.

Kahveci, T., A. Singh, and A. Gurel (2002). "Similarity Searching for Multi-Attribute Sequences." In: *Proceedings 14th International Conference on Scientific and Statistical Database Management*. 14th International Conference on Scientific and Statistical Database Management. Edinburgh, UK: IEEE Comput. Soc, pp. 175–184. ISBN: 978-0-7695-1632-5. DOI: 10.1109/SSDM.2002.1029718. URL: http://ieeexplore.ieee.org/document/1029718/ (visited on 05/26/2024).

Kaidor (2013). *Global Circulation of Earth's Atmosphere Displaying Hadley Cell, Ferrell Cell and Polar Cell.* URL: https://commons.wikimedia.org/wiki/File:Earth_Global_Circulation_-_en.svg (visited on 08/21/2023).

Keller, J. H., S. C. Jones, J. L. Evans, and P. A. Harr (2011). "Characteristics of the TIGGE Multimodel Ensemble Prediction System in Representing Forecast Variability Associated with Extratropical Transition." In: *Geophysical Research Letters* 38.12. ISSN: 1944-8007. DOI: 10.1029/2011GL047275. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2011GL047275 (visited on 08/02/2023).

Kern, Michael, Tim D. Hewson, Filip Sadlo, Rüdiger Westermann, and Marc Rautenhaus (2018). "Robust Detection and Visualization of Jet-Stream Core Lines in Atmospheric Flow." In: *IEEE Transactions on Visualization and Computer Graphics* 24.1, pp. 893–902. ISSN: 10772626. DOI: 10.1109/TVCG.2017.2743989.

Kern, Michael, Tim D. Hewson, Andreas Schäfler, Rüdiger Westermann, and Marc Rautenhaus (2019). "Interactive 3D Visual Analysis of Atmospheric Fronts." In: *IEEE Transactions on Visualization and Computer Graphics* 25.1, pp. 1080–1090. ISSN: 19410506. DOI: 10.1109/TVCG.2018.2864806.

Keyser, Daniel and M. A. Shapiro (1986). "A Review of the Structure and Dynamics of Upper-Level Frontal Zones." In: *Monthly Weather Review* 114.2, pp. 452–499. ISSN: 0027-0644. DOI: 10.1175/1520-0493(1986)114<0452:AROTSA>2.0.CO;2. URL: http://journals.ametsoc.org/doi/10.1175/1520-0493(1986)114%3C0452:AROTSA%3E2.0.CO;2.

Kovesi, Peter (2010). "Fast Almost-Gaussian Filtering." In: *Proceedings - 2010 Digital Image Computing: Techniques and Applications, DICTA 2010*, pp. 121–125. DOI: 10.1109/DICTA.2010.30.

Kumler-Bonfanti, Christina, Jebb Stewart, David Hall, and Mark Govett (2020). "Tropical and Extratropical Cyclone Detection Using Deep Learning." In: *Journal of Applied Meteorology and Climatology* 59.12, pp. 1971–1985. ISSN: 1558-8424, 1558-8432. DOI: 10.1175/JAMC-D-20-0117.1. URL: https://journals.ametsoc.org/view/journals/apme/59/12/jamc-d-20-0117.1.xml (visited on 08/17/2023).

Kumpf, Alexander, Bianca Tost, Marlene Baumgart, Michael Riemer, Rüdiger Westermann, and Marc Rautenhaus (2018). "Visualizing Confidence in Cluster-Based Ensemble Weather Forecast Analyses." In: *IEEE Transactions* *on Visualization and Computer Graphics* 24.1, pp. 109–119. ISSN: 10772626. DOI: `10.1109/TVCG.2017.2745178`.

Lawrence Livermore National Laboratory (2009). *Raleigh-Taylor Instability*. VisIt Home. URL: `https://visit-dav.github.io/visit-website/examples/ex02/` (visited on 10/12/2023).

Lloyd, S. (1982). "Least Squares Quantization in PCM." In: *IEEE Transactions on Information Theory* 28.2, pp. 129–137. ISSN: 1557-9654. DOI: `10.1109/TIT.1982.1056489`.

Locatelli, John D., Jonathan E. Martin, and Peter V. Hobbs (1994). "A Wide Cold-Frontal Rainband and Its Relationship to Frontal Topography." In: *Quarterly Journal of the Royal Meteorological Society* 120.516, pp. 259–275. ISSN: 00359009. DOI: `10.1002/qj.49712051603`. URL: `https://onlinelibrary.wiley.com/doi/10.1002/qj.49712051603`.

Locatelli, John D., Mark T. Stoelinga, and Peter V. Hobbs (2005). "Re-Examination of the Split Cold Front in the British Isles Cyclone of 17 July 1980." In: *Quarterly Journal of the Royal Meteorological Society* 131.612, pp. 3167–3181. ISSN: 00359009. DOI: `10.1256/qj.04.157`. URL: `http://doi.wiley.com/10.1256/qj.04.157`.

Lorensen, William E. and Harvey E. Cline (1987). "Marching Cubes: A High Resolution 3D Surface Construction Algorithm." In: *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '87*. Vol. 21. New York, New York, USA: ACM Press, pp. 163–169. ISBN: 0-89791-227-6. DOI: `10.1145/37401.37422`. URL: `http://portal.acm.org/citation.cfm?doid=37401.37422`.

MacQueen, J. (1967). "Some Methods for Classification and Analysis of Multivariate Observations." In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. Vol. 5.1. University of California Press, pp. 281–298. URL: `https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fifth-Berkeley-Symposium-on-Mathematical-Statistics-and/chapter/Some-methods-for-classification-and-analysis-of-multivariate-observations/bsmsp/1200512992` (visited on 06/01/2023).

Madonna, Erica, Heini Wernli, Hanna Joos, and Olivia Martius (2014). "Warm Conveyor Belts in the ERA-Interim Dataset (1979–2010). Part I: Climatology and Potential Vorticity Evolution." In: *Journal of Climate* 27.1, pp. 3–26. ISSN: 0894-8755. DOI: `10.1175/JCLI-D-12-00720.1`. URL: `http://journals.ametsoc.org/doi/10.1175/JCLI-D-12-00720.1`.

Martínez-Alvarado, Oscar, Laura H. Baker, Suzanne L. Gray, John Methven, and Robert S. Plant (2014). "Distinguishing the Cold Conveyor Belt and Sting Jet Airstreams in an Intense Extratropical Cyclone." In: *Monthly Weather Review* 142.8, pp. 2571–2595. ISSN: 0027-0644. DOI: `10.1175/MWR-D-13-00348.1`. URL: `http://journals.ametsoc.org/doi/10.1175/MWR-D-13-00348.1`.

McCormick, Bruce H., Thomas A. DeFantis, and Maxine D. Brown (1987). *Visualization in Scientific Computing*.

Met Office (2023). *What Is an Ensemble Forecast?* Met Office. URL: https://www.metoffice.gov.uk/research/weather/ensemble-forecasting/what-is-an-ensemble-forecast (visited on 06/13/2023).

Met3d.documentation (2024). *Met.3D – Documentation: User Documentation*. URL: https://met3d.readthedocs.io/ (visited on 06/16/2022).

Met3d.homepage (2024). *Met.3D – Homepage: Interactive 3D Visualization of Meteorological Simulations*. URL: https://met3d.wavestoweather.de (visited on 06/16/2024).

Meyer, Marcel, Iuliia Polkova, Kameswar Rao Modali, Laura Schaffer, Johanna Baehr, Stephan Olbrich, and Marc Rautenhaus (2021). "Interactive 3-D Visual Analysis of ERA5 Data: Improving Diagnostic Indices for Marine Cold Air Outbreaks and Polar Lows." In: *Weather and Climate Dynamics* 2.3, pp. 867–891. ISSN: 2698-4016. DOI: 10.5194/wcd-2-867-2021. URL: https://wcd.copernicus.org/articles/2/867/2021/.

Miltenberger, Annet K., S. Pfahl, and H. Wernli (2013). "An Online Trajectory Module (Version 1.0) for the Nonhydrostatic Numerical Weather Prediction Model COSMO." In: *Geoscientific Model Development* 6.6, pp. 1989–2004. ISSN: 1991-9603. DOI: 10.5194/gmd-6-1989-2013. URL: https://gmd.copernicus.org/articles/6/1989/2013/.

Möller, Tomas and Ben Trumbore (1997). "Fast, Minimum Storage Ray-Triangle Intersection." In: *Journal of Graphics Tools* 2.1, pp. 21–28. ISSN: 1086-7651. DOI: 10.1080/10867651.1997.10487468. URL: https://doi.org/10.1080/10867651.1997.10487468 (visited on 05/24/2023).

"DTW-Based Motion Comparison and Retrieval" (2007). In: *Information Retrieval for Music and Motion*. Ed. by Meinard Müller. Berlin, Heidelberg: Springer, pp. 211–226. ISBN: 978-3-540-74048-3. DOI: 10.1007/978-3-540-74048-3_10. URL: https://doi.org/10.1007/978-3-540-74048-3_10 (visited on 05/26/2024).

Müller, Meinard, Henning Mattes, and Frank Kurth (2006). "An Efficient Multiscale Approach to Audio Synchronization." In:

Neiman, Paul J., M. A. Shapiro, and L. S. Fedor (1993). "The Life Cycle of an Extratropical Marine Cyclone. Part II: Mesoscale Structure and Diagnostics." In: *Monthly Weather Review* 121.8, pp. 2177–2199. ISSN: 0027-0644. DOI: 10.1175/1520-0493(1993)121<2177:TLCOAE>2.0.CO;2. URL: http://journals.ametsoc.org/doi/10.1175/1520-0493(1993)121%3C2177:TLCOAE%3E2.0.CO;2.

Nellikkattil, Arjun Babu, Danielle Lemmon, Travis Allen O'Brien, June-Yi Lee, and Jung-Eun Chu (2024). "Scalable Feature Extraction and Tracking (SCAFET): A General Framework for Feature Extraction from Large Climate Data Sets." In: *Geoscientific Model Development* 17.1, pp. 301–320. ISSN: 1991-959X. DOI: 10.5194/gmd-17-301-2024. URL: https://gmd.copernicus.org/articles/17/301/2024/ (visited on 05/07/2024).

Neu, Urs, Mirseid G. Akperov, Nina Bellenbaum, Rasmus Benestad, Richard Blender, Rodrigo Caballero, et al. (2013). "IMILAST: A Community Effort to

Intercompare Extratropical Cyclone Detection and Tracking Algorithms." In: *Bulletin of the American Meteorological Society* 94.4, pp. 529–547. DOI: `10.1175/BAMS-D-11-00154.1`. URL: `https://journals.ametsoc.org/view/journals/bams/94/4/bams-d-11-00154.1.xml` (visited on 07/19/2023).

Neuhauser, Christoph, Maicon Hieronymus, Michael Kern, Marc Rautenhaus, Annika Oertel, and Rüdiger Westermann (2023). "Visual Analysis of Model Parameter Sensitivities along Warm Conveyor Belt Trajectories Using Met.3D (1.6.0-Multivar1)." In: *Geoscientific Model Development* 16.16, pp. 4617–4638. ISSN: 1991-959X. DOI: `10.5194/gmd-16-4617-2023`. URL: `https://gmd.copernicus.org/articles/16/4617/2023/` (visited on 08/21/2023).

Niebler, Stefan, Annette K. Miltenberger, Bertil Schmidt, and Peter Spichtinger (2022). "Automated Detection and Classification of Synoptic-Scale Fronts from Atmospheric Data Grids." In: *Weather and Climate Dynamics* 3.1, pp. 113–137. ISSN: 2698-4016. DOI: `10.5194/wcd-3-113-2022`. URL: `https://doi.org/10.5194/wcd-2021-27`.

Niebler, Stefan, Bertil Schmidt, Holger Tost, and Peter Spichtinger (2023). "Automated Identification and Location of Three Dimensional Atmospheric Frontal Systems." In: *Computational Science – ICCS 2023*. Ed. by Jiří Mikyška, Clélia de Mulatier, Maciej Paszynski, Valeria V. Krzhizhanovskaya, Jack J. Dongarra, and Peter M.A. Sloot. Cham: Springer Nature Switzerland, pp. 3–17. ISBN: 978-3-031-36021-3. DOI: `10.1007/978-3-031-36021-3_1`.

Niennattrakul, Vit and Chotirat Ann Ratanamahatana (2007). "On Clustering Multimedia Time Series Data Using K-Means and Dynamic Time Warping." In: *2007 International Conference on Multimedia and Ubiquitous Engineering (MUE'07)*. 2007 International Conference on Multimedia and Ubiquitous Engineering (MUE'07). Seoul, Korea: IEEE, pp. 733–738. ISBN: 978-0-7695-2777-2. DOI: `10.1109/MUE.2007.165`. URL: `http://ieeexplore.ieee.org/document/4197360/` (visited on 05/26/2024).

NinJo Workstation: NinJo Workstation. URL: `http://www.ninjo-workstation.com/ninjo-workstation.0.html` (visited on 07/14/2023).

Nyberg, A., E. Palmén, and T. Bergeron (1942). "Synoptisch—Aerologische Bearbeitung Der Internationalen Registrierballonaufstiege in Europa in Der Zeit 17.–19. Oktober 1935." In: *Geografiska Annaler* 24.1-2, pp. 51–93. ISSN: 2001-4422. DOI: `10.1080/20014422.1942.11880709`. URL: `https://doi.org/10.1080/20014422.1942.11880709` (visited on 02/27/2023).

Oertel, Annika, Maxi Boettcher, Hanna Joos, Michael Sprenger, Heike Konow, Martin Hagen, and Heini Wernli (2019). "Convective Activity in an Extratropical Cyclone and Its Warm Conveyor Belt – a Case-Study Combining Observations and a Convection-Permitting Model Simulation." In: *Quarterly Journal of the Royal Meteorological Society* 145.721, pp. 1406–1426. ISSN: 1477870X. DOI: `10.1002/qj.3500`.

Oertel, Annika, Maxi Boettcher, Hanna Joos, Michael Sprenger, and Heini Wernli (2020). "Potential Vorticity Structure of Embedded Convection in a Warm Conveyor Belt and Its Relevance for Large-Scale Dynamics." In: *Weather and*

Climate Dynamics 1.1, pp. 127–153. ISSN: 2698-4016. DOI: 10.5194/wcd-1-127-2020. URL: https://wcd.copernicus.org/articles/1/127/2020/.

Orf, Leigh, Robert Wilhelmson, Bruce Lee, Catherine Finley, and Adam Houston (2017). "Evolution of a Long-Track Violent Tornado within a Simulated Supercell." In: *Bulletin of the American Meteorological Society* 98.1, pp. 45–68. ISSN: 0003-0007. DOI: 10.1175/BAMS-D-15-00073.1. URL: https://journals.ametsoc.org/doi/10.1175/BAMS-D-15-00073.1.

Pfaffelmoser, Tobias, Matthias Reitinger, and Rüdiger Westermann (2011). "Visualizing the Positional and Geometrical Variability of Isosurfaces in Uncertain Scalar Fields." In: *Computer Graphics Forum* 30.3, pp. 951–960. ISSN: 1467-8659. DOI: 10.1111/j.1467-8659.2011.01944.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2011.01944.x (visited on 06/01/2023).

Pfahl, Stephan and Heini Wernli (2012). "Quantifying the Relevance of Cyclones for Precipitation Extremes." In: *Journal of Climate* 25.19, pp. 6770–6780. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-11-00705.1. URL: https://journals.ametsoc.org/view/journals/clim/25/19/jcli-d-11-00705.1.xml (visited on 02/27/2023).

BBC News (2023). *Power Cuts and Schools Closed as Storm Otto Hits*. BBC News. URL: https://www.bbc.com/news/uk-scotland-64662300 (visited on 06/05/2023).

Quandt, Lisa-Ann, Julia H. Keller, Olivia Martius, and Sarah C. Jones (2017). "Forecast Variability of the Blocking System over Russia in Summer 2010 and Its Impact on Surface Conditions." In: *Weather and Forecasting* 32.1, pp. 61–82. ISSN: 1520-0434, 0882-8156. DOI: 10.1175/WAF-D-16-0065.1. URL: https://journals.ametsoc.org/view/journals/wefo/32/1/waf-d-16-0065_1.xml (visited on 08/02/2023).

Raible, Christoph C., Joaquim G. Pinto, Patrick Ludwig, and Martina Messmer (2021). "A Review of Past Changes in Extratropical Cyclones in the Northern Hemisphere and What Can Be Learned for the Future." In: *WIREs Climate Change* 12.1, e680. ISSN: 1757-7799. DOI: 10.1002/wcc.680. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.680 (visited on 08/17/2023).

Rand, William M. (1971). "Objective Criteria for the Evaluation of Clustering Methods." In: *Journal of the American Statistical Association* 66.336, pp. 846–850. ISSN: 0162-1459. DOI: 10.1080/01621459.1971.10482356. URL: https://www.tandfonline.com/doi/abs/10.1080/01621459.1971.10482356 (visited on 10/10/2023).

RandomInfinity17 (2023). *Track of Storm Otto of the 2022–23 European Windstorm Season*. URL: https://commons.wikimedia.org/wiki/File:Otto_2023_track.png (visited on 06/05/2023).

Rasp, Stephan, Tobias Selz, and George C. Craig (2016). "Convective and Slantwise Trajectory Ascent in Convection-Permitting Simulations of Midlatitude Cyclones." In: *Monthly Weather Review* 144.10, pp. 3961–3976. ISSN: 15200493. DOI: 10.1175/MWR-D-16-0112.1.

Rautenhaus, Marc (2015). “Interactive 3D Visualization of Ensemble Weather Forecasts.” Technische Universität München. URL: `https://mediatum.ub.tum.de/1253343`.

Rautenhaus, Marc, Michael Bottinger, Stephan Siemen, Robert Hoffman, Robert M. Kirby, Mahsa Mirzargar, Niklas Rober, and Rudiger Westermann (2018). “Visualization in Meteorology—A Survey of Techniques and Tools for Data Analysis Tasks.” In: *IEEE Transactions on Visualization and Computer Graphics* 24.12, pp. 3268–3296. ISSN: 1077-2626. DOI: `10.1109/TVCG.2017.2779501`. pmid: `29990196`. URL: `https://ieeexplore.ieee.org/document/8126857/`.

Rautenhaus, Marc, Christian M. Grams, Andreas Schäfler, and Rüdiger Westermann (2015a). “Three-Dimensional Visualization of Ensemble Weather Forecasts - Part 2: Forecasting Warm Conveyor Belt Situations for Aircraft-Based Field Campaigns.” In: *Geoscientific Model Development* 8.7, pp. 2355–2377. ISSN: 19919603. DOI: `10.5194/gmd-8-2355-2015`.

Rautenhaus, Marc, Michael Kern, Andreas Schäfler, and Rüdiger Westermann (2015b). “Three-Dimensional Visualization of Ensemble Weather Forecasts - Part 1: The Visualization Tool Met.3D (Version 1.0).” In: *Geoscientific Model Development* 8.7, pp. 2329–2353. ISSN: 19919603. DOI: `10.5194/gmd-8-2329-2015`.

Renard, Robert J. and Leo C. Clarke (1965). “Experiments in Numerical Objective Frontal Analysis.” In: *Monthly Weather Review* 93.9, pp. 547–556. ISSN: 0027-0644. DOI: `10.1175/1520-0493(1965)093<0547:einofa>2.3.co;2`.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox (2015). “U-Net: Convolutional Networks for Biomedical Image Segmentation.” In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 234–241. ISBN: 978-3-319-24574-4. DOI: `10.1007/978-3-319-24574-4_28`.

Rousseeuw, Peter J. (1987). “Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis.” In: *Journal of Computational and Applied Mathematics* 20, pp. 53–65. ISSN: 0377-0427. DOI: `10.1016/0377-0427(87)90125-7`. URL: `https://www.sciencedirect.com/science/article/pii/0377042787901257` (visited on 06/02/2023).

Rubner, Yossi, Carlo Tomasi, and Leonidas J. Guibas (2000). “The Earth Mover’s Distance as a Metric for Image Retrieval.” In: *International Journal of Computer Vision* 40.2, pp. 99–121. ISSN: 1573-1405. DOI: `10.1023/A:1026543900054`. URL: `https://doi.org/10.1023/A:1026543900054` (visited on 07/21/2023).

Russ, Reff, Glenn Davis, Steve Emmerson, Cathy Cormack, John Caron, Robert Pincus, Ed Hartnett, Dennis Heimbigner, Lynton Appel, and Ward Fisher (1989). “Unidata NetCDF.” In: *UCAR/NCAR - Unidata*. DOI: `10.5065/D6H70CW6`. URL: `http://www.unidata.ucar.edu/software/netcdf/` (visited on 07/21/2023).

Saikia, H. and T. Weinkauf (2017). “Global Feature Tracking and Similarity Estimation in Time-Dependent Scalar Fields.” In: *Computer Graphics Forum* 36.3, pp. 1–

11. ISSN: 1467-8659. DOI: 10.1111/cgf.13163. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.13163 (visited on 05/22/2024).

Saito, T., H.N. Miyamura, M. Yamamoto, H. Saito, Y. Hoshiya, and T. Kaseda (2005). "Two-Tone Pseudo Coloring: Compact Visualization for One-Dimensional Data." In: *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.* IEEE Symposium on Information Visualization, 2005. INFOVIS 2005. Pp. 173–180. DOI: 10.1109/INFVIS.2005.1532144.

Sakoe, H. and S. Chiba (1978). "Dynamic Programming Algorithm Optimization for Spoken Word Recognition." In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26.1, pp. 43–49. ISSN: 0096-3518. DOI: 10.1109/TASSP.1978.1163055.

Sakoe, Hiroaki and Seibi Chiba (1971). "A Dynamic Programming Approach to Continuous Speech Recognition | BibSonomy." In: Proceedings of the Seventh International Congress on Acoustics. Vol. 3. Budapest: Akadémiai Kiadó, pp. 65–69.

Sanders, Frederick and Charles A. Doswell (1995). "A Case for Detailed Surface Analysis." In: *Bulletin of the American Meteorological Society* 76.4, pp. 505–521. ISSN: 0003-0007. DOI: 10.1175/1520-0477(1995)076<0505:ACFDSA>2.0.CO;2. URL: http://journals.ametsoc.org/doi/10.1175/1520-0477(1995)076%3C0505:ACFDSA%3E2.0.CO;2.

Schemm, Sebastian, Michael Sprenger, and Heini Wernli (2018). "When during Their Life Cycle Are Extratropical Cyclones Attended by Fronts?" In: *Bulletin of the American Meteorological Society* 99.1, pp. 149–165. ISSN: 0003-0007. DOI: 10.1175/BAMS-D-16-0261.1. URL: https://journals.ametsoc.org/doi/10.1175/BAMS-D-16-0261.1.

Schroeder, Will, Ken Martin, and Bill Lorensen (2018). *The Visualization Toolkit An Object-Oriented Approach To 3D Graphics*. Edition 4.1.

Schultz, David M., Lance F. Bosart, Brian A. Colle, Huw C. Davies, Christopher Dearden, Daniel Keyser, et al. (2019). "Extratropical Cyclones: A Century of Research on Meteorology's Centerpiece." In: *Meteorological Monographs* 59, pp. 16.1–16.56. ISSN: 0065-9401. DOI: 10.1175/amsmonographs-d-18-0015.1.

Schultz, David M., Daniel Keyser, and Lance F. Bosart (1998). "The Effect of Large-Scale Flow on Low-Level Frontal Structure and Evolution in Midlatitude Cyclones." In: *Monthly Weather Review* 126.7, pp. 1767–1791. ISSN: 0027-0644. DOI: 10.1175/1520-0493(1998)126<1767:TEOLSF>2.0.CO;2. URL: http://journals.ametsoc.org/doi/10.1175/1520-0493(1998)126%3C1767:TEOLSF%3E2.0.CO;2.

Schultz, David M. and Geraint Vaughan (2011). "Occluded Fronts and the Occlusion Process: A Fresh Look at Conventional Wisdom." In: *Bulletin of the American Meteorological Society* 92.4, pp. 443–466. ISSN: 0003-0007. DOI: 10.1175/2010BAMS3057.1. URL: https://journals.ametsoc.org/doi/10.1175/2010BAMS3057.1.

Schäfler, Andreas, George C. Craig, Heini Wernli, Philippe Arbogast, James D. Doyle, Ron Mctaggart-Cowan, et al. (2018). "The North Atlantic Waveguide and Downstream Impact Experiment." In: *Bulletin of the American Meteorological Society* 99.8, pp. 1607–1637. ISSN: 00030007. DOI: `10.1175/BAMS-D-17-0003.1`.

Shahapure, Ketan Rajshekhar and Charles Nicholas (2020). "Cluster Quality Analysis Using Silhouette Score." In: *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*. 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), pp. 747–748. DOI: `10.1109/DSAA49011.2020.00096`.

Shapiro, M. A. and Daniel Keyser (1990). "Fronts, Jet Streams and the Tropopause." In: *Extratropical Cyclones* 1955, pp. 167–191. DOI: `10.1007/978-1-944970-33-8_10`.

Shields, Christine A., Jonathan J. Rutz, Lai-Yung Leung, F. Martin Ralph, Michael Wehner, Brian Kawzenuk, et al. (2018). "Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Project Goals and Experimental Design." In: *Geoscientific Model Development* 11.6, pp. 2455–2474. ISSN: 1991-959X. DOI: `10.5194/gmd-11-2455-2018`. URL: `https://gmd.copernicus.org/articles/11/2455/2018/` (visited on 07/19/2023).

Sinclair, Victoria A., Mika Rantanen, Päivi Haapanala, Jouni Räisänen, and Heikki Järvinen (2020). "The Characteristics and Structure of Extra-Tropical Cyclones in a Warmer Climate." In: *Weather and Climate Dynamics* 1.1, pp. 1–25. DOI: `10.5194/wcd-1-1-2020`. URL: `https://wcd.copernicus.org/articles/1/1/2020/` (visited on 08/17/2023).

Slffea and Mysid (2006). *An Example of an Unstructured Grid*. URL: `https://commons.wikimedia.org/wiki/File:Unstructured_grid.svg` (visited on 10/09/2023).

Steinley, Douglas (2006). "K-Means Clustering: A Half-Century Synthesis." In: *British Journal of Mathematical and Statistical Psychology* 59.1, pp. 1–34. ISSN: 2044-8317. DOI: `10.1348/000711005X48266`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1348/000711005X48266` (visited on 05/26/2024).

Stolfi, Jorge (2009). *Illustration of the Cartesian Coordinate System for 3D*. URL: `https://commons.wikimedia.org/wiki/File:Coord_system_CA_0.svg` (visited on 10/09/2023).

UK Met Office (2023). *Storm Otto Named by Danish Met Service*. UK Met Office. URL: `https://www.metoffice.gov.uk/about-us/press-office/news/weather-and-climate/2023/storm-otto-named-by-danish-met-service` (visited on 06/05/2023).

Stull, Roland B. (2017). *Practical Meteorology: An Algebra-Based Survey of Atmospheric Science*. Version 1.02b. Vancouver, BC, Canada: Dept. of Earth, Ocean & Atmospheric Sciences, University of British Columbia. ISBN: 978-0-88865-283-6.

Tavenard, Romain, Johann Faouzi, Gilles Vandewiele, Felix Divo, Guillaume Androz, Chester Holtz, et al. (2020). "Tslearn, A Machine Learning Toolkit for Time Series Data." In: *Journal of Machine Learning Research* 21.118, pp. 1–6. ISSN: 1533-7928. URL: `http://jmlr.org/papers/v21/20-091.html` (visited on 05/30/2023).

Telea, Alexandru (2015). *Data Visualization: Principles and Practice*. Second edition. Boca Raton: CRC Press, Taylor & Francis Group. 605 pp. ISBN: 978-1-4665-8526-3.

New VIS Conference (2021). *Things Are Changing in 2021: The New VIS Conference*. Things are Changing in 2021: The New VIS Conference. URL: `https://ieeevis.org/year/2024/blog/things-are-changing-2021` (visited on 05/25/2024).

Thomas, Carl M. and David M. Schultz (2019a). "Global Climatologies of Fronts, Airmass Boundaries, and Airstream Boundaries: Why the Definition of "front" Matters." In: *Monthly Weather Review* 147.2, pp. 691–717. ISSN: 15200493. DOI: `10.1175/MWR-D-18-0289.1`.

– (2019b). "What Are the Best Thermodynamic Quantity and Function to Define a Front in Gridded Model Output?" In: *Bulletin of the American Meteorological Society* 100.5, pp. 873–896. ISSN: 00030007. DOI: `10.1175/BAMS-D-18-0137.1`.

Wetterpate (2023). *Tiefdruckgebiete 2023*. Wetterpate - Tiefdruckgebiete 2023. URL: `https://www.wetterpate.de/namenslisten/tiefdruckgebiete/tief2023/index.html` (visited on 06/05/2023).

Ullrich, Paul A. and Colin M. Zarzycki (2017). "TempestExtremes: A Framework for Scale-Insensitive Pointwise Feature Tracking on Unstructured Grids." In: *Geoscientific Model Development* 10.3, pp. 1069–1090. ISSN: 1991-959X. DOI: `10.5194/gmd-10-1069-2017`. URL: `https://gmd.copernicus.org/articles/10/1069/2017/` (visited on 07/19/2023).

Ullrich, Paul A., Colin M. Zarzycki, Elizabeth E. McClenny, Marielle C. Pinheiro, Alyssa M. Stansfield, and Kevin A. Reed (2021). "TempestExtremes v2.1: A Community Framework for Feature Detection, Tracking, and Analysis in Large Datasets." In: *Geoscientific Model Development* 14.8, pp. 5023–5048. ISSN: 1991-959X. DOI: `10.5194/gmd-14-5023-2021`. URL: `https://gmd.copernicus.org/articles/14/5023/2021/` (visited on 08/14/2023).

Velichko, V. M. and N. G. Zagoruyko (1970). "Automatic Recognition of 200 Words." In: *International Journal of Man-Machine Studies* 2.3, pp. 223–234. ISSN: 0020-7373. DOI: `10.1016/S0020-7373(70)80008-6`. URL: `https://www.sciencedirect.com/science/article/pii/S0020737370800086` (visited on 05/26/2023).

Wallace, John M. and Peter Victor Hobbs (2006). *Atmospheric Science: An Introductory Survey*. 2nd ed. International Geophysics Series v. 92. Amsterdam ; Boston: Elsevier Academic Press. 483 pp. ISBN: 978-0-12-732951-2.

Wandel, Jan, Christina Wisotzky, Florian Pantillon, Bernhard Mühr, Florian Becker, Daniel Friederich, Jonas Straub, and Susanna Mohr (2018). *Wintersturm FRIEDERIKE*, pp. 1–12.

Wang, Junpeng, Subhashis Hazarika, Cheng Li, and Han-Wei Shen (2019). "Visualization and Visual Analysis of Ensemble Data: A Survey." In: *IEEE Transactions on Visualization and Computer Graphics* 25.9, pp. 2853–2872. ISSN: 1941-0506. DOI: `10.1109/TVCG.2018.2853721`.

Ward, Garry (2020). *Frontal Systems & Types of Fronts Found on Weather Charts*. Meteorology101. URL: https://meteorology101.com/fronts/ (visited on 06/16/2023).

Washington, W. M., B. T. O'Lear, J. Takamine, and D. Robertson (1968). "The Application of CRT Contour Analysis to General Circulation Experiments." In: *Bulletin of the American Meteorological Society* 49.9, pp. 882–889. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/1520-0477-49.9.882. URL: https://journals.ametsoc.org/view/journals/bams/49/9/1520-0477-49_9_882.xml (visited on 10/14/2023).

Wenande, Christian (2023). *The Copenhagen Post*. The Copenhagen Post. URL: https://cphpost.dk/2023-02-16/news/powerful-storm-bearing-down-on-denmark/ (visited on 06/05/2023).

Wetter Und Klima (2023). Wetter und Klima - Deutscher Wetterdienst - Numerical weather prediction models. URL: https://www.dwd.de/EN/research/weatherforecasting/num_modelling/01_num_weather_prediction_modells/num_weather_prediction_models_node.html (visited on 10/22/2023).

Wetter und Klima - Deutscher Wetterdienst - NinJo - Meteorologischer Arbeitsplatz. URL: https://www.dwd.de/DE/forschung/wettervorhersage/met_fachverfahren/met_arbeitsplatz/ninjo/ninjo_node.html (visited on 07/14/2023).

Whitaker, Ross T., Mahsa Mirzargar, and Robert M. Kirby (2013). "Contour Boxplots: A Method for Characterizing Uncertainty in Feature Sets from Simulation Ensembles." In: *IEEE Transactions on Visualization and Computer Graphics* 19.12, pp. 2713–2722. ISSN: 1941-0506. DOI: 10.1109/TVCG.2013.143. URL: https://ieeexplore.ieee.org/document/6634129 (visited on 06/11/2024).

Wilks, Daniel S. (2019). *Statistical Methods in the Atmospheric Sciences*. Fourth edition. Amsterdam, Netherlands ; Cambridge, MA: Elsevier. 818 pp. ISBN: 978-0-12-815823-4.

Williams, James G. and And Others (1995). "Visualization." In: *Annual Review of Information Science and Technology (ARIST)* 30, pp. 161–207. ISSN: 0066-4200.

Kachelmannwetter (2023). *Windböen 1std Messwerte Schottland vom 17.02.2023, 06:00 Uhr*. Kachelmann GmbH. URL: https://kachelmannwetter.com/de/messwerte/schottland/windboeen/20230217-0500z.html (visited on 06/05/2023).

XantaCross (2011). *Diffference in Matching between Euclidean and Dynamic Time Warping*. URL: https://commons.wikimedia.org/wiki/File:Euclidean_vs_DTW.jpg (visited on 05/30/2023).

Zwatz-Meise, V and G Mahringer (1988). "Use of Satellite Imagery, Combined with Numerical Model Diagnostics, to Locate Fronts and Predict Their Activity: Methods and Examples." In: Satellite and Radar Imagery Interpretation, Preprints for a Workshop in Reading, England, 20–24 July 1987. Vol. 143. Eumetsat, p. 162.

VERSICHERUNG AN EIDES STATT

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

Hamburg, June 2024

Andreas A. Beckert

Andreas Alexander Beckert