Analysis and evaluation of cold season characteristics and their future development: A case study for the region of Hamburg

Dissertation

zur Erlangung des Doktorgrades

an der Fakultät für Mathematik, Informatik und Naturwissenschaften

Fachbereich Erdsystemwissenschaften

der Universität Hamburg

vorgelegt von

Louisa Marie Bell

Hamburg, 2024

Fachbereich Erdsystemwissenschaften

Datum der Disputation: 16.10.2024

Gutachter/Gutachterinnen der Dissertation: Prof. Dr. K. Heinke Schlünzen
Dr. Kevin Sieck

Zusammensetzung der Prüfungskommission: Prof. Dr. K. Heinke Schlünzen
Dr. Kevin Sieck
Prof. Dr. Jörn Behrens
Prof. Dr. Bernd Leitl
Prof. Dr. Lars Kutzbach

Vorsitzender des Fach-Promotionsausschusses
Erdsystemwissenschaften: Prof. Dr. Hermann Held

Dekan der Fakultät MIN: Prof. Dr. Ing. Norbert Ritter

Zusammenfassung

In dieser Arbeit wird die Kältesaison für eine urbane Region charakterisiert, Eisepisoden definiert und deren zukünftige Entwicklung im Zusammenhang mit atmosphärischen Mustern untersucht. Kälteereignisse können insbesondere für Städte verheerende Folgen haben. Zu den Folgen zählen beispielsweise Stromausfälle, Verkehrsbehinderungen, Schließung öffentlicher Einrichtungen, Schäden an Straßen, Schienen, Gebäuden und Versorgungsleitungen sowie Ernteausfälle. Um deren zukünftige Entwicklung und die damit verbundenen Risiken einschätzen zu können, werden hier Zukunftsprojektionen von regionalen Klimamodellen (RCM) analysiert.

Da es bislang keine einheitliche Definition von Kälteereignissen gibt und in der Literatur häufig unterschiedliche Begriffe verwendet werden, wird in dieser Arbeit eine Definition für anhaltende Kälteereignisse mit Temperaturen unter dem Gefrierpunkt vorgeschlagen. Dafür wird der Begriff Eisepisode eingeführt, die Definition dafür lautet: Eine Eisepisode ist ein Ereignis von mehr als fünf aufeinander folgenden Eistagen. Ein Eistag ist ein sogenannter Schwellwert-basierter Klimaindex, bei dem die maximale Temperatur eines Tages unter dem Gefrierpunkt verbleibt. Der Schwellwert ist also 0°C. Um die Zuverlässigkeit der Schwellwert-basierten Klimaindizes zu beurteilen, musste zunächst ermittelt werden, inwieweit ihre Ergebnisse durch die Datenunsicherheit (z.B. Temperaturdaten) beeinflusst werden. Es konnte gezeigt werden, dass integrierende Indizes, die auf aufeinanderfolgenden Tagen und nicht auf der Zählung einzelner Tage beruhen, bei kleinen Änderungen des Schwellenwerts nur geringe Änderungen der resultierenden Anzahl von Tagen aufweisen. Die Kältesumme einer gesamten Kältesaison zeigt keine nennenswerten Änderungen und ist daher am wenigsten anfällig für Datenunsicherheit. Die hier definierten Eisepisoden sind weniger anfällig für Datenunsicherheit als die Gesamtanzahl der Eistagen.

Damit belastbare Aussagen über die Zukunft dieser Eisepisoden für die Zielregion gemacht werden können, wurde geprüft wie realistisch die Ergebnisse der RCMs für einen Evaluationszeitraum in der Vergangenheit sind. Außerdem wurden die während dieser Eisepisoden vorherrschenden atmosphärischen Zirkulationsmuster in Reanalysedaten identifiziert. Zusätzlich wurde überprüft, ob die RCMs die (thermo-)dynamischen Prozesse, die zu Eisepisoden in der Zielregion führen, generell richtig darstellen. Dafür wurde untersucht, inwieweit die RCMs die in der Reanalyse identifizierten atmosphärischen Zirkulationsmuster reproduzieren und ob diese dann auch mit Eisepisoden einhergehen. Dafür wurde eine Pattern-Matching Methode entwickelt und angewendet. Die Ergebnisse dieser Methode bestätigen, dass die hier analysierten RCMs generell in der Lage sind, die vorherrschenden Zirkulationsmuster zu reproduzieren und dass diese Muster mit Eisepisoden in der

Zielregion einhergehen.

Die zukünftige Entwicklung der Eisepisoden wurde unter der Annahme des Emissionsszenarios RCP8.5 gemacht. Dieses geht von hohen CO_2 Emissionen und damit von einem starken Anstieg der global gemittelten Temperatur aus. Dieses Szenario gibt somit eine Obergrenze für die mögliche Reduktion bzw. eine Untergrenze für das Auftreten von Eisepisoden in der Zukunft an. Es konnte gezeigt werden, dass Eisepisoden in Hamburg auch am Ende dieses Jahrhunderts noch auftreten können. Trotz anhaltender globaler Erwärmung werden diese Ereignisse mit mittleren Temperaturen von -4.6 °C einhergehen, wobei die Minimalwerte bis auf -13.2 °C absinken können. Ebenso konnte gezeigt werden, dass die in der Reanalyse identifizierten atmosphärischen Zirkulationsmuster in zwei Drittel aller Fälle mit einer Eisepisode in Zusammenhang gebracht werden können. Dieses Verhältnis bleibt nahezu konstant von der jüngsten Vergangenheit (1971–2000) über die nahe Zukunft (2030–2060) bis in die ferne Zukunft (2065–2095). Allerdings nimmt die absolute Anzahl der Eisepisoden in der fernen Zukunft im Vergleich zur jüngsten Vergangenheit um 80 % ab.

Die Ergebnisse dieser Dissertation zeigen, dass Eisepisoden trotz des starken globalen Temperaturanstiegs auch in Zukunft in den mittleren Breiten auftreten können und daher bei langfristigen Planungen (z.B. Planung von Anpassungsmaßnahmen für Städte und Infrastruktur) nicht vernachlässigt werden sollten.

Abstract

In this study, the cold season has been characterised for an urban area, ice episodes were defined and their future development was analysed in association with atmospheric circulation patterns. Cold weather events can have severe consequences, particularly for cities. These include power outages, traffic disruptions, the closure of public facilities, damage to roads, railways, buildings and supply lines as well as crop failures. In order to be able to assess the future development of these events and the associated risks, future projections from regional climate models (RCMs) are analysed here.

As there is no standard definition of cold weather events to date and different terms are often used in the literature, this paper proposes a definition for prolonged cold weather events with temperatures below the freezing point. For this the term ice episode is introduced and the definition is as follows: An ice episode is defined as an event with more than five consecutive ice days. An ice day is a so-called threshold-based climate index where the maximum temperature of a day remains below the freezing point. The threshold value is therefore 0°C. In order to assess the reliability of threshold-based climate indices, it was necessary first to determine the extent to which their results are affected by data uncertainty. It was demonstrated that integrating indices, which are based on consecutive days rather than counting single days, show only small changes in the resulting number of days for small changes in the threshold value. The coldsum of an entire cold season shows negligible changes and is hence least susceptible to data uncertainty. The ice episodes defined here are less susceptible to data uncertainty than the total number of ice days.

Prior to making robust statements about the future of these ice episodes for the target region, the reliability of RCM results was examined for an evaluation period in the past. Furthermore, the prevailing atmospheric circulation patterns during these ice episodes were identified in reanalysis data. Additionally, it was investigated whether the RCMs are able to reproduce the atmospheric circulation patterns identified in the reanalysis and whether these are then also associated with ice episodes. A pattern-matching method was developed and applied for this purpose. The results of this method confirmed that the RCMs analysed here are generally capable of reproducing the prevailing atmospheric circulation patterns and that these patterns are associated with ice episodes in the target region.

The future development of the ice episodes was analysed under the assumption of the RCP8.5 emissions scenario. This assumes high CO_2 emissions and therefore a large increase in global mean temperature. This scenario therefore sets an upper limit for the possible reduction and a lower lower limit for the occurrence of ice episodes in futuree.

It has been shown that ice episodes can still occur in Hamburg by the end of this century. Despite the continued global warming, these events will be accompanied by mean temperatures of -4.6 °C, with the minimum values dropping to -13.2 °C. It has also been shown that the atmospheric circulation patterns identified in the reanalysis can be associated with an ice episode in two-thirds of all cases. This ratio remains almost constant from the recent past (1971–2000) through the near future (2030–2060) to the far future (2065–2095). However, the absolute number of ice episodes in the far future decreases by 80 % compared to the recent past.

The results of this thesis show that mid-latitude ice episodes may still occur in the future despite a large global temperature increase and should therefore not be neglected in long-term planning (e.g. adaptation planning for cities and infrastructure).

List of publications

Bell, L. M. and K. H. Schlünzen and K. Sieck (2023): Influence of data uncertainty on cold season threshold-based climate indices. *Meteorologische Zeitschrift* 32(3), 195–206. http://dx.doi.org/10.1127/metz/2023/1158.

Bell, L. M. and K. H. Schlünzen and K. Sieck: Reproducibility of local cold season characteristics, ice episodes and prevailing circulation patterns in RCMs. *Quarterly Journal of the Royal Meteorological Society*, in review.

Bell, L.M. and K. H. Schlünzen and K. Sieck: Local cold season characteristics, ice episodes and their circulation patterns in a changing climate. Prepared to be submitted to *Quarterly Journal of the Royal Meteorological Society*.

Contents

x

List of Figures

List of Tables

1 Introduction

Climate change — one of the most pressing global challenges of our time — has far-reaching impacts on natural systems and human activities. Global warming caused by climate change leads to changes in atmospheric and oceanic circulation, affecting weather patterns and increasing the frequency and intensity of extreme events (INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC), 2021). In recent years, the impacts of climate change have already been felt in different parts of the world.

Looking at individual regions, it is clear that warming is progressing at different rates. Europe is one of the fastest warming mid-latitude regions in the world. The temperature increase in Europe since the pre-industrial era (1850–1900) is about 2.2 °C, which is about 1 °C above the global average (COPERNICUS CLIMATE CHANGE SERVICE (C3S), 2023, last updated on 20 April 2023). According to the *State of the Climate in Europe in 2022* report by the WORLD METEOROLOGICAL ORGANIZATION (WMO) (2022), the year 2022 was one of the warmest years in Europe. However, according to the same report, there were also several cold weather events: In December 2022, a cold wave hit northern and western Europe. Extremely low temperatures were recorded in Reykjavik, Iceland, making it the coldest December in the last 100 years. Athens, Greece, closed schools, public facilities, and private businesses in January due to remarkably low temperatures and heavy snowfall. The severe conditions left some households without power for days and caused transport disruption. Widespread frost in France, Germany, Spain and Austria caused agricultural losses (WORLD METEOROLOGICAL ORGANIZATION (WMO), 2022).

Cold weather events pose specific challenges for urban areas. They can result in a multitude of consequences: In addition to the closure of public facilities, power outages, and transport disruptions, as illustrated by the Athens example, extreme cold can also lead to an increased demand for heating and electricity, as described in RANASINGHE et al.

(IPCC report 2021, Working Group I, Chapter 12) and highlighted by STUIVENVOLT-ALLEN and WANG (2019). Furthermore, temperatures below the freezing point can cause water pipes to burst and damage roads, railway tracks, or buildings, as noted by UNDERWOOD et al. (2017). The occurrence of cold waves early in the winter season can impact human health, as individuals have not yet adapted to the cold weather (RANAS-INGHE et al., 2021, IPCC report, Working Group I, Chapter 12). Another effect is that of temperature-attributable deaths, with cold-related mortality exceeding heat-related mortality in urban areas, as highlighted by GASPARRINI et al. (2015); RYTI et al. (2016); CHENG et al. (2017). Prolonged cold weather events can increase the risk of hypothermia, which can lead to a range of cold-related health problems, particularly for the homeless and the elderly (BAUMGARTNER et al., 2008; WANG et al., 2016; LEI et al., 2022; RICHARD et al., 2023). Due to the demographic shift occurring in Europe, more elderly people will experience cold-related health problems and this will likely result in more winter-related deaths in the future, as discussed in RANASINGHE et al. (IPCC report 2021, Working Group I, Chapter 12). Although global temperatures are increasing, the frequency and intensity of cold waves are decreasing and will continue to do so in the future (SENEVIRATNE et al., 2021, IPCC report, Working group I, Chapter 11), it is still possible for cold weather events to periodically occur and affect urban areas and their infrastructure (DODMAN et al., 2023, IPCC report, Working Group II, Chapter 6). Consequently, cold extremes will remain a potentially hazardous factor for urban areas in the future.

More and more people are choosing to live in cities. The proportion of the urban population is increasing worldwide and it is expected that this trend will persist in the future (UNITED NATIONS, 2019). As extreme weather events will increase in frequency and intensity in the future, it is very important to be able to make reliable statements for urban regions, or to be able to make statements about the effects of such extreme events with or without adaptation measures. Studies on the future development of such extreme weather events and their impact on urban regions often deal with heavy precipitation and heat waves in the summer. There are fewer studies for the cold season. This may be due to the fact that for Northern hemisphere mid- and high-latitude regions it is not yet clear how a weakened north-south temperature gradient (due to faster warming of the Arctic compared to other regions) will influence cold extremes, as discussed in (SENEVIRATNE et al., 2021, IPCC report, Working Group I, Chapter 11). It is the subject of current research and is not yet certain. However, uncertainty regarding the occurrence of cold weather events does not rule out the possibility of such events happening in the future. Therefore, it is crucial to take into account the cold season when examining possible future developments.

In order to be able to examine the future development of cold weather events, such events are often defined by threshold-based climate indices. If, for example, the daily minimum temperature falls below 0 °C, this day is a frost day according to the definition proposed by the EXPERT TEAM ON CLIMATE CHANGE DETECTION AND INDICES (ETCCDI) (2020). For more extreme situations, the daily maximum temperature is used instead of the daily minimum temperature: If the daily maximum temperature does not exceed 0 °C, this day is, according to the correspondingly defined climate index, an ice day. These indices thus allow complex climate data to be translated into easily understandable, concrete quantities. An analysis of these indices then enables an objective assessment of changes over time. The definitions of the indices examined in this study are taken from the EUROPEAN CLIMATE ASSESSMENT AND DATASET (2024) climate index dictionary, which provides indices specifically for Europe.

1.1   Hamburg as target region

In order to investigate how cold weather events will change in the future for a European urban region, Hamburg was chosen as the target region. This medium-sized city in north-western Europe is situated within a warm temperate climate region, characterised by a fully humid climate with warm summers (KOTTEK et al., 2006). The city of Hamburg is considered as an example of a warm temperate climate region in the mid-latitudes, where cold weather events also occur. It is the largest city in northern Germany and, as a Hanseatic city on the River Elbe, it is an important hub. As mentioned above, there are several studies on extreme events in summer but few for winter. In the Hamburg climate report (VON STORCH et al., 2017), the word "Sommer" (summer) is used 54 times in the chapter *Stadtklima Hamburg* (urban climate of Hamburg SCHLÜNZEN et al., 2018), while the word "Winter" is used only 18 times. The word "Kälte" (cold) or the climate indices "Frosttag" and "Eistag" (frost day/ice day) are not mentioned at all. This shows that there is an imbalance between studies on summer and winter. There is a need for research on cold season characteristics in cities, including Hamburg, to close the research gap.

In a comprehensive study by SCHLÜNZEN et al. (2010), the analysis of meteorological in-situ data from standard sites, showed that the annual mean daily temperature increased significantly by 0.07 °C/decade (1891–2007). Extending the analysis up to 2023 (own analysis of station data from HH-Fuhlsbüttel provided by GERMAN METEOROLOGICAL SERVICE (2020)) shows that the linear temperature trend has increased from 0.07 °C/decade to 0.11 °C/decade. Despite this trend, temperatures below freezing in the cold season have not been uncommon over the last 30 years, as shown in Figure 1.1. For

example, the lowest value of daily minimum temperatures in the last 30 years of the analysed timeseries (1993–2022) was $-19.3\,°\text{C}$, which is lower than the value in 1891–1920 ($-18.4\,°\text{C}$, first 30 years of analysed timeseries). This suggests that temperatures around and below freezing might also continue to occur regularly until the end of this century. Against this background, the city of Hamburg is very well suited as an example for the planned investigations of the cold season characteristics for an urban area.
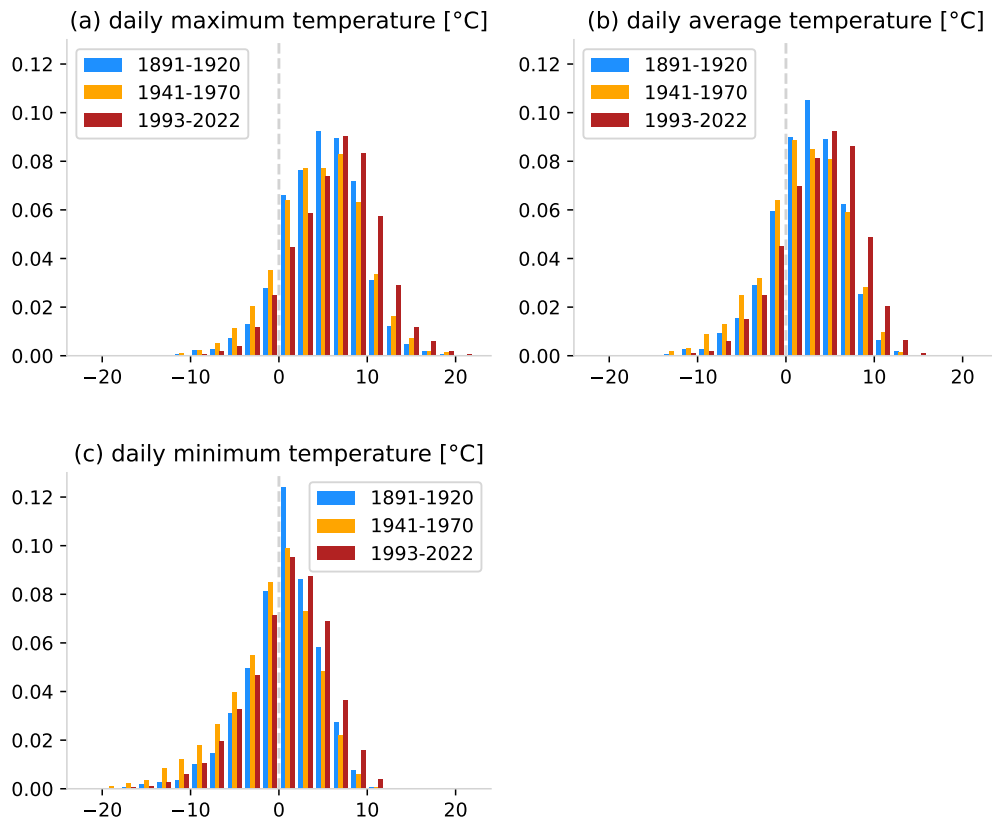
Figure 1.1. Histograms of daily (a) maximum, (b) average, (c) minimum temperatures in the cold season (November–March) for three different climate periods. Grey dashed vertical lines indicate the freezing point. Data are taken from station measurements at Hamburg-Fuhlsbüttel provided by GERMAN METEO-ROLOGICAL SERVICE (2020)

1.2   Research questions

In order to address the identified research gap, three guiding questions are formulated, each of which is addressed in a separate paper and presented here in the following three chapters.

For a general characterisation of the current cold seasons in Hamburg, a reference data set (derived from station data) can be used. To quantify the future development of cold season characteristics, long-term climate projections are required. These are typically

calculated using global climate models (GCMs) up to the year 2100, assuming different future scenarios. The focus in this thesis is on the RCP8.5 emission scenario (RIAHI et al., 2011), which projects the highest level of CO_2 emissions compared to the RCP2.6 (VAN VUUREN et al., 2011) or the RCP4.5 (SAMUELSSON et al., 2011) scenarios, and is associated with strong climate change. By analysing the high emission RCP8.5 scenario, an upper limit of possible changes can be determined.

GCMs developed for long-term climate projections have a too coarse resolution (grid size 50–100 km) to provide data that is sufficiently detailed for regional and local considerations. Therefore, they are refined in a dynamic process using regional climate models (RCMs). In the Coordinated Regional Downscaling Experiment (CORDEX GIORGI and GUTOWSKI, 2015) initiative of the WCRP, GCM data from the 5th Coupled Model Intercomparison Project (CMIP5 TAYLOR et al., 2012) have been refined using a number of different RCMs.

To assess how reliable statements about future cold seasons in Hamburg based on RCM data are, the climate model's ability to simulate cold season characteristics is evaluated. For this assessment, several climate indices for the cold season are calculated based on RCM data and the results are compared with reference data. However, the reference data, which can be reanalysis data or gridded data sets based on in-situ measurements, are not perfect representations of reality. It is therefore necessary to ask how "perfect" the RCMs should reproduce the results of the reference data, if the ground truth is not known. In order to determine the extent to which RCMs should be allowed to deviate from reference data, it is essential to investigate the influence of both reference data uncertainty and RCM data uncertainty on the results of the climate index calculation. The research question leading to the study presented in chapter 2 is

RQ1. How does data uncertainty affect cold season threshold-based climate indices?

For more impact-related information, the calculation of climate indices such as the number of frost days or the number of ice days is not sufficient, since single-day events are not as damaging as cold weather events of longer duration. As there is no uniform definition of cold weather events in the literature reviewed and different terms are often used (e.g. cold wave, cold spell, cold weather event, extreme cold event), a definition of prolonged cold events needs to be given first. The study presented in chapter 3 introduces the term ice episode and proposes a definition.

Against the background of the discussion about the influence of Arctic warming on winter weather in the mid-latitudes and the question of the future development of ice episodes in Hamburg, it is important to investigate the atmospheric conditions during these ice episodes. The question is whether there is a typical condition, i.e. a very typical atmo-

spheric circulation pattern, or whether the occurrence of ice episodes cannot be linked to specific atmospheric conditions at all.

In order to assess the future development of ice episodes using the RCM results for Hamburg, the relationship between these ice episodes and the atmospheric state in the RCMs must be examined. The question is whether the ice episodes in the RCMs can be associated with the same patterns as those in the reference data. If a different relationship is found in the RCMs, this would indicate that the relevant physical processes are insufficiently represented in the climate models and that no reliable statements can be made for the future. Whether it is possible for the RCMs to reproduce the relationship identified in the reanalysis data is investigated in the work presented in Chapter 3 and is guided by the following research question

RQ2. Can regional climate models reproduce local cold season characteristics, ice episodes and prevailing atmospheric patterns?

It is questionable whether ice episodes will occur at all in the future due to global warming. In particular, under the assumption of RCP8.5 as the upper limit of possible changes, it is of interest to see whether global warming influences the amospheric dynamics to such an extent that ice episodes no longer occur in the mid-latitudes, or whether the air in the Arctic warms up to such an extent that cold air outbreaks from the Arctic are no longer cold enough to lead to an ice episode in Hamburg. It is important to investigate whether ice episodes will continue to play a role in the future and whether the relationship between ice episodes and atmospheric conditions will change. These aspects are explored in chapter 4 and are guided by the following research question

RQ3. How will the frequency of local ice episodes and associated atmospheric patterns change in future climate periods?

Chapter 5 provides answers to the research questions, discusses the limitations of this work, recommends future research directions, and ends with some closing words.

2

Influence of data uncertainty on cold season threshold-based climate indices

Preface

This chapter is published as:

The layout and the numbering of the manuscript were adopted to fit this thesis. The references are combined with all references used in this thesis and can be found at the end of this thesis in References. The acknowledgments and the supporting information that were published together with this manuscript can be found in a separate section of the Appendix (First paper).

K. Heinke Schlünzen has contributed to the conceptualisation. K. Heinke Schlünzen and Kevin Sieck have contributed to the discussion of the results.

Abstract

Climate indices are used to reduce the complex climate system and its changes to simple measures. The data basis — whether observational data or climate model data — to which the climate indices are applied, is usually subject to uncertainties. For threshold-based climate indices, the data uncertainty influences the threshold value, and, hence, the uncertainty can influence the values for the climate index. What the actual impacts of these uncertainties are on threshold-based climate indices is examined in this paper. The focus is not only on the climate model uncertainty, but also on the observational data uncertainty. The general sensitivity of each of the chosen climate indices to arbitrary changes in the threshold is studied. This shows a higher sensitivity of indices assessing extremes (ice days, heavy precipitation days) to changes in the threshold than indices that integrate a quantity over a given time interval (coldsum, consecutive days). For assessing an ensemble of climate model data with respect to their ability to reproduce the index values for current climate, the reference data uncertainty is applied to the chosen threshold-based climate indices by changing their threshold value by its corresponding uncertainty. It is shown that the climate model uncertainty can be within the range of the reference data uncertainty. When using threshold-based climate indices to assess changes in future climate periods, uncertainties should always be taken into account and ideally corrected in an appropriate way. This is especially important for indices that assess extremes.

2.1   Introduction

The climate system is of a chaotic nature and is dominated by complex spatial and temporal variability. With the help of climate indices, this complexity can be reduced to simple measures (e.g. STENSETH et al., 2003). Climate indices are diagnostic tools that enable one to monitor and describe the state and the changes of the climate system. Examples are mean temperatures or precipitation sums; these are simple climate indices from which essential information on the state of the climate system and its changes can be derived. Nevertheless, these indices give little insight into changes in the frequency of specific meteorological situations. To further investigate, understand, and quantify these changes, related climate indices are used.

The CCI/CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETC-CDI) recommends a core set of 26 climate indices, the ECA&D Indices Dictionary (ECA&D) provides another 50 indices specifically for Europe. These indices cover a wide range of areas of application as well as complexity and spatial/temporal scales. Large-scale atmospheric circulation climate indices, such as ENSO, NAO, AO, SOI, or PDO, are used to better understand local ecological effects (STENSETH et al., 2003; HALLETT et al., 2004). Small-scale and more complex indices include, for example, the Tourism Climatic Index (MIECZKOWSKI, 1985; DUBOIS et al., 2016) or the Universal Thermal Climate Index (BRÖDE et al., 2012; JENDRITZKY et al., 2012), one of the many human thermal indices (DE FREITAS and GRIGORIEVA, 2017; FISCHEREIT and SCHLÜNZEN, 2018). In agriculture, relevant indices are, for example, growing degree days, growing season length, and first fall and last spring frost day (e.g. KUKAL and IRMAK, 2018). These agro-meteorological indices are less complex than the other aforementioned indices, since they are based on only one climate variable — in this case near, surface air temperature. In addition, they are based on fixed thresholds: A day where the minimum temperature is below 0 °C is defined as frost day.

One important application of climate indices is in climate change studies. For many areas of society, the knowledge about future changes of those indices is crucial in order to adapt properly to climate change. As we cannot get this information from measurements, we have to use climate models to assess future changes in climate indices. Because climate models only represent the real world in simplified form, they are subject to uncertainty (TEBALDI and KNUTTI, 2007). A common way to deal with, and quantify, climate model uncertainty is to use a multi-model ensemble (TEBALDI and KNUTTI, 2007; KNUTTI et al., 2010; HAWKINS and SUTTON, 2011). Nevertheless, an important step before using climate models to assess climate change is the evaluation of such models. This is typically done by comparing climate models to observations during a

historical/reference period (GLECKLER et al., 2008; CADULE et al., 2010; KOTLARSKI et al., 2014; KATRAGKOU et al., 2015).

A particular challenge that arises when comparing climate models and observations is the mismatch in the spatial and/or temporal resolution of both data-sets. On the one hand, model based data is typically given on some sort of grid, where the actual value represents a spatial and temporal mean/sum for a box of the grid. On the other hand, observations are often taken from a single point in space and time. To overcome this mismatch, observations are often extrapolated to a grid in order to make the comparison between climate models and observations easier.

The process of extrapolating measurement data to a gridded observational data-set involves several sources of uncertainty. As ZUMWALD et al. (2020) have shown, there are generally three types of uncertainties in these data-sets: (1) The uncertainty that arises in the generation of such data-sets, e.g., on account of the statistical model that is used. (2) The uncertainty that arises because of biased samples, e.g., from inaccurate measurements. (3) The uncertainty that occurs on account of the choice of abstract properties, such as the resolution or a particular metric. To get a measure of uncertainty in gridded observational data-sets, the ensemble approach can also be adopted here. For example, according to ZUMWALD et al. (2020), there are structural ensembles (equivalent to the multi-model ensemble for climate models) where several observational data-sets are considered. Another is the parametric ensemble (equivalent to a perturbed-physics ensemble for climate models), where multiple realisations of a data-set are generated. One such data-set is E-OBS, which is used for this study.

Being aware of all these sources of uncertainty, it quickly becomes clear that fixed thresholds cause a problem. The data — whether measurement data at specific sites, gridded observational data, or climate model data — are subject to uncertainties on account of the various reasons mentioned above. For the calculation of frost days where the threshold value is 0 °C, it is possible that the temperature value of 0 °C in the data does not correspond to 0 °C in the "real world". Since impact studies often consider these threshold-based indices, it is very important to take uncertainties of climate models, as also of observations, into account, so to make these studies more transparent and thereby more reliable.

The present paper, therefore, focuses on the impact of the data uncertainty on calculated values of threshold-based climate indices and on the impact of data uncertainty on the comparison between gridded observational data and climate model data.

One way to overcome climate model uncertainty is bias correction. HOFFMANN et al. (2018) presented a method, wherein thresholds of several climate indices were bias-

corrected for an ensemble of climate models. This means that each climate model is assigned its own threshold value, depending on how large the bias is relative to the reference (gridded observational) data. We have applied this method to the climate model data used in this work. Nevertheless, the focus of this paper is on the influence of data uncertainty on threshold-based climate indices, not on the correction of any data. A full analysis of the results obtained using the bias-correction method is, therefore, beyond the scope of this paper and will not be shown.

Here, we present and apply a new method to estimate the influence of data uncertainty on a set of threshold-based climate indices. In the present study, we answer the following questions: 1) How can data uncertainty be incorporated in threshold-based climate indices? 2) What is the influence of data uncertainty on threshold-based climate indices in current and 3) future climates under different emission scenarios?

For the application of our method, we focus on the cold season. Temperatures close to the freezing point occur frequently within the cold season and lead to challenging situations, e.g. icy roads. How often these situations occur in a future climate is of relevance not only for urban planning but also for e.g. infrastructure maintenance in the future. As an example of its use, the new method is applied to the north-western European city of Hamburg (Germany) and its rural surroundings, an area of $\sim 70 \times 70$ km^2. Currently, daily minimum temperatures of 0 °C are of high frequency in Hamburg during the cold season. SCHLÜNZEN et al. (2010) showed that winter temperatures in this city have already increased. A warming, therefore, can have a strong influence on e.g. frost duration.

This paper is structured as follows: This introduction (sec. 2.1) is followed by sec. 2.2, with information on the data used. The description of our method as well as some necessary data preparation steps can be found in sec. 2.3. The application of our method for the area of interest is presented in sec. 2.4. Conclusions are drawn in sec. 2.5.

2.2   Database

2.2.1   Climate model data

For the analysis, the results of regional climate models (RCMs) are used. These are taken from the Coordinated Downscaling Experiment (CORDEX, GIORGI and GUTOWSKI (2015)), at 0.11° horizontal resolution for the European domain (EURO-CORDEX, JACOB et al. (2014)). An ensemble of model results is available, provided by the regional climate modeling community. They downscaled global climate model (GCM) simulations from CMIP5 (TAYLOR et al., 2012) for different climate scenarios. These scenarios

Table 2.1. Global climate model (GCM) - regional climate model (RCM) combinations, used GCM ensemble member (Ens Mem), modeling group providing the data and calendar of each GCM-RCM result.

GCM	RCM	Ens Mem	Modeling Group	Calendar
EC-EARTH	CCLM4-8-17	r12i1p1	CLM-Community	Proleptic Gregorian
	REMO2015	r12i1p1	GERICS	Proleptic Gregorian
	RACMO22E	r12i1p1	KNMI	standard
	RCA4	r12i1p1	SMHI	standard
	HIRHAM5	r3i1p1	DMI	Proleptic Gregorian
MIROC5	REMO2015	r1i1p1	GERICS	Proleptic Gregorian
HadGEM2	REMO2015	r1i1p1	GERICS	360 day
	RACMO22E	r1i1p1	KNMI	360 day
	RCA4	r1i1p1	SMHI	360 day
MPI-ESM-LR	REMO2009	r1i1p1	MPI-M and CS2.0	Proleptic Gregorian
	REMO2009	r2i1p1	MPI-M and CS2.0	Proleptic Gregorian
	RCA4	r1i1p1	SMHI	Proleptic Gregorian
	CCLM4-8-17	r1i1p1	CLM-Community	Proleptic Gregorian
	WRF361H	r1i1p1	UHOH	standard

start in 2006 and last until 2099 or 2100, depending on the duration of the GCM simulation. In the present paper, the years 1971–2000 are chosen as the reference period for comparison with observational (reference) data (sec. 2.2.2), using historical runs from the very same RCM-GCM combinations.

For future changes, the results of two climate model ensembles are analysed; one ensemble is based on the representative concentration pathway (RCP) 2.6 scenario (VAN VUUREN et al., 2011) and one ensemble is based on the RCP8.5 scenario (RIAHI et al., 2011). The two climate model ensembles comprise the results of four different GCM simulations downscaled by seven RCMs. The reason for choosing this particular ensemble of models is that these simulations were available for both RCP scenarios at the time we started our research (January 2019). The model simulation-combinations used are given in Table 2.1. The ensemble member (Ens Mem) indicates the realisation of the GCM; there may be more than one simulation from one GCM. The rip-nomenclature (e.g. r1i1p1) indicates which realisation (r: could be started from different initial states), which initialisation (i: could be started with different initialisation methods), and which physics (p: could be started with different physics details, such as parameterisation constants) were used.

2.2.2   Reference data

As reference E-OBS, a gridded daily data-set produced by interpolating station data is used. E-OBS is used in its ensemble version for temperature (daily minimum, mean, and maximum values) as well as precipitation (v20.0e, CORNES et al. (2018)). This gridded data-set is frequently used as reference for the evaluation of EURO-CORDEX data (e.g. JACOB et al., 2014; KOTLARSKI et al., 2014; HOFFMANN et al., 2018). E-OBS gridded daily fields have a spatial resolution of 0.1°.

The ensemble version consists of 100 ensemble members for each daily field to capture the uncertainty range of the statistical model used for interpolation. The ensemble mean is provided as "best guess" fields, while the ensemble spread is calculated as the difference between the 5th and 95th percentiles of the ensemble. Hence, the ensemble version of E-OBS provides a measure which indicates the 90 % uncertainty range. This uncertainty range relates to the interpolation uncertainty, which is closely related to station density. It is used and denoted in the present paper as reference data uncertainty (sec. 2.3.3).

2.3   Calculation of cold season climate indices and their uncertainty

For the analysis, we concentrate on the time period from November 1971 to March 2100. Here, the days between the 1st of November and the 31st of March of the following year are defined as the cold season. We hereby follow the German Meteorological Service (DWD), which defined the coldsum, one of the climate indices we use, for this period. As shown in Table 2.1, the different GCM-RCM combinations have different calendars: Proleptic Gregorian or standard with 365/366 days per year including leap years and a 360-day calendar where every month has 30 days. To have a consistent ensemble of model simulations every 31st of December, January and March as well as every 29th of February (if existing) was deleted. The same was applied to the reference data-set. Hence, a cold season contains 148 instead of 151/152 days.

2.3.1   Preparation of data for area of interest

As mentioned in the introduction, one aim of the present study is to investigate climate changes by means of threshold-based climate indices. Here, the area of interest is the region of Hamburg, located in the northern part of Germany (Fig. 2.1). Hamburg experiences a mild maritime climate with average (1971–2000) temperatures of 2.9 °C and average minimum temperatures around zero (0.1 °C) in the cold season (Station data for

Hamburg-Fuhlsbüttel accessed via OpenData of DWD). SCHLÜNZEN et al. (2010) investigated long-term changes in temperature and precipitation based on station data (1891–2007). They found a significant annual mean temperature increase by 0.7 °C/century. The winter (December, January, February) temperature increase — based on climate periods 1891–1917 to 1978–2007 — was found to be 0.6 °C. It is, therefore, of great interest to investigate cold season changes especially for climate indices with threshold values close to the average temperature in the reference period.
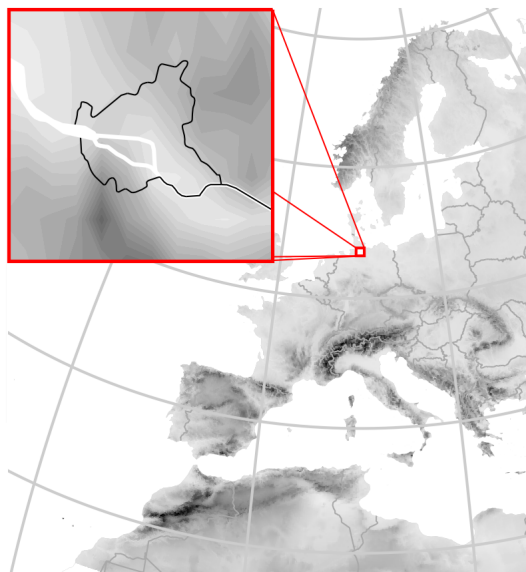
Figure 2.1. Location of area of interest, the city of Hamburg and its rural surroundings.

The area of interest lies at 53.25 °N – 53.85 °N and 9.48 °E – 10.53 °E. It covers the area of Hamburg and its surroundings and is located in the north-western part of Europe (Fig. 2.1). Climate model data and reference data were not interpolated onto a common grid for the comparison, but each data-set was analysed in its native grid to preclude the introduction of further uncertainties by additional interpolation or aggregation. Since all data were selected for the same domain but each data-set has a different spatial resolution, the data-sets for the area of interest consist of a different number of grid points: Climate model data = 36 grid points, reference data = 60 grid points.

2.3.2   Cold season threshold-based climate indices

The calculation of the threshold-based climate indices (Table 2.2) is based on mean values of the relevant variables for the area of interest. The calculation was also tested for extreme values within the domain (not shown), but since reference and climate model data have different spatial resolutions, differences between the data-sets were large and, so, the interpretation of the results was not straightforward.

Table 2.2. Definition of used threshold-based climate indices. Squared brackets in the calculation stand for the Iverson bracket that converts any logical proposition into the numbers 1 or 0. n: total number of days within a cold season; i,j:first and last day of a period of consecutive frost days or ice days. mod. = moderately.

Index	Variable	Threshold	Calculation
frost days[1]	T_{min}	0 °C	$\sum_{d=1}^{n}[T_{min,d} < 0\ °C]$
mnoc frost days[1]	T_{min}	0 °C	$\max(\sum_{i}^{j}[T_{min,i} < 0\ °C])$
ice days[1]	T_{max}	0 °C	$\sum_{d=1}^{n}[T_{max,d} < 0\ °C]$
mnoc ice days[1]	T_{max}	0 °C	$\max(\sum_{i}^{j}[T_{max,i} < 0\ °C])$
coldsum[2]	T_{avg}	0 °C	$\sum_{d=1}^{n}|T_{avg,d}|[T_{avg,d} < 0\ °C]$
wet days[1]	pr	1 mm	$\sum_{d=1}^{n}[pr_d \geq 1\ mm]$
heavy pr days[1]	pr	10 mm	$\sum_{d=1}^{n}[pr_d \geq 10\ mm]$
coldsum categories:	< 98	mild	
	98 to 196	mod. warm	
	196 to 294	mod. cold	
	> 294	harsh	

[1] Definition from ECA&D

[2] Definition from DWD

Values for the threshold-based climate indices used in this work are calculated as the sum of days over a cold season. The number of frost days is the sum of all days per cold season with a minimum temperature (T_{min}) value below 0 °C (Table 2.2). Frost days that occur one after another are consecutive frost days. Of the numbers of consecutive frost days, the maximum number of consecutive (mnoc) frost days is found per cold season. The same procedure applies to the number of ice days and the maximum number of consecutive (mnoc) ice days, using daily maximum temperature (T_{max}) values.

The coldsum is calculated from daily average temperatures (T_{avg}) as the sum of all absolute values of negative temperatures within a cold season. Based on the coldsum, the cold season falls into one of four defined categories: Mild, moderately warm, moderately cold, or harsh. Here, a cold season contains 148 instead of 151/152 days; hence, the category thresholds given in Table 2.2 are used. They are adjusted from the original category thresholds provided by DWD.

Indices derived from daily precipitation (pr) amounts are the number of wet days and the number of heavy precipitation days per cold season. A day on which the precipitation amount is equal to, or exceeds, 1 mm (10 mm) is considered a wet day (heavy precipitation day).

2.3.3 Determination of data uncertainty

Reference data uncertainty

As mentioned in sec. 2.2.2, the reference data-set is based on an ensemble of 100 calculations; it consists of an ensemble mean ("best guess") and an ensemble spread value indicating the 90 % ensemble range. Here, the latter is used as the estimate for the reference data uncertainty; it is denoted as uncertainty hereafter for simplicity. For each threshold used in the indices of Table 2.2, the uncertainty of that threshold value is determined by calculating the median uncertainty across all days and all cold seasons (November to March) within the reference time period 1971–2000.
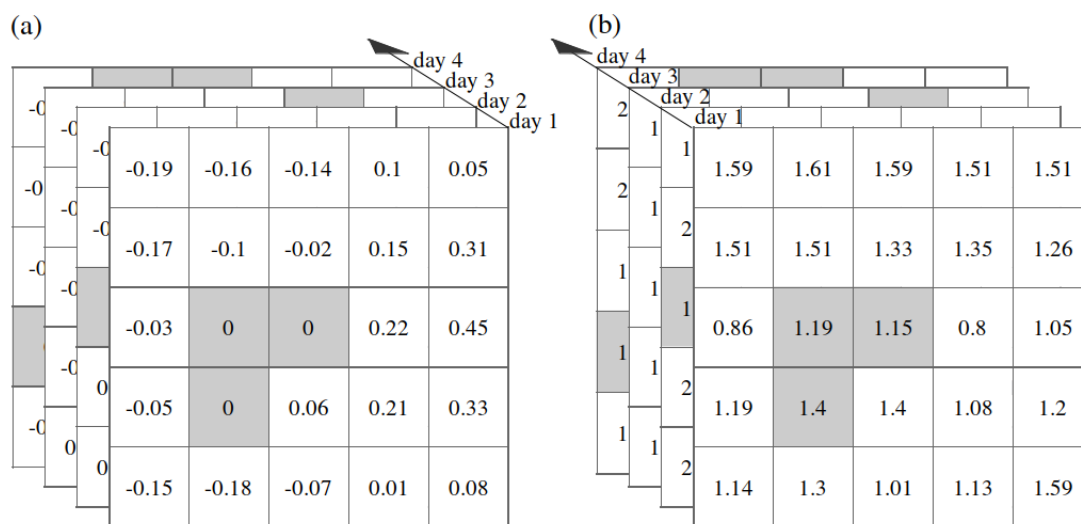
Figure 2.2. Illustration of (a) E-OBS "best guess" field and (b) E-OBS 90 % ensemble range (uncertainty) field of the minimum daily temperature (T_{min}) for four arbitrary days. Grey marked cells indicate (a) where $T_{min} = 0$ °C and (b) the corresponding uncertainty.

This procedure for estimating the uncertainty is not based on area averages, since the aim is to derive an uncertainty estimate for the exact threshold (e.g. 0.0 °C). By averaging over the area of interest, one would loose the exact threshold value that is necessary to determine the uncertainty estimate. Therefore, each grid cell within the area of interest is considered separately: In the "best guess" field, all grid cells are considered in which the value is equal to the threshold sought. This is depicted in Fig. 2.2 as an example for frost days for which the threshold in daily minimum temperatures is 0 °C (grey in Fig. 2.2 (a)). The median uncertainty is then found from all the grid cells considered in the uncertainty field (grey in Fig. 2.2 (b)) of all days within the reference period.

The median uncertainty was also separately calculated for cold seasons within the entire

Table 2.3. E-OBS (reference data) uncertainty for cold season threshold-based climate indices. pr: precipitation.

Index	frost day	ice day	coldsum	wet day	heavy pr day
Variable	T_{min} [°C]	T_{max} [°C]	T_{avg} [°C]	pr [mm/day]	pr [mm/day]
Threshold	0	0	0	1	10
Uncertainty	± 0.86	± 0.76	± 0.64	± 0.8	± 2.8

reference data time period (January 1950 to July 2019), 10-year time slices, and for each cold season month to investigate the variability and stationarity of the uncertainty estimate. No detailed result is presented here. Across the months, no variability was found. However, the last two 10-year time slices showed a slightly larger spread than the time slices between 1950 and 2000. This might be on account of a change in station density over time. Therefore, the reference time period 1971 to 2000 was selected without the last two decades available in the data-set. Across the period 1971–2000, the median uncertainty was calculated for T_{min}, T_{max}, T_{avg} at 0 °C and for pr at 1 and 10 mm/day. The calculated median uncertainty determined for each threshold is further used in the calculation of the climate indices. More precisely, under the assumption that the uncertainty is normally distributed, half of the uncertainty is added to and half of the uncertainty is subtracted from the original threshold for a sensitivity study. These uncertainty values can be found in Table 2.3.

Model data uncertainty

As stated in previous sections, the EURO-CORDEX multi-model ensemble is used as model data in this work. The ensemble consists of 14 different simulations, each based on a different GCM-RCM model combination (Table 2.1). To assess the ensemble uncertainty, a common method is to define the bandwidth of the results (e.g. DÉQUÉ et al., 2007). Here, we define the model bandwidth as the ensemble spread between the 5th and the 95th percentiles. In addition, we analyse the ensemble median (instead of the ensemble mean) to preclude possible outliers from having too strong an influence on the result.

2.4 Results

2.4.1 Sensitivity of threshold-based climate indices to chosen threshold

The simplest way to include data uncertainty in the calculation of threshold-based climate indices is to vary the threshold by the uncertainty.

As a first step, the general sensitivity of the climate index values to changes in the threshold is assessed. Hence, the number of days determined by each of the threshold-based climate indices is calculated for several different thresholds above and below the originally defined threshold (Table 2.2) for all cold season days within the reference period. The total number of cold season days within the reference period is 4,440 days.

The total numbers of days calculated within the reference period are compared to the total number of days calculated with the originally defined threshold. This dependency is depicted in Fig. 2.3 for frost days and ice days as the percentage change in the number of days as a function of the selected threshold relative to the number of days calculated with the originally defined threshold.
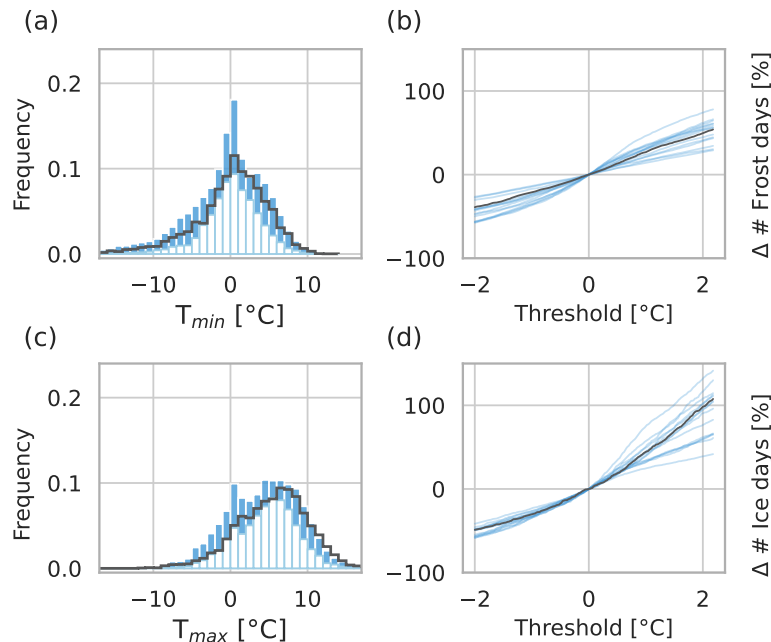
Figure 2.3. Distribution of (a) daily minimum and (c) daily maximum temperature between 1971-2000. Percentage changes in the number of (b) frost days and (d) ice days are shown as function of the threshold. Blue: 5th to 95th percentile of climate model ensemble (a,c), all 14 model results (b,d); black: reference data.

When changing the threshold for frost days from 0 °C to values smaller than zero or values greater than zero, changes in the number of frost days are apparent and of similar

size for increased and decreased thresholds (Fig. 2.3b). When changing the threshold for frost days to e.g. −2 °C and 2 °C, relative changes in the number of days are around ±45 % which translates to (on average) ±30 days per cold season.

Doing the same for ice days reveals a higher sensitivity towards thresholds greater than the originally defined one: An increase in the threshold to 2 °C leads to an increase in the number of ice days of nearly 100 % (+18 days). A decrease in the threshold to −2 °C leads to a decrease of 50 % (−9 days).

The difference in the sensitivity between frost days and ice days is caused by the distance between the position of the threshold and the mode of the underlying temperature distribution. Fig. 2.3a, c show the distributions for daily minimum and daily maximum temperatures for the area of interest. If the originally defined threshold is in accordance to the mode of the underlying distribution, a change in the threshold in both directions causes changes in the number of days of similar magnitude. This is the case for frost days (Fig. 2.3a, b). If the threshold is not at the mode of the underlying distribution, a change towards the mode causes a far greater change in the number of days than a change away from the mode. Hence, the relative change curve would increase until the mode of the underlying distribution is reached.

Changes in the coldsum caused by changes in the threshold are small compared to the other climate indices (Appendix A, Fig. A.1). When changing the threshold to −2 or 2 °C, the coldsum changes by ±15 %.

The precipitation-based indices show a high sensitivity towards a decrease in the threshold: When decreasing the threshold for wet days by 1 mm/day, the number of wet days increases by 149 % (+90 days) (Appendix A, Fig. A.2). An increase by 1 mm/day results in a decrease in the number of wet days by 23 % (−13 days). For heavy precipitation days, a decrease in the threshold from 10 mm/day to 5 mm/day results in an increase in the number of days by +223 % (+15 days) whereas an increase by 5 mm/day results in a decrease in the number of days by 71 % (−5 days). This behaviour can again be explained by the underlying frequency distribution: the precipitation distribution is skewed with frequent light precipitation events while heavy precipitation events are rare (SCHLÜNZEN et al., 2010).

This sensitivity study shows that the number of days derived from a threshold-based climate index strongly depends on the underlying distribution of the climate variable and on the distance between the threshold and the mode. A larger distance between the threshold and the mode leads to a larger percentage change in the number of days when the threshold is changed (e.g. ice days vs frost days). A change in the threshold towards the mode generally leads to a larger percentage change in the number of days than a

change away from the mode.

2.4.2 Consideration of data uncertainty in determined values of threshold-based climate indices

To not just take an arbitrary number for the uncertainty of the thresholds, the uncertainty estimates derived from the reference data (sec. 2.3.3) are used to vary the threshold of each climate index. As an example, for frost days, the varied thresholds result to T^+_{min} = 0.86 °C and T^-_{min} = −0.86 °C (Table 2.3). The number of frost days based on the originally defined threshold is compared with the number of frost days based on the varied thresholds. The number of days per cold season is calculated between 1971 and 2000 and then averaged.
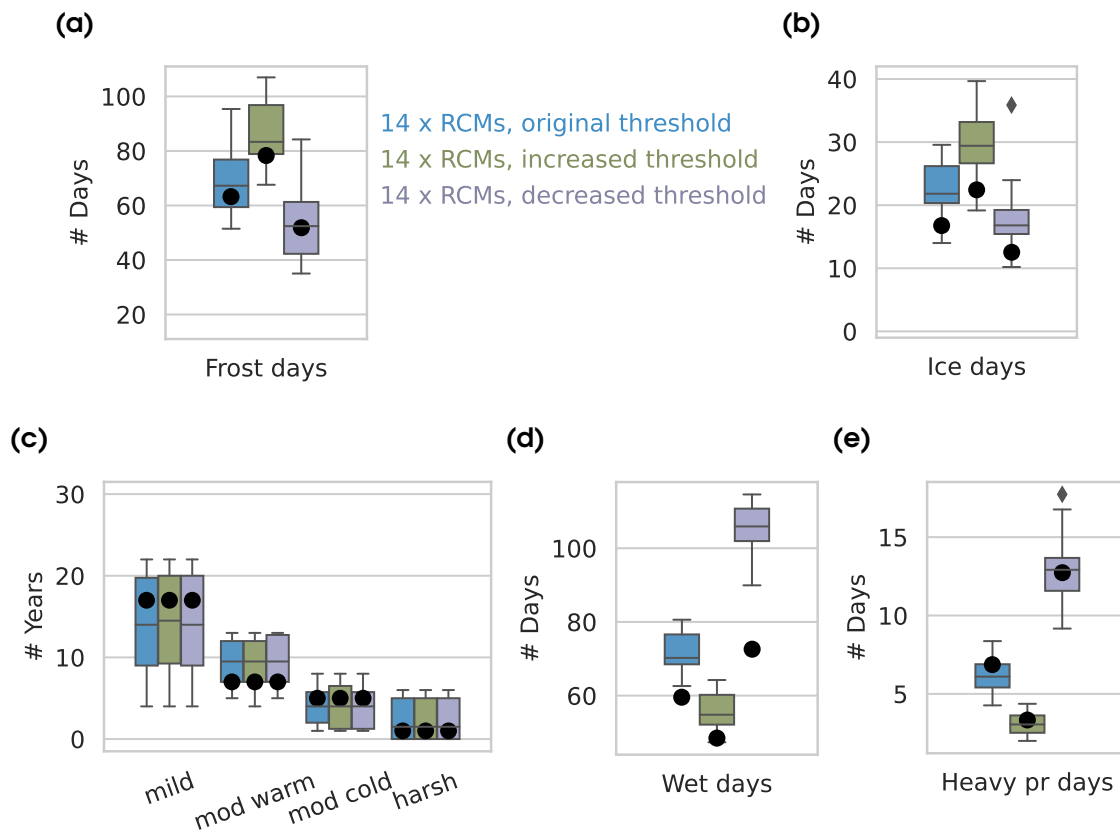
Figure 2.4. Number of (a) frost days, (b) ice days, (c) years of each coldsum category, (d) wet days and (e) heavy precipitation days as mean cold season value between 1971-2000. From left to right: climate model ensemble bandwidth for the originally defined threshold, climate model ensemble bandwidth for increased threshold, climate model ensemble bandwidth for decreased threshold. Black dots represent reference data values for the corresponding thresholds.

The reference data-set is considered the "truth" here, and the truth — as we have seen — is not exact, but is subject to uncertainty. If we let the "truth" vary by its uncertainty, the baseline is changed. For a fair comparison between reference and climate model data,

the baseline in the climate model data needs to be changed as well. Consequently, when calculating the number of frost days in the climate model data, the thresholds are varied by the same values as used for the reference data. The resultant three groups of 14 values for the climate model data and the three values for the reference data are shown in Fig. 2.4 for the number of frost days (Fig. 2.4a), the number of ice days (Fig. 2.4b), the coldsum (Fig. 2.4c), the number of wet days (Fig. 2.4d), and the number of heavy precipitation days ((Fig. 2.4e). Each box in this figure shows the first quartile (25th percentile), the median, and the third quartile (75th percentile) of the distribution. The central 50 % of the data are within this interquartile range, the rest of the distribution is shown by the whiskers. Outliers are determined using a method that is a function of the interquartile range.

For frost days (Fig. 2.4a), a change in the threshold by the reference data uncertainty leads to an overall range of possible numbers of frost days per cold season of between 35 (lower whisker of climate model ensemble with decreased threshold) and 108 (upper whisker of climate model ensemble with increased threshold). This is a range of roughly one to three months of frost caused by an uncertainty in the daily minimum temperature. The range based on reference data values is between 52 and 80 days per cold season and is roughly the same as the median of the climate model data. The maximum number of consecutive (mnoc) frost days (Appendix A, Fig. A.3a) varies from 15 days to 40 days in the climate model data with varied thresholds. The values derived from the reference data-set are located below the 25th percentile of the climate model data distribution. This shows that the models, in general, overestimate the duration of frost periods.

For ice days (Fig. 2.4b), a change in the threshold by the reference data uncertainty ($\pm$ 0.76, Table 2.3) leads to a range in the number of days of 10 to 40 days. In general, the climate model ensemble tends to result in too many ice days independent of the chosen threshold, since the models are generally too cold in this region and for this period (e.g. Fig. 2.3c). A lowering of the threshold in the climate model data, therefore, leads to a better agreement between the range of reference data values and the climate model ensemble compared to the original threshold.

For the mnoc ice days (Appendix A, Fig. A.3b), the range with varied thresholds is not as large as for the number of ice days, which is caused by the lower number of mnoc ice days in general. The reference data value is generally at the first quartile of the climate model data distribution independent of the chosen threshold. As for the number of ice days, a lowering of the threshold in the climate model data improves the agreement between climate model data and reference data.

For the coldsum (Fig. 2.4c), a change in the threshold does not result in a noticeable

change. Here, daily average values below the chosen threshold are summed up per cold season (148 days). Regardless of whether the threshold is 0 °C, -0.64 °C, or 0.64 °C, in the end, the large negative values are decisive for the coldsum of a cold season. And these large values are counted irrespective of a changed threshold. As a result, the inclusion of uncertainty in the threshold leads to marginal changes in the coldsum. Therefore, the categorisation based on the coldsum seems to be robust against uncertainties in the reference data.

For wet days (Fig. 2.4d), a decrease in the threshold by 0.8 mm/day (Table 2.3) leads to a large increase in the number of days in the model results. It is much larger than the corresponding reference data value. This is probably because of the fact that very small precipitation amounts cannot be measured and result in zero precipitation. The climate models produce precipitation in this "barely measurable" range. Precipitation rates up to 1 mm/day (without 0 mm/day) are more frequent in the climate model data compared to the reference data. A change in the threshold towards these small precipitation amounts leads to a stronger increase in wet days for the climate model data compared to the reference data. The overall climate model data range with changed thresholds ranges from 47 to 118 days. In general, the models overestimate the number of wet days since the models are too wet in this region (KOTLARSKI et al., 2014, shown for DJF in their Figure 3). Although the increase in the threshold, and, so, the consideration of a possible bias in the reference data in this comparison, leads to a better agreement between the range of reference data values and the climate model ensemble, an overestimation of wet days is still evident.

Changing the threshold by $\pm$ 2.8 mm/day (Table 2.3) for heavy precipitation days (Fig. 2.4e) leads to an overall range of 3 to 13 days. The climate model ensemble median and the reference data are in very good agreement for all chosen thresholds. Therefore, a change in the threshold neither improves nor leads to a deterioration in the agreement between climate models and reference data.

2.4.3 Influence of data uncertainty on threshold-based climate indices for future climate periods

The number of days derived from threshold-based climate indices are analysed for the near future (2031–2060) and the far future (2071–2099) for the emission scenarios RCP2.6 and RCP8.5, based on regional climate model results. The number of days is calculated for three different thresholds: the originally defined threshold and the thresholds increased and decreased by the uncertainty estimate derived from the reference data-set. Changes in future climate periods are considered relative to the reference period (1971–
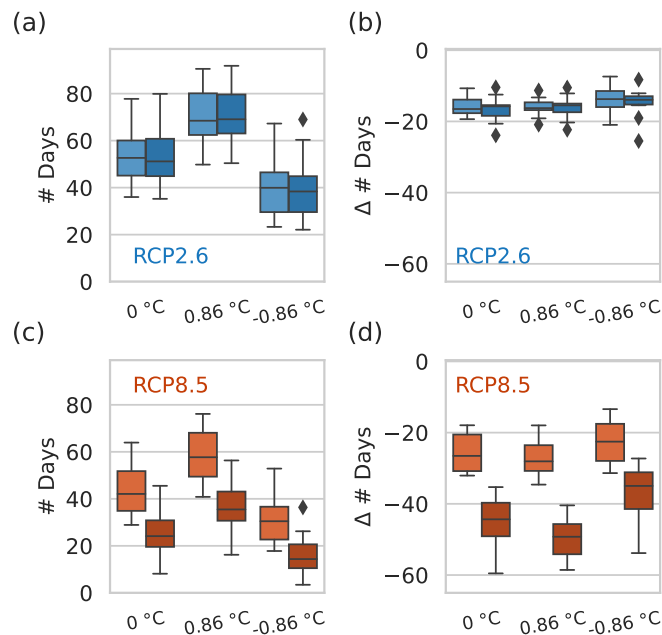
Figure 2.5. Number of frost days for three different thresholds (x-axis) as mean cold season value in near future (2031–2060, light shading) and far future (2071–2099, dark shading) based on climate model data. (a,c): Total number of days, (b,d): Change in the number of days relative to reference period 1971–2000. Numbers are shown for two climate scenarios, (a,b) RCP2.6 in blue and (c,d) RCP8.5 in red.

2000) for all climate indices. The relative change between two future climate periods is always considered on the basis of values calculated with exactly the same threshold.

In general, the number of frost days per cold season is projected to decrease in the future. For RCP2.6, the decrease relative to the reference period is around 18 days (Fig. 2.5 a, b) for the different thresholds with minimal difference between the two future periods. In addition, the climate model ensemble spread is small and remains similar for different thresholds and future periods. For RCP8.5 (Fig. 2.5 c, d), the reduction of the number of days as well as the ensemble spread is larger in near future compared to RCP2.6 and further increases in the far future.

Very similar results are found for the maximum number of consecutive frost days per cold season in the near and far future (Appendix A, Fig. A.4): The decrease for RCP2.6 in the near future is around 5 days and slightly larger for the far future. For RCP8.5, again a stronger reduction in the number of days is found which is larger for the far future. Differences in the number of days are small for the different thresholds.

The results for the number of ice days and the maximum number of consecutive ice days (Appendix A, Fig. A.5 and A.6) are very similar to the results for the number of frost days and the maximum number of consecutive frost days. Since the absolute values in the reference period are already small (5 to 17), the change in future climate periods cannot be greater than these values. For RCP8.5, the maximum number of consecutive ice days
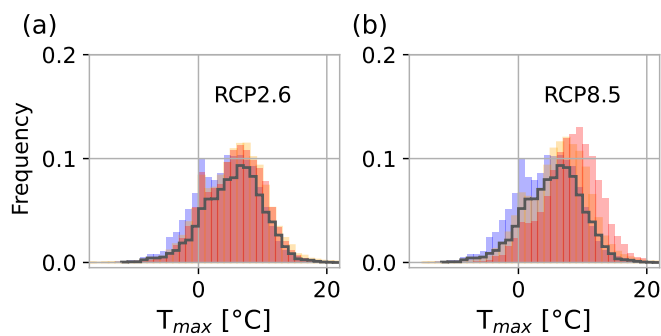
Figure 2.6. Distribution of cold season daily maximum temperatures in the area of interest for (a) RCP2.6 and (b) RCP8.5. Shown are distributions for three different climate periods: the reference period (1971–2000, blue), near future (2031–2060, orange) and far future (2071–2099, red). Shown is the 95th percentile of the climate model ensemble. Black: reference data values for the reference period.

is close to zero in the far future, which means that the changes relative to the reference period cannot increase much after 2100.

For all temperature threshold-based climate indices, it is evident that an increase in the threshold leads to a higher number of days in the reference period but also to a stronger decrease in future climate periods compared to the number of days based on the original or decreased threshold. The larger differences in the changes relative to the reference period for the three different thresholds, found for RCP8.5 compared to RCP2.6, may be caused by the fact that the underlying distribution of the climate variable changes its shape and shifts to a higher mean and higher mode (Fig. 2.6). For RCP2.6, the frequency distribution shifts only slightly to higher values. Hence, a change in the threshold has a small effect on the relative change of the number of days. The larger the change of the shape of the distribution for future climate periods, the larger the difference in the number of days between the three thresholds when looking at future changes relative to the reference period.

As shown in sec. 2.4.2, changes in the threshold lead to marginal changes in the coldsum. Therefore, the categorisation based on the coldsum seems to be robust against changes in the threshold. This can also be seen for future periods, when looking at the total number of years shown in Fig. 2.7a, c, and also when looking at relative changes shown in Fig. 2.7b, d. In future periods, mild winters are projected to increase whereas winters of all other categories occur less frequently. Again, in RCP2.6, there is no further change from the near to the far future. In RCP8.5, around 10 more years of mild winters are projected in the near future and around 15 more years of mild winters are projected for the 30-year period in the far future. Moderately cold as well as harsh winters will become very rare.

For wet days, there is little difference between the two scenarios and the models do not
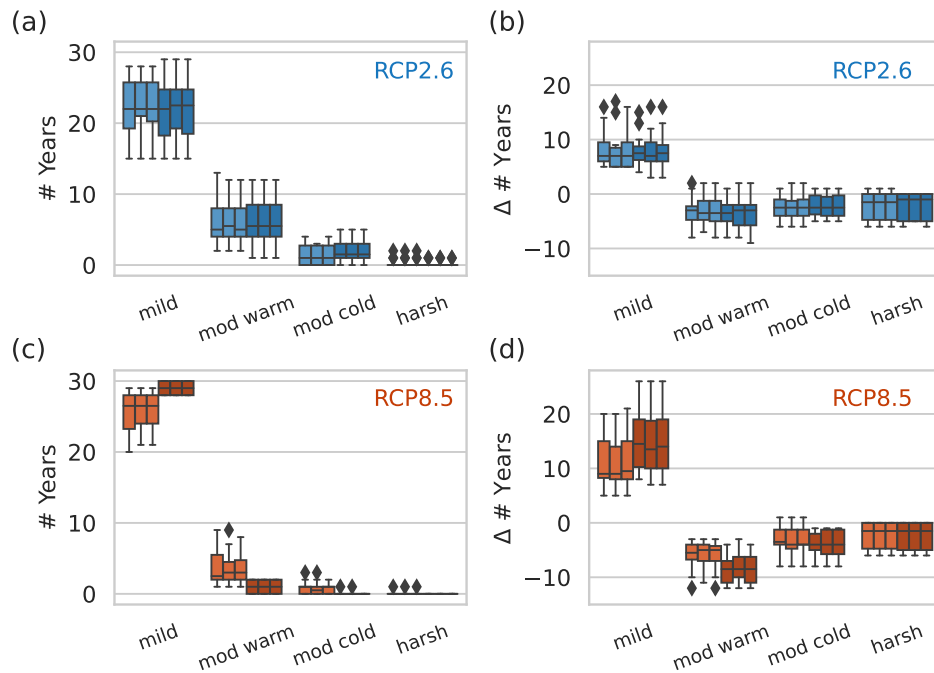
Figure 2.7. Categorisation of future winters based on the coldsum. Shown are the number of years of each category in two future climate periods. Light shading: 2031–2060, dark shading: 2071–2099. Three boxes of the same shading represent values for three different thresholds (f.l.t.r.: original, increased threshold, decreased threshold). (a,c): Total number of days, (b,d): Change in the number of days relative to reference period 1971–2000. Numbers are shown for two climate scenarios, (a,b) RCP2.6 in blue and (c,d) RCP8.5 in red.

agree on the direction of change (Appendix A, Fig. A.7). Some simulate an increase in the number of wet days while some others simulate a decrease in the number of wet days. Consequently, the overall changes are small and in the range of ± 5 days. As for the reference period, a decrease in the threshold by the uncertainty estimate leads to a large increase in the number of wet days which is not visible when one looks at changes relative to the reference period.

For heavy precipitation days, the climate models project an increase in the near and the far future for RCP8.5 (Fig. 2.8). For RCP2.6, the climate models do not agree on the direction of change and are mostly in the range of ± 1 day. Similarly as for the temperature threshold-based climate indices, a decrease in the threshold leads to a larger number of days in the reference period but also to a stronger increase in future climate periods compared to the number of days based on the original, or the increased, threshold.

2.5   Summary and conclusions

The influence of data uncertainty on cold season threshold-based climate indices is investigated in this work. In addition to a general sensitivity study, wherein thresholds of
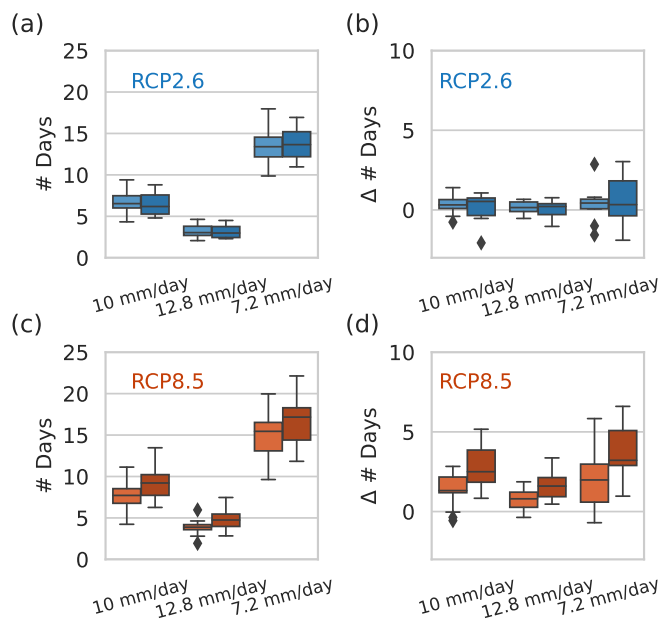
Figure 2.8. Number of heavy precipitation days for three different thresholds (x-axis) as mean cold season value in near future (2031–2060, light shading) and far future (2071–2099, dark shading) based on climate model data. (a,c): Total number of days, (b,d): Change in the number of days relative to reference period 1971–2000. Numbers are shown for two climate scenarios, (a,b) RCP2.6 in blue and (c,d) RCP8.5 in red.

a set of climate indices are varied arbitrarily, uncertainties derived from a reference data-set have been used to vary the defined thresholds in a defined range of uncertainty. The application of our method has been demonstrated by analysing future climatic changes for a north-western European city and its rural surroundings. Based on our results, we conclude the following:

1. Data uncertainty can be included by changing the threshold by a data uncertainty estimate that is preferably based on known uncertainties — here, the ensemble spread of the E-OBS data-set in its ensemble version.

2. In general, each threshold-based climate index responds differently to changes in the threshold. This behaviour mostly depends on the frequency distribution of the climate variable, especially on the distance between the defined threshold and the mode of the distribution. Indices based on extreme values are more sensitive to variations of the threshold.

3. For RCP2.6, changes can be reliably assessed when the same threshold is used for current and future climate periods. Relative changes are very similar for all chosen thresholds.
   For RCP8.5, relative changes for the far future depend on the chosen threshold when compared to the current climate period. Differences across the different thresholds are most pronounced for the number of frost/ice days, the maximum

number of consecutive frost/ice days, and heavy precipitation days. A reason for these large differences is most likely the change in the frequency distribution of the corresponding data.

In addition, it has been shown that climate model uncertainty can be within the range of (interpolation) uncertainties of the reference data. Hence, it is important to account for these uncertainties when evaluating climate models. This might also be of help for climate change impact analyses, e.g. when threshold-based indices from climate models are used to force impact models. However, it should always be checked if the frequency distribution of the analysed data remains similar for future climate periods.

The categorisation based on the coldsum seems to be robust against changes in the threshold, and can, therefore, be used to gain insights into future changes. Indices that integrate a quantity over a given time interval (such as coldsum or consecutive days) are less susceptible to uncertainties in the threshold. When applying the bias-correction method proposed by HOFFMANN et al. (2018) to the thresholds, it became clear that integrating indices benefit little from it. This supports our conclusion that integrating indices show little response to changes in the threshold, and, so, can be used for future assessments. Other indices can be used, but uncertainties should always be taken into account, and ideally corrected in an appropriate way in order to assess climate change.

3

Reproducibility of local cold season characteristics, ice episodes and prevailing circulation patterns in RCMs

Preface

This chapter was submitted for publication as:

Bell, L. M. and K. H. Schlünzen and K. Sieck: Reproducibility of local cold season characteristics, ice episodes and prevailing circulation patterns in regional climate models. *Quarterly Journal of the Royal Meteorological Society*, in review.

The references are combined with all references used in this thesis and can be found at the end of this thesis in References. The supporting information, the acknowledgments and other information that were submitted together with this manuscript can be found in a separate section of the Appendix (Second paper).

K. Heinke Schlünzen has contributed to the conceptualisation and contributed some ideas for analysis. K. Heinke Schlünzen and Kevin Sieck have contributed to the discussion of the results.

Abstract

This study investigates the cold season characteristics and the occurrence of ice episodes for a target region in northern Germany. Additionally, it identifies the typical atmospheric patterns associated with ice episodes using reanalysis data. A pattern-matching method is developed, that uses the structural similarity index to determine whether regional climate models (RCMs) can reproduce the ice episode-specific atmospheric patterns identified from the reanalysis data. The results show that the frequency of ice days is overestimated, but not the frequency of frost days. This suggests a daytime cold bias in the RCMs in this region, as maximum temperatures are too low. Moreover, most of the analysed RCMs simulate too many ice episodes (> 5 consecutive ice days). The developed pattern-matching demonstrated, based on reanalysis data, that ice episodes in the target region can be associated with a variety of different atmospheric patterns. Specifically, longer ice episodes (lasting 14–29 consecutive ice days) in the target region are associated with a blocking pattern over Iceland/the British Isles. This is the most frequent pattern observed during ice episodes in the target region. The least frequent patterns show a system of low GPH over the European continent and are associated with shorter ice episodes (6–8 consecutive ice days) in the target region. The RCMs can reproduce these patterns and their frequency well, regardless of their forcing, and these patterns are also associated with an ice episode in the RCMs. Furthermore, it can be concluded that the "typical" blocking pattern is not a reliable indicator for ice episodes in the target region.

3.1 Introduction

A cold spell is an event characterized by low temperatures that persist for several days. Cold spells can have a variety of impacts on society. They can be the cause of severe damage and loss of life (PFAHL, 2014; WEILNHAMMER et al., 2021), they can cause transport disruption and have a direct impact on energy demand (VAN DER WIEL et al., 2019). Therefore, cold spells pose a considerable threat to society.

The characteristics of such events can be studied using observational or reanalysis data. The future development of cold spells can be analysed using climate model projections. To examine these changes for a specific region, such as Europe, projections from regional climate models (RCMs) can be used. RCMs are limited-area climate models that typically have a finer resolution than global climate models (GCMs) and are frequently used for impact studies (see GIORGI (2019) for a summary). However, in order to derive reliable statements about the future development of cold spells from the RCM data, it is crucial to evaluate the RCMs' ability to reproduce the current cold season characteristics. Thus, the first research question is: **Are RCMs able to reproduce cold season characteristics for a specific target region?**

As there is no standard definition for cold spells (MENG et al., 2022, e.g.), here the term "ice episode" is introduced. An ice episode is defined here as an event consisting of more than five consecutive ice days. Therefore, this study analyses events where the maximum daily temperature is below freezing for at least six consecutive days. This duration is chosen based on the understanding that certain atmospheric conditions associated with cold events persist for about 7 to 10 days (PFAHL, 2014). To assess the ability of RCMs to reproduce the frequency of ice episodes in a specific target region, we have selected an urban area as a use case. The second question to answer is **Are RCMs able to reproduce the frequency of ice episodes for a specific target region?**

As use case the Hamburg metropolitan region in northern Germany is selected. Due to its geographical location, the cold season in this region is strongly influenced by weather patterns originating in the North Atlantic. Cold extreme events in many parts of Europe are generally associated with similar atmospheric conditions in the North Atlantic region (KAUTZ et al., 2022). A weakened air pressure gradient between Greenland and the Azores can cause the jet stream to meander and become unstable. This can interrupt the predominant westerly flow and result in meridional air mass transport. When a high-pressure system separates and remains stable in the same location for an extended period, it "blocks" the westerly flow, which is known as atmospheric blocking (REX, 1950). For central Europe, BRUNNER et al. (2018) could show, that up to 70 % of cold spells can be associated a blocking anywhere between 60°W and 30°E. Therefore, to investigate

the future development of ice episodes using RCM data, it is necessary to examine the extent to which the RCMs can reproduce typical atmospheric patterns and the relationship between those patterns and ice episodes in the target region. The third question to answer is **Can RCMs accurately reproduce the atmospheric patterns prevailing during ice episodes?**

To evaluate the ability of RCMs to reproduce the prevailing atmospheric patterns, these patterns are first identified in the reanalysis data and then searched for in the RCM data using a pattern matching method developed in this study. This method of pattern matching utilises the Structural Similarity Index (SSIM) (WANG et al., 2004), which is commonly used in image processing. The developed method allows to identify the atmospheric patterns prevailing during local ice episodes and to search for these ice episode specific patterns in RCM data. Subsequently, it can be examined whether the ice episode specific patterns found in RCM data lead to a local ice episode in the target region.

Following this introduction, sec. 3.2 describes the data used for the study. A detailed description of the workflow and methods is given in sec. 3.3. The results are then presented and discussed in sec. 3.4. Sec. 3.5 closes with a summary and conclusions.

3.2   Database

The database consists of a gridded observational data-set, a reanalysis as well as the output of RCMs. Variables that are analysed are: daily 2 meter temperature (T2m), daily minimum and maximum 2 meter temperatures (T2m$_{min}$, T2m$_{max}$) and daily geopotential height (GPH) at 500 hPa (Zh$_{500hPa}$). Two threshold-based climate indices derived are from T2m$_{min}$ and T2m$_{max}$, the number of frost days and the number of ice days.

3.2.1   Gridded data-set

The E-OBS data-set is a gridded data-set based on interpolated station data across Europe (CORNES et al., 2018). It is provided with a spatial resolution of 0.1° (used here) and 0.25°, and has a temporal resolution of 1 day. The E-OBS data-set version used here (v27.0e), consists of an ensemble of 20 members for temperature (daily minimum, mean and maximum values), precipitation, sea level pressure, wind speed, relative humidity and global radiation. The ensemble methodology, based on the work of HUTCHINSON and GESSLER (1994), takes into account the uncertainty of the underlying statistical model, resulting in larger model residuals leading to a wider spread within the ensemble. Variables from E-OBS used here are T2m, T2m$_{min}$ and T2m$_{max}$. Besides the "best guess" (ensemble mean) value, the ensemble version provides an uncertainty estimate. This

uncertainty estimate represents the 90 % uncertainty range (5th to 95th percentile) of the ensemble. E-OBS is commonly used as reference for the evaluation of EURO-CODEX model data (e.g. JACOB et al., 2014; KOTLARSKI et al., 2014; HOFFMANN et al., 2018) and is also used as reference for target region temperatures in this work.

3.2.2 Reanalysis

The reanalysis used in this work is the global reanalysis ERA5 (HERSBACH et al., 2020). It is the successor of ERA-Interim reanalysis (DEE et al., 2011) and has a finer horizontal resolution (31 km), a higher temporal resolution (hourly) and a longer duration period (1950-present). In this work, surface variables (HERSBACH et al., 2018b) as well as variables on the 500 hPa pressure level (HERSBACH et al., 2018a) are used to calculate daily values. Daily T2m$_{min}$ and daily T2m$_{max}$ are calculated from the hourly instantaneous T2m values, as recommended by ECMWF (2022). ERA5 is used as the reference for the atmospheric circulation patterns in this work.

3.2.3 Model simulations

RCM results used here were generated within the frame of the Coordinated Downscaling Experiment (CORDEX, GIORGI and GUTOWSKI (2015)). They have a spatial resolution of 0.11° for the European domain (EURO-CORDEX, JACOB et al. (2014)). Here, historical runs as well as evaluation runs from RCMs are analysed. The historical runs are forced with GCM simulations from the Coordinated Model Intercomparison Project, Phase 5 (CMIP5, TAYLOR et al. (2012)) models at the boundaries and have therefore higher degrees of freedom than the evaluation runs forced with ERA-Interim (DEE et al., 2011). A general evaluation of the ERA-Interim forced runs can be found in KOTLARSKI et al. (2014).

The model-simulation combinations used in this work are listed in Table 3.1. Our selection criteria are, first, that an evaluation run, a historical run and a future projection based on RCP8.5 (forced by any GCM) are available, second, that the evaluation run covers at least the period 1989–2005 and, third, that in addition to the surface variables, also the geopotential (height) at 500 hPa is provided. These criteria are met by eight evaluation runs and 29 GCM-driven RCM simulations. This selection provides a variety of forcing GCMs with different performance in reproducing, for example, blocking frequencies in the northern hemisphere. Performance information is taken from the circulation-based performance atlas of CMIP5 and CMIP6 models presented by BRANDS (2022).

The CCLM4-8-17 (ROCKEL et al., 2008), HIRHAM5 (BØSSING CHRISTENSEN et al., 2007), WRF331F (SKAMAROCK et al., 2008) and REMO2009 (JACOB et al., 2001)

Table 3.1. Climate model simulations used in this work. Forcing from global climate model (GCM) or reanalysis (ERA-Interim) data. The regional climate model (RCM) used and the used GCM-realisation for the forcing. The different calendars used for the simulations and the duration of the simulations lead to a different number of total cold season days (days) within 1979–2005.

Forcing	RCM	Realisation	Calendar	total days
ERA-Interim	CCLM4-8-17	r1i1p1	proleptic Gregorian	2419
	ALADIN53	r1i1p1	proleptic Gregorian	3932
	HIRHAM5	r1i1p1	proleptic Gregorian	2419
	REMO2015	r1i1p1	proleptic Gregorian	3932
	WRF331F	r1i1p1	proleptic Gregorian	2419
	RACMO22E	r1i1p1	standard	3932
	REMO2009	r1i1p1	proleptic Gregorian	2419
	RCA4	r1i1p1	standard	3871
CanESM2	CCLM4-8-17	r1i1p1	365-day	3925
	REMO2015	r1i1p1	proleptic Gregorian	3932
CNRM-CM5	CCLM4-8-17	r1i1p1	proleptic Gregorian	3932
	ALADIN53	r1i1p1	proleptic Gregorian	3932
	REMO2015	r1i1p1	proleptic Gregorian	3932
	RACMO22E	r1i1p1	standard	3932
	RCA4	r1i1p1	standard	3932
EC-Earth	CCLM4-8-17	r12i1p1	proleptic Gregorian	3932
	HIRHAM5	r3i1p1	proleptic Gregorian	3932
	REMO2015	r12i1p1	proleptic Gregorian	3932
	RACMO22E	r3i1p1	standard	3932
	RCA4	r12i1p1	standard	3932
IPSL-CM5A-MR	WRF331F	r1i1p1	proleptic Gregorian	3932
	RACMO22E	r1i1p1	standard	3925
	RCA4	r1i1p1	standard	3925
MIROC5	CCLM4-8-17	r1i1p1	365 day	3925
	REMO2015	r1i1p1	proleptic Gregorian	3932
HadGEM2-ES	CCLM4-8-17	r1i1p1	360 day	3900
	RACMO22E	r1i1p1	360 day	3900
	RCA4	r1i1p1	360 day	3900
	REMO2015	r1i1p1	360 day	3900
MPI-ESM-LR	CCLM4-8-17	r1i1p1	proleptic Gregorian	3932
	REMO2015	r3i1p1	proleptic Gregorian	3932
	RACMO22E	r1i1p1	standard	3932
	REMO2009	r2i1p1	proleptic Gregorian	3932
	RCA4	r1i1p1	proleptic Gregorian	3932
NorESM1-M	HIRHAM5	r1i1p1	proleptic Gregorian	3932
	REMO2015	r1i1p1	proleptic Gregorian	3932
	RACMO22E	r1i1p1	365-day	3925

evaluation runs start in 1989, the RCA4 (SAMUELSSON et al., 2011) evaluation run

starts in 1980, ALADIN53 (COLIN et al., 2010), REMO2015 (JACOB et al., 2001) and
RACMO22E (VAN MEIJGAARD et al., 2008) evaluation runs start in 1979. Therefore,
the evaluation runs have different numbers of total cold season days within the chosen
analysis period of 1979–2005 (see Table 3.1, last column).

3.3 Methodology

The driving questions for the development of the methodology are: How to determine ice
episodes for the target region? How to determine the atmospheric patterns that prevail
during these events? How to find these patterns in RCM data? To be able to answer these
questions, a data processing/analysis workflow is established. The methodologies for the
different analysis steps within the workflow are shown in Fig. 3.1 and described below.
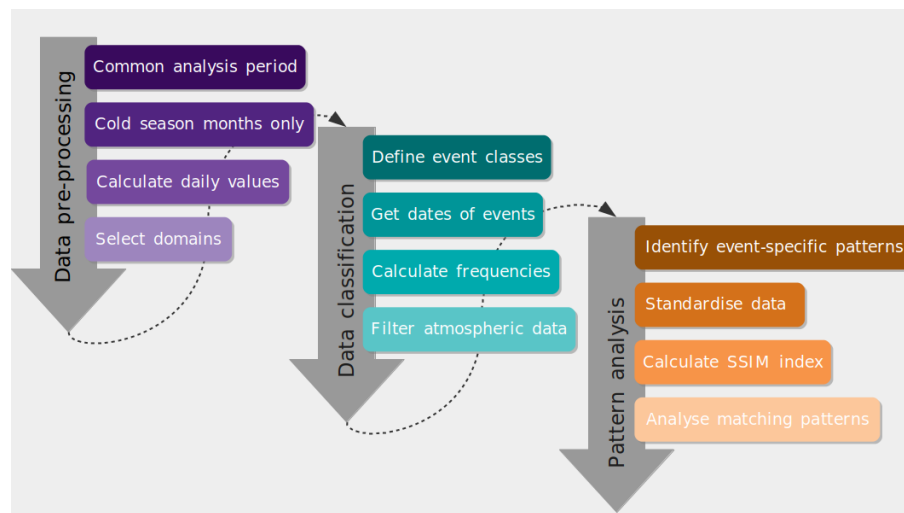
Figure 3.1. Workflow established to answer research questions. From data pre-processing, through data
classification, to atmospheric pattern analysis. For more details see text.

3.3.1 Data pre-processing

All data-sets are pre-processed to meet the following requirements:

1. A common analysis period. The longest overlapping period of the data-sets used
   in this work is 1979–2005.

2. The consideration of cold season months only. As cold season, we define all days
   of a year between the 1st of November and the 31st of March of the following year.
   This results in the final time period from the 1st of November 1979 til the 31st of
   March 2005. Hence, 26 cold seasons in total. As shown in Table 3.1, the different
   model simulations were generated using different internal calendars: The proleptic
   Gregorian or standard calendar with 365/366 days per year including leap years;

the 360 day calendar where each month has exactly 30 days; the 365 day calendar excluding leap years. To ensure a fair comparison of, for example, the number of frost days, results are always given relative to the total number of cold season days in the respective data-set.

3. All data is converted into daily average or daily minimum/maximum values. This temporal frequency is needed for the calculation of the climate indices.

4. Two domains are selected. A target region is selected for which the ice episodes are calculated. The target region covers the area of Hamburg (53.55°N and 10°E) and its surroundings and its size is roughly 70 km × 70 km (53.25°N – 53.85°N and 9.48°E – 10.53°E) (Fig. 3.2). For the data sets used here, this results in an area of 60 grid cells for E-OBS, 12 grid cells for ERA5 and 36 grid cells for the RCMs. For ERA5, a larger north European domain is selected to investigate the atmospheric circulation during the ice episodes. This larger domain corresponds to the EURO-CORDEX domain with the boundaries 27°N – 72°N and −22°E – 45°E (Fig. 3.3).

This data-processing methodology can be applied to other data-sets or target regions. The large domain should be large enough to ensure that atmospheric circulation patterns are captured. The target region should cover several climate model grid cells to enable the calculation of average values from more than one value (grid cell). Averaging over several grid cells is crucial as models lack grid point precision, and the kinetic energy in the system is too low for a single grid point (based on work by SKAMAROCK (2004)). Consequently, averaging over several grid cells results in a more representative value.

3.3.2   Data classification

In this paper, cold season characteristics are studied for the target region. For this purpose, the classes shown in Table 3.2 are defined, into which all the data used in this work are classified.

Classes 1, 2 and 3 are warm, frost and ice days, respectively. Three sub-classes of ice days are defined to account for different durations of ice episodes. The short duration class (3–Ds) corresponds to the duration of average atmospheric blocking events that can cause cold events (see also the Introduction). The long duration class (3–Dl) corresponds to the persistence of atmospheric blocking events during extreme cold events. All ice episodes with durations between average and extreme persistence of atmospheric blocking events belong to the medium duration class (3–Dm).

For classes 1–3 it is checked for each day whether the daily maximum and minimum
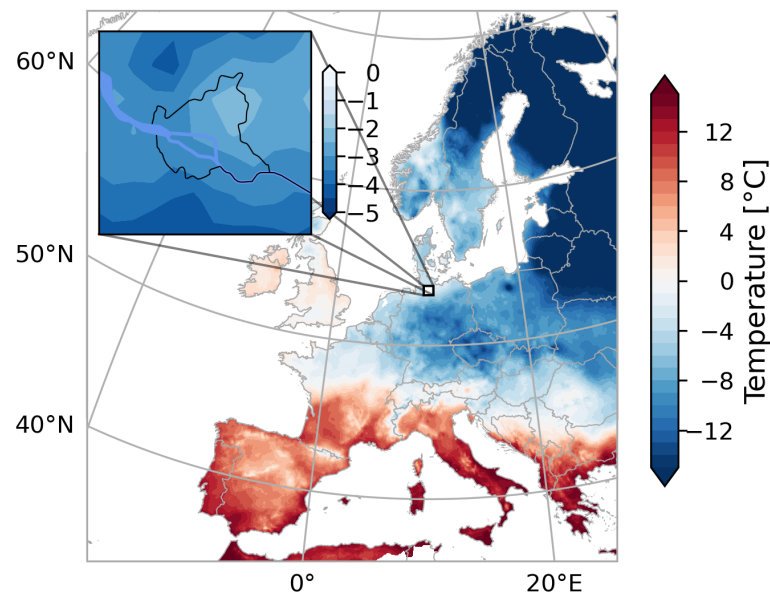
Figure 3.2. Target region: Hamburg and its surroundings. Shown is the daily average 2m temperature on the 1996-12-24 based on the E-OBS data-set. This date is the 5th day of the longest ice episode identified in E-OBS within 1979–2005. Note the different scalings.
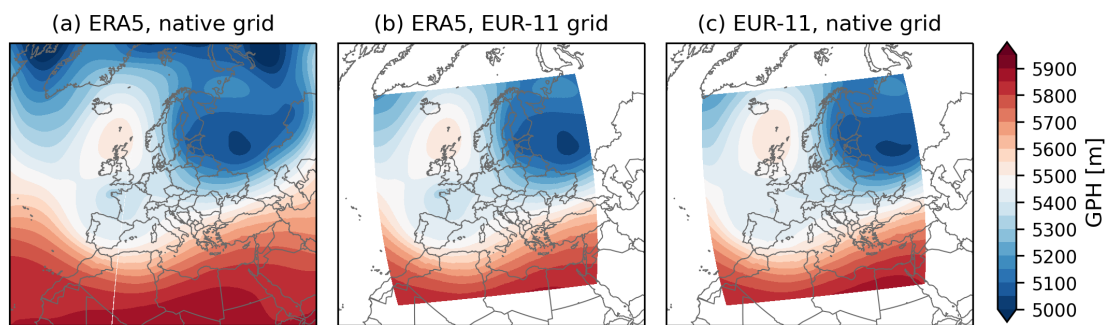
Figure 3.3. Remapping of ERA5 data to fit the EURO-CORDEX domain. As an example, the geopotential height at 500 hPa pressure level is shown for three different data-sets on different grids on the 1996-12-24. The different data-sets are: (a) ERA5 on its native grid. (b) ERA5 remapped to the exact EURO-CORDEX domain and resolution. (c) A EURO-CORDEX RCM simulation on its native grid.

Table 3.2. Defined cold season characteristic classes into which data is divided. Temperatures (T2m$_{min}$, T2m$_{max}$) are considered as target region averages.

Class	Cold season characteristic	Definition
1	warm day	T2m$_{min}$ > 0 °C
2	frost day	T2m$_{min}$ < 0 °C and T2m$_{max}$ > 0 °C
3	ice day	T2m$_{max}$ < 0 °C
Sub-classes		
3–Ds	short duration	6–10 consecutive ice days
3–Dm	medium duration	11–15 consecutive ice days
3–Dl	long duration	> 15 consecutive ice days

37

temperatures are below 0 °C. Depending on the result of the test, the date of the day is assigned to the corresponding class. For example, the daily minimum temperature is below 0 °C but the daily maximum temperature is not. Then, the date of this day is assigned to the frost-day class. In this way, a time mask is created for each class.

For the ice episode classes, the ice days are analysed separately per season. If the condition for an ice day is met, the next day is checked. This continues until the condition is no longer met. The duration of the event is then determined and all dates of all days within the event are allocated to the appropriate duration class. Again, a time mask is created for each class. In addition, the start date and duration of the considered ice episode are stored. The same procedure is then carried out for the remaining days within the analysis period.

For each class, the frequency of occurrence is calculated relative to all cold season days within the analysis period. For E-OBS and ERA5 there are 3932 cold season days within 1979–2005. Since the GCMs and RCMs have different calendars (see Table 3.1 and sec. 3.3.1, point 2.), the number of cold season days in the RCM simulations is smaller. The numbers range from 2419 to 3932 and can be found in Table 3.1.

For E-OBS, the frequency of occurrence is calculated for the defined threshold of 0 °C as well as for varying thresholds based on the observational uncertainty provided by the E-OBS ensemble version (see sec. 3.2.1). The uncertainty estimates for the thresholds have been derived in BELL et al. (2023). This provides an observational uncertainty range for warm days, frost days and ice days. This uncertainty estimate was not used to calculate the frequency for the ice episode classes. The reason for this is that a change in the threshold may lead to longer events on the one hand, but also to fewer events in total on the other. These compensating effects would lead to an ambiguous result that would be difficult to interpret.

In the final step, the created time masks are used to filter the GPH at the 500 hPa pressure level.

All these steps are applied to the reference data-set E-OBS, the reanalysis ERA5 and the RCM data.

3.3.3 Pattern analysis

In order to better estimate when and under what conditions certain ice episodes occur, it is helpful to know the state of the atmosphere during these events. As described e.g. in MESSORI et al. (2017), the GPH at 500 hPa is a good proxy for the large-scale atmospheric circulation in the North Atlantic region. It is a key component of synoptic-

meteorology and is often used for the analysis of weather patterns, e.g. the calculation of atmospheric blocking (e.g. SCAIFE et al., 2010; DUNN-SIGOUIN and SON, 2013; FRANCIS et al., 2022). Therefore, the GPH at 500 hPa is used here to determine typical patterns for each of the classes defined.

In a first step, mean atmospheric patterns are calculated for each of the defined classes. This is done by averaging over all days within the class under consideration. This is done both for ERA5 as a reference and for the RCM simulations. In this way, the mean patterns for each class can be compared.

In the second step, the focus is shifted to the individual events in the duration classes. It is examined whether the RCMs are able to reproduce the event-specific patterns identified for the reference. The information whether and if so, which cold event occurs in the target region in the RCM is neglected. In the RCM simulations, therefore, only the event-specific atmospheric reference pattern is searched for. How this search works and which individual steps are important is described below:

1. The ERA5 reanalysis data are used as reference for the atmospheric circulation patterns against which the RCM data are compared. Thus, event-specific patterns are derived from the ERA5 reanalysis data and are referred to as reference patterns. For each event within the different duration classes, an event-specific reference pattern is obtained. For the entire analysis period, there are 17 events for class 3–Ds, so 17 individual patterns are derived. For class 3–Dm there are 4 events (= 4 patterns) and for class 3–Dl there are also 4 events (= 4 patterns). See Table B.1 in Appendix B for the number of events derived from E-OBS, ERA-Interim and ERA5.

2. In order to find the event-specific reference patterns in the RCM data, a method often used in image processing is applied: The structural similarity index (SSIM) (WANG et al., 2004). This index describes the similarity between two input signals (images). The first signal is the original, which is assumed to be perfect. The SSIM can then be used as a quantitative measure of the similarity of the second signal. This index divides the similarity measure into three separate comparisons: Luminance (l), contrast (c) and structure (s). Combining the three comparison functions gives the following equation:

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right)_l \cdot \left(\frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right)_c \cdot \left(\frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \right)_s \tag{3.1}$$

$$= \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_2)} \tag{3.2}$$

where $\mathbf{x}$, $\mathbf{y}$ are the two signals. $\mu_{x,y}$ is the mean intensity (luminance), $\sigma_{x,y}$ is the standard deviation (contrast) and σ_{xy} is the correlation (structure) of the two images. l, c and s indicate the luminance, contrast and structure comparison functions. The constants C_1 and $C_2 = 2C_3$ ensure stability of the equation when the denominator is close to 0. The structure comparison function s (last factor of eq. 3.1) is similar to the anomaly correlation coefficient (ACC) calculation. In the field of weather and climate prediction, the ACC is widely used to determine the skill of a forecast (e.g. WILKS, 2019). For the complete derivation of the SSIM with all the details, we refer to the original work by WANG et al. (2004).

To apply the SSIM index to the data used in this paper, the implementation within the Python library *scikit-image* (VAN DER WALT et al., 2014) is used.

HOFFMANN et al. (2021) used this index to derive a weather persistency index from it. In their work, they were able to show an increase in persistent weather conditions that are associated with hydro-climatic risks. While HOFFMANN et al. (2021) looked at the persistence of weather patterns in ERA5 data, the approach presented here uses the SSIM index to detect event-specific reference patterns in the RCM data.

3. To ensure a fair comparison between event-specific reference patterns and patterns of the RCM data, a sliding window (temporal) is applied to the RCM data within which average patterns are calculated. This sliding window is of the same length as the duration of the reference event that is being considered. For example: The cold season reference event consists of 14 consecutive ice days, so a sliding window of 14 days is used to calculate average patterns in the RCM data time series. Each day i that is compared to the ERA5 pattern contains an average pattern calculated from days [i:i+14]. Thus, a comparison is made between average patterns that are calculated from the same number of days.

4. Before being able to calculate the SSIM index between the two average patterns, further pre-processing steps are required. First, the RCM data and the ERA5 data have to be made comparable. This is done by remapping (`cdo remapycon`
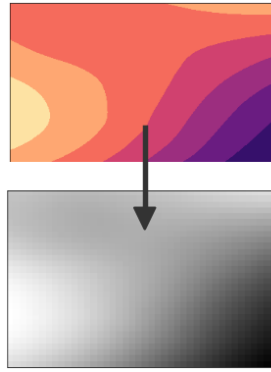
Figure 3.4. Geopotential height conversion into greyscale image [0–255].

(SCHULZWEIDA, 2023)) the event-specific reference (ERA5) patterns, which are searched for in RCM data, to the RCM native (EUR-11) grid (Fig. 3.3). Second, the two average patterns have to be converted into 1-channel greyscale images (Fig. 3.4). For the conversion, each average pattern is normalised to a range of 0–255. For the normalisation, the minimum and maximum values from the ERA5 time series (cold season months only) are estimated once for the whole period of 1979–2005 and used for the normalisation. The normalisation is derived as follows:

$$P_{o,\text{gray}} = \frac{P_o - \min_{erai}}{\max_{erai} - \min_{erai}} \cdot 255$$

where P is the average ice episode pattern and o is the origin (data source), which can be either ERA5 or RCM data.

5. A calibration of the SSIM index implementation was performed to find the appropriate configuration for our use case. For further details on configuration options of the SSIM, we refer the reader to VAN DER WALT et al. (2014). In addition, an appropriate threshold is set above which the samples are considered to be very similar, i.e. a match. The most suitable setup for our application was found to be a window size of 61x61 pixels (pixel size ≈ 12.5 km × 12.5 km) without Gaussian weights for the SSIM calculation and a threshold of 0.85 for finding the event-specific reference patterns. The selected window size is approximately equal to the horizontal length scale of the synoptic scale systems.

6. With these settings, event-specific reference patterns are searched for in the RCM data. If a match is found, details of the RCM in question, such as the average temperature during the identified event within the target region, are stored in order to make a statement about whether there was an ice episode present during the identified atmospheric pattern being sought.

3.4 Results

3.4.1 Cold season characteristics

The frequency of the defined classes is determined in relation to all cold season days within 1979–2005. The cold season characteristics are evaluated for the target region (see Fig. 3.2). The results are shown in Fig. 3.5 and described in detail below.
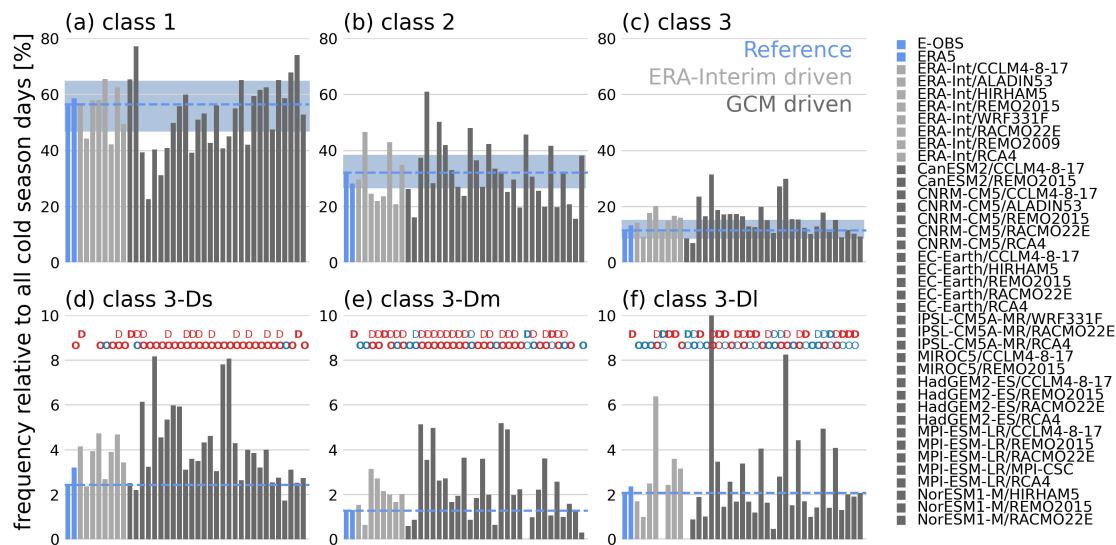
Figure 3.5. Frequency of different classes, describing cold season characteristics (Class 1: warm days, Class 2: frost days, Class 3: ice days). Ice days are further divided into three sub-classes representing ice episodes (class 3–Ds: 6 to 10 consecutive ice days, class 3–Dm: 11 to 15 consecutive ice days, class 3–Dl: more than 15 consecutive ice days). Blue dashed lines indicate the E-OBS value, blue shading indicates the E-OBS uncertainty range. Red/blue letters in lower panel indicate overestimation (red) or underestimation (blue) of duration (D) and occurrence (O) of ice episodes. Boldface indicates whether D or O has a greater effect on over- or underestimation of ice episode frequencies. Frequencies are given relative to all cold season days within the respective data set.

The reference data show that warm days (class 1, Fig. 3.5(a)) occur with a frequency of 57%. Including the uncertainty (as described in sec. 3.3.2), the frequency range is between 47% and 65% (blue shading). The ERA5 reanalysis data are in the range of the E-OBS results. Just over half of the RCM simulations result in warm day frequencies that are within the uncertainty range, regardless of whether runs driven by reanalysis data (light grey bars) or GCM data (dark grey bars) are considered. The RCMs ALADIN53 and RACMO22E generally underestimate the frequency of warm days regardless of the forcing. Simulations forced by the GCMs CNRM-CM5 and MIROC5 underestimate the frequency of warm days, while those forced by NorESM1-M overestimate it. The overestimation of the driving GCM NorESM1-M counteracts the general tendency of RACMO22E to underestimate this frequency, resulting in a frequency close to the reference for this particular GCM/RCM combination.

The frequency of frost days (class 2, Fig. 3.5(b)) is 32% (27–38%) in the reference data. The result from the reanalysis is within the uncertainty range of the reference data. The RCMs show both over- and underestimation of frequency, with most models showing underestimation. For the ERA-Interim driven runs, two of the eight runs are within the uncertainty range of the reference data, four underestimate the frequency and two are above the uncertainty range. For the 29 GCM-driven runs, 12 are within the uncertainty range, 10 show an underestimation and seven show an overestimation. The runs that show an overestimation of frost days are always the ALADIN53 and RACMO22E runs, regardless of the forcing. For the GCM/RCM combination NorESM1-M/RACMO22E, the opposite tendencies again result in a frost day frequency which is close to the reference.

Ice days (class 3, Fig. 3.5(c)) occur with a frequency of 11.5% (9–15%) in the reference data. The reanalysis data, again, do not differ much from this value, hence, ERA5 is within the uncertainty range of the reference data. The majority of the RCM data show an overestimation of the frequency. For the ERA-Interim driven runs, four are within the uncertainty range and the other four show higher frequencies above the uncertainty range of the reference data. A similar behaviour is observed for the GCM-driven runs. 11 runs are within the uncertainty range, 16 runs overestimate and 2 underestimate it the frequency of ice days. In the two runs driven by the MIROC5 GCM, ice days occur more than twice as often as in the reference data (about 30% compared to 11.5%). In this case the forcing appears to have a major influence on the results.

Based on the reference data, ice days occur most frequently ($\approx$ 50%) as very short events (1-5 consecutive ice days, not shown). Ice episodes of short duration (6–10 days, Fig. 3.5(d)) as well as of long duration ($>$ 15 days, Fig. 3.5(f)) occur similarly frequent. Ice episodes of medium duration (11–15 days, Fig. 3.5(e)) have the lowest frequency. This is not always reproduced by the different RCMs.

The short period duration class (class 3–Ds, Fig. 3.5(d)) contains 13 ice episodes for the reference data E-OBS (see B, Table B.1 for details). For ERA5 there are 17 ice episodes. The RCMs overestimate the frequency of this class, with two exceptions of underestimation. The overestimation is mostly due to an overestimation of the number of ice episodes (O), not necessarily an overestimation of the duration of each ice episode (D).

In the reference data there are four ice episodes in the medium period duration class (class 3–Dm, Fig. 3.5(e)), the same is true for ERA5. Most of the RCMs overestimate the frequency of this class, too. And, as in the case of class 3–Ds ice episodes, this is mostly due to an overestimation of the number of ice episodes (O). Among the 29 GCM-

driven runs, there is one where there is no ice episode in this class. This does not seem to be a typical behaviour of any particular model, but rather a coincidence.

There are four ice episodes in the long period duration class (class 3–Dl, Fig. 3.5(f)) in the reference data, with durations of 20 days, 17 days, 21 days and one with a duration of 23 days. These are also found in ERA5 with slightly different durations and starting dates (± one day). One of the ice periods in this class is much longer with a duration of 29 days. In the reference, there are two ice episodes during these 29 days. There is one ice episode in class 3–Ds (7 days) and one in class 3–Dm (17 days). More than half of the ERA-Interim driven runs overestimate the number of ice episodes, three underestimate the frequency. For the GCM-driven runs, more than half show an underestimation and 10 show an overestimation. The largest overestimations in Fig. 3.5(f) come from the REMO2015 simulations, which overestimate the number of ice episodes (O) and their duration (D). In the ERA-Interim driven run of REMO2015, there are 10 ice episodes with durations ranging from 17 to 31 days and one very long ice episode of 53 days. To reproduce the correct frequency of this type of event, the forcing seems to play a rather minor role, at least in terms of the number of occurrences.

In general, there is no better agreement between RCM data and reference data when reanalysis (ERA-Interim) data is used as forcing compared to a GCM data forcing.

3.4.2 Circulation patterns of cold season characteristic classes

Typical patterns for each cold season characteristic class are derived from the GPH at 500 hPa as the mean of all days within each class. The data from the ERA5 reanalysis are used as pattern reference, since only surface variables are available for the reference data used so far (E-OBS). The typical patterns are shown in Fig. 3.6.

During warm days, a maritime flow from the west is present, resulting from a stable north-south GPH gradient. For frost days, the large-scale flow is more meandering, which is the result of a less pronounced GPH gradient. The flow into the target region (cross in Fig. 3.6(a)–(g)) comes from a north-westerly direction. For ice days, the gradient over the target region is more east-west oriented with high GPH over the British Isles and low GPH over north-eastern Europe. Cold air is advected from the north towards the target region.

A similar pattern to the ice-days pattern is found for class 3–Ds (short ice episodes) in the target region. For ice episodes of medium duration (class 3–Dm), large parts of central Europe are covered by an area of the same GPH (5300–5350 m), leading to lower GPH values also south of this area. For ice episodes of long duration (3–Dl), a blocking system
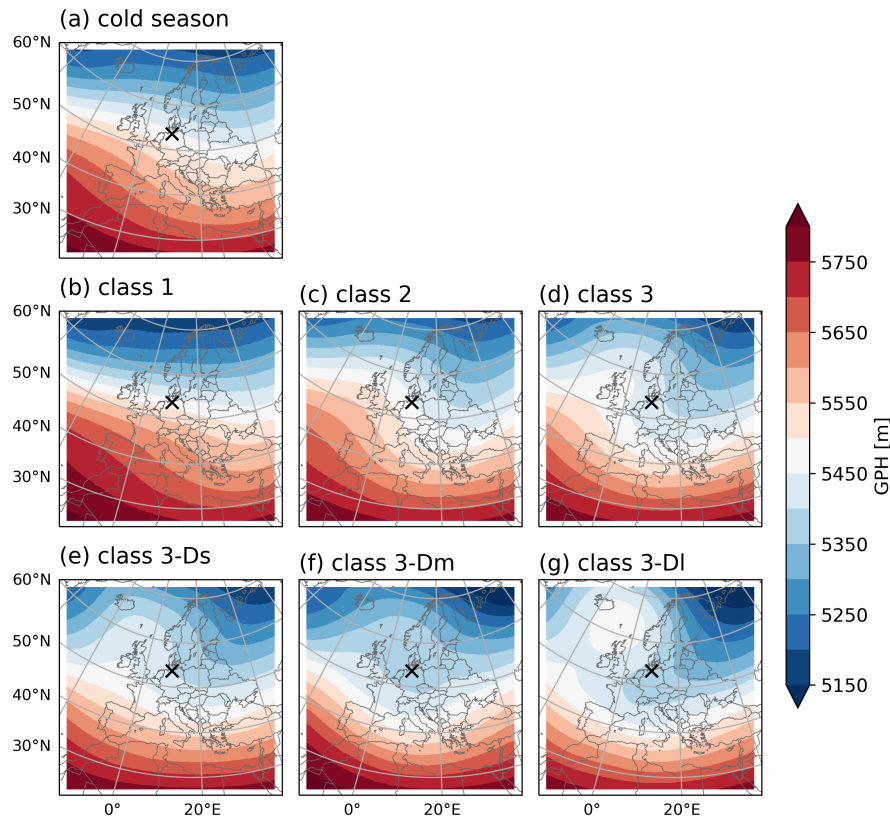
Figure 3.6. Geopotential height at 500 hPa for the defined cold season characteristic classes based on ERA5. The time period for computing average class patterns is 1979–2005. The black cross indicates the target region.

prevails in the North Atlantic, located between Iceland and the British Isles, leading to a flow from the north-east into the target region, bringing cold air from the Arctic. As described in BUEHLER et al. (2011) and PFAHL (2014), cold spells need some time to evolve during blocking situations. The development of a longer-lasting cold event (like class 3–Dl events) is therefore more probable during long-lasting blocking events.

3.4.3 Circulation patterns generated by RCMs

Mean circulation

As VAN ULDEN and VAN OLDENBORGH (2006) describe, differences in the mean circulation can indicate that climate models may have some shortcomings. Therefore, the first step is to look at the mean circulation in the models during the cold season and compare it with the pattern reference (ERA5). Fig. 3.7 shows these differences (RCM minus ERA5) for each of the RCM simulations considered here. Since the RCM ALADIN53 uses a slightly different grid, we excluded all ALADIN53 simulations from the further analysis.

For the ERA-Interim-driven runs, the GPH differences are less than $\pm$ 25 m (except
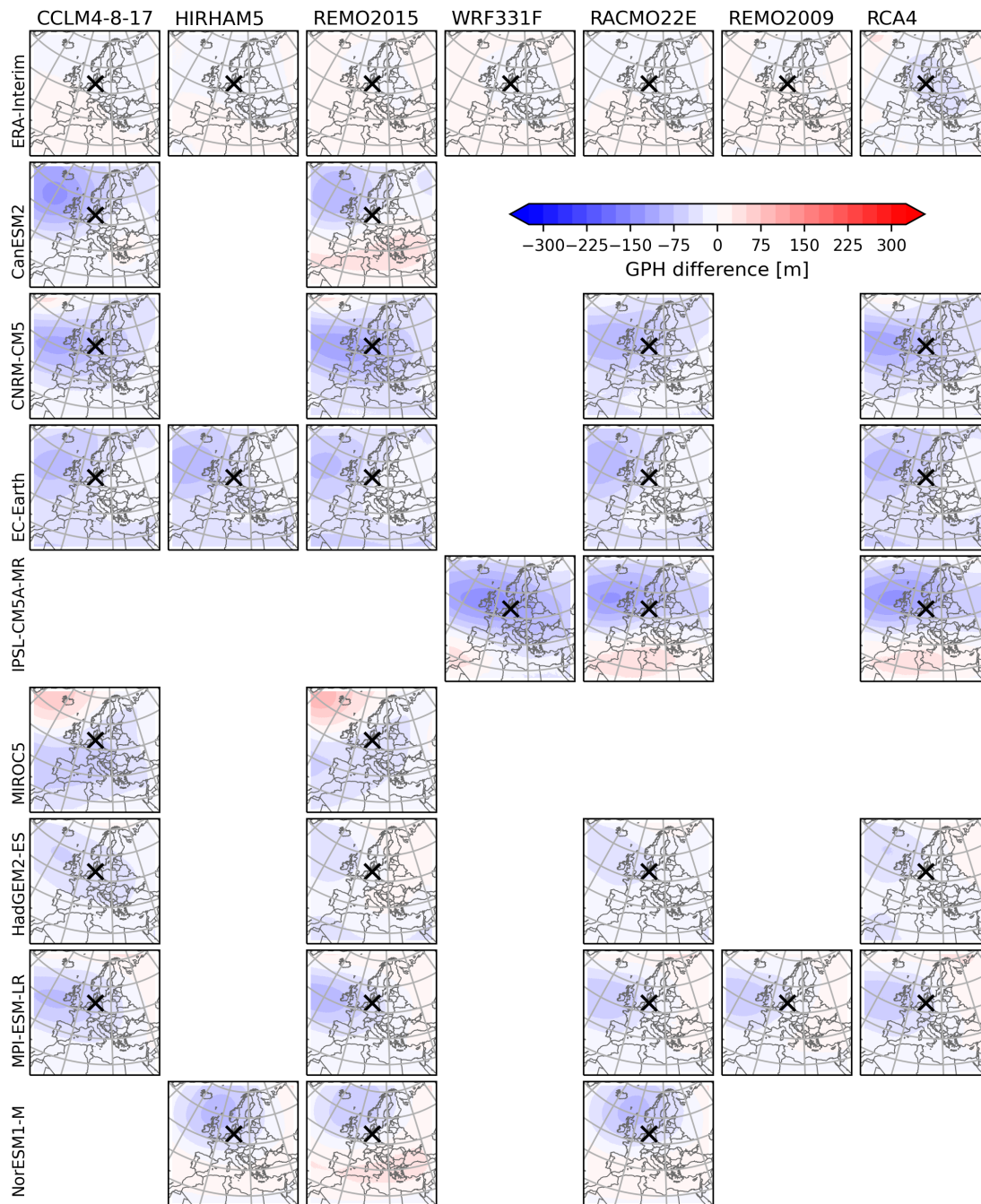
Figure 3.7. Difference in mean (1979–2005) cold season GPH at the 500 hPa level between RCM simulations and ERA5. Columns indicate the RCM, rows indicate the forcing of the RCM simulation (either ERA-Interim or GCM).

RCA4 with up to 50 m). The mean GPH (north) west of the British Isles over the North Atlantic (Fig. 3.6(a)) is underestimated in all GCM-driven RCM simulations, thus, leading to a stronger north-south GPH gradient. The underestimation is strongest for the CanESM2 and IPSL-CM5A-MR forcings with GPH difference values of up to 125 m. An exception are the RCM simulations forced by MIROC5 — the lower mean GPH values in the north are overestimated and the higher mean GPH values in the south are

underestimated. Therefore, the gradients between the southerly higher and the northerly lower GPH are reduced (GPH differences of up to ± 50 m). This highlights the fact that the mean cold season large-scale atmospheric circulation produced by the RCMs is strongly influenced by the forcing and may result in different gradients at the 500 hPa level than found in ERA5. This could lead to slightly wrong flow directions and, thus, prevent the models to e.g. produce cold spells. Overall, the GPH differences of the 500 hPa level are well below 50% of the ERA5 gradients (3.6(a)).

In order to examine the differences between the RCMs and the reference for each of the defined classes, average patterns are now calculated separately for each of the classes (Fig. 3.8). For convenience of the reader, the average class patterns for the reference (Fig. 3.6) are repeated here as the first row. For classes 1–3, the patterns for ERA-Interim-driven runs are similar to the reference patterns (Fig. 3.8, columns 1–3). The patterns for duration classes 3–Ds, and 3–Dl are also similar to the reference, with slight differences. However, there are features that are common to all RCMs. For class 3–Ds events, there is a west to east GPH gradient over the target region and an area of low GPH over (south-) eastern Europe, resulting in cold air advection from the north/northeast into the target region. For class 3–Dl events, not all models show a blocking between Iceland and the British Isles, but there is at least an area of higher GPH in this region. There is no event in this class in the WRF331F evaluation run. For class 3–Dm events, the RCMs produce different results that do not resemble the reference pattern for this class. The majority (5/8) shows a pattern with an area of high GPH over northwestern Europe and an area of lower GPH over southeastern Europe, some (2/8) show a pattern with low GPH over most of Europe and one RCM shows a pattern with high GPH over most of Europe. Overall, the "typical" patterns identified here for the ice episode classes may not be representative of individual ice episodes within the classes.

The patterns of the GCM-driven runs are similar to the reference patterns for the classes 1 and 2 (warm days, frost days). For all ice day classes there are substantial differences, which appear to be mainly caused by the forcing (i.e. the GCM). RCMs driven by the same GCM show similar patterns for all ice day classes. Figures for the GCM-driven RCM patterns can be found in the Appendix (Figures B.1–B.8).

Ice episode specific patterns

From the analysis presented above, it is evident that the RCMs produce different mean patterns for ice episodes of different durations (Class 3–Ds,–Dm,–Dl) compared to the reference. Therefore, in the following it is examined whether the RCMs are able to reproduce the event-specific reference patterns, regardless of whether an ice episode occurred in the target region of the RCM during this time or not. Thus, it is only a matter of repro-
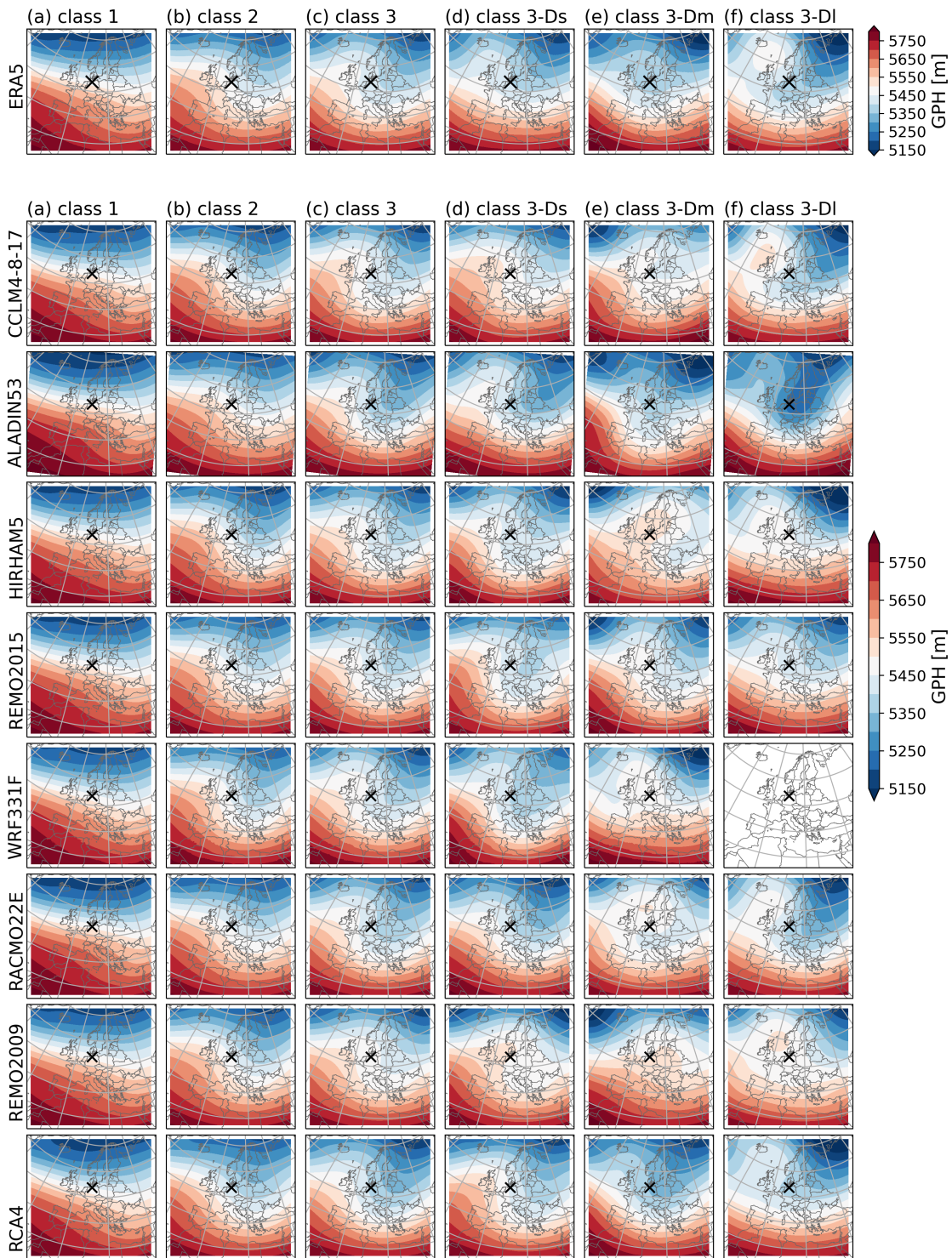
Figure 3.8. Mean class patterns derived from the GPH at 500 hPa. Shown are climate model data resulting from RCM simulations with ERA-Interim forcing. The RCMs considered are (from top to bottom): CCLM4-8-17, ALADIN53, HIRHAM5, REMO2015, WRF331F, RACMO22E, REMO2009, RCA4. The top row shows mean class patterns for ERA5.

ducing the event-specific patterns identified in the pattern reference (ERA5), independent of the timing. It is important to focus on longer-lasting events (ice episodes), as these are potentially dangerous for society and can have far-reaching consequences.

In order to be able to investigate reproducibility, the event-specific reference patterns (Fig. 3.9) are searched for in the RCMs. For this purpose, the SSIM index (described in sec. 3.3.3) is used to measure the similarity of the patterns. In addition, an ERA5 self-test is performed to provide a best case measure. The analysis considers total fields rather than anomalies. The reason for this was shown in sec. 3.4.3: The RCMs, driven by GCM simulations, produce a mean cold season flow that shows differences from the mean reference flow (up to 100 m for the GPH at 500 hPa, Fig. 3.7). If this "wrong" mean flow were subtracted to obtain anomalies, the exact features sought might be lost.
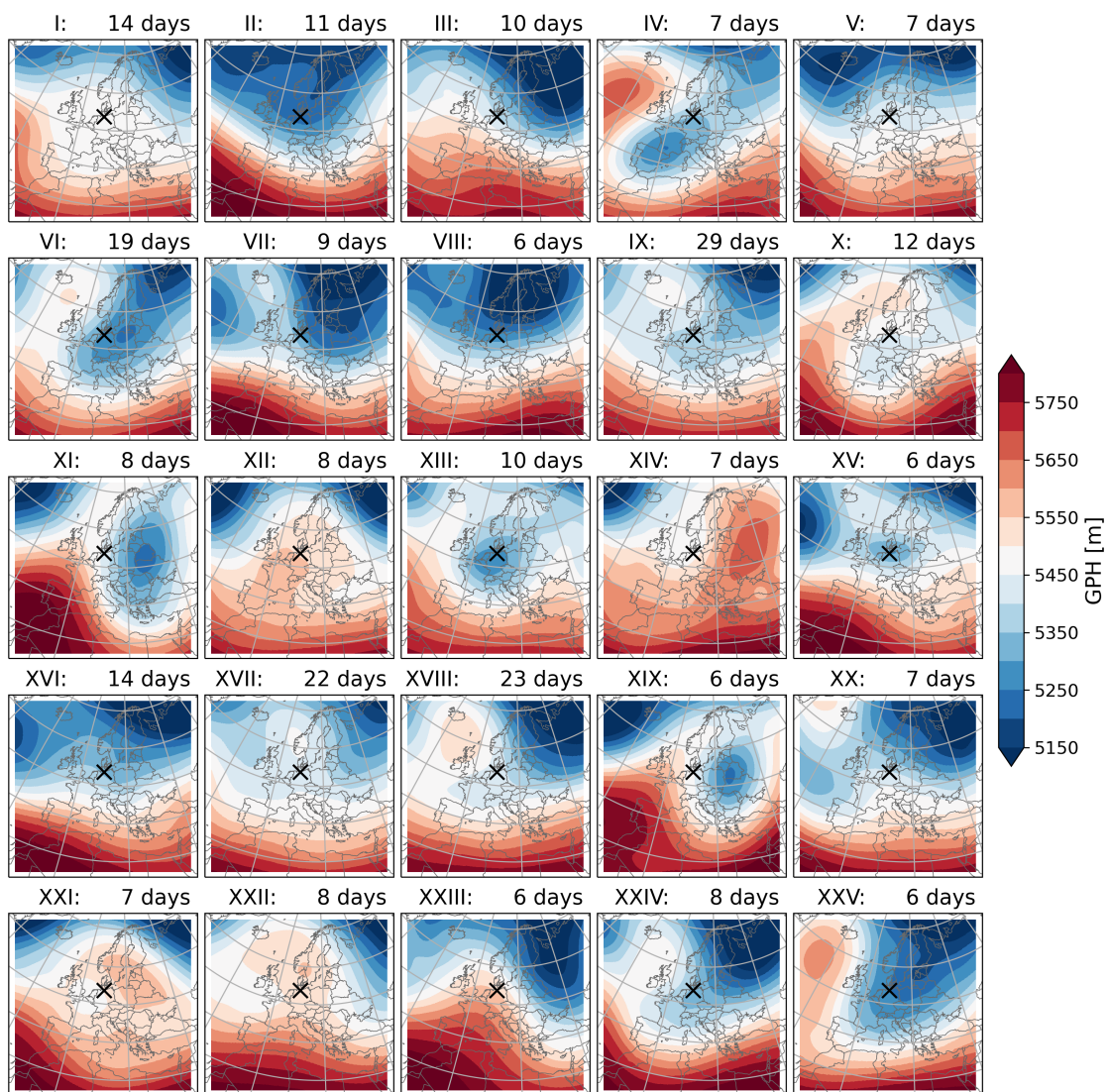
Figure 3.9. Event-specific patterns identified from ERA5 reanalysis data.

The results of the method presented here show that the event-specific reference patterns generally occur with a similar frequency in the RCM simulations as in the ERA5 self-test. However, the patterns found are not necessarily associated with extreme cold events. For many of the events found, the maximum daily temperature in the target region is above freezing, and thus it is not an ice day in the target region. This is true for both

the ERA-Interim and the GCM-driven RCM simulations as well as for the ERA5 self-test. Therefore, a second filtering stage is required to obtain only those pattern matches associated with a cold event in the target region. The generated pattern matches were therefore filtered so that only those pattern matches are returned where the daily maximum temperature in the target region within the event under consideration is below 0 °C on average. After this second filtering stage, about one third of the pattern matches remain (not shown). In a final step, the remaining pattern matches were used to derive the number of matches of individual events. This is important because in the search for the same ERA5 pattern, one event in the RCM (i.e. several average patterns on consecutive days) can theoretically generate several pattern matches and the search for longer events could artificially generate too many pattern matches. The resulting numbers of event-specific pattern matches are given as ensemble mean values using all RCMs in Table 3.3. Numbers are given for the ERA-Interim driven RCM ensemble (evaluation) and for the GCM-driven ensemble (historical), both as ensemble mean values ($\overline{\Sigma}$)). In addition, numbers are given for the ERA5 self-test (ERA5 Σ).

Table 3.3 shows that all event-specific reference patterns can be reproduced by both the evaluation RCM ensemble and the historical RCM ensemble, even when the target-region temperature is used as an additional selection criterion. Events IV (7 days), XI (8 days) and XIV (7 days) generate two pattern matches in the ERA5 self-test, indicating that these event-specific patterns occur very rarely in ERA5 during 1979–2005. The small number of pattern matches generated for these events is reflected within the assumed uncertainty ($\pm$ 2 pattern matches) in both the evaluation ensemble and the historical RCM ensemble. Individually, this does not apply to every RCM (see Appendix B, Tables B.5–B.11): Some cannot reproduce these patterns or have a too high temperature within the target region (0 pattern matches) and others generate significantly more than 2 pattern matches (maximum of 10 generated pattern matches for event pattern IV in MIROC5/CCLM4-8-17, see Appendix Table B.9). Event patterns IV and XI both show a closed system with low GPH over the European continent. Other event patterns may share similarities, but typically the low-GPH system is smaller in these cases. Examples of these include events XV (6 days) and XIX (6 days), for which the GCM-driven RCM ensemble generates the lowest number of pattern matches. Event XIV (7 days), which generates the lowest number of matches for the ERA-driven RCM ensemble and the GCM-driven RCM ensemble, is dissimilar to other event patterns and is extremely rare. This is indicated by the number of pattern matches generated in the ERA5 self-test. This event pattern shows a blocking system over Eastern Europe and the adjacent Russian areas.

The most frequent patterns are those of events XVI (14 days) and XVII (22 days). These patterns are also the most frequent in the two RCM ensembles. Similarly frequent in the

Table 3.3. Number of event-specific reference (ERA5) pattern matches in the ERA5 self-test (ERA5 Σ), the ERA-Interim driven RCM ensemble (evaluation $\overline{\Sigma}$) and the GCM-driven RCM ensemble (historical $\overline{\Sigma}$). Events considered here are ice episodes (class 3–Ds, class 3–Dm and class 3–Dl events). The red and blue numbers indicate the highest and lowest numbers per column. Rows in bold indicate event-specific reference patterns that are well reproduced by the RCMs (deviation of < 2 in the number of matches compared to the ERA5 self-test).

№	ERA5 event duration	ERA5 Σ	evaluation $\overline{\Sigma}$	historical $\overline{\Sigma}$
I	14	11	10.6	7.5
II	11	8	9.1	21.8
III	10	10	10.0	7.0
IV	**7**	**2**	**0.6**	**2.3**
V	7	5	4.7	8.2
VI	**19**	**6**	**5.1**	**6.0**
VII	**9**	**6**	**4.7**	**6.3**
VIII	6	8	4.6	12.9
IX	**29**	**10**	**8.3**	**9.1**
X	12	11	6.9	6.2
XI	**8**	**2**	**0.7**	**1.9**
XII	8	4	6.3	4.4
XIII	10	5	1.7	4.5
XIV	**7**	**2**	**1.0**	**0.6**
XV	**6**	**3**	**1.4**	**1.7**
XVI	14	12	9.0	14.2
XVII	**22**	**12**	**10.3**	**10.6**
XVIII	23	11	6.1	4.6
XIX	**6**	**3**	**2.3**	**1.8**
XX	7	7	4.4	3.3
XXI	7	9	6.9	3.8
XXII	8	6	6.9	3.6
XXIII	**6**	**3**	**3.9**	**3.8**
XXIV	8	8	5.6	7.6
XXV	6	7	6.6	10.5

ERA-Interim driven ensemble are events I (14 days) and III (10 days), which are also the second and third most frequent in the ERA5 self-test. The GCM-driven RCM ensemble generates the most pattern matches for event pattern II (11 days) and significantly overestimates its frequency (8 in the ERA5 self-test to 22 in the GCM/RCM ensemble). This pattern shows an area of low GPH extending far to the south of Europe without any distinctive features. Following event pattern XVI, which generates the second-highest number of pattern matches in the GCM/RCM ensemble and the highest number in the ERA5 self-test, event pattern VIII generates the third-highest number of pattern matches. This pattern is similar to event pattern II, but shows slightly higher GPH over Iceland. The area of very low GPH is limited to Scandinavia.

In general, the event patterns that are least overestimated or underestimated (deviation

of less than two matches) are those with a high GPH over (or east of) Iceland and/or the British Isles and a low GPH over the rest of Europe (blocking patterns in IV, VI, VII, IX, XVII) or a low GPH system in southwestern (IV), central (XV) or eastern Europe (XI, XIX). Event XXIII is also one of the least overestimated or underestimated event patterns by the RCMs. It shows a strongly meandering flow.

The results indicate a general tendency towards longer ice episodes in the target region during an Iceland/British Isles blocking. This is well reproduced by the RCMs regardless of the forcing. For shorter ice episodes, atmospheric blocking is not a reliable indicator, as short-lived low GPH systems tend to play a role in these events. This is also well reproduced by the RCMs.

3.5   Summary and conclusions

In this work, the cold season characteristics of a target region and their representation in RCMs were investigated. The cold season was divided into three characteristic classes, namely warm days, frost days and ice days, with the ice day class being further divided into three sub-classes covering ice episodes of different durations. The frequency of the classes within 1979–2005 was determined and compared between reference (E-OBS) data and RCM data.

The first research question **Are RCMs able to reproduce cold season characteristics for a specific target region** can be answered by this analysis as follows: The majority of the RCMs appear to have a cold bias for the target region. This is consistent with the analyses of other authors (e.g. KOTLARSKI et al., 2014; LHOTKA and KYSELÝ, 2018). As a result, the number of warm days is underestimated by the majority of the RCMs analysed here. On the other hand, the number of ice days is overestimated. For the number of frost days, there is no uniform behaviour of the RCMs analysed here. It can therefore be assumed that the cold bias only affects the daily maximum temperatures, but not the daily minimum temperatures.

The second research question **Are RCMs able to reproduce the frequency of ice episodes for a specific target region?** can be answered als follows: RCMs seem to produce too many ice episodes of short and medium duration. For ice episodes of long duration no clear tendency is found.

In addition to analysing the temperature based cold season characteristics, atmospheric patterns were analysed. In agreement with many other studies, it was shown that the mean cold season large-scale atmospheric circulation produced by the RCMs is strongly influenced by the forcing and that this influences the frequency of warm vs. ice days.

Mean patterns for the different cold season characteristic classes were determined and compared between reference (ERA5) and RCM data. For warm and frost days, the average class patterns in the RCMs are similar to the reference. For the ice episode classes, the average class patterns differ from the reference and are unique for each RCM simulation. Therefore, the average class patterns are not representative of ice episodes and it should be emphasized that these patterns are a possible but not a necessary condition for the occurrence of certain ice episodes.

The fact that the mean class patterns are not representative of ice episodes was again confirmed by analysing the event-specific patterns in ERA5. For the 25 ice episodes identified in the target region, the atmosphere (GPH at 500 hPa) shows many different patterns. However, there are patterns that occur more frequently in cold seasons within 1979–2005 and are associated with longer ice episodes. These are atmospheric blockings over Iceland/the British Isles. Patterns that are very rare in cold seasons within 1979–2005 and are associated with shorter ice episodes usually show a system of low GPH over the European continent.

In order to answer the third research question **Can RCMs accurately reproduce the atmospheric patterns prevailing during ice episodes?**, an image processing method was used to assess the similarity of the patterns. As with the results obtained, it was found that the RCMs are not only able to reproduce the patterns identified in ERA5, but also to reproduce their frequency without major deviations, with a few exceptions. The association of the patterns with an ice period in the target region is also reproduced by the RCMs.

Overall, the RCMs seem to be able to reproduce cold season characteristics for the target region in the ensemble average. Each RCM produces slightly different results and, therefore, individual results should be treated cautiously. In particular, results for future projections must be interpreted with caution, as the physical mechanisms that cause these events may not be perfectly represented in climate models.

Regarding the further application of the pattern-matching method presented here, it can be said that it works well in cases where one wants to search for a frequently occurring feature (e.g. blocking event) or wants to know how persistent certain features are (HOFFMANN et al., 2021). It can also be used to assess the ability of climate models, whether global or regional, to reproduce certain atmospheric features (as presented here for RCMs). All that is needed is a representative reference pattern to search for.

Limiting aspects of the work presented here include the analysis being limited to the present state. Another important aspect would be to see how the frequency of the patterns identified here changes in future climate periods. PFAHL (2014) argues that an increase in

the frequency of atmospheric blocking and a tendency towards and increased occurrence of the negative NAO phase leads to a weakening of the prevailing westerly flow. This is accompanied by increased advection of cold Arctic air masses into Europe. Accordingly, the frequency of the events analysed here is expected to increase in the future. However, there is an ongoing scientific debate about the underlying mechanisms, in particular the influence of Arctic warming on mid-latitude cold extremes and how these will change in the future (COHEN et al., 2020; BLACKPORT and SCREEN, 2020). Further research is needed to better understand these processes and how they can be better represented in (global) climate models.

Furthermore, in this work events have been analysed only in terms of their temperature. In general, winter events that are perceived as extreme, at least in the Hamburg region, are usually associated with strong winds and related storm surges or heavy snowfall. Therefore, it would also be important to analyse other variables (e.g. wind, precipitation) and their prevailing patterns.

4

Local cold season characteristics, ice episodes and their circulation patterns in a changing climate

Preface

This chapter is prepared for submission to the *Quarterly Journal of the Royal Meteorological Society*.

The references are combined with all references used in this thesis and can be found at the end of this thesis in References. The supporting information and the acknowledgments that will be submitted together with this manuscript can be found in a separate section of the Appendix (Third paper).

K. Heinke Schlünzen has contributed to the conceptualisation and contributed some ideas for analysis. K. Heinke Schlünzen and Kevin Sieck have contributed to the discussion of the results.

Abstract

This study investigates cold season characteristics and the frequency of local ice episodes in the Hamburg region of northern Germany under the RCP8.5 emission scenario using regional climate model (RCM) simulations. Three cold season characteristic classes — warm days, frost days, and ice days — are analysed for the recent past, the near future, and the far future. In addition to investigating the frequency of ice episodes for future climate periods, atmospheric patterns associated with these ice episodes are identified in the recent past using ERA5 data. The future changes of the identified atmospheric patterns are assessed using a newly developed pattern-matching method. The results indicate that despite climate change, local ice episodes will persist into the far future, with a total of eight events (ensemble mean) projected to occur within the 30 years of the far future climate period. Moreover, the frequency of atmospheric circulation patterns associated with these events is not expected to change significantly, suggesting a consistent association between identified patterns and ice episode occurrences. Notably, no definitive "typical" ice episode pattern could be identified, highlighting the complexity of atmospheric dynamics governing ice episodes. These findings provide insights into the persistence and characteristics of future ice episodes, facilitating informed adaptation and mitigation strategies in affected regions.

4.1   Introduction

Cold weather events in winter affect regions in the mid-latitudes of the Northern Hemisphere. Such events have an impact on various sectors, including critical infrastructure, agriculture, and human health. Cold weather events can disrupt transport networks and increase energy demand for heating and electricity (VAN DER WIEL et al., 2019; RANASINGHE et al., 2021; STUIVENVOLT-ALLEN and WANG, 2019). At the same time, buildings and, for example, supply pipes may be damaged (UNDERWOOD et al., 2017). Prolonged cold weather events can increase the risk of hypothermia, which can lead to a range of cold-related health problems, particularly for the elderly and the homeless (BAUMGARTNER et al., 2008; WANG et al., 2016; LEI et al., 2022; RICHARD et al., 2023).

How cold weather events will change in the future under climate change is the subject of current research and controversial debate (COHEN et al., 2020; SENEVIRATNE et al., 2021). Observations indicate that the strong warming of the Arctic leads to a weakening of the pressure gradient between the North Pole and the equator, which makes the prevailing westerly flow more unstable (more meandering) and cold air outbreaks from the Arctic could occur more frequently (e.g. FRANCIS and VAVRUS, 2012; PFAHL, 2014). On the other side, modelling studies analysing climate models find no link between Arctic warming induced sea ice loss and mid-latitude cooling through more frequent cold air outbreaks (e.g. MCCUSKER et al., 2016; BLACKPORT et al., 2019) or only a weak link between future sea ice loss and changes in winter atmospheric circulation (e.g. SMITH et al., 2022; YE et al., 2024). It is questionable whether the climate change induced warming, which is largest in the Arctic, will also warm the cold air in the Arctic to such an extent that Arctic air outbreaks will not be as extreme as they are today in the mid-latitudes. The report *The Next Frontier for Climate Change Science: Insights from the authors of the IPCC 6th Assessment Report on knowledge gaps and priorities for research* by BEDNAR-FIEDL et al. (2024) states that efforts should be made to fill the knowledge gap regarding the influence of arctic warming on mid-latitude winter weather, as it is crucial for informed policy-making in all sectors to improve resilience, sustainability and preparedness.

In previous studies on the future development of cold weather events, the term cold spell is usually used. A cold spell has no uniform definition, but usually refers to days that are below a certain percentile of the temperature (e.g. VIHMA et al. (2020); MENG et al. (2022)). As fresh water freezes when the temperature falls below 0 °C which can cause damage to e.g. plants, buildings or pipes, it is useful for a risk assessment to analyse ice days ($T_{max} < 0$ °C) and to determine events consisting of consecutive ice

days which last longer than a specified number of ice days. In this study ice episodes are determined, which are events of more than five consecutive ice days (BELL et al., 2024). The analysis of ice episodes is of significant importance for impact assessments, as the frost only develops its effect after a certain period of time in which the temperatures remain constantly below the freezing point. The current study focuses on ice episodes that occur in a local, spatially limited target region, as ice episodes do not always occur regionally at the same time and with the same duration everywhere e.g. in Europe and not even in a region, like Northern Germany. To investigate future projections that are detailed enough for local consideration, the study uses the results of regional climate model (RCM) simulations. Based on the above aspects, the following research question is posed: **Will local ice episodes still play a role at the end of this century?**

To better understand the relationship between such ice episodes and certain atmospheric conditions and whether this relationship will change in a projected future climate, the atmospheric circulation patterns that prevail during local ice episodes are identified and their characteristics are analysed. Thus, the second research question of this study is: **Will atmospheric circulation patterns associated with local ice episodes change towards the end of this century?**

Previous studies have analysed changes in the prevailing westerly flow or certain circulation types, using various data sets (Observations or Reanalysis data (FARANDA et al., 2023), global climate model (GCM) results (HANSEN et al., 2023), RCM results) and methods (Großwetterlagen classification (BECK et al., 2007; KUČEROVÁ et al., 2017; HUGUENIN et al., 2020), Grosswetterlagen classification with deep-learning approach MITTERMEIER et al. (2022) or Lamb-weather type, i.e. Jenkinson-collision classifications (HERRERA-LORMENDEZ et al., 2022) or self-organising map cluster analysis (HORTON et al., 2015; FRANCIS et al., 2023)). These studies focus solely on the atmospheric state and use it to derive a general conclusion for an entire region, such as Central Europe.

In contrast to these previous studies, the current study assesses ice episodes from a ground-level, local perspective as a starting point and analyses how the atmosphere behaves during such events. This study is unique in that it focuses not on the statistics of cold extremes that are single-day events, but on prolonged events with temperatures below 0 °C. BELL et al. (2024) has identified prevailing atmospheric circulation patterns during ice episodes based on the ERA5 reanalysis (HERSBACH et al., 2020). In addition, (BELL et al., 2024) investigated the reproducibility of the atmospheric circulation patterns identified in ERA5 in RCMs using a pattern-matching approach. The work presented here builds on these results and applies a further developed pattern-matching method to future climate periods. The aim is to determine whether the prevailing at-

mospheric circulation patterns identified in ERA5 (representing the recent climate) also occur in RCM projections of future climate periods and whether they are associated with local ice episodes. It is also investigated whether the projected global warming in the RCM reduces the effect of future Arctic cold air outbreaks and ice episodes no longer occur in the target region.

The data used are described in section 4.2, an overview of the methods is given in section 4.3. The findings are then presented in section 4.4, followed by a discussion in section 4.5. Finally, a brief summary and answers to the posed research questions are provided in section 4.6.

4.2   Data

The findings presented here are based on an analysis of the results of RCM simulations. These RCM simulations are a refinement of GCM simulations. As RCM simulations are more computationally intensive due to their higher resolution, they are calculated for smaller areas rather than globally. The EURO-CORDEX simulations (JACOB et al., 2014) analysed here are an ensemble of RCMs that refine the data of the CMIP5 GCM simulations for the European domain. This study analyses historical simulations for a recent past climate period (1971–2001) and future projections for two future climate periods, the near future (2030–2060) and the far future (2065–2095), assuming the RCP8.5 emission scenario (RIAHI et al., 2011). Table 4.1 provides a list of the GCM-RCM model chain simulations used.

As observations, E-OBS (CORNES et al., 2018) is used as a reference in the recent past climate period for ground level variables. E-OBS is based on a statistical model that converts station data into a gridded data set. The data set provides "best guess" values as mean of the 20-member ensemble that the full data set consists of. In addition to this "best guess" value, the ensemble spread is provided and taken into account here as an uncertainty in the calculation of ice days. The procedure of including this uncertainty in the calculation of cold season climate indices is described in BELL et al. (2023). In addition, the ERA5 renalysis is used as a second reference, specifically for the comparison of atmospheric circulation patterns for which the daily geopotential height at 500 hPa (HERSBACH et al., 2018a) is used.

Table 4.1. Climate model simulations used in this study. The driving global climate models (**GCM**) combined with the regional climate model (**RCM**) simulations use different internal **calendar**s. This results in a different number of **total cold season days**. The recent past and far future climate periods have the same number of total cold season days, while the near future has one day less in the case of the standard/Gregorian calendar due to one leap year less.

GCM	RCM	Calendar	total cold season days
CanESM2	CCLM4-8-17	365-day	4530
	REMO2015	proleptic Gregorian	4538
CNRM-CM5	CCLM4-8-17	proleptic Gregorian	4538
	ALADIN53	proleptic Gregorian	4538
	REMO2015	proleptic Gregorian	4538
	RACMO22E	standard	4538
	RCA4	standard	4538
EC-Earth	CCLM4-8-17	proleptic Gregorian	4538
	HIRHAM5	proleptic Gregorian	4538
	REMO2015	proleptic Gregorian	4538
	RACMO22E	standard	4538
	RCA4	standard	4538
IPSL-CM5A-MR	WRF331F	proleptic Gregorian	4538
	RACMO22E	standard	4530
	RCA4	standard	4530
MIROC5	CCLM4-8-17	365 day	4530
	REMO2015	proleptic Gregorian	4538
HadGEM2-ES	CCLM4-8-17	360 day	4500
	RACMO22E	360 day	4500
	RCA4	360 day	4500
	REMO2015	360 day	4500
MPI-ESM-LR	CCLM4-8-17	proleptic Gregorian	4538
	REMO2015	proleptic Gregorian	4538
	RACMO22E	standard	4538
	REMO2009	proleptic Gregorian	4538
	RCA4	proleptic Gregorian	4538
NorESM1-M	HIRHAM5	proleptic Gregorian	4538
	REMO2015	proleptic Gregorian	4538
	RACMO22E	365-day	4530

4.3 Methods

4.3.1 Cold season characteristics

The cold season is defined here as the months of November, December, January, February and March. To examine the temperature characteristics of the cold season within the target region and possible changes in future climate periods, all cold season days are

classified according to their temperature. The characteristic classes into which all cold season days are assigned to are ice days ($T_{max} < 0$ °C), frost days excluding ice days ($T_{min} < 0$ °C and $T_{max} \geq 0$ °C) and warm days which are the remaining cold season days. These characteristic classes are referred to as class 1 (warm days), class 2 (frost days) and class 3 (ice days).

To analyse the frequency of ice episodes of different duration, three further classes are defined, which represent subclasses of class 3. Class 3-Ds contains short ice episodes of 6–10 days, while class 3-Dm contains medium ice episodes with durations of 11–15 days. All longer events are assigned to class 3-Dl. The nomenclature of these classes is derived from their subclassification of class 3. The designation "D" — which stands for the "duration" — is followed by a letter indicating the specific duration of the ice episode, which can be classified as "short", "medium" or "long". This nomenclature was proposed in (BELL et al., 2024).

Since the GCMs and RCMs use different internal calendars and therefore may not include leap years, the individual simulations have different numbers of cold season days (Table 4.1). For this reason, the frequencies of the defined classes for each GCM-RCM simulation are given as a proportion relative to all cold season days occurring in the respective model simulation.

4.3.2 Target region

Maximum temperatures below 0 °C for several consecutive days are particularly problematic for cities and their critical infrastructure. As such temperature events, defined here as ice episodes, tend to occur locally rather than regional or continental, they are identified here for an urban target region. The Hamburg region in northern Germany was chosen as a mid-latitude target region. The Hamburg region is affected by frost and ice episodes due to its location and is therefore a representative starting point for investigating the behaviour of the atmospheric circulation in the North Atlantic/European sector during local ice episodes. For the identification of e.g. ice days, T_{max} values are calculated as spatial mean across the target region domain.

4.3.3 Prevailing circulation patterns

Numerous studies suggest that the geopotential (or geopotential height (GPH)) at the 500 hPa level is a good indicator of how weather systems move below this layer. For simplicity, the GPH at the 500 hPa level is denoted GPH_{500} hereafter. As the ECMWF, 2024 writes about their GPH_{500} charts, the contours show the main tropospheric waves that affect our (Europe) weather, with low GPH_{500} indicating troughs and cyclones in the
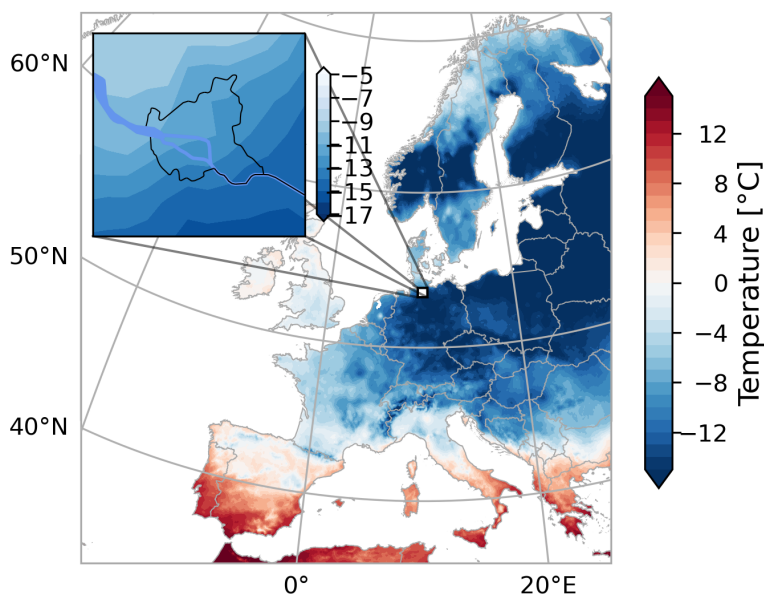
Figure 4.1. Daily average 2m temperature on the 1996-12-31 based on the E-OBS data-set. This date is the 12th day of the longest ice episode (23 days) identified in E-OBS within 1979–2005. The square marks the target region also in the enlarged figure. Note the different scalings.

middle troposphere, and high GPH_{500} indicating ridges and anticyclones. Wind speeds are proportional to the distance between contours. The denser the contours, the stronger the winds.

The GPH_{500} is used in the current study to identify representative atmospheric circulation patterns that prevail during the ice episodes that occur in the target region. The ice episodes were previously identified for the target region in BELL et al. (2024), based on ERA5 data for the analysis period of 1979–2005. The methodology for identifying representative event patterns for these ice episodes is explained below.

The identification of representative event patterns is based on the concept of finding atmospheric circulation patterns that, after applying a 3-day smoothing, are most representative of the entire ice episode identified in ERA5. A 3-day smoothing is selected in accordance with the definition of a Grosswetterlage as outlined in BAUR et al. (1944), who defined a Grosswetterlage as an atmospheric pattern that persists for at least three consecutive days. To test for representativeness, the similarity of the individual atmospheric patterns within each ERA5 ice episode was determined using a matching criterion. The matching criterion used is the structural similarity index (SSIM) (WANG et al., 2004). This index focuses on features (i.e. the structure) and outperforms other metrics such as the RMSE, which is often used to assess image quality. The SSIM can produce values between 0 and 1, where 1 means "the images are the same" and 0 means "the images are completely different".
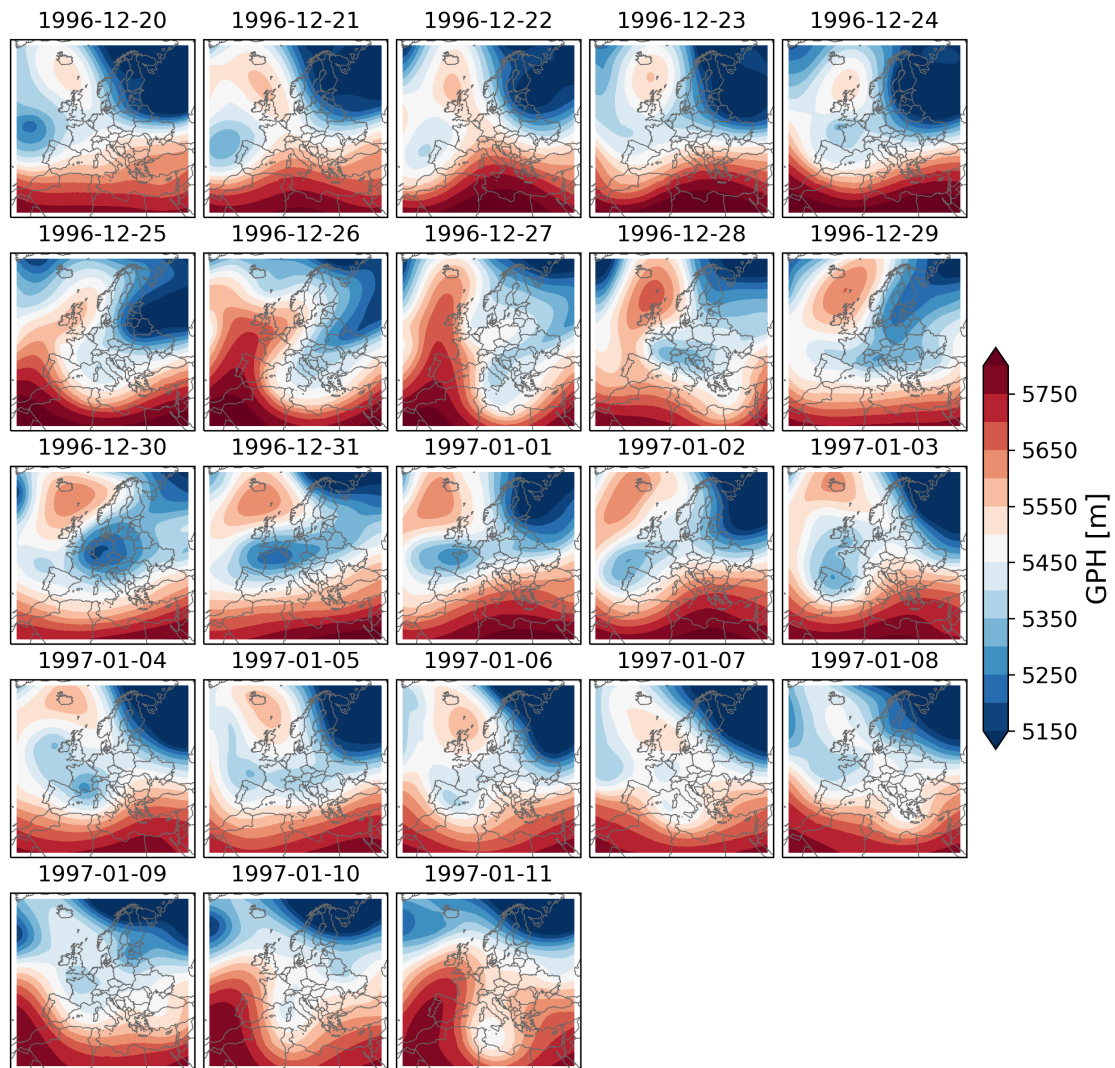
Figure 4.2. Ice episode XVIII identified for ERA5 emerged on the 20th December 1996 and lasted 23 days in the target region. Shown is the 3-day smoothed geopotential height at the 500 hPa pressure level for each day within the ice episode. Day 19 (1997-01-07) was determined as the representative event pattern.

To illustrate the atmospheric state during an ice episode, Figure 4.2 shows the GPH_{500} (3-day smoothed) for all days within a 23-day ice episode (Event XVIII). Note that for the later comparison with RCM data, the ERA5 data have been remapped to the EURO-CORDEX grid using CDO's (SCHULZWEIDA, 2023) function `remapycon`. For each day of this event, the corresponding atmospheric pattern was tested as a representative event pattern. To determine the representativeness of each day's pattern for the event in question, the similarity between it and each of the 22 other patterns within the event was calculated. The day on which the atmospheric pattern with the highest number of intra-event matches occurred was selected as the representative event pattern, in this example day 19. Details of the representative event patterns identified are given in Table 4.2. If two patterns generated the same number of intra-event matches, the minimum SSIM required to generate a match with all other days within the event was used as a second selection

Table 4.2. Details of the determined representative event patterns for ERA5 ice episodes. The ice episode number (№) and its duration (**dur**) are given in the first two columns. The day of occurrence of the identified representative event pattern within the ice episode (denoted as representative day **rd**), the persistence of this pattern within the corresponding ice episode (denoted as pattern persistence **pp**), the intra-event **mean SSIM** index and the intra-event **min SSIM** index values are given in the other columns. Numbers in parentheses in the **pp** column indicate more than one occurrence of the pattern within the corresponding ice episode. The date of the day on which the representative event pattern occurred is given in the last column (**date of rd**). See text for details of the methodology used to determine a representative event pattern.

№	dur	rd	pp	mean SSIM	min SSIM	date of rd
I	14	10	9	0.78	0.56	1980-01-17
II	11	7	9	0.86	0.77	1981-12-20
III	10	5	7	0.83	0.65	1982-01-10
IV	7	4	5	0.95	0.91	1983-02-11
V	7	3	4	0.86	0.73	1983-12-13
VI	19	10	11	0.75	0.49	1985-01-11
VII	9	4	7	0.88	0.78	1985-02-10
VIII	6	3	4	0.93	0.85	1985-12-29
IX	29	17	17(3+14)	0.77	0.63	1986-02-18
X	12	7	5	0.74	0.55	1987-01-16
XI	8	4	6	0.92	0.81	1987-03-04
XII	8	3	6	0.91	0.82	1989-12-30
XIII	10	6	7	0.87	0.72	1991-02-11
XIV	7	3	5	0.89	0.79	1993-11-23
XV	6	2	4	0.95	0.89	1994-02-20
XVI	14	7	11	0.86	0.72	1995-12-30
XVII	22	9	10(2+8)	0.75	0.55	1996-01-28
XVIII	23	19	21	0.83	0.75	1997-01-07
XIX	6	2	4	0.92	0.84	1998-12-07
XX	7	3	5	0.91	0.84	2000-12-23
XXI	7	3	5	0.94	0.85	2001-01-17
XXII	8	4	6	0.94	0.89	2002-12-12
XXIII	6	3	4	0.93	0.86	2002-12-21
XXIV	8	4	6	0.92	0.81	2003-01-07
XXV	6	3	4	0.95	0.89	2005-03-03

criterion. This ensures that a pattern with the highest similarity to all other patterns within the considered ice episode is selected as the representative event pattern. In general, the atmospheric patterns of the last two days within an event look quite different from the other patterns and are therefore not relevant to the onset or persistence of the ice episode at ground level. Therefore, the last two days are neglected when looking at the minimum SSIM.

The threshold for the similarity criterion was set to 0.75 when searching for representative event patterns. When looking at intra-event matches in ERA5 event patterns, the 0.75 threshold seems to generate matches for atmospheric patterns that appear similar enough to the sought representative event pattern to theoretically have the same impact on the

target region, which is a purely subjective judgment.

In order to investigate the extent to which atmospheric patterns that prevail during ice episodes in the target region will change in future climate periods, the patterns found as representative event patterns in ERA5 are searched for in the RCM data. In this way, it can be determined whether the representative event patterns are also characteristic of ice episodes in the GCM-driven RCM simulations and whether they will remain characteristic of these events in future climate periods. For this purpose, the pattern matching method described in BELL et al. (2024) is applied in a slightly modified form.

Similar to the changes in the methodology for identifying the representative event patterns in ERA5, the comparison between ERA5 and RCM patterns has also been modified. As the ERA5 based representative event patterns consist of a 3-day average, a 3-day smoothing is also used for the comparison with the RCM data. The day for comparison is always the centre of the averaging (i.e. $\pm$ one day). If the comparison between the representative event pattern and the 3-day average pattern of the considered RCM results in an SSIM index ≥ 0.75, this is considered as a match. If the pattern of the RCM day under consideration produces matches with several ERA5 event patterns, all matches belonging to that specific RCM day are distributed proportionally to the corresponding ERA5 event patterns. For example, if 12 January 1987 generated a match with ERA5 event patterns III, XII, XVI and XIX in any RCM simulation, 0.25 of a match is added to each of these 4 patterns, rather than a full match. Thus, each day for which at least one match is found is equally weighted, ensuring that the total number of RCM days that generated a match is equal to the total number of matches in the considered RCM. For each match generated in the model simulations, a further step is to check whether the day that generated the match falls within an ice episode in the target region. This sequence of steps is repeated for each of the 25 ERA5 event patterns against each of the 24 model simulations for the three climate periods. The result is then analysed as an ensemble mean.

Note that only 24 of the 29 available RCM simulations were used in the pattern matching. ALADIN53 data are provided on a slightly different grid, so values were missing at some boundaries of the EURO-CORDEX domain, and the HadGEM2-ES GCM uses a 360-day calendar with each month having 30 days, which caused problems when using `python datetime` functions in the developed algorithm. Therefore, the simulation of CNRM-CM5/ALADIN53 and all HadGEM2-ES forced RCM simulations were neglected.

In order to be able to deduce from the number of matches how, for example, the ratio of one pattern to the other will develop in future climate periods, in addition to the total number of matches (any day or within an ice episode), the percentage share of this pattern in the total number of matches (any day or within an ice episode) was calculated. This is

done separately for each model simulation and the result is then combined and analysed as an ensemble mean. In order to derive a tendency of change from this ensemble mean, it is necessary to define conditions under which a change can be seen as a tendency. The conditions formulated here are based on the assumption that all patterns occur with equal frequency (equal distribution). For each of the 25 identified ERA5 patterns, this would result in a frequency of 4 %. A relevant change, i.e. a tendency, is found if, for example the frequency of a pattern in the near future differs by 10 % of the 4 % assumption compared to the frequency of this pattern in the recent past. So a tendency is found, when the change in frequency is greater than or equal to 0.4 %. There are two cases to consider:

(1)   if $|\Delta_{near-recent}| < |\Delta_{far-recent}|,$ then $|\Delta_{near-recent}|, |\Delta_{far-near}| \geq 0.4\%$

(2)   if $|\Delta_{near-recent}| \geq |\Delta_{far-recent}|,$ then $|\Delta_{near-recent}| \geq 0.4\%$ and $|\Delta_{far-recent}| \geq 0.8\%$

where $\Delta_{near-recent}$ is the change in frequency in the near future relative to the recent past, $\Delta_{far-near}$ is the change in frequency in the far future relative to the near future, and $\Delta_{far-recent}$ is the change in frequency in the far future relative to the recent past. In case (1), the change in frequency in the far future $\Delta_{far-recent}$ is greater than the change in the near future $\Delta_{near-recent}$, so a steady change in frequency can be seen. A tendency is then found if the change in frequency in the near future $\Delta_{near-recent}$ is greater than (or equal to) 0.4 % and the change in frequency in the far future (relative to the near future) $\Delta_{far-near}$ is also greater than (or equal to) 0.4 %. In case (2), the change in frequency in the far future $\Delta_{far-recent}$ is less than (or equal to) the change in frequency in the near future $\Delta_{near-recent}$, so no steady change can be seen. However, a tendency is found, if the change in frequency in the near future $\Delta_{near-recent}$ is greater than (or equal to) 0.4 %, and the change in frequency in the far future $\Delta_{far-recent}$ is at least as large as the total change in case (1), so must be at least 0.8 %. Finally, the last criterion for a tendency is that both changes relative to the recent past must be in the same direction. Changes in opposite directions are indicative of (multi-) decadal variability, and tendencies cannot be clearly identified in these cases. The term tendency is used because no statistical significance tests have been carried out.

The entire pattern-matching method was also applied to ERA5 itself (ERA5 self-test) to see how the identified patterns are distributed across all ice episode days in ERA5. The results of the self-test are shown in Figure 4.6 and described in section 4.4.2.

4.4   Results

4.4.1   Frequency of cold season classes and number of ice episodes

Figure 4.3 shows the frequencies of the cold season classes, i.e. the cold season temperature characteristics, for three climate periods — the recent past, the near future and the far future — based on RCM results assuming the high emission RCP8.5 scenario. The box plot shows the RCM ensemble distribution, where the black horizontal line represents the ensemble median and the box represents the interquartile range (IQR) from the first quartile to the third quartile (25th to 75th percentile) of all RCM results. The whiskers indicate a distance of $1.5 \times$ IQR from the first and third quartiles and represent the ensemble range. All other data points outside this range are shown as outliers. For comparison, the E-OBS values are shown by the blue dashed lines and their uncertainty is shown by the blue shading, both for the recent past.

Not surprisingly, the frequency of warm days (class 1) increases in the future under the assumption of the RCP8.5 scenario (Figure 4.3a). While E-OBS (blue dashed line in Figure 4.3a) shows a proportion of 58 % warm days in the recent past for the target region, the GCM-driven RCM simulations show a slightly lower value (ensemble median), but this is within the E-OBS uncertainty range. The range of the RCM ensemble is 58 % and lies between 21 % and 79 %. The proportion of warm days increases in the ensemble median to 70 % in the near future and to 80 % in the far future, while the range is slightly reduced (50 % and 30 %, respectively).

The proportion of frost days (Figure 4.3b) is similar in E-OBS and in the RCMs (ensemble median) in the recent past. The ensemble median decreases from 30 % to 25 % in the near future and to 15 % in the far future. With 42 % for the near future and 35 % for the far future, the ensemble range is somewhat larger than in the recent past climate period (32 %).

Ice days cover the smallest proportion of all cold season days, and are somewhat overestimated in the RCMs (Figure 4.3c, see also Bell et al. (2024)). The ensemble median decreases from 18 % to 10 % in the near future and to 5 % in the far future. The ensemble range decreases from 20 % in the recent past to 10 % in the near future and to 8 % in the far future. This means that the RCM results for the proportion of ice days are closer together than for the proportion of warm and frost days.

The division of ice episodes into three classes with different ice episode durations, shows a general decrease of ice episodes. While class 3-Ds events accounted for 4 % of all cold

season days ($\approx$ 182 days), in the ensemble median this class accounts for 2 % of days in the near future and for 1 % of days in the far future, with 1 % of days corresponding to about 45 days (Figure 4.3d). The proportion of class 3-Dm events was just over 2 % in the recent past. This class accounts for about 1 % of all cold season days in the near future and 0.5 % of days in the far future (Figure 4.3e). The results for class 3-Dl events are very similar, although the range within the recent past is larger (Figure 4.3f).
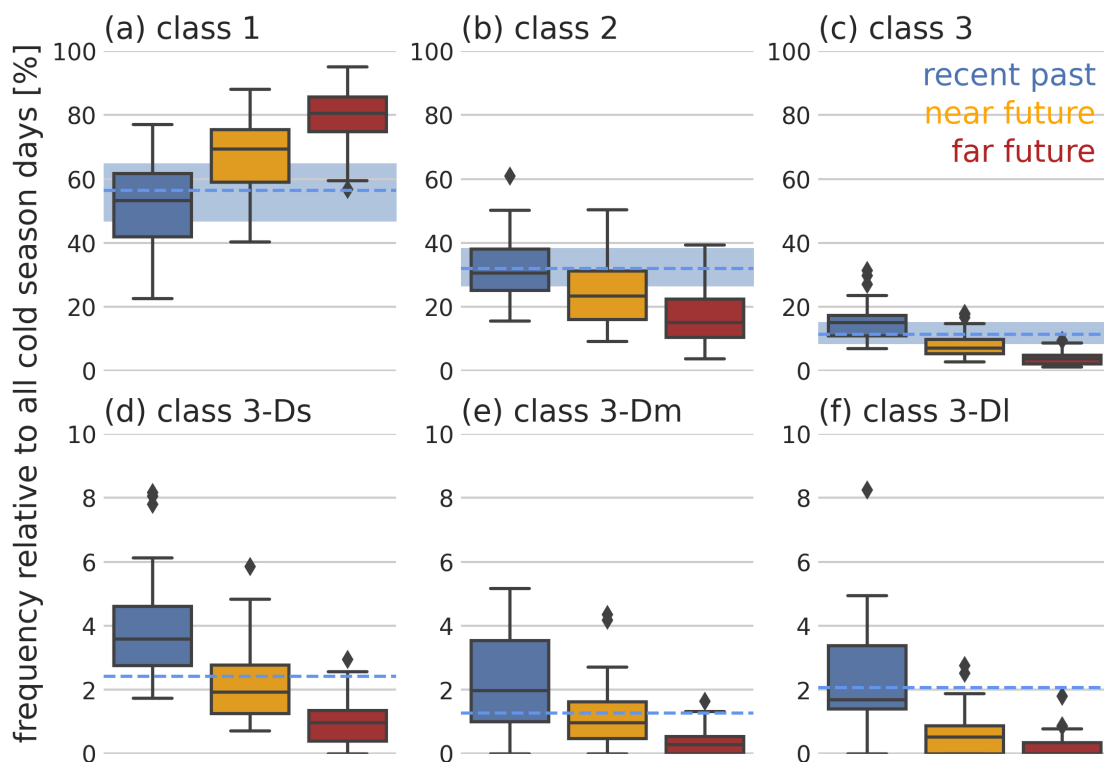
Figure 4.3. Days of the cold season classified into (a) warm days, (b) frost days, and (c) ice days. The ice day class is further divided into classes of different ice episode lengths: (d) Ds: 6–10 consecutive ice days, (e) Dm: 11–15 consecutive ice days and (f) Dl: more than 15 consecutive ice days. The frequencies are always given as proportion relative to all cold season days within the respective climate period. Blue dashed lines indicate the E-OBS value for the recent past, blue shading indicates the E-OBS uncertainty range.

Looking at the number of ice episodes for each RCM simulation individually (Table 4.3), it is clear that the driving GCM has an influence on the total number and also on its change. For example, the change in the number of ice episodes in all RCM simulations with CanESM2 or EC-Earth forcing is > 60 % in the near future. At the same time, the number of ice episodes in the recent past is much lower in RCM simulations with CanESM2 forcing ($\approx$ 15) than with EC-Earth forcing ($\approx$ 44). However, it is clear that the RCMs themselves also have an influence on the number of ice episodes. For example, the RCMs RACMO22E and RCA4 show the general tendency to have the lowest number of ice episodes compared to the other RCMs using the same GCM as forcing. The RCMs CCLM4-8-17, REMO2015 and HIRHAM5 show the general tendency to have the highest

number of ice episodes.

While there were 41 ice episodes in the ensemble mean in the recent past (minimum 15, maximum 87), the number is reduced to 21 in the near future (minimum 6, maximum 52) and to 8 ice episodes in the far future (minimum 1, maximum 23). All model simulations analysed here therefore show at least one ice episode within the far future, assuming RCP8.5.

Table 4.3. Number of ice episodes in the recent past, near future and far future climate periods together with the percentage change relative to the recent past climate period. The last row shows the ensemble mean values.

GCM	RCM	hist	near	diff [%]	far	diff [%]
CanESM2	CCLM4-8-17	17	7	-59	3	-82
	REMO2015	21	6	-71	3	-86
CNRM-CM5	CCLM4-8-17	60	43	-28	21	-65
	ALADIN53	34	22	-35	8	-76
	REMO2015	87	52	-40	23	-74
	RACMO22E	52	31	-40	8	-85
	RCA4	41	28	-32	11	-73
EC-Earth	CCLM4-8-17	50	18	-64	15	-70
	HIRHAM5	56	17	-70	11	-80
	REMO2015	44	17	-61	15	-66
	RACMO22E	37	6	-84	2	-95
	RCA4	31	8	-74	12	-61
IPSL-CM5A-MR	WRF331F	56	37	-34	10	-82
	RACMO22E	42	13	-69	1	-98
	RCA4	25	10	-60	1	-96
MIROC5	CCLM4-8-17	79	45	-43	20	-75
	REMO2015	86	41	-52	12	-86
HadGEM2-ES	CCLM4-8-17	41	18	-56	3	-93
	REMO2015	36	20	-44	4	-89
	RACMO22E	30	8	-73	3	-90
	RCA4	31	11	-65	2	-94
MPI-ESM-LR	CCLM4-8-17	32	22	-31	8	-75
	REMO2015	54	13	-76	6	-89
	RACMO22E	24	12	-50	1	-96
	REMO2009	42	25	-40	9	-79
	RCA4	20	12	-40	7	-65
NorESM1-M	HIRHAM5	28	20	-29	9	-68
	REMO2015	24	17	-29	9	-62
	RACMO22E	18	16	-11	8	-56
ensemble mean		41	21	-49	8	-80

4.4.2 ERA5 ice episode patterns

The atmospheric circulation patterns prevailing during local ice episodes are identified in ERA5 based on the GPH_{500}. In general, ice episodes in the target region can be associated with a variety of different atmospheric circulation patterns. As can be seen in Figure 4.4, there are 25 more or less distinct atmospheric circulation patterns for the 25 local ice episodes identified in ERA5. These 25 atmospheric circulation patterns are hereafter referred to as representative patterns. To assess how similar the 25 representative patterns are to each other, an ERA5 self-test matrix is generated (Figure 4.5).
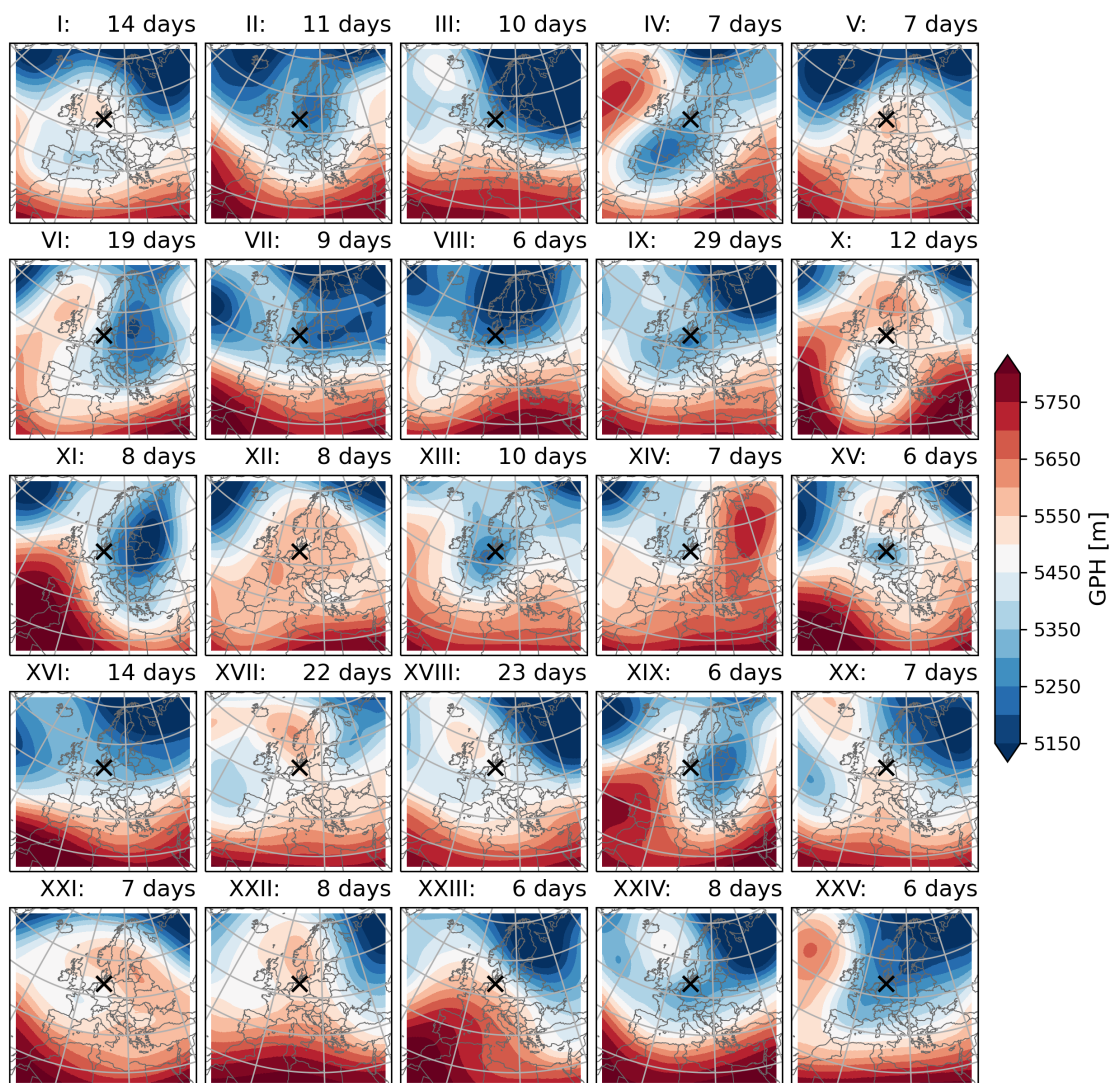
Figure 4.4. Representative patterns identified in ERA5, that prevail during local ice episodes in the target region. Shown is the geopotential height at the 500 hPa pressure level. The cross (×) marks the target region.

In principle, local ice episodes can be associated with atmospheric circulation patterns in which the prevailing westerly flow is somehow disrupted. The most frequent representative patterns (Figure 4.6, all) are the representative patterns of ice episode V (high

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX	XXI	XXII	XXIII	XXIV	XXV
I	1				0.78													0.83							
II		1														0.75									
III			1				0.75									0.81		0.76		0.81			0.78	0.75	
IV				1																					0.84
V	0.78				1																			0.75	
VI						1													0.82						
VII			0.75				1		0.79							0.88		0.78						0.84	
VIII								1								0.75									
IX							0.79		1				0.76			0.84	0.75	0.86			0.77	0.78		0.9	0.79
X										1															
XI											1								0.89						
XII												1									0.77				
XIII									0.76				1												0.79
XIV														1											
XV															1						0.78				
XVI		0.75	0.81				0.88	0.75	0.84							1		0.84		0.79				0.82	
XVII									0.75								1	0.79		0.75		0.83			
XVIII	0.83		0.76				0.78		0.86							0.84	0.79	1			0.81	0.76	0.82	0.85	
XIX						0.82					0.89								1						
XX			0.81													0.79	0.75	0.81		1		0.76			
XXI									0.77			0.77			0.78			0.76			1	0.78			
XXII									0.78								0.83	0.82		0.76	0.78	1		0.75	
XXIII			0.78	0.75																			1		
XXIV			0.75				0.84		0.9							0.82		0.85				0.75		1	
XXV				0.84					0.79				0.79												1

Figure 4.5. Resulting SSIM index values from applying the pattern-matching method to ERA5 as a self-test. The 25 representative patterns identified in ERA5, that prevail during local ice episodes in the target region are compared. The resulting SSIM value, i.e. their similarity, is noted for each pair of comparisons.

GPH_{500} with weak gradients over most of Europe) and ice episode XXIII (meandering character, GPH_{500} gradients in southwest to northeast direction without special features), but can rarely be associated with an ice episode in the target region (Figure 4.6, within ie). Due to their meandering character, there is a similarity between these two representative patterns (Figure 4.5).

The representative patterns most frequently associated with an ice episode in the target region are the representative patterns of ice episode IX (low GPH_{500} over most of Europe, with slightly larger GPH_{500} between the British Isles and Iceland) and ice episode XXI (large GPH_{500} over north-eastern Europe, low GPH_{500} over south-western Europe). A pattern with larger GPH_{500} located somewhere between the British Isles and Iceland, as

ie pattern		matches	
№	ie dur	all	within ie
I	14	4.6	5.1
II	11	8.0	4.3
III	10	2.4	5.1
IV	7	1.3	2.1
V	7	17.3	3.2
VI	19	2.6	5.0
VII	9	0.7	3.3
VIII	6	7.0	2.8
IX	29	1.9	6.5
X	12	1.7	3.0
XI	8	2.0	2.1
XII	8	4.0	4.7
XIII	10	3.6	3.7
XIV	7	1.4	2.0
XV	6	1.0	3.6
XVI	14	4.3	5.8
XVII	22	0.9	3.1
XVIII	23	2.2	5.6
XIX	6	3.8	3.4
XX	7	1.3	2.1
XXI	7	3.3	6.9
XXII	8	3.6	6.0
XXIII	6	15.3	1.6
XXIV	8	0.9	4.0
XXV	6	4.8	5.0

Figure 4.6. Number of pattern matches resulting from the ERA5 self-test. The representative patterns of the 25 ice episodes (ie pattern) are searched for in ERA5. The number () of each ice episode and its duration (ie dur) are given in the first two columns, the pattern matches are given in the last two columns. The number of all matches is given as a percentage of all cold season days (all), and the number of matches within an ice episode (within ie) is given as a percentage of all ice episode days. Colours indicate the highest and lowest number of matches per column, with reds indicating higher values and blues indicating lower values.

in the representative pattern of ice episode IX, is known as a blocking pattern, which can cause extreme weather events at different locations in Europe (KAUTZ et al., 2022). The representative pattern of ice episode XVIII is another good example of such a blocking situation. In addition to the similarity between the representative patterns of ice episode IX and XVIII, eight other ice episodes are associated with a similar blocking pattern (I, III, VII, XVI, XVII, XX, XXII, XXIV). However, these 10 representative patterns have different GPH_{500} values and slightly different locations of the blocking position, resulting in different gradient orientations and magnitudes.

The representative patterns of ice episodes VI, XI and XIX are similar to each other. These representative patterns consist of a ridge over the western coast of Europe extending northwards to the British Isles and a trough/cyclone extending southwards over eastern Europe. Another group of similar patterns includes the representative patterns of ice episodes IV and XXV, which also show a ridge-trough pattern. However, this is located more westerly: the ridge is located west of Europe and extends further north to Iceland, and the trough is located across northern and central Europe. The representative

pattern of ice episode XXV is also similar to the representative patterns of ice episodes IX and XIII.

The representative patterns of ice episodes II and VIII show similarities with the representative pattern of ice episode XVI, but not with each other. Both show low GPH_{500} extending towards the Alpine region and large GPH_{500} in the southern Mediterranean region, hence a north-south gradient. As with the patterns just described, the representative patterns of ice episodes XII and XV show similarities with the representative pattern of ice episode XXI, but not with each other. In these three patterns, areas of low GPH_{500} are located far to the north (except for the area towards Greenland) and larger GPH_{500} are found across Europe. The pattern of ice episode XV shows a small system of low GPH_{500} over the target region.

Two patterns that bear no resemblance to any of the other patterns, nor to each other, are the representative patterns of ice episode X and XIV. The representative pattern of ice episode X shows a high GPH_{500} system over Scandinavia and a low GPH_{500} system south of it (northwestern Mediterranean area). The representative pattern of ice episode XIV shows a ridge across the Russian border and south-eastern Europe with a relatively high GPH_{500} in the centre.

Overall, the frequency of each of the 25 patterns associated with ice episodes is much more evenly distributed (1.6–6.9 %, Figure 4.6, within ie) than the overall frequency of the patterns in the cold season (0.9–17.3 %, Figure 4.6, all).

4.4.3   Ice episode patterns in future climate periods

For the recent past, the near future and the far future, the representative event patterns were searched for in GCM-driven RCM simulations. The results are shown for each event pattern relative to the total number of matches per climate period and are summarised in Fig. 4.7. These are described in more detail below.

Certain representative patterns occur more frequently in the GCM-driven RCM simulations, while others occur less frequently. Similar characteristics were found in the ERA5 self-test (Figure 4.6). As in ERA5, the representative patterns of ice episodes V and XXIII occur frequently during the cold season (Figure 4.7, all matches, past). In addition to these two patterns, those of ice episodes II, VIII and XVI are also frequently found in the cold season of the RCMs (Figure 4.7, all matches, past). Patterns that are found least frequently in the RCMs during the cold season ($\leq 1\%$) are the representative patterns of ice episodes VII, X, XIV, XVII, and XX. A similar characteristic was also found for ERA5. The characteristics of the recent past frequency distribution of the representative

ie pattern		all matches			within ie				percentage of RCMs		
№	ie dur	past	near	far	past	near	far		past	near	far
I	14	1.9	1.8	1.7	1.7	1.9	1.8		91.7	79.2	62.5
II	11	12.3	12.2	11.5	10.0	8.6	5.8	↓	100.0	100.0	83.3
III	10	2.2	2.2	2.1	3.8	3.1	4.4		100.0	95.8	58.3
IV	7	1.1	1.2	1.1	2.1	1.8	3.7		95.8	79.2	83.3
V	7	12.4	11.0	11.0	5.0	5.1	3.3		100.0	100.0	75.0
VI	19	3.3	3.2	3.0	5.3	3.8	3.4	↓	100.0	100.0	83.3
VII	9	1.0	1.0	1.0	2.6	3.8	3.2		100.0	100.0	66.7
VIII	6	11.7	12.3	12.7	↑	5.9	4.6	4.3	100.0	95.8	87.5
IX	29	2.2	2.0	1.8	3.7	4.1	4.9	↑	100.0	95.8	91.7
X	12	1.0	0.9	1.0	3.0	2.7	3.4		100.0	83.3	79.2
XI	8	2.5	2.8	2.6	4.9	6.8	5.3		100.0	100.0	70.8
XII	8	2.4	2.3	2.3	2.4	2.2	3.1		83.3	70.8	58.3
XIII	10	4.2	4.9	5.0	3.6	3.8	4.6		100.0	95.8	79.2
XIV	7	0.6	0.9	0.9	0.5	0.1	0.4		66.7	37.5	25.0
XV	6	1.5	1.3	1.2	2.9	2.0	2.7		95.8	83.3	70.8
XVI	14	9.4	9.3	9.3	8.5	8.4	7.4		100.0	100.0	79.2
XVII	22	0.4	0.5	0.4	0.7	0.8	0.8		87.5	79.2	45.8
XVIII	23	1.6	1.5	1.4	2.3	2.1	2.3		100.0	91.7	75.0
XIX	6	3.5	3.8	3.8	4.9	5.1	4.4		100.0	100.0	87.5
XX	7	0.8	1.0	1.0	1.0	0.7	0.5		100.0	58.3	41.7
XXI	7	3.3	3.1	3.2	4.6	4.2	6.3		100.0	83.3	83.3
XXII	8	3.3	3.4	3.6	5.6	5.1	5.5		100.0	95.8	83.3
XXIII	6	11.4	10.9	11.7	0.8	1.1	0.8		83.3	70.8	41.7
XXIV	8	1.8	2.0	1.9	5.0	9.4	6.4	↑	100.0	100.0	87.5
XXV	6	4.3	4.5	4.6	9.6	8.6	11.3		100.0	95.8	95.8

Figure 4.7. Number of pattern matches in the GCM-driven RCM ensemble for three climate periods (recent past, near future, far future) given as percentage for each pattern relative to the sum of all matches within the respective climate period. The ensemble mean value of percentages in the RCMs is provided for all matches and for those matches that occurred within a ice episode (within ie). Colors indicate the highest and lowest number of matches per column. Reds indicate higher values and blues indicate lower values. For the percentage of RCMs that resulted in at least one pattern match within a ice episode (last three columns) white does indicate a higher number of models, dark grey does indicate a lower number of models.

patterns found in the RCMs do not change in the two future climate periods. An exception is the representative event pattern of the ice episode VIII (one of the more frequent ones), which shows a tendency to increase towards the far future (11.7 % to 12.7 %). In general, most matches are found in five patterns with little change in the near and far future (∼ 56 % of all matches). The remaining 20 patterns have an average share of 2.1 % (0.4–5 %). In summary, it can be said that the representative event patterns are found in the GCM-driven RCM simulations with a similar frequency within the cold season as in ERA5, and that this frequency remains stable towards the end of this century.

If only those matches are analysed in which an ice episode (ie) occurred in the RCM simulation for the target region, a slightly different characteristic of the frequencies can be observed (Figure 4.7, within ie). In the recent past, the representative patterns of ice episodes II, XVI and XXV are the most frequent. While the representative patterns of ice episode XVI and XXV are still among the most frequent patterns in the far future, the representative pattern of ice episode II almost halves its frequency in the far future and thus shows a tendency to decrease. The three least frequent patterns, as in the analysis of all matches, are the representative patterns of ice episodes XIV, XVII and XX, which

are still the least frequent patterns in the far future and are found in the fewest RCMs (25–46% of RCMs in the far future).

One pattern that is one of the most frequent patterns when looking at all matches and one of the least frequent when looking at matches within an ice episode is the representative pattern of ice episode XXIII. This was also found for ERA5. Just over a third of all pattern matches within an ice episode show counter-directional changes in the near future and the far future. As this is not the case for the frequency of all pattern matches, it can be assumed that the occurrence of ice episodes in the target region is subject to decadal or multi-decadal variability.

By comparing the number of ice episode days with the number of ice episode days that generate a match with the representative patterns, it can be shown that over two-thirds of all ice episode days in the RCMs can be associated with the representative patterns identified in ERA5. This proportion is consistent for the recent past, the near future and the far future (see Table 4.4). The patterns that are projected to occur during ice episodes in Hamburg in the far future, as simulated by the RCMs under the RCP8.5 scenario, are presented in the Appendix C.

Table 4.4. Number of ice episode (ie) days vs. number of matches within ie (ie days with match). The left over days without match are given and the proportion of ie days with a match relative to all ie days is given and denoted as captured ie characteristic.

	ERA5	recent past	near future	far future
ie days	270	460	206	85
ie days with match	256	307	135	59
ie days without match	14	153	71	26
captured ie characteristic	94.8 %	66.7 %	65.5 %	69.4 %

4.5   Discussion

As mentioned in the introduction, the field of classification of atmospheric circulation patterns (or types) and their changes is very diverse in terms of the selection of underlying data and methods. For example, HUGUENIN et al. (2020) use the results of a CMIP5 GCM ensemble (23 members) and the CESM12-LE large ensemble (84 members) for their analysis. They classified GPH$_{500}$ in the main circulation types (CTs). No clear change in the individual large-scale weather types for the winter (DJF) (2070–2099 compared to 1988–2017) could be identified. MITTERMEIER et al. (2022) classified the data of the SMHI-LENS large ensemble (50 members, based on the GCM EC-Earth3) for the winter half-year (ONDJFM) into 29 subjective circulation types (Grosswetterlagen after Hess and Brezowsky (e.g. GERSTENGARBE and WERNER, 2005)) using a deep learning

approach. Assuming the SSP37.0 scenario, they found a significant increase in an easterly flow type ("High Fennoscandia") and a southerly flow type ("Low British Isles") and a decrease in the cyclonic and anticyclonic south-easterly flow types (2071–2100 compared to 1991–2020). Overall, a significant change was observed in 34% of all CTs during the winter season. HANSEN et al. (2023) also analysed data from the SMHI-LENS. In their study, CTs were determined based on six SSP scenarios with 50 members each, based on the sea level pressure. The clustering method SANDRA (Simulated Annealing and Diversified Randomisation, see PHILIPP et al. (2010)) was used for this purpose. No significant frequency changes were found for CTs in winter (2081–2100 compared to 1995–2014). OZTURK et al. (2022) analysed EURO-CORDEX RCM simulations, as in the study presented here, to investigate the changes in atmospheric circulation over Europe for different global warming levels. They employed a pattern-scaling approach and analysed a number of variables, including the monthly mean GPH_{500}. They found a robust increase in the GPH_{500} in winter over the Mediterranean and Central Europe (2080–2099 compared to 1985–2004). This suggests an intensification of the anticyclonic circulation.

The above summary of these studies demonstrates that different studies arrive at different conclusions. This is due to the differing classification methods and the similarity criteria employed (HANSEN et al., 2023). Additionally, all studies use different data sets as well as different reference and future climate periods, which also differ in their duration. The pattern matching method developed in the present study represents a novel approach that employs the SSIM index as a similarity criterion, specifically designed for the comparison of features within images. Furthermore, it considers the magnitude of the pressure gradients by using a contrast comparison function. The pressure gradient magnitude is not considered, for example, in the Grosswettertypen classification, which makes it difficult to estimate the influence of CTs on other variables such as temperature and precipitation (HANSEN and BELUŠIĆ, 2021; HANSEN et al., 2023).

All of these studies take into account the atmospheric circulation at each point in time for a specific analysis period, allowing them to make statements about general CT changes based on the model data used. The approach presented here considers only days on which ice episodes appear for the recent past and future climate periods and classifies only their corresponding atmospheric circulation patterns. Therefore, this work cannot be directly compared with other studies. When all days are classified, as is done in the aforementioned and many other studies, the change in circulation patterns of the rarely occurring ice episodes in a future climate might be superimposed by the overall winter circulation type trend and thus not be seen. Therefore, considering only days on which some kind of extreme weather occurs will help to understand how the associated

atmospheric circulation patterns change. Furthermore, an event definition is chosen here that is physically relevant due to the 0 °C limit.

In general, the results presented here depend on the representativeness of the ice episode patterns. The approach presented here focusing on a relatively small target region and analysing how the atmosphere behaves during certain events has advantages, precisely because it was found here that the atmosphere for local ice episodes varies greatly and there is therefore no typical pattern. To categorise a whole season for a whole region (a whole country or a whole continent), it would be sufficient to look for the atmospheric pattern that represents the negative phase of the North Atlantic Oscillation (NAO$^-$). The NAO$^-$ is accompanied by colder and drier winters than normal almost everywhere in Europe (MARSHALL et al., 2001; HURRELL et al., 2003). As demonstrated by VIHMA et al. (2020) in their large-scale pattern analysis, the NAO$^-$ correlates with the divergence of dry-static energy transport and is associated with cold spells. In their study cold spells correspond to extreme temperature anomaly events and are defined as the detrended temperature anomaly $\leq -2\times$ standard deviation for at least four consecutive days).

The NAO index (taken from NOAA, CLIMATE PREDICTION CENTER, analysis not shown) is not a reliable indicator for the occurrence of ice episodes in the target region considered in the present study, as the NAO index also takes positive values up to a value of 0.8 as the mean value within an ice episode. As previously stated, the NAO index is more effective for categorising an entire cold season in a region as mild or severe, or for predicting the trend of the upcoming cold season with seasonal prediction systems, as demonstrated in e.g. SCAIFE et al. (2014) and DOBRYNIN et al. (2018).

4.6 Summary and Conclusions

The occurrence of cold events has a significant impact on critical infrastructure, agriculture and human health. In order to estimate the influence of such events in the future, this study analyses the occurrence of local ice episodes in future climate periods, taking global warming into account. The analysis is based on the results of RCM simulations under the assumption of the RCP8.5 emission scenario. The analysis focuses on the region of Hamburg in northern Germany to determine the frequency of local ice episodes in the future. For the purposes of this analysis, the days of the cold season were divided into three classes: warm days, frost days and ice days. The frequency of each of these classes was then analysed. In addition to the recent past climate period, the frequency was analysed for the near future and the far future. The ice days were divided into three further subclasses, covering different durations of ice episodes. Furthermore, atmospheric patterns were identified in ERA5 that prevail during the local ice episodes.

A pattern-matching method was employed to analyse the RCM data in order to identify the extent to which the ice episode patterns identified in ERA5 will change in the future and to ascertain whether they can still be associated with local ice episodes.

Based on our analysis, the posed research questions can be answered as follows:

Will local ice episodes still play a role at the end of the century?

All RCM simulations analysed here show the occurrence of at least one local ice episode within 2065–2095, assuming the RCP8.5 emission scenario. The ensemble average indicates that eight ice episodes occur in the target region within 2065–2095.

Will atmospheric circulation patterns associated with local ice episodes change towards the end of the century?

The use of the developed pattern-matching method has revealed that the representative patterns of the ice episodes identified in the ERA5 data will continue to occur during ice episodes with small variations in frequency in future climate periods compared to the recent past, as simulated by the RCMs.

Furthermore, it was found that the representative patterns of the identified ice episodes exhibit distinct features and, thus, lack a definitive "typical pattern", such as a blocking pattern or the NAO^- pattern. Furthermore, the identified ice episode patterns in ERA5 can be associated with two-thirds of all ice episode days in the RCMs, a proportion that remains consistent over the recent past, near future, and the far future, suggesting a consistent association between identified patterns and ice episode occurrences.

5 Overall assessment and concluding remarks

In the following sections, the posed research questions are answered. Based on the developed methods and using the obtained results, general conclusions are drawn. In addition, the limitations of this thesis are identified and areas for further research are highlighted.

5.1 Answers to research questions

In summary, the research questions identified can be answered as follows:

RQ1. How does data uncertainty affect cold season threshold-based climate indices?

The influence of data uncertainty on several threshold-based climate indices was investigated for the cold season. These include the number of frost days, the number of ice days, the maximum number of consecutive frost days and the maximum number of consecutive ice days, the number of wet days and the number of heavy precipitation days. The influence of data uncertainty was also investigated for the categorisation of cold seasons into mild, moderately warm, moderately cold and harsh cold seasons based on the coldsum.

To assess the influence of data uncertainty on cold season threshold-based climate indices, the threshold value of each index was first varied by an arbitrary range. This showed that the response, i.e. the percentage change in the number of days, depends on the distribution of the underlying climate variable. A threshold change towards the mode of the distribution leads to a larger percentage change in the number of days than a threshold change towards the tails of the distribution. This is particularly evident for indices describing more extreme situations (e.g. ice days or heavy precipitation days). For example, the percentage change in the number of ice days resulting from varying the threshold of 0 °C by an arbitrary range, depends on the distribution of T_{max}. Taking Hamburg as an example (Figure 1.1), where the most frequent values in the T_{max} dis-

tribution are around 7 °C, increasing the threshold causes a more pronounced increase (percentage change) in the number of ice days, compared to frost days. This is because the most frequent values in the T_{min} distribution are around 0°C, which corresponds to the defined threshold value. Wet days and heavy precipitation days show a much higher sensitivity to threshold changes than the temperature-based indices considered, due to the skewed precipitation distribution in Hamburg, characterised by frequent light precipitation events and rare heavy precipitation events. Indices that integrate quantities over time intervals (e.g. coldsum or consecutive days) show less sensitivity to threshold changes. Categorisation based on the coldsum appears to be robust to threshold changes, making it a useful tool for gaining insight into possible future changes.

A threshold uncertainty estimate was then derived from the E-OBS reference data set to avoid arbitrary threshold changes. This uncertainty estimate was determined separately for each threshold of the considered climate indices and used in calculations with both E-OBS and RCM data, allowing the RCMs to vary by the observational uncertainty. Using this uncertainty range and the same threshold as for the recent past, the results of future projections were analysed. For the example of Hamburg, it was shown that data uncertainty has only a small impact on projected future changes when analysing future changes relative to the recent past, especially for the RCP2.6 scenario. For the RCP8.5 scenario, a dependence on threshold changes was observed, especially in the far future. This dependence is likely due to changes in the distribution of the underlying climate variable.

Overall, when using threshold-based climate indices to assess future change, and especially for impact studies, it is essential to take account of uncertainties and, ideally, to correct for them appropriately. This is particularly important for indices that assess extremes. This highlights the importance of understanding and accounting for data uncertainties in climate change assessments.

RQ2. Can regional climate models reproduce local cold season characteristics, ice episodes and prevailing atmospheric patterns?

To obtain cold season characteristics for Hamburg, all cold season days were classified as either ice days, frost days or days that are neither ice days nor frost days, which are summarised as warm days. To assess their reproducibility by the RCMs, the frequency of the cold season characteristic classes was calculated for a recent past climate period (1979–2005) and compared with reference data sets (E-OBS and ERA5). The RCMs show a tendency to underestimate warm days and overestimate ice days, indicating a cold bias. No consistent tendency to over- or underestimate frost days was found, suggesting that the cold bias mainly affects daily maximum temperatures, indicating a daytime cold

bias.

In addition to characterising the cold season, it is important to consider cold weather events of longer duration, as they are potentially more damaging than temperature-related single-day events. As there is no uniform definition of cold weather events, the term ice episode is introduced and defined as more than five consecutive ice days. The frequency of ice episodes is calculated. The results of the RCMs show a tendency towards slightly more ice episodes compared to the reference data sets E-OBS and ERA5 for the period 1979–2005, which is in line with the general overestimation of ice days.

To gain a better understanding of the atmospheric conditions under which an ice episode occurs in Hamburg, the study identified prevailing atmospheric circulation patterns based on ERA5 data. It was shown that longer ice episodes are generally associated with blocking patterns over Iceland/the British Isles, while shorter ice episodes are generally associated with lower geopotential heights (GPH_{500} troughs or cyclones) extending towards southern Europe. Nevertheless, no definitive "typical" ice episode pattern could be identified, highlighting the complexity of atmospheric dynamics governing ice episodes.

To enable comparison of the identified atmospheric patterns in the ERA5 data with the RCM data, a pattern-matching method was developed that uses the Structural Similarity Index (SSIM) as a matching criterion to measure reproducibility. The SSIM index was originally developed for image processing by WANG et al. (2004) and, to the best of our knowledge (as of May 2024), has only been used in two other peer-reviewed publications in the field of climate data analysis. In the other studies, however, the SSIM index was not used to find certain patterns of one dataset in another, as here. DOAN et al. (2021) used the SSIM index as a replacement for the more traditional Euclidean distance measure, to group data into different weather types. HOFFMANN et al. (2021) used the SSIM index to investigate the persistence of certain weather patterns. Using the developed pattern-matching method it was found in this thesis that the RCMs are able to reproduce the identified atmospheric circulation patterns and their frequencies well, regardless of whether they are forced with GCM or reanalysis data at their boundaries.

Overall, the results indicate that RCMs can reproduce local ice episodes in the Hamburg region. However, biases, such as the too low daily maximum temperatures, need to be addressed in order to conduct more in-depth studies of local impacts.

RQ3. How will the frequency of local ice episodes and associated atmospheric patterns change in future climate periods?

Despite the projected increase in global mean temperatures under the high emission scenario RCP8.5, local ice episodes are expected to continue into the future. The RCMs

project 6–45 ice episodes (21 in the ensemble mean) within the near future (2030–2060) and 1–23 ice episodes (eight in the ensemble mean) within the far future (2065–2095). This represents a reduction in the ensemble mean of 49 % in near future and of 80 % in the far future compared to the recent past (1971–2000).

The developed pattern-matching method shows that the representative patterns of local ice episodes identified in ERA5 data persist in the RCMs into future climate periods with reduced frequency, due to the reduced frequency in ice episodes. However, the frequency relative to all ice episode days in the respective climate period, shows minimal variation. The representative patterns remain associated with two-thirds of all ice episode days in the RCMs over the recent past, near future, and far future climate periods. This suggests a stable association between the identified patterns and the occurrence of ice episodes in Hamburg.

Despite the projected increase in global mean temperature, cold events are expected to continue to occur in Hamburg under the assumption of the RCP8.5 emission scenario. However, they will occur less frequently. The probability of experiencing an ice episode, defined here as more than five consecutive ice days, is projected to be about twice every three cold seasons (0.7 events/cold season) in the near future (2030–2060) and once every four cold seasons (0.26 events/cold season) in the far future (2065–2095) under the high emission scenario RCP8.5. As this high emissions scenario represents strong climate change, the resulting number of future ice events presented here can be seen as the upper limit of potential reductions. Consequently, it can be seen as the lower limit of ice episode occurrences in the future. During ice episodes between 1979 and 2005, the average temperature in Hamburg was -6.3°C. This average ice episode temperature is projected to increase to -5.2°C in the near future and to -4.6°C in the far future.

Overall, the results of this thesis provide valuable insights into the persistence and characteristics of future cold season characteristics and ice episodes in the Hamburg region, facilitating informed adaptation and mitigation strategies to address their impacts on critical infrastructure, agriculture, and human health.

5.2 Limitations

The data analysis and methods developed in this thesis were applied solely to the cold season and an urban area. However, they can be applied to any point on Earth. Two points should be considered: (1) More than one grid cell should be chosen as the target region. This is because a more robust result can be obtained by averaging values from several grid cells of a numerical model instead of considering only a single value. (2)

The climate indices used are only relevant in certain regions. In regions situated near the equator, for example, cold does not pose a threat because the influence of cold polar air does not extend as far and temperatures do not drop as much. Thus other climate indices should be considered. However, in regions situated in the northern hemisphere mid-latitudes, such as Hamburg, a study at the freezing point is crucial, as the results can influence, for example, the planning of the depth of pipe installations and the choice of construction materials for buildings. Furthermore, the impact of successive temperature changes above and below the freezing point on vegetation is also relevant. Consequently, research into agricultural cultivation areas should be a priority alongside that of urban regions.

Another aspect to consider is the choice of climate indices. In this thesis, the focus was on temperature-based indices. As demonstrated in chapter 2, under the assumption of RCP8.5 in the far future (2060–2090), an increase in heavy precipitation days is projected for the cold season by the RCMs. Therefore, in addition to temperature-based indices, it would be important to consider precipitation in particular for detailed studies within the city. However, the simulation of precipitation is a challenge for the RCMs. In particular, heavy precipitation events occur at scales below the resolution of the RCMs. Since the output of the RCM data provides daily sums of precipitation, it is not possible to derive the actual impact from a heavy precipitation day from the data. For example, if this precipitation occurs within 30 minutes, it is a much more devastating event for the city than if the precipitation is spread over 24 hours. It would also be important to consider the potential threat of winter storms to Hamburg. However, due to the resolution of the RCMs, they cannot accurately simulate surface winds, which are strongly influenced by the orography. Consequently, the more detailed analysis (see chapter 3) was limited to temperature-based indices.

Another issue related to climate indices is the choice of the underlying climate variable. If T_{max} is used to calculate an ice day, observations and (climate) models may produce different results for this climate variable. Models have short computational timesteps, but provide their results at relatively long output intervals (e.g. 6 hourly or daily) in order to avoid exceeding storage capacities. According to the metadata in the EURO-CORDEX model data, the daily T_{max} is the most extreme value of all computational timesteps within the output interval (e.g. one day). In the case of observational data, the instantaneous value has traditionally been obtained every hour after changing from mercury to electical thermometers. It is unclear whether this instantaneous value represents the mean value for the hour or a short-term value; extremes within the hour may or may not be taken into account. Daily extreme values are derived from these hourly temperature values, which may result in the de facto minimum and maximum values being missed. Consequently,

the challenging questions arises to "What should be compared with what?", and "To what extent should climate model results align with observations or reanalyses, given that the methodologies for determining such extremes are not identical?". In the course of this thesis, a better agreement of ice episodes, in terms of start date and duration, was found between ERA-Interim (based on a model) and E-OBS (based on station data), by calculating the daily T_{max} as the maximum of the 6 hourly T_{avg} for ERA-Interim and comparing it with the given daily T_{max} from E-OBS.

As has been shown, the large-scale atmospheric circulation generated by the RCMs during the cold season is strongly forcing-dependent. RCMs can therefore only be as reliable as their boundary values, which are either provided by reanalysis or GCM data. As demonstrated by BRANDS (2022), there are considerable differences in the performance of the large-scale circulation in GCMs. In general, the CMIP6 models perform better on average than the CMIP5 models, which can be attributed to the enhanced resolution. Consequently, it would be beneficial to repeat the analysis presented in this thesis for the CMIP6 simulations in order to reach a more robust conclusion. However, the downscaling of CMIP6 simulations using RCMs was not fully completed at the end of this thesis, and since GCMs are generally only of limited use for regional or even local studies, the results of the CMIP6 GCMs were not considered for the analysis of local ice episodes.

5.3 Future directions – recommendations

In addition to further research into the relationship between Arctic warming and winter weather in mid-latitudes of the Northern hemisphere, it is necessary to focus on cities. A considerable proportion of the global population lives in urban areas, thus certain weather events influence many people if hitting a city. It is therefore crucial (1) to improve the representation of cities in GCMs and RCMs, as cities develop their own urban climate, which can influence the regional and the global climate, and (2) to conduct high resolution city simulations to estimate the impact of certain weather events on particularly vulnerable parts of an urban area. The characteristics of cities in the context of climate change are subject of the planned IPCC Special Report on Climate Change and Cities, for which a first scoping meeting was held in April 2024 (IPCC, 2024).

In the following sections the representation of cities in climate models is examined by presenting examples from current research. In addition, the possibility of high-resolution city simulations is discussed and areas of further research needs in this context are identified.

5.3.1   Representation of cities in climate models

As stated in the previous section, a repetition of the analysis presented here with EURO-CORDEX simulations based on CMIP6 would be beneficial in order to facilitate more comprehensive statements. Furthermore, the use of results from the new generation of convection-permitting RCMs with a grid spacing of 2–5 km (in contrast to the 12.5 km RCM data used here), would enable the inclusion of precipitation in this analysis, which can also fall as snow during ice episodes. Moreover, the representation of the city and its effects on the regional climate are more precise at a 2–5 km resolution, if appropriate parameterisations have been included in the models.

With regard to the representation of cities, a pilot study is currently being conducted to investigate the influence of urban regions on regional climate and the influence of regional climate on cities in CORDEX models (FPS URB-RCC URB-RCC, officially endorsed by WCRP). In order to achieve this, coordinated RCM simulations of two extreme situations in the city of Paris are currently being carried out by various research institutes. The objective is to quantify the influence of the city on the regional and local climate and to investigate how different this influence is in the various models. Differences arise not least due to the fact that these models use different parameterisation to account for city-related effects. The results will then be used to derive how complex such a city parameterisation in RCMs must be in order to map the city effect on the local and regional climate. Subsequently, further coordinated simulations will be conducted with the improved parameterisations for a selection of cities under future scenarios. This is important with regard to the development of climate services for health aspects, risk management, urban planning and others.

Another area of research concerns the further development of GCMs towards finer resolutions, by resolving more complex processes. As an example, the Earth system model ICON (JUNGCLAUS et al., 2022) is briefly discussed. Among other new features, this model uses an icosahedral grid, i.e. triangular grid cells, in contrast to square grid cells typically used in previous climate models. Further developments, such as ICON-Sapphire (HOHENEGGER et al., 2023), demonstrate the potential of conducting a globally coupled 5 km simulation for an entire year with today's computing capacities. Additionally, a model configuration can be selected in which an atmosphere-only simulation is conducted for a smaller area, analogous to the refinement of GCM simulations with RCMs. Further grid refinements are then applied towards the centre of the area, using a uniform grid with inside nests. These inner nests then reach grid spacings of 620 m and 308 m. The authors demonstrate that ICON-Sapphire can be employed for regional climate and process studies. However, it is unclear from referenced publications whether

and, if so, how cities and their effects are incorporated.

The aforementioned efforts to resolve more details in GCMs and RCMs are promising. However, the computing effort increases with increasing resolution. Consequently, even ICON-Sapphire in global configuration can only be run for a simulation of one year or, in the case of large-eddy simulations, only for a small area and for a comparatively short period of time. For simulations to be carried out for a future climate, all these models require boundary values from the long-term projections of GCMs, which were developed for exactly this specific purpose. In this respect, it is essential that the processes are better understood and better modelled in the GCMs without increasing the computational effort too much.

5.3.2 Simulations of extreme events within the city

It is not possible to investigate the impact of an ice episode on the heterogeneity within Hamburg using regional climate model data, as these models are not sufficiently detailed with a grid sizes of 12.5 km, or 2–5 km in the case of convection-permitting RCMs. Furthermore, the city itself is, if at all, represented in the model by a very simple parameterisation with surface properties equal to those of a rock. Models that can resolve detailed processes within an urban area are obstacle resolving microscale models (ORMs). These are computationally intensive due to their high resolution, which currently limits their applicability to small domain sizes (several 10 km^2) and short integration periods. To simulate realistic impacts of an ice episode, these models require boundary values for the small model domain. However, there is a large gap in resolution (scale jump) between the RCM data (12.5 km) and the ORM ($\sim$500 m at the boundaries for non-equidistant grids BÖTTCHER (2023)). This scale jump leads to two key issues: (1) The phase velocities of the phenomena on the grids of varying size are different, which leads to reflections of the waves at the boundaries and thus introduces numerical disturbances (SCHLÜNZEN et al., 2011). (2) The simulation of a large area with an ORM covering more than one RCM grid cell (i.e. more than 12.5 km^2) is computationally intensive and, thus, too expensive. The scale jump, namely the difference in resolution between the ORM and the boundary values, should not exceed a factor of 4–6 (SCHROEDER and SCHLÜNZEN, 2009; BÖTTCHER, 2023). However, in order to resolve small-scale details within the urban area, the resolution required to resolve processes is of the order of 10 m. Thus a dynamic downscaling is not possible, even with boundary values of 1 km resolution. An ORM simulation should not use a grid spacing of less than 200 m (factor 5). Consequently, a method must be developed that overcomes the scale jump without introducing numerical disturbances and at the same time does not exceed the computing capacities of today's supercomputers.

The nesting represents a major challenge for the ORM community. There is a limited availability of observational or reliable model data at such high resolutions that could possibly be used as boundary values. One potential solution of obtaining such data would be the use of mesoscale models, which, however, also require time dependent boundary values. Furthermore, conducting such a mesoscale model simulation, to obtain a high-resolution data set to be used by an ORM, would also require considerable computing power. Consequently, current research is seeking to identify methods to circumvent this intermediate step at the mesoscale level without causing numerical disturbances, while ensuring that the coarser data used as boundary values are sufficiently detailed.

A mesoscale interface (INIFOR) was developed for the ORM PALM (MARONGA et al., 2020), which derives realistic initial and boundary values from the operational weather model COSMO-DE/D2 (2.8 km grid spacing, BALDAUF et al. (2011)). This allows simulations to be calculated for realistic heterogeneous areas under changing synoptic conditions. Nevertheless, precipitation events cannot be simulated with this ORM, as this requires the inclusion of the relevant processes within the model, which necessitates further model development. By incorporating precipitation relevant processes into the ORM MITRAS (SALIM et al., 2018), FERNER et al. (2023) were able to demonstrate first results of rainfall heterogeneity within an urban neighbourhood. This kind of model development is essential for the estimation of the risks, especially those associated with compound events, which are defined as events comprising several extremes at the same time (e.g. winter storms and extreme precipitation).

A more promising and resource-efficient approach for developing a scale interface that generates data sets of sufficient detail is that of machine learning. For example, SINGH et al. (2023) use a convolutional neural network (CNN), which learns patterns from gridded data sets, to create a high-resolution precipitation climatology from remote sensing precipitation data. The authors have applied an iterative super-resolution convolutional neural network (SRCNN) and downscaled the ~10 km satellite precipitation data to a grid spacing of ~300 m. However, this resolution is still coarser than in-urban resolutions. They argue that their method may not be able to identify hyper-local convective rain events, but that it is likely to be most effective for larger-scale frontal rain events to improve the spatial information at the local scale. Nevertheless, validating such high-resolution results is challenging due the lack of observational reference data sets. In a recently published review on deep learning in statistical downscaling for deriving high spatial resolution gridded meteorological data (SUN et al., 2024), it was highlighted that deep learning has outperformed traditional methods in various areas. Deep learning has the potential to establish a complex mapping between large-scale and local-scale meteorological data, and this could be of significant benefit to downscaling, particularly with

super-resolution networks.

For planned RCM simulations, it would be desirable to consider the possibility of storing results for e.g. heavy precipitation events, at a higher frequency. This would at least improve the temporal resolution for mesoscale or ORM applications. An internal automated criterion could be used (based on the idea proposed by BUNGERT (2008)) to determine whether a change is large or small and whether the output frequency should be increased.

In conclusion, it can be stated that there is no universal method for simulating extreme events that is optimal for all use cases. The choice of method depends on the specific use case (Which variable? Which area? Which time period?) and the resources available. In most cases, a combination of different techniques is a good approach.

5.4   Closing words

This work is the first to define and analyse local ice episodes, to determine the prevailing atmospheric circulation patterns for them and to investigate how both might change in the future under the assumption of the high emission scenario RCP8.5. Based on the GCM-driven RCM projections, it was shown that ice episodes will become less frequent, but may still occur in Hamburg in the far future despite the strong global warming. Furthermore, it was shown that ice episodes are generally associated with a variety of different large-scale atmospheric patterns and cannot be associated with a single typical pattern. The results of this work demonstrate the need to consider the possibility of ice episodes occurring towards the end of the century in future planning (urban development, infrastructure, etc.), despite existing uncertainties regarding the climate scenario that humanity will follow in the future. At the same time, it is essential to improve models or develop methods that can accurately simulate the impact of small-scale urban effects on the global climate, and provide a more comprehensive understanding of the relationship between Arctic warming or sea-ice retreat and winter weather in the mid-latitudes of the Northern hemisphere.

Acknowledgments

Die Promotion war ein langer Weg, der rückblickend betrachtet schnell zu Ende ging. Der Weg war nicht immer einfach, aber es gab auch Phasen der Leichtigkeit. Die Reise brachte Geselligkeit durch die vielen neuen Kontakte, aber auch Einsamkeit durch die Pandemie. Während der Reise wurde neues Leben geboren und von anderem Leben musste für immer Abschied genommen werden. Die Reise war ein Auf und Ab der Gefühle, eine Reise der Gegensätze. Sie hat mich wachsen lassen und mich zu dem Menschen gemacht, der ich heute bin. Deshalb möchte ich mich bei allen bedanken, die mich auf diesem Weg begleitet haben!

Als allererstes möchte ich mich bei Prof. Dr. K. Heinke Schlünzen bedanken. Danke, dass Du dich bereiterklärt hast, die Betreuung auf Uni-Seite für dieses HICSS Projekt zu übernehmen. Danke, dass Du uns auch nach Deinem wohl verdientem Ruhestand weiterhin betreut hast. Ich weiß, dass ich mich glücklich schätzen kann, Dich als Doktormutter zu haben. Du hast Dir immer viel Zeit genommen, Sachen kritisch hinterfragt und damit mein Bewusstsein für viele Dinge geschärft. In zahlreichen Gesprächen, seit der Pandemie ausschließlich online aber nicht weniger persönlich, habe ich Deine Ratschläge, ob fachlich oder auch privat, immer zu schätzen gewusst.

Als Zweitbetreuer möchte ich Dr. Kevin Sieck danken. Danke, dass Du zusammen mit Heinke dieses Projekt ins Leben gerufen hast und Dich immer für uns am GERICS eingesetzt hast. Danke für die regelmäßigen Treffen und Dein andauerndes Verständnis, wenn bei mir mal wieder nicht alles ganz so rund lief. Danke auch für die Gespräche abseits der Fachthemen, die mir persönlich weitergeholfen haben und meinen Blick für die Zukunft geschärft haben.

Ich möchte mich auch bei Prof. Dr. Jörn Behrens bedanken. Jörn, vielen Dank, dass Du den Vorsitz meines Betreuungspanels übernommen hast. Danke für Deine ehrlichen Worte, persönlichen Ratschläge und Deine konstruktive Kritik in den zahlreichen Panel-Meetings. Ich bin froh, dass Du meinen Doktoranden-Weg begleitet hast!

Danke an alle Kolleginnen und Kollegen vom GERICS. Ein besonderes Dankeschön

A   First paper

This section was published together with chapter 2 as:

Bell, L. M. and K. H. Schlünzen and K. Sieck (2023): *Influence of data uncertainty on cold season threshold-based climate indices*. Meteorologische Zeitschrift, http://dx.doi.org/10.1127/metz/2023/1158.

A.1   Supporting information
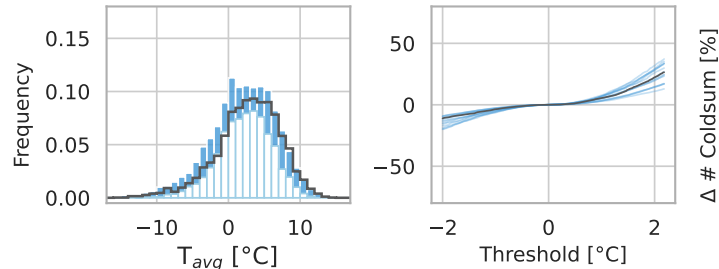
Figure A.1. Distribution of daily average temperatures between 1971–2000 (left) and percentage changes of the coldsum as function of the chosen threshold (right). Blue: 5th to 95th percentile of the climate model ensemble (a, both panels), all 14 model results (b,c, both panels); black: reference data.
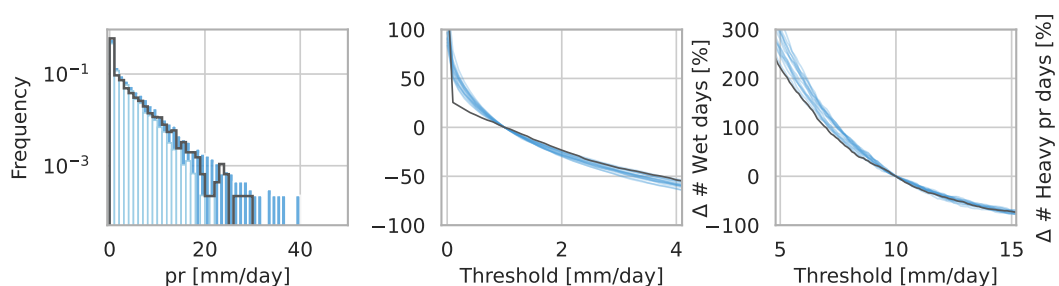

Figure A.2. Distribution of daily precipitation (left) and percentage changes of the number of wet days (middle) and the number of heavy precipitation days (right) as function of the chosen threshold. Blue: 5th to 95th percentile of the climate model ensemble (a, both panels), all 14 model results (b,c, both panels); black: reference data.

**(a)**   **(b)**
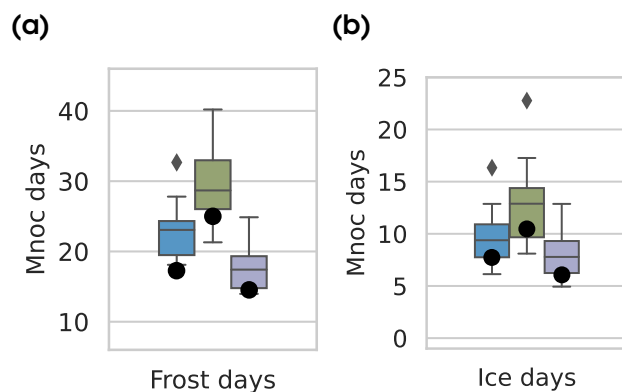
Figure A.3. Maximum number of consecutive (a) frost days and (b) ice days as mean cold season value between 1971-2000. From left to right: climate model ensemble bandwidth for the originally defined threshold, model ensemble bandwidth for increased threshold, model ensemble bandwidth for decreased threshold. Black dots represent reference data values for the corresponding thresholds.
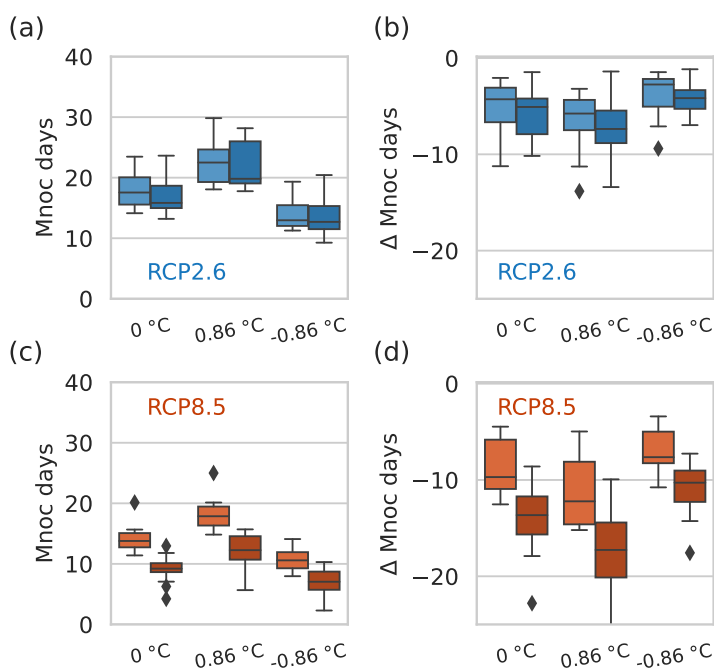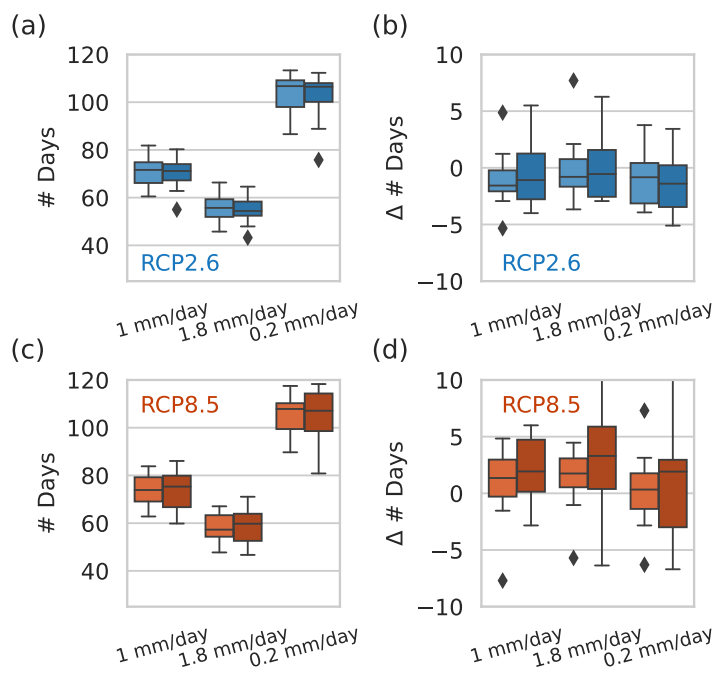

Figure A.4. Maximum number of consecutive frost days for three different thresholds (x-axis) as mean cold season value in near future (2031–2060, light shading) and far future (2071–2099, dark shading) based on the climate model ensemble. (a,c): Total number of days, (b,d): Change in the number of days relative to reference period 1971–2000. Numbers are shown for two climate scenarios, (a,b) RCP2.6 in blue and (c,d) RCP8.5 in red.
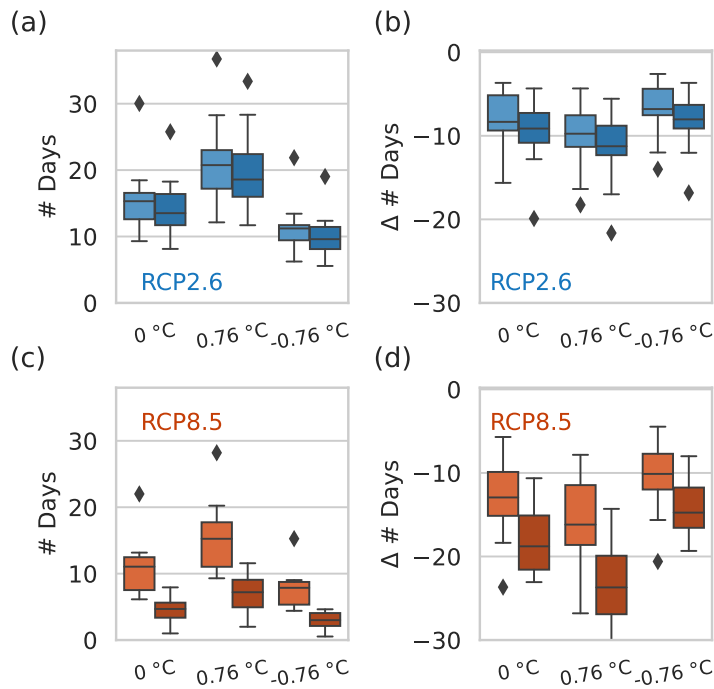
Figure A.5. Number of ice days for three different thresholds (x-axis) as mean cold season value in near future (2031–2060, light shading) and far future (2071–2099, dark shading) based on the climate model ensemble. (a,c): Total number of days, (b,d): Change in the number of days relative to reference period 1971–2000. Numbers are shown for two climate scenarios, (a,b) RCP2.6 in blue and (c,d) RCP8.5 in red.
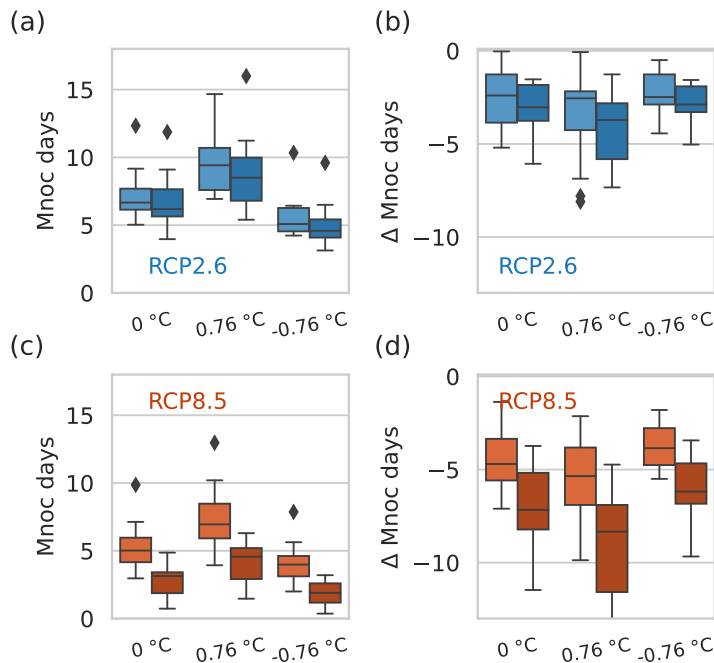
Figure A.6. Maximum number of consecutive (mnoc) ice days for three different thresholds (x-axis) as mean cold season value in near future (2031–2060, light shading) and far future (2071–2099, dark shading) based on the climate model ensemble. (a,c): Total number of days, (b,d): Change in the number of days relative to reference period 1971–2000. Numbers are shown for two climate scenarios, (a,b) RCP2.6 in blue and (c,d) RCP8.5 in red.

Figure A.7. Number of wet days for three different thresholds (x-axis) as mean cold season value in near future (2031–2060, light shading) and far future (2071–2099, dark shading) based on the climate model ensemble. (a,c): Total number of days, (b,d): Change in the number of days relative to reference period 1971–2000. Numbers are shown for two climate scenarios, (a,b) RCP2.6 in blue and (c,d) RCP8.5 in red.

A.2 Acknowledgements

B   Second paper

This section was submitted together with chapter 3 as:

Bell, L. M. and K. H. Schlünzen and K. Sieck: *Reproducibility of local cold season characteristics, ice episodes and prevailing circulation patterns in regional climate models*. Quarterly Journal of the Royal Meteorological Society, *under review*.

B.1   Supporting information

Comparison between ERA5 and ERA-Interim

A comparison between ERA5 and its successor ERA-Interim, which was used as forcing for the EURO-CORDEX evaluation runs, is presented. The identified ice episodes are listed in Table B.1 for ERA-Interim, ERA5 and, for comparison with a gridded observational dataset, also for E-OBS. The ice episodes are given for the period 1979 to 2005, for ERA-Interim only for 1989 to 2005. This is due to technical problems associated with the provision of ERA-Interim data. As a result, ERA-Interim data were only available for this short period.

It is obvious that there is a large agreement between ERA5 and E-OBS, whereas ERA-Interim fails to capture the same ice episodes. This might be due to two different reasons: (1) As ERA-Interim has the lowest resolution, values are averaged over a larger area that represent an extreme value (either maximum or minimum). Higher resolutions capture more detail, hence, ice episodes might be better captured. (2) In ERA-Interim, daily maximum Temperature values are selected from the 6 hourly maximum temperature values. For ERA5 it is recommended to calculated daily maximum (or minimum) temperature values from the provided on hourly instantaneous T2m values (https://confluence.ecmwf.int/display/CKB/ERA5%3A+2+metre+temperature). Therefore, ERA5s daily maximum values are similarly obtained as is done for station data on which E-OBS is based.

In addition, the proposed pattern-matching method is applied to search for ERA5 patterns in ERA-Interim. For this, the same methodology, as described in sec. 2.3 is used. The resulting matches can be found in Table B.2. In addition, matches for a carried out ERA5 self-test are provided. The results clearly show that there is no exact agreement between ERA5 and ERA-Interim in terms of the number of pattern matches. One reason for this may be that the daily maximum temperature in the target region differs in the two reanalyses, as it is calculated in different ways (see previous paragraph). Another reason could be due to the relatively coarse resolution of ERA-Interim, small-scale features are

Table B.1. Details of ice episodes (classes 3–Ds, 3–Dm, 3–Dl) for E-OBS, ERA-Interim and ERA5. The details given are the start date of each event and the duration (D). As ERA5 events are used as reference in the analysis of atmospheric patterns, each ice episode is given a number (№), according to its time of occurence within 1979–2005. Due to technical issues, ERA-Interim data was only available for 1989–2005.

Class	E-OBS start date	D	ERA-Interim start date	D	№:	ERA5 start date	D
3–Ds	1982-01-06	10	N.A	N.A	III:	1982-01-06	10
	-	-	N.A	N.A	IV:	1983-02-08	7
	1983-12-11	7	N.A	N.A	V:	1983-12-11	7
	1985-02-07	9	N.A	N.A	VII:	1985-02-07	9
	-	-	N.A	N.A	VIII:	1985-12-27	6
	1986-02-05	7	N.A	N.A		-	-
	-	-	N.A	N.A	XI:	1987-03-01	8
	1989-12-28	6	1989-12-28	6	XII:	1989-12-28	8
	1991-02-06	9	1991-02-06	9	XIII:	1991-02-06	10
	1993-11-21	7	-	-	XIV:	1993-11-21	7
	1994-02-19	6	-	-	XV:	1994-02-19	6
	-	-	1996-02-05	6		-	-
	1998-12-06	6	-	-	XIX:	1998-12-06	6
	-	-	2000-12-21	6	XX:	2000-12-21	7
	2001-01-15	6	-	-	XXI:	2001-01-15	7
	2002-12-09	8	2002-12-08	9	XXII:	2002-12-09	8
	2002-12-20	6	-	-	XXIII:	2002-12-19	6
	2003-01-04	8	2003-01-04	8	XXIV:	2003-01-04	8
	-	-	-	-	XXV:	2005-03-01	6
3–Dm	1980-01-08	14	N.A.	N.A.	I:	1980-01-08	14
	1981-12-14	11	N.A.	N.A.	II:	1981-12-14	11
	1987-01-10	12	N.A.	N.A.	X:	1987-01-10	12
	1995-12-25	13	1995-12-24	14	XVI:	1995-12-24	14
3–Dl	1985-01-02	20	N.A.	N.A.	VI:	1985-01-02	19
	1986-02-13	17	N.A.	N.A.	IX:	1986-02-02	29
	1996-01-21	21	1996-01-18	17	XVII:	1996-01-20	22
	1996-12-20	23	1996-12-20	23	XVIII:	1996-12-20	23

simply not present even after interpolating to the much finer EURO-CORDEX grid for the comparison.

Table B.2. Number of matches generated by applying the pattern-matching method to ERA-Interim to find event-specific ERA5 patterns (vs. ERA-Interim). In addition, results for the ERA5 self-test are provided (vs. ERA5). Considered are atmospheric circulation patterns associated with ice episodes of short (5 to 10 consec. ice days), medium (11 to 15 consec. ice days) and long duration (> 15 consec. ice days) only.

| **ERA5 event**		**vs. ERA5**	**vs. ERA-Interim**
startdate	duration	Σ	
1989-12-28	8	3	3
1991-02-06	10	1	1
1993-11-21	7	1	1
1994-02-19	6	2	2
1995-12-24	14	5	4
1996-01-20	22	7	6
1996-12-20	23	7	4
1998-12-06	6	2	1
2000-12-21	7	2	2
2001-01-15	7	4	5
2002-12-09	8	4	4
2002-12-19	6	2	2
2003-01-04	8	4	2
2005-03-01	6	4	2

ERA5 ice episode temperatures

Table B.3. ERA5 mean target region temperature values for the identified ice episodes. Temperatures provided are the event average daily average temperature ($\overline{T_{avg}}$), the event average daily maximum temperature ($\overline{T_{max}}$), the event maximum daily maximum temperature ($(T_{max})_{min}$) and the event minimum daily minimum temperature ($(T_{min})_{min}$).

ERA5 event		ERA5 temperature			
№	duration	$\overline{T_{avg}}$	$\overline{T_{max}}$	$(T_{max})_{max}$	$(T_{min})_{min}$
I	14	-6.2	-3.9	-0.5	-13.7
II	11	-8	-5.6	-0.9	-16
III	10	-9.6	-5.6	-2.4	-16.5
IV	7	-4.2	-1.1	-0.1	-14.5
V	7	-4.4	-2.2	-0.8	-9.2
VI	19	-8.3	-5.7	-0.3	-16.1
VII	9	-7.4	-4	-0.4	-14.3
VIII	6	-3.8	-2.4	-0.7	-8.5
IX	29	-5.9	-2.7	-0.1	-21.4
X	12	-10.7	-8.3	-1.4	-19.7
XI	8	-5.2	-2.1	-0.1	-12.3
XII	8	-1.7	-1.1	-0.2	-2.7
XIII	10	-5.4	-3.4	-0.1	-11.9
XIV	7	-4.1	-2.5	-0.9	-7.5
XV	6	-5.1	-2.2	-0.5	-13.5
XVI	14	-6	-4	-0.2	-12.5
XVII	22	-6.7	-4	-0.1	-14.8
XVIII	23	-6.9	-4.5	-0.8	-21
XIX	6	-4.5	-2.2	-0.9	-10.7
XX	7	-2.8	-1.2	-0.2	-8.7
XXI	7	-2.7	-1.3	-0.5	-6.2
XXII	8	-5.5	-3	-1.5	-10.6
XXIII	6	-2.3	-1.3	-0.3	-4.9
XXIV	8	-7.4	-3.9	-1	-15.6
XXV	6	-3.4	-0.9	-0.2	-11.1

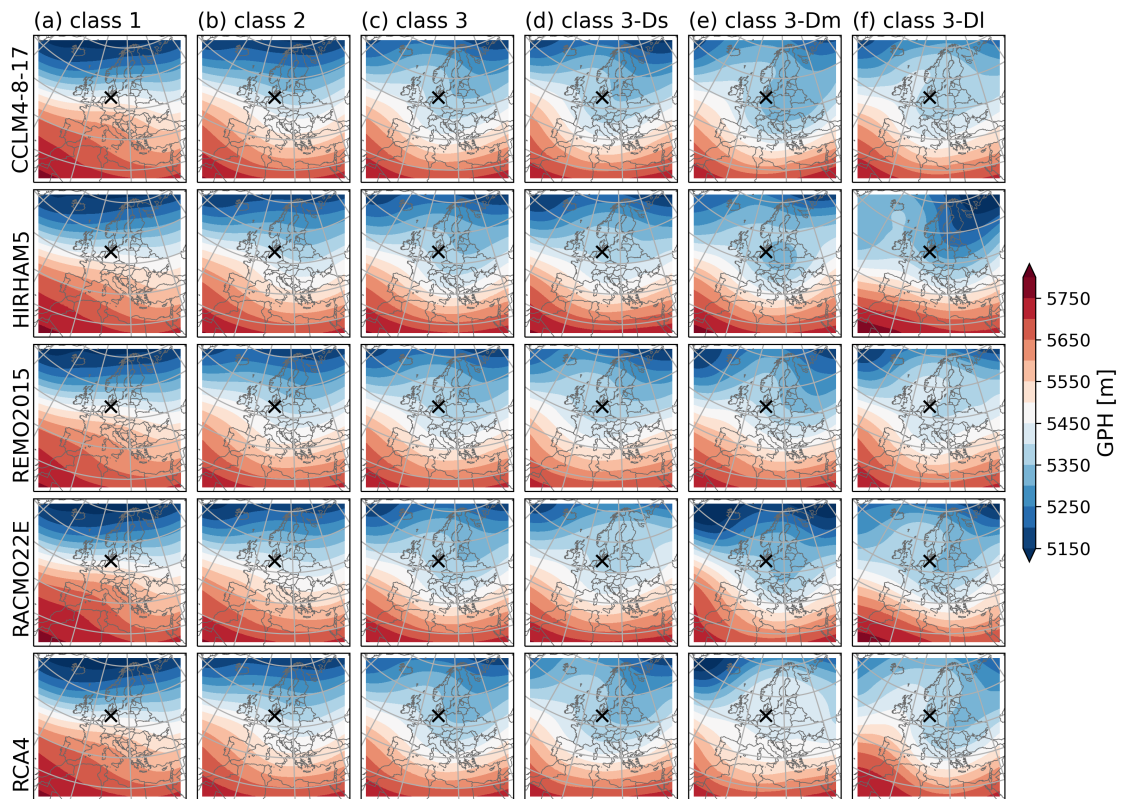## Mean class patterns for RCMs with GCM forcing

Figures displaying the average cold season characteristic class patterns for each RCM simulation are presented. Each figure represents a different GCM forcing.
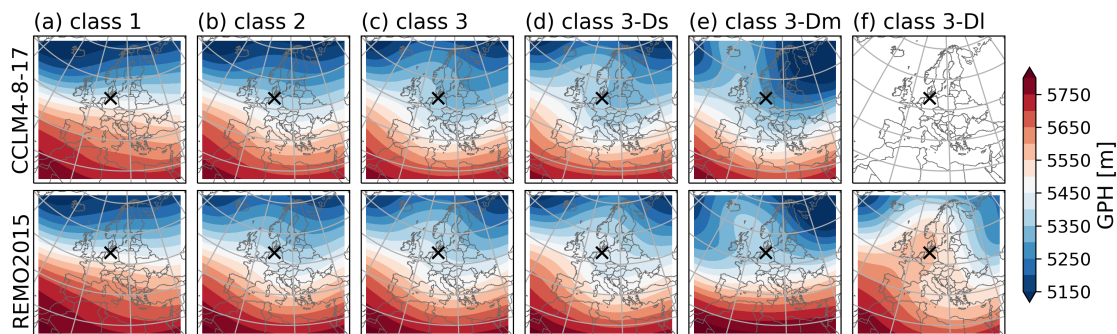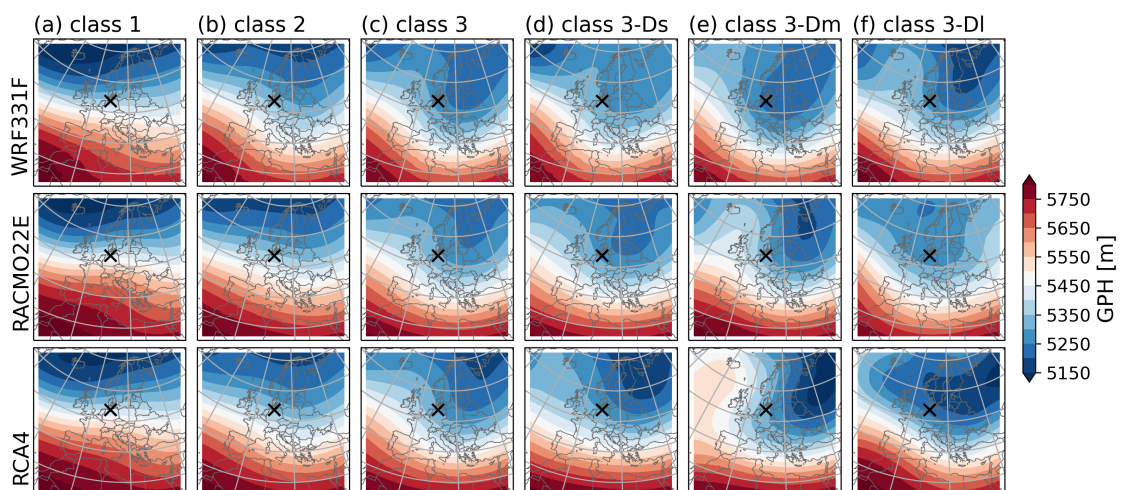

Figure B.1. Mean class patterns derived from the geopotential height at 500 hPa. Shown are climate model data resulting from RCM simulations with **CanESM2** forcing. The RCMs considered are (from top to bottom): CCLM4-8-17, ALADIN53, REMO2015.
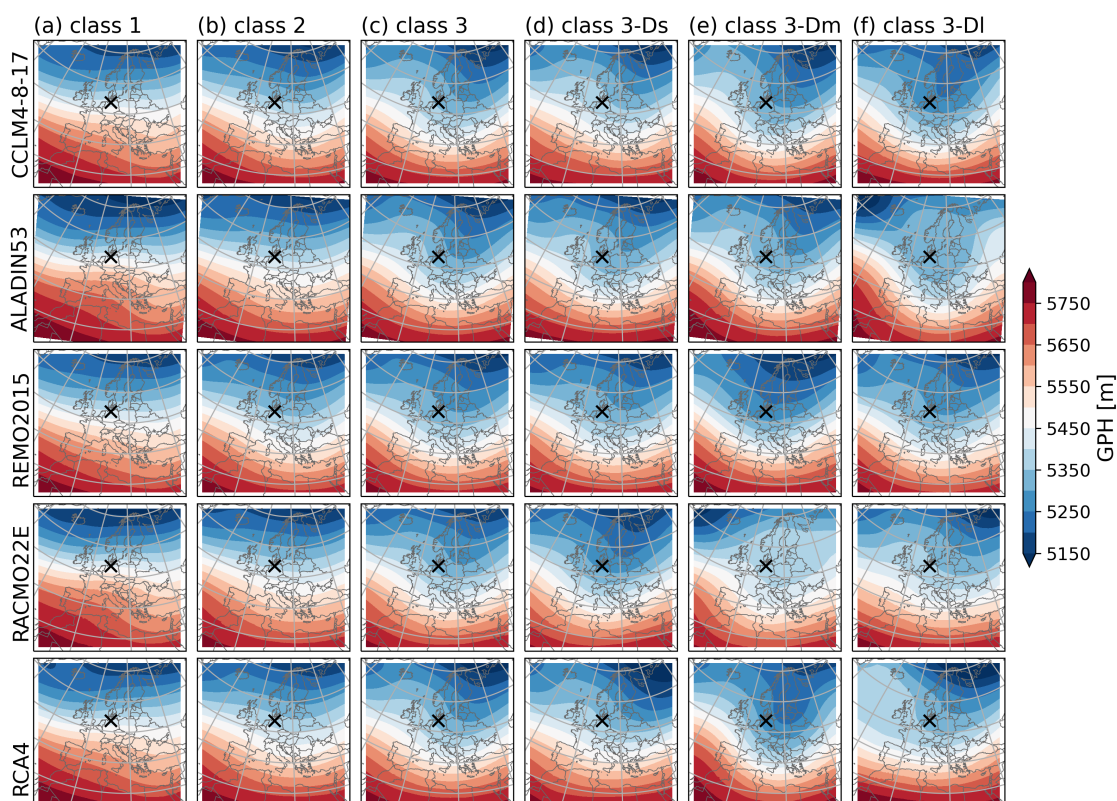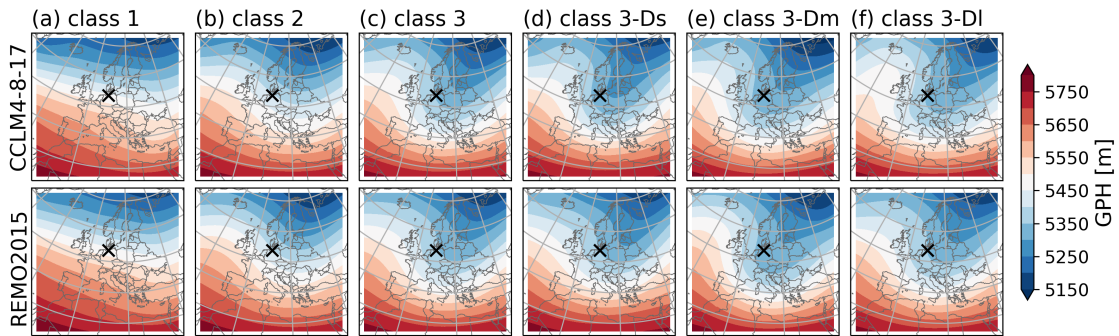

Figure B.2. Mean class patterns derived from the geopotential height at 500 hPa. Shown are climate model data resulting from RCM simulations with **CNRM-CM5** forcing. The RCMs considered are (from top to bottom): CCLM4-8-17, REMO2015, RACMO22E, RCA4.
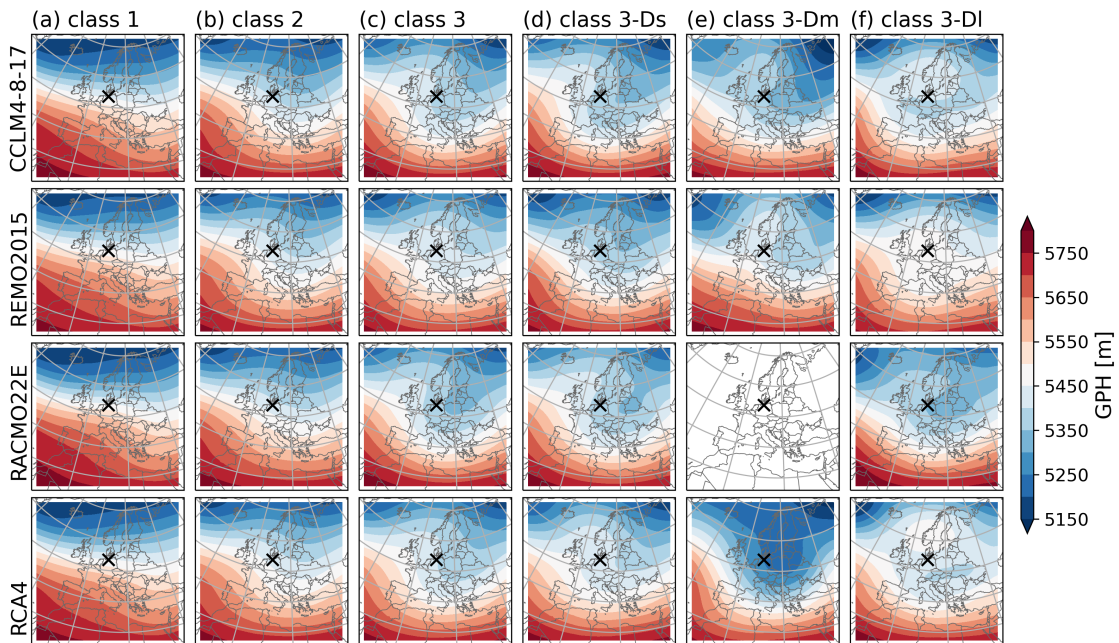
Figure B.3. Mean class patterns derived from the geopotential height at 500 hPa. Shown are climate model data resulting from RCM simulations with **EC-Earth** forcing. The RCMs considered are (from top to bottom): CCLM4-8-17, HIRHAM5, REMO2015, RACMO22E, RCA4.

Figure B.4. Mean class patterns derived from the geopotential height at 500 hPa. Shown are climate model data resulting from RCM simulations with **IPSL-CM5A-MR** forcing. The RCMs considered are (from top to bottom): WRF331F, RACMO22E, RCA4.
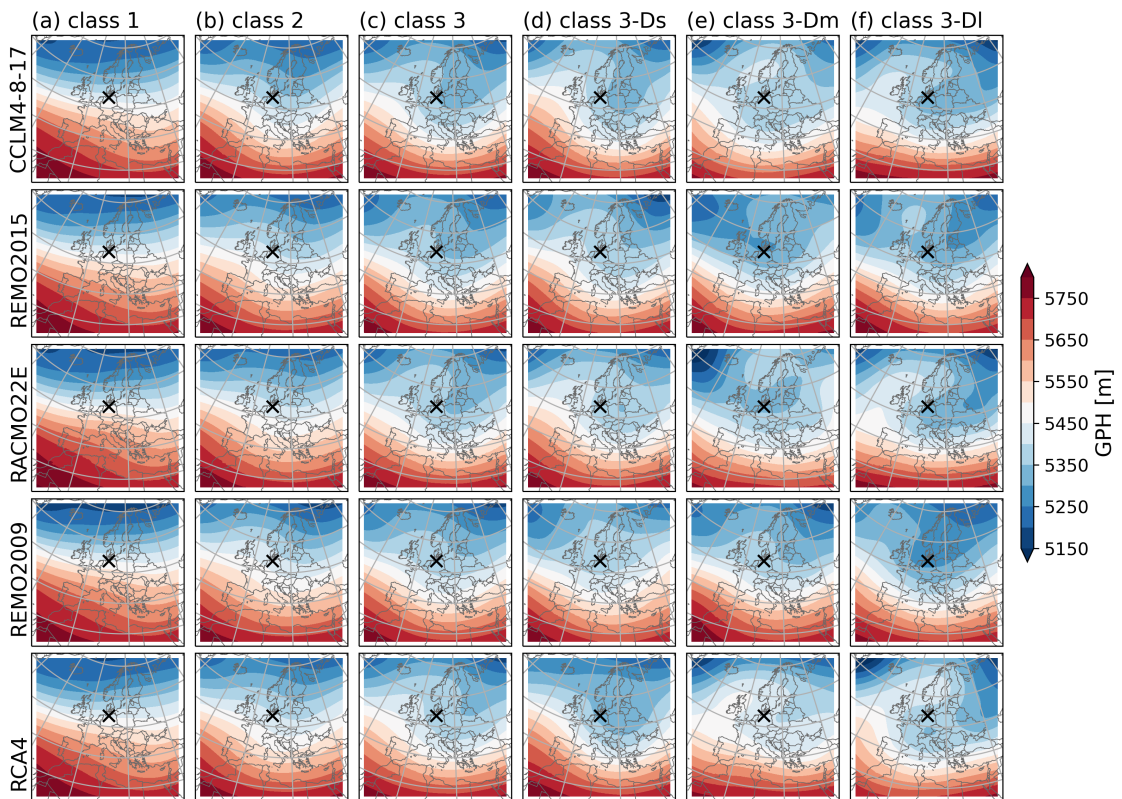
Figure B.5. Mean class patterns derived from the geopotential height at 500 hPa. Shown are climate model data resulting from RCM simulations with **MIROC5** forcing. The RCMs considered are (from top to bottom): CCLM4-8-17, REMO2015.
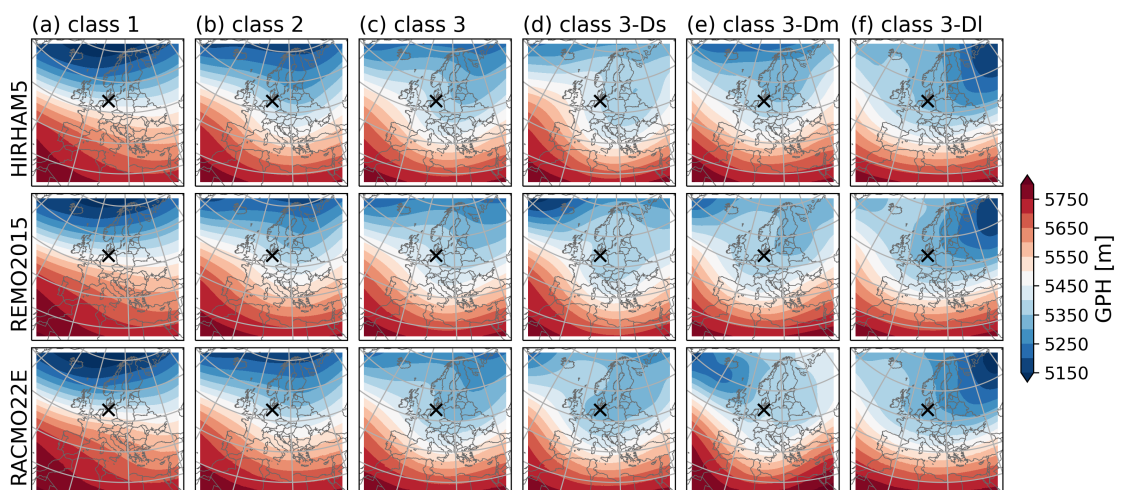

Figure B.6. Mean class patterns derived from the geopotential height at 500 hPa. Shown are climate model data resulting from RCM simulations with **HadGEM2-ES** forcing. The RCMs considered are (from top to bottom): CCLM4-8-17, REMO2015.

Figure B.7. Mean class patterns derived from the geopotential height at 500 hPa. Shown are climate model data resulting from RCM simulations with **MPI-ESM-LR** forcing. The RCMs considered are (from top to bottom): CCLM4-8-17, REMO2015, RACMO22E, REMO2009, RCA4.

Figure B.8. Mean class patterns derived from the geopotential height at 500 hPa. Shown are climate model data resulting from RCM simulations with **NorESM1-M** forcing. The RCMs considered are (from top to bottom): HIRHAM5, REMO2015, RACMO22E.

Pattern matches between ERA5 and GCM forced RCM simulations

The following Tables provide the resulting numbers from applying the pattern-matching for each RCM individually. Each Table represents a different forcing, which can be either reanalysis (ERA-Interim) or GCM data.

Table B.4. Number of event-specific reference (ERA5) pattern matches (Σ) in ERA-Interim-driven RCM simulations. Letters stand for the following RCM simulations: R: Reference (ERA5), A: CCLM4-8-17, B: HIRHAM5, C: REMO2015, D: WRF331F, E: RACMO22E, F: REMO2009, G: RCA4. In addition, an RCM mean (avg) number of matches is provided in the last column. Events considered are ice episode of short (class 3–Ds), medium (class 3–Dm) and long duration (class 3–Dl) only.

№	duration	R Σ	A	B	C	D	E	F	G	avg Σ̄
I	14	11	9	10	16	6	11	9	13	10.6
II	11	8	5	9	14	9	10	5	12	9.1
III	10	10	8	9	12	4	12	7	18	10.0
IV	7	2	1	0	1	0	1	0	1	0.6
V	7	5	2	3	9	2	7	4	6	4.7
VI	19	6	3	3	8	2	10	3	7	5.1
VII	9	6	2	3	8	3	6	3	8	4.7
VIII	6	8	2	5	8	1	6	3	7	4.6
IX	29	10	5	7	13	4	11	5	13	8.3
X	12	11	4	6	12	4	10	4	8	6.9
XI	8	2	1	0	1	1	1	0	1	0.7
XII	8	4	8	6	9	3	7	7	4	6.3
XIII	10	5	1	1	2	1	2	1	4	1.7
XIV	7	2	1	1	0	1	2	1	1	1.0
XV	6	3	1	1	2	1	2	1	2	1.4
XVI	14	12	4	8	16	3	15	8	9	9.0
XVII	22	12	5	12	15	7	12	9	12	10.3
XVIII	23	11	3	4	8	3	10	4	11	6.1
XIX	6	3	2	2	3	1	3	2	3	2.3
XX	7	7	5	5	4	1	4	3	9	4.4
XXI	7	9	8	7	10	3	7	6	7	6.9
XXII	8	6	4	7	10	5	8	6	8	6.9
XXIII	6	3	3	3	5	1	8	3	4	3.9
XXIV	8	8	3	5	9	2	9	5	6	5.6
XXV	6	7	5	6	9	4	9	5	8	6.6

Table B.5. Number of event-specific reference (ERA5) pattern matches (Σ) in RCM simulations forced with **CanESM**. Letters stand for the following RCM simulations: R: Reference (ERA5), A: CCLM4-8-17, B: HIRHAM5, C: REMO2015, D: WRF331F, E: RACMO22E, F: REMO2009, G: RCA4. In addition, an RCM mean (avg) number of matches is provided in the last column. Events considered are ice episodes of short (class 3–Ds), medium (class 3–Dm) and long duration (class 3–Dl) only.

№	duration	R Σ	A	C	avg $\overline{\Sigma}$
I	14	11	7	3	5
II	11	8	15	6	10.5
III	10	10	2	2	2
IV	7	2	0	0	0
V	7	5	9	5	7
VI	19	6	2	0	2
VII	9	6	4	3	3.5
VIII	6	8	13	0	6.5
IX	29	10	1	1	1
X	12	11	5	3	4
XI	8	2	0	0	0
XII	8	4	4	7	5.5
XIII	10	5	4	1	2.5
XIV	7	2	0	0	0
XV	6	3	1	1	1
XVI	14	12	8	5	6.5
XVII	22	12	5	4	4.5
XVIII	23	11	0	0	0
XIX	6	3	0	0	0
XX	7	7	2	2	2
XXI	7	9	4	4	4
XXII	8	6	0	2	1
XXIII	6	3	0	0	0
XXIV	8	8	2	3	2.5
XXV	6	7	3	1	2

Table B.6. Number of event-specific reference (ERA5) pattern matches (Σ) in RCM simulations forced with **CNRM-CM5**. Letters stand for the following RCM simulations: R: Reference (ERA5), A: CCLM4-8-17, B: HIRHAM5, C: REMO2015, D: WRF331F, E: RACMO22E, F: REMO2009, G: RCA4. In addition, an RCM mean (avg) number of matches is provided in the last column. Events considered are ice episodes of short (class 3–Ds), medium (class 3–Dm) and long duration (class 3–Dl) only.

ERA5 event pattern		R	A	C	E	G	avg
№	duration	Σ					$\overline{\Sigma}$
I	14	11	12	14	11	8	11.3
II	11	8	42	59	29	33	40.8
III	10	10	18	20	13	13	16.0
IV	7	2	5	3	2	3	3.3
V	7	5	16	18	18	13	16.3
VI	19	6	8	9	2	3	5.5
VII	9	6	11	17	9	13	12.5
VIII	6	8	21	34	26	24	26.3
IX	29	10	13	21	9	10	13.3
X	12	11	8	6	4	2	5.0
XI	8	2	2	4	3	0	2.3
XII	8	4	13	7	7	10	9.3
XIII	10	5	5	9	1	1	4.0
XIV	7	2	0	1	2	0	0.8
XV	6	3	4	4	5	5	4.5
XVI	14	12	35	39	25	27	31.5
XVII	22	12	21	17	10	12	15.0
XVIII	23	11	8	8	3	5	6.0
XIX	6	3	1	6	2	3	3.0
XX	7	7	5	5	6	5	5.3
XXI	7	9	5	8	4	4	5.3
XXII	8	6	7	7	6	5	6.3
XXIII	6	3	11	8	4	11	8.5
XXIV	8	8	16	14	5	10	11.3
XXV	6	7	18	18	14	14	16.0

Table B.7. Number of event-specific reference (ERA5) pattern matches (Σ) in RCM simulations forced with **EC-Earth**. Letters stand for the following RCM simulations: R: Reference (ERA5), A: CCLM4-8-17, B: HIRHAM5, C: REMO2015, D: WRF331F, E: RACMO22E, F: REMO2009, G: RCA4. In addition, an RCM mean (avg) number of matches is provided in the last column. Events considered are ice episodes of short (class 3–Ds), medium (class 3–Dm) and long duration (class 3–Dl) only.

№	duration	R Σ	A	B	C	E	G	avg $\overline{\Sigma}$
I	14	11	12	16	9	8	7	10.4
II	11	8	18	19	15	14	12	15.6
III	10	10	6	9	6	5	3	5.8
IV	7	2	1	0	4	0	2	1.4
V	7	5	10	14	7	16	5	10.4
VI	19	6	5	6	6	2	5	4.8
VII	9	6	3	10	3	8	4	5.6
VIII	6	8	6	11	6	9	10	8.4
IX	29	10	8	9	7	5	4	6.6
X	12	11	9	8	7	7	6	7.4
XI	8	2	0	2	1	1	0	0.8
XII	8	4	8	8	5	2	7	6.0
XIII	10	5	2	2	9	4	3	4.0
XIV	7	2	1	0	0	2	2	1.0
XV	6	3	2	2	3	2	1	2.0
XVI	14	12	15	21	13	12	9	14.0
XVII	22	12	13	11	9	12	8	10.6
XVIII	23	11	7	3	2	2	3	3.4
XIX	6	3	2	3	2	2	2	2.2
XX	7	7	1	3	1	4	2	2.2
XXI	7	9	9	6	4	3	6	5.6
XXII	8	6	2	7	3	4	2	3.6
XXIII	6	3	3	4	4	4	2	3.4
XXIV	8	8	7	6	8	5	6	6.4
XXV	6	7	10	3	10	2	9	6.8

Table B.8. Number of event-specific reference (ERA5) pattern matches (Σ) in RCM simulations forced with **IPSL-CM5A-MR**. Letters stand for the following RCM simulations: R: Reference (ERA5), A: CCLM4-8-17, B: HIRHAM5, C: REMO2015, D: WRF331F, E: RACMO22E, F: REMO2009, G: RCA4. In addition, an RCM mean (avg) number of matches is provided in the last column. Events considered are ice episodes of short (class 3–Ds), medium (class 3–Dm) and long duration (class 3–Dl) only.

ERA5 event pattern		R	D	E	G	avg
№	duration	Σ				$\overline{\Sigma}$
I	14	11	4	4	2	3.3
II	11	8	29	35	25	29.7
III	10	10	4	9	6	6.3
IV	7	2	1	2	1	1.3
V	7	5	2	5	5	4.0
VI	19	6	7	3	2	4.0
VII	9	6	10	7	6	7.7
VIII	6	8	12	14	11	12.3
IX	29	10	10	7	4	7.0
X	12	11	4	3	2	3.0
XI	8	2	2	1	2	1.7
XII	8	4	3	1	1	1.7
XIII	10	5	0	1	1	0.7
XIV	7	2	1	0	0	0.3
XV	6	3	1	0	1	0.7
XVI	14	12	12	15	7	11.3
XVII	22	12	10	6	4	6.7
XVIII	23	11	4	2	2	2.7
XIX	6	3	2	1	1	1.3
XX	7	7	2	2	3	2.3
XXI	7	9	1	0	0	0.3
XXII	8	6	0	4	3	2.3
XXIII	6	3	3	3	3	3.0
XXIV	8	8	11	6	6	7.7
XXV	6	7	9	8	7	8.0

Table B.9. Number of event-specific reference (ERA5) pattern matches (Σ) in RCM simulations forced with **MIROC5**. Letters stand for the following RCM simulations: R: Reference (ERA5), A: CCLM4-8-17, B: HIRHAM5, C: REMO2015, D: WRF331F, E: RACMO22E, F: REMO2009, G: RCA4. In addition, an RCM mean (avg) number of matches is provided in the last column. Events considered are ice episodes of short (class 3–Ds), medium (class 3–Dm) and long duration (class 3–Dl) only.

ERA5 event pattern		R	A	C	avg
№	duration	Σ			$\overline{\Sigma}$
I	14	11	19	8	13.5
II	11	8	28	32	30
III	10	10	20	15	17.5
IV	7	2	10	6	8
V	7	5	11	10	10.5
VI	19	6	23	21	22
VII	9	6	13	7	10
VIII	6	8	21	26	23.5
IX	29	10	26	20	23
X	12	11	19	17	18
XI	8	2	9	2	5.5
XII	8	4	0	6	3
XIII	10	5	11	10	10.5
XIV	7	2	0	0	0
XV	6	3	3	1	2
XVI	14	12	23	19	21
XVII	22	12	26	20	23
XVIII	23	11	19	15	17
XIX	6	3	3	3	3
XX	7	7	9	11	10
XXI	7	9	6	3	4.5
XXII	8	6	10	4	7
XXIII	6	3	5	9	7
XXIV	8	8	23	16	19.5
XXV	6	7	35	37	36

Table B.10. Number of event-specific reference (ERA5) pattern matches (Σ) in RCM simulations forced with **MPI-ESM-LR**. Letters stand for the following RCM simulations: R: Reference (ERA5), A: CCLM4-8-17, B: HIRHAM5, C: REMO2015, D: WRF331F, E: RACMO22E, F: REMO2009, G: RCA4. In addition, an RCM mean (avg) number of matches is provided in the last column. Events considered are ice episodes of short (class 3–Ds), medium (class 3–Dm) and long duration (class 3–Dl) only.

№	ERA5 event pattern duration	R Σ	A	C	E	F	G	avg $\overline{\Sigma}$
I	14	11	7	8	6	3	8	6.4
II	11	8	14	19	12	17	8	14
III	10	10	1	0	1	3	1	1.2
IV	7	2	3	1	0	2	2	1.6
V	7	5	7	8	5	8	1	5.8
VI	19	6	7	7	4	6	4	5.6
VII	9	6	5	7	4	5	2	4.6
VIII	6	8	10	5	10	6	8	7.8
IX	29	10	11	11	9	11	6	9.6
X	12	11	7	4	6	4	4	5
XI	8	2	2	1	4	3	0	2
XII	8	4	3	6	2	6	4	4.2
XIII	10	5	4	3	5	5	4	4.2
XIV	7	2	1	0	1	1	1	0.8
XV	6	3	1	2	0	1	0	0.8
XVI	14	12	12	8	10	10	5	9
XVII	22	12	12	12	11	8	6	9.8
XVIII	23	11	6	1	4	3	3	3.4
XIX	6	3	0	0	1	1	0	0.4
XX	7	7	3	4	2	0	3	2.4
XXI	7	9	6	5	3	4	4	4.4
XXII	8	6	2	6	3	2	3	3.2
XXIII	6	3	1	0	3	3	0	1.4
XXIV	8	8	4	4	4	4	3	3.8
XXV	6	7	6	7	3	6	7	5.8

Table B.11. Number of event-specific reference (ERA5) pattern matches (Σ) in RCM simulations forced with **NorESM1-M**. Letters stand for the following RCM simulations: R: Reference (ERA5), A: CCLM4-8-17, B: HIRHAM5, C: REMO2015, D: WRF331F, E: RACMO22E, F: REMO2009, G: RCA4. In addition, an RCM mean (avg) number of matches is provided in the last column. Events considered are ice episodes of short (class 3–Ds), medium (class 3–Dm) and long duration (class 3–Dl) only.

ERA5 event pattern		R	B	C	E	avg
№	duration	Σ				$\overline{\Sigma}$
I	14	11	5	4	3	4.0
II	11	8	15	11	14	13.3
III	10	10	5	4	3	4.0
IV	7	2	1	0	1	0.7
V	7	5	3	3	3	3.0
VI	19	6	2	3	2	2.3
VII	9	6	2	1	1	1.3
VIII	6	8	6	9	10	8.3
IX	29	10	4	6	6	5.3
X	12	11	6	4	6	5.3
XI	8	2	4	2	1	2.3
XII	8	4	0	0	0	0.0
XIII	10	5	9	4	8	7.0
XIV	7	2	1	0	0	0.3
XV	6	3	0	1	1	0.7
XVI	14	12	5	6	6	5.7
XVII	22	12	5	6	6	5.7
XVIII	23	11	5	2	1	2.7
XIX	6	3	4	3	2	3.0
XX	7	7	1	1	1	1.0
XXI	7	9	3	1	1	1.7
XXII	8	6	3	3	2	2.7
XXIII	6	3	2	1	3	2.0
XXIV	8	8	6	6	6	6.0
XXV	6	7	6	6	6	6.0

B.2 Acknowledgments

B.3 Conflict of Interest

The authors declare that they have no conflict of interest.

B.4 Data availability

The RCM data that support the findings of this study are available in ESGF at `https://esgf-data.dkrz.de/projects/esgf-dkrz/`. These data were derived from the following resources available in the public domain: `https://esgf-data.dkrz.de/search/cordex-dkrz/`.

ERA5 and E-OBS data that support the findings of this study are openly available through the Climate Data Store (`https://cds.climate.copernicus.eu/#!/home`) at

http://doi.org/10.24381/cds.adbb2d47 (ERA5 hourly data on single levels), http://doi.org/10.24381/cds.bd0915c6 (ERA5 hourly data on pressure levels) and http://doi.org/10.24381/cds.151d3ec6 (E-OBS data).

C   Third paper

This section will be submitted together with chapter 4.

C.1   Supporting information

Prevailing circulation patterns during ice episodes in far future
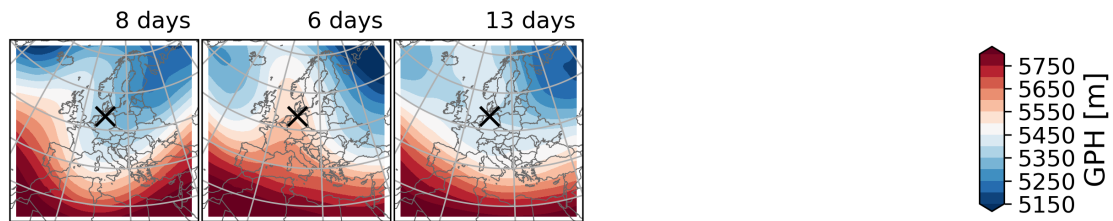
Figure C.1. Ice episode patterns (GPH_{500}) in far future based on CCLM4-8-17 with CanESM2 forcing. The number of days in the header indicates the duration of the respective ice episode.
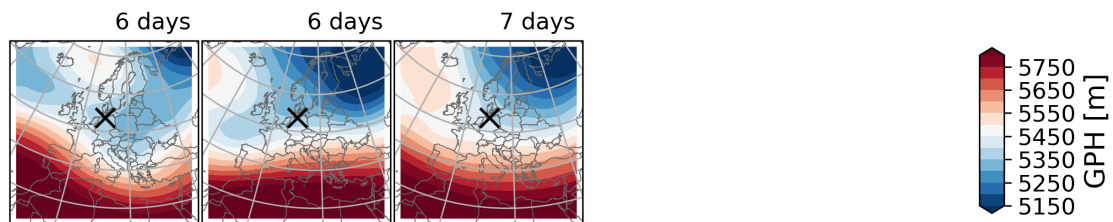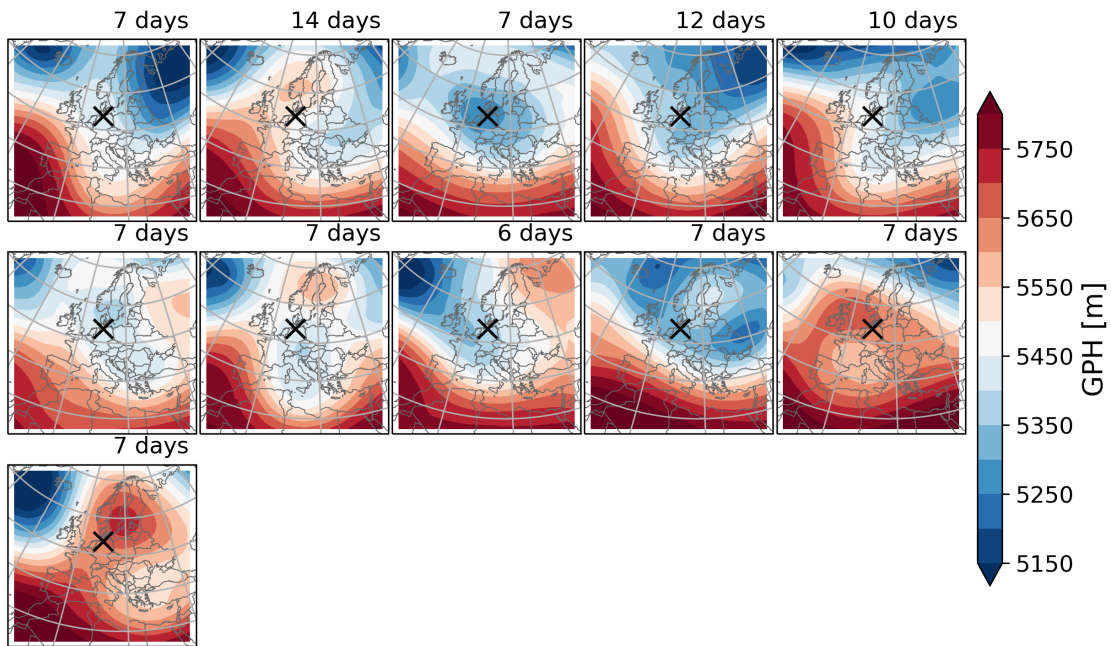

Figure C.2. Ice episode patterns (GPH_{500}) in far future based on REMO2015 with CanESM2 forcing. The number of days in the header indicates the duration of the respective ice episode.
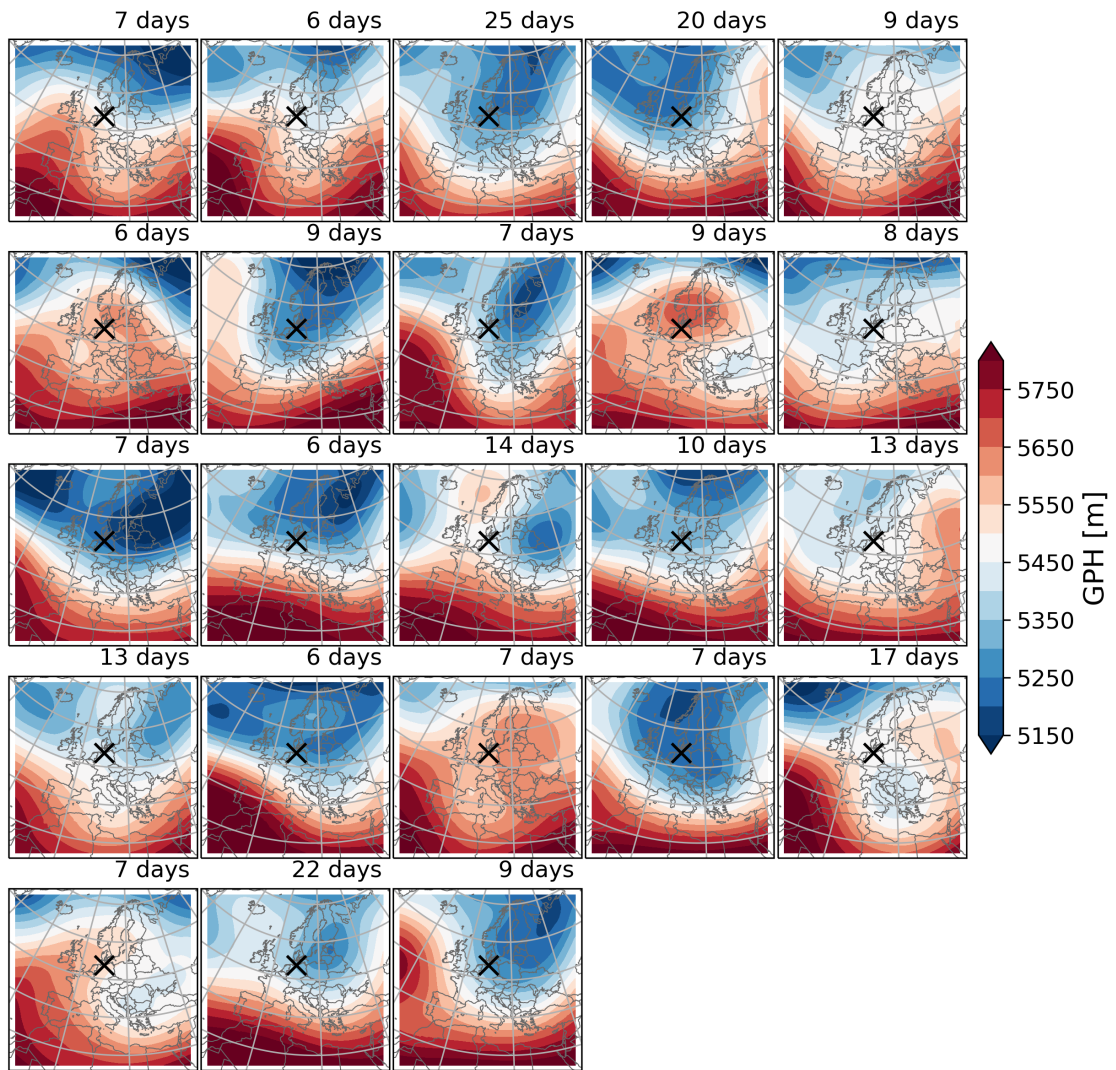
Figure C.3. Ice episode patterns (GPH_{500}) in far future based on CCLM4-8-17 with CNRM-CM5 forcing. The number of days in the header indicates the duration of the respective ice episode.
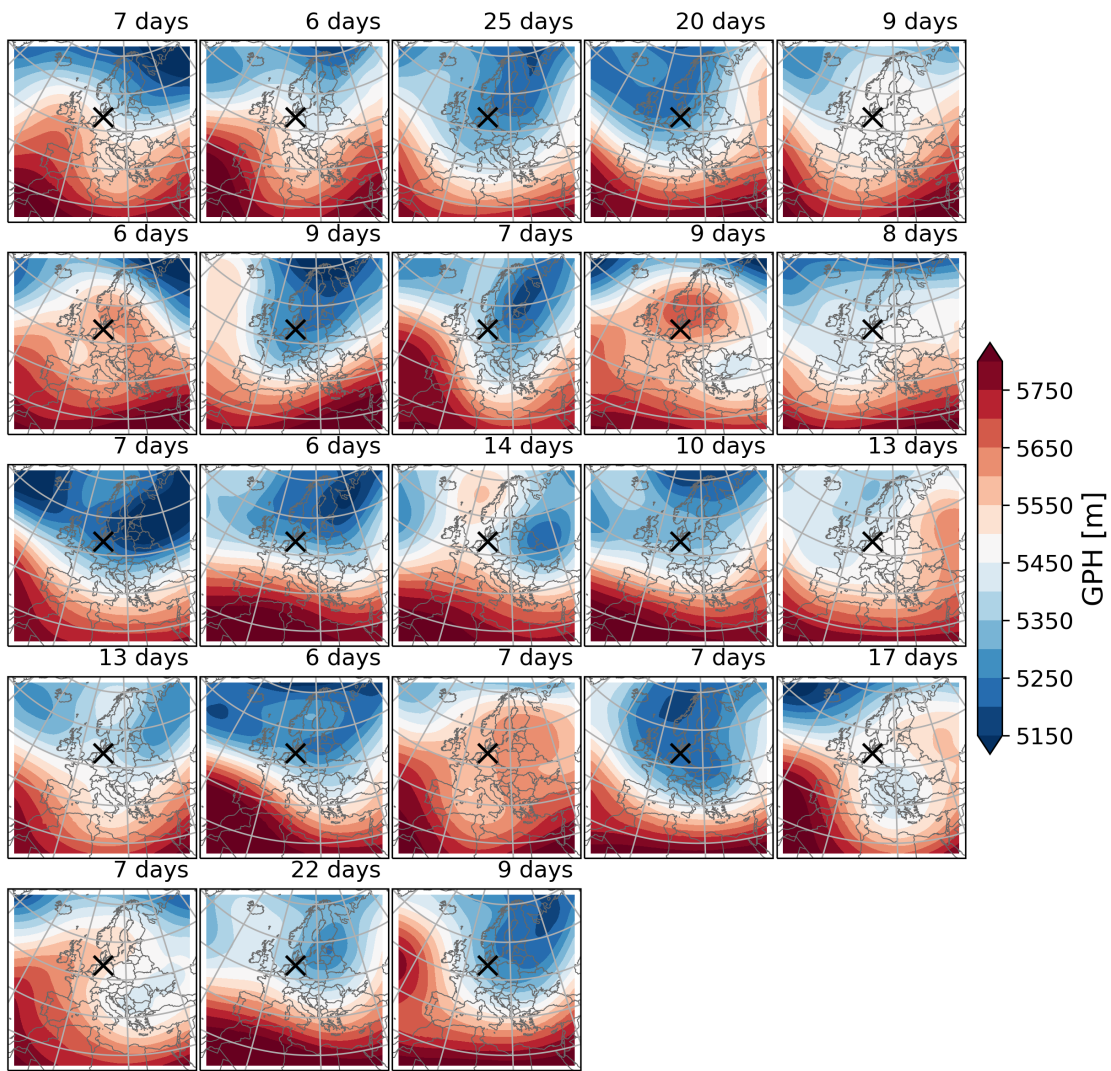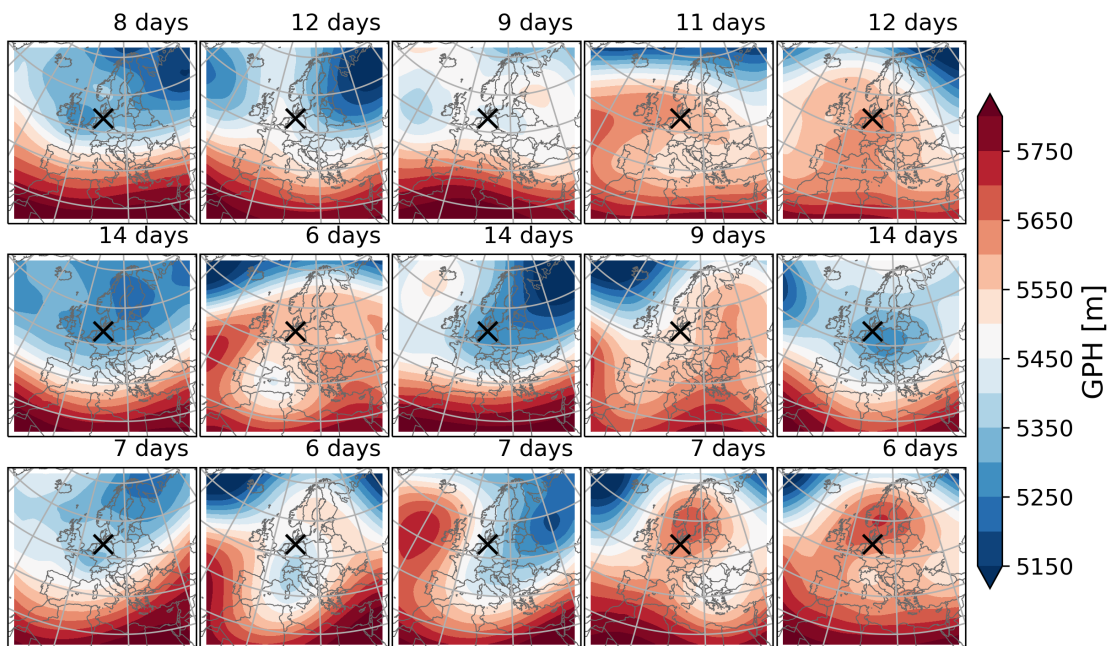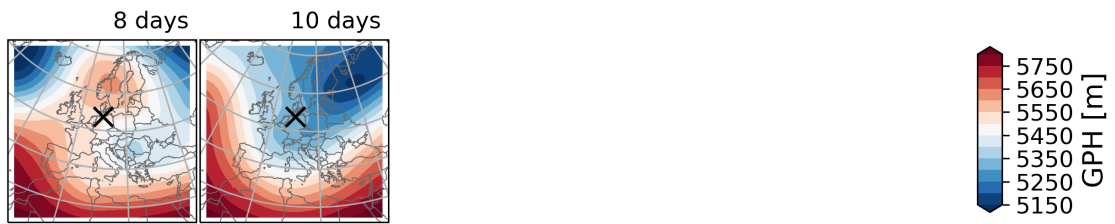
Figure C.4. Ice episode patterns (GPH$_{500}$) in far future based on REMO2015 with CNRM-CM5 forcing. The number of days in the header indicates the duration of the respective ice episode.
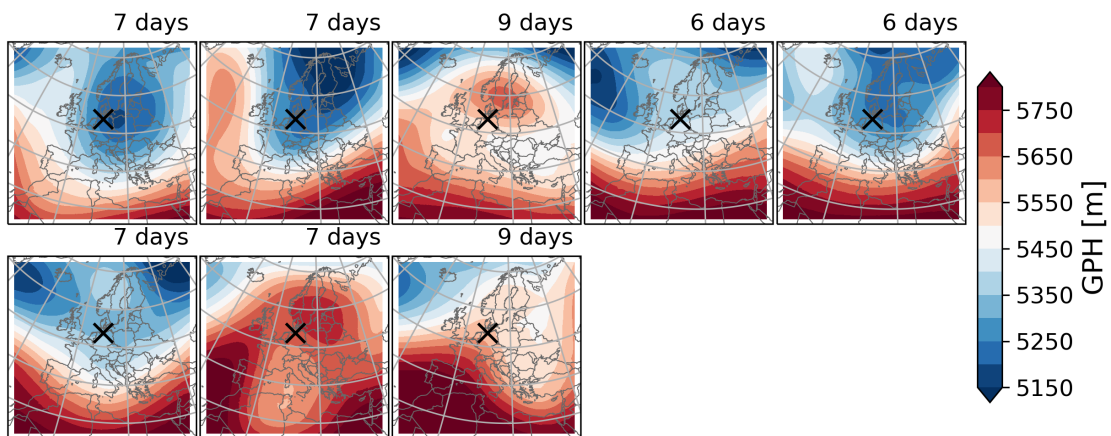
Figure C.5. Ice episode patterns (GPH$_{500}$) in far future based on RACMO22E with CNRM-CM5 forcing. The number of days in the header indicates the duration of the respective ice episode.
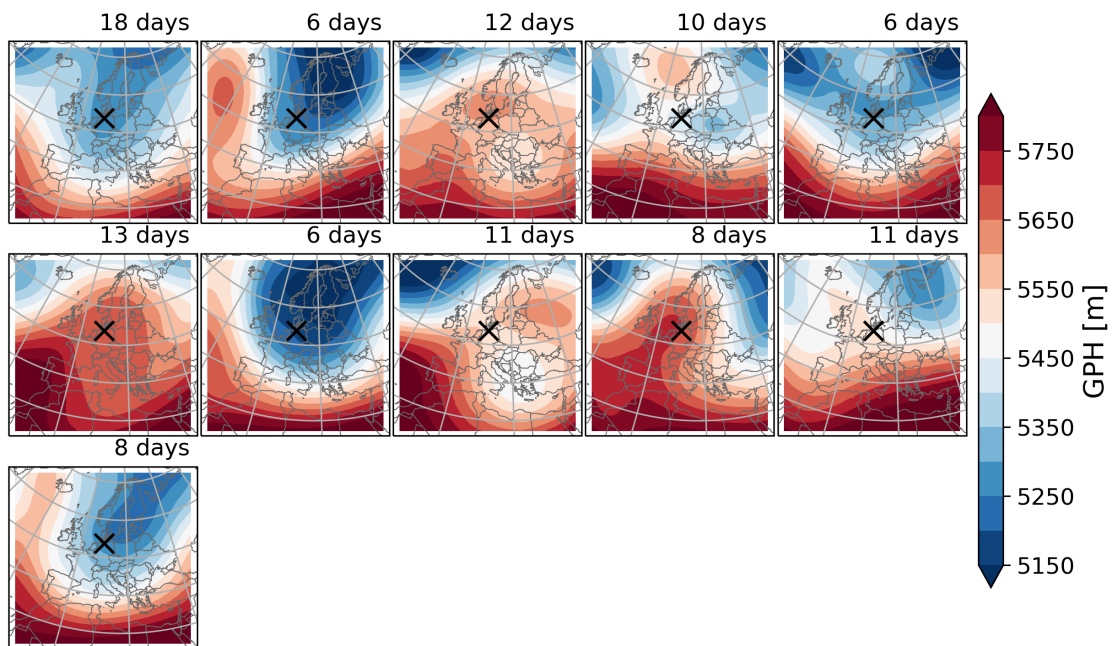
Figure C.6. Ice episode patterns (GPH$_{500}$) in far future based on RCA4 with CNRM-CM5 forcing. The number of days in the header indicates the duration of the respective ice episode.
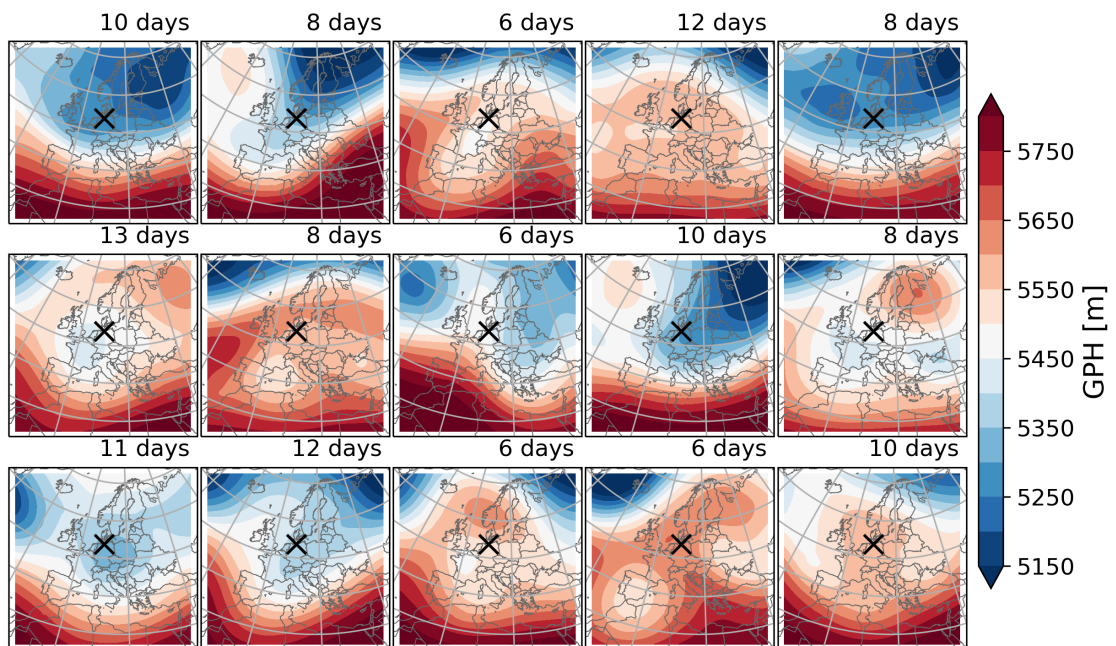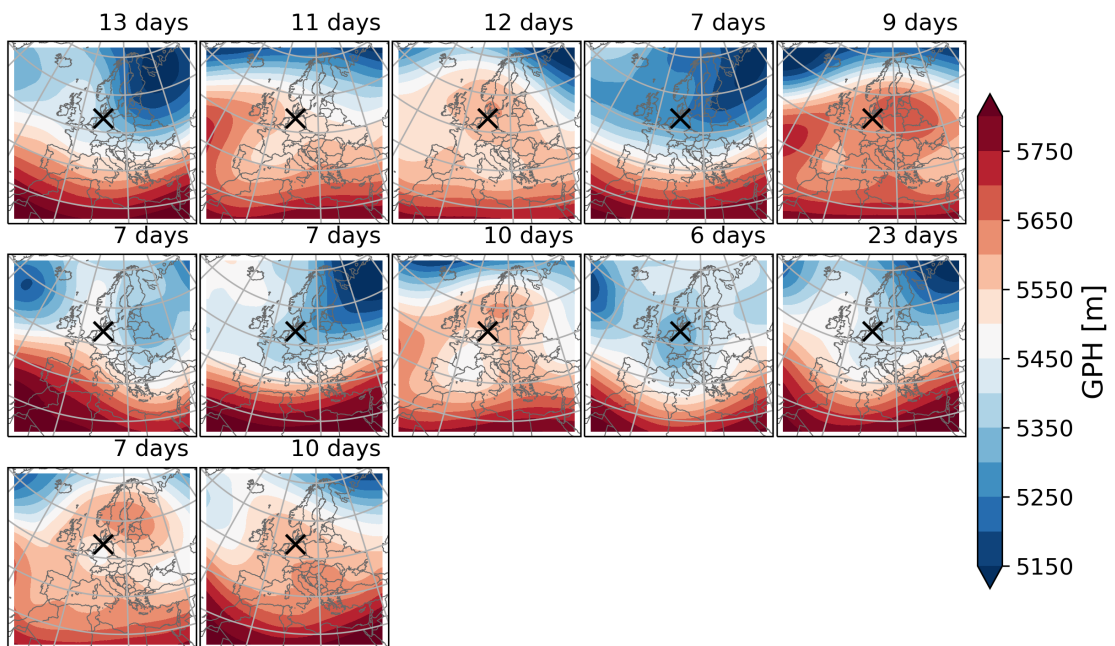

Figure C.7. Ice episode patterns (GPH$_{500}$) in far future based on CCLM4-8-17 with EC-Earth forcing. The number of days in the header indicates the duration of the respective ice episode.

Figure C.8. Ice episode patterns (GPH$_{500}$) in far future based on HIRHAM5 with EC-Earth forcing. The number of days in the header indicates the duration of the respective ice episode.

Figure C.9. Ice episode patterns (GPH$_{500}$) in far future based on REMO2015 with EC-Earth forcing. The number of days in the header indicates the duration of the respective ice episode.
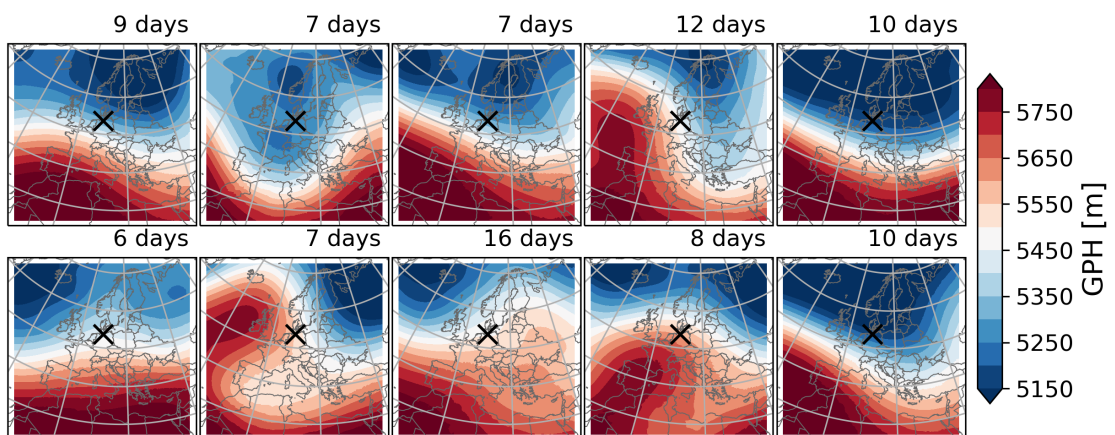
Figure C.10. Ice episode patterns (GPH$_{500}$) in far future based on RACMO22E with EC-Earth forcing. The number of days in the header indicates the duration of the respective ice episode.

Figure C.11. Ice episode patterns (GPH$_{500}$) in far future based on RCA4 with EC-Earth forcing. The number of days in the header indicates the duration of the respective ice episode.
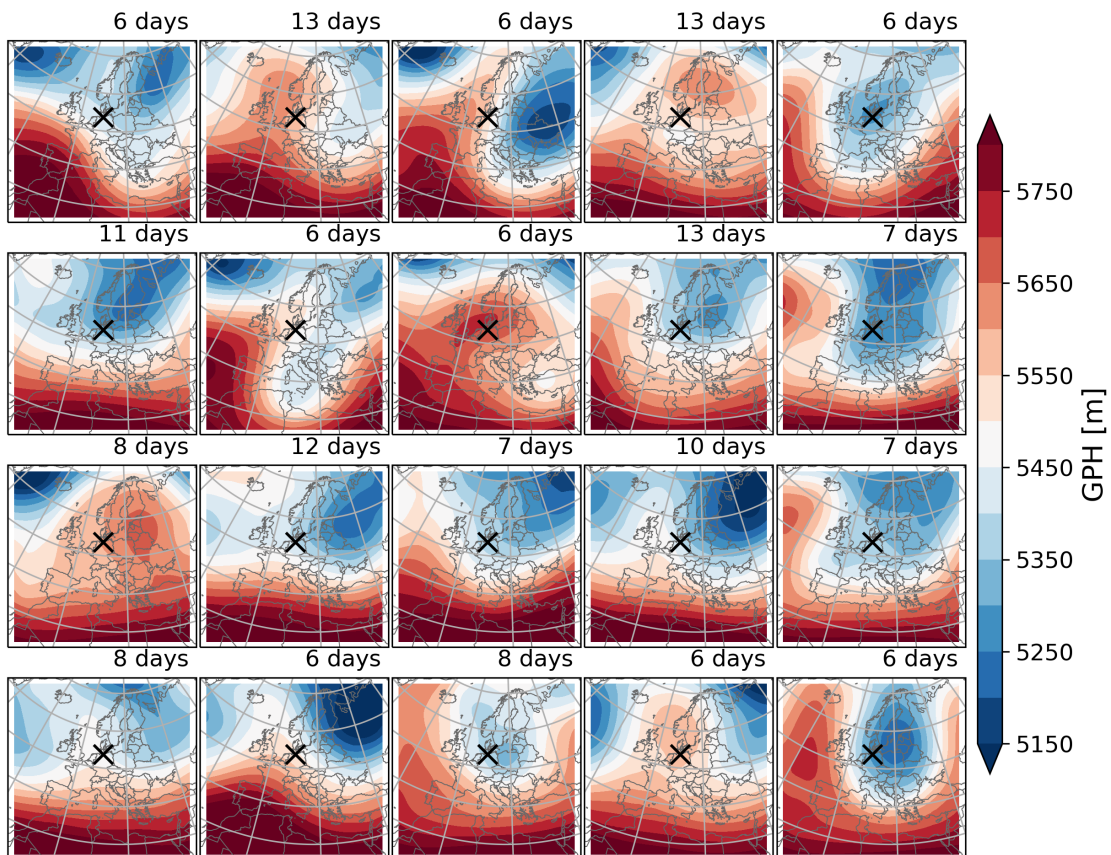

Figure C.12. Ice episode patterns (GPH$_{500}$) in far future based on WRF331F with IPSL-CM5A-MR forcing. The number of days in the header indicates the duration of the respective ice episode.
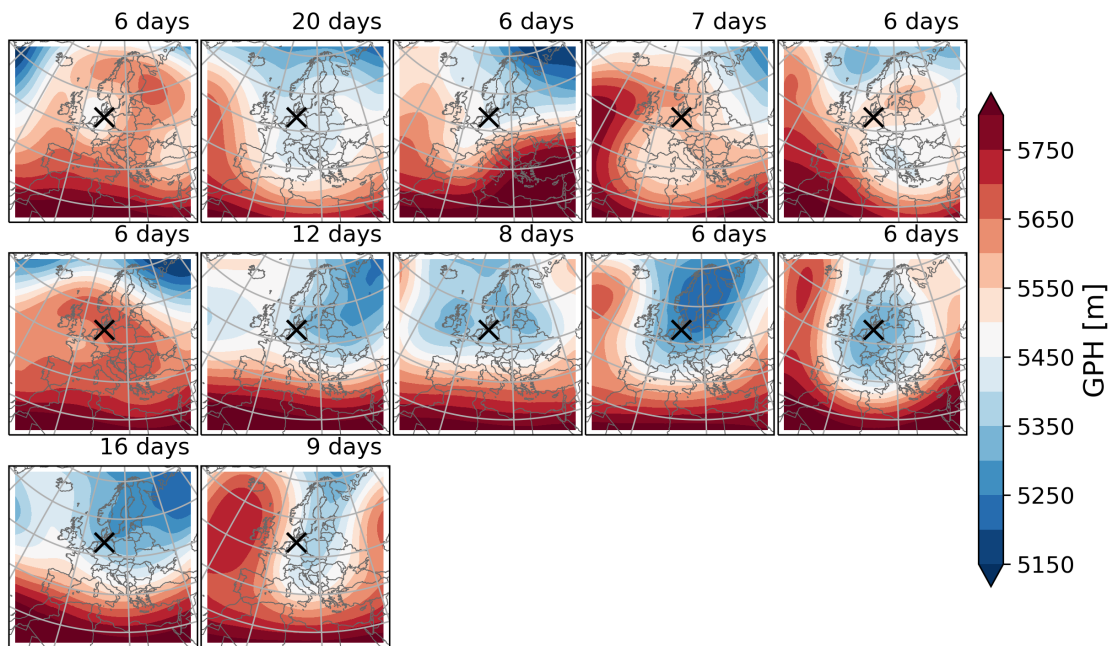
Figure C.13. Ice episode patterns (GPH$_{500}$) in far future based on RACMO22E with IPSL-CM5A-MR forcing. The number of days in the header indicates the duration of the respective ice episode.
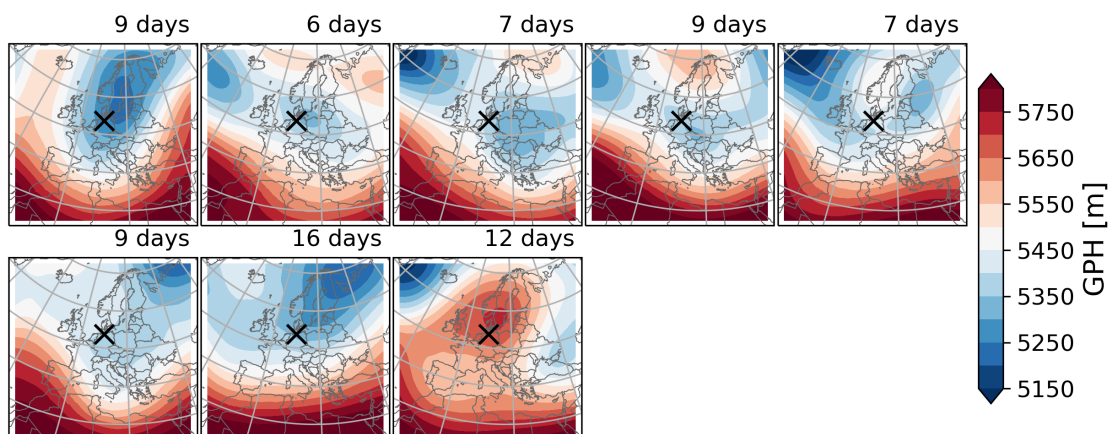

Figure C.14. Ice episode patterns (GPH$_{500}$) in far future based on RCA4 with IPSL-CM5A-MR forcing. The number of days in the header indicates the duration of the respective ice episode.
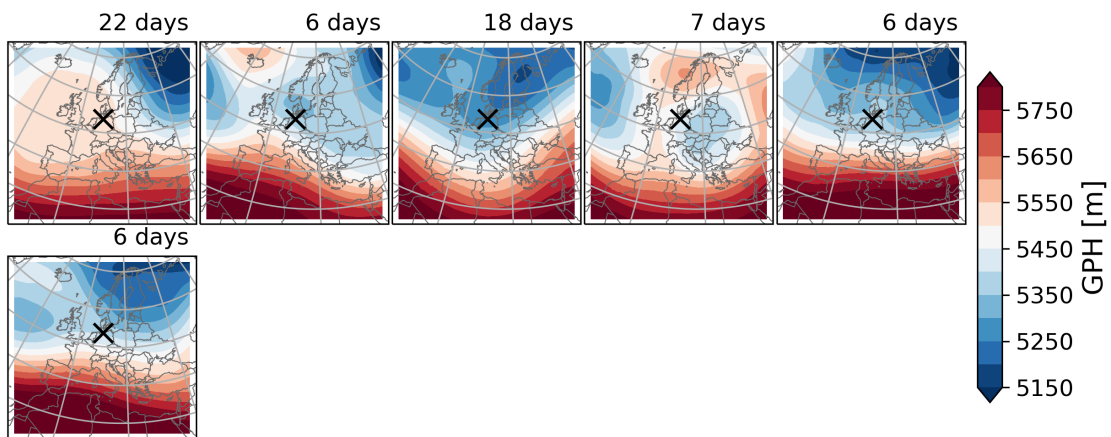

Figure C.15. Ice episode patterns (GPH$_{500}$) in far future based on CCLM4-8-17 with MIROC5 forcing. The number of days in the header indicates the duration of the respective ice episode.

Figure C.16. Ice episode patterns (GPH$_{500}$) in far future based on REMO2015 with MIROC5 forcing. The number of days in the header indicates the duration of the respective ice episode.
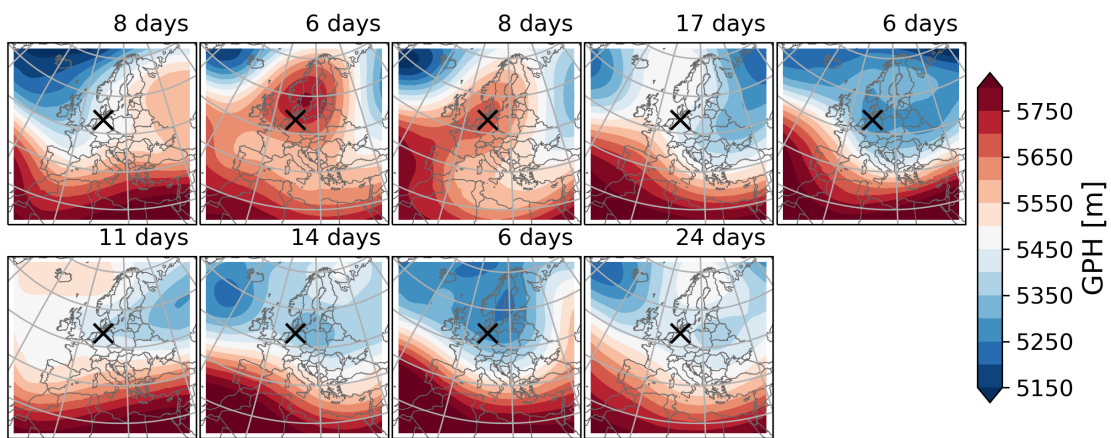

Figure C.17. Ice episode patterns (GPH$_{500}$) in far future based on CCLM4-8-17 with MPI-ESM-LR forcing. The number of days in the header indicates the duration of the respective ice episode.
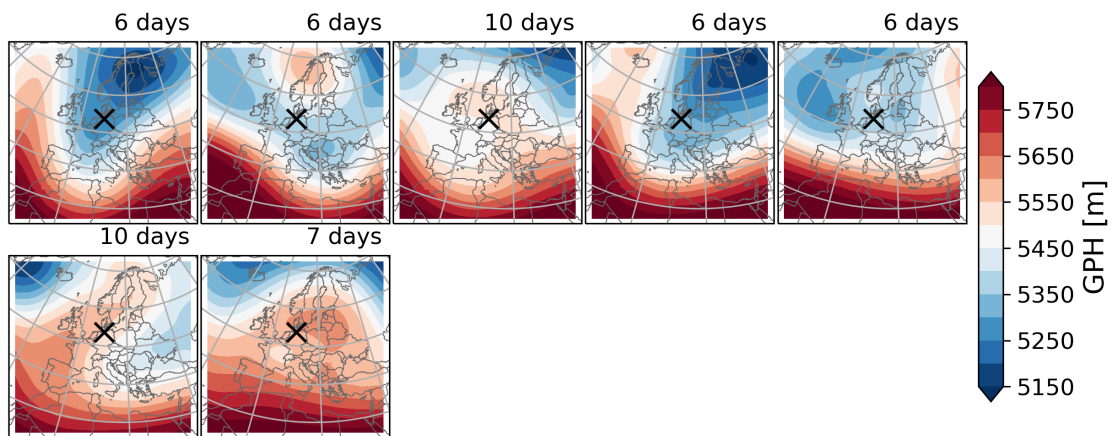
Figure C.18. Ice episode patterns (GPH$_{500}$) in far future based on REMO2015 with MPI-ESM-LR forcing. The number of days in the header indicates the duration of the respective ice episode.
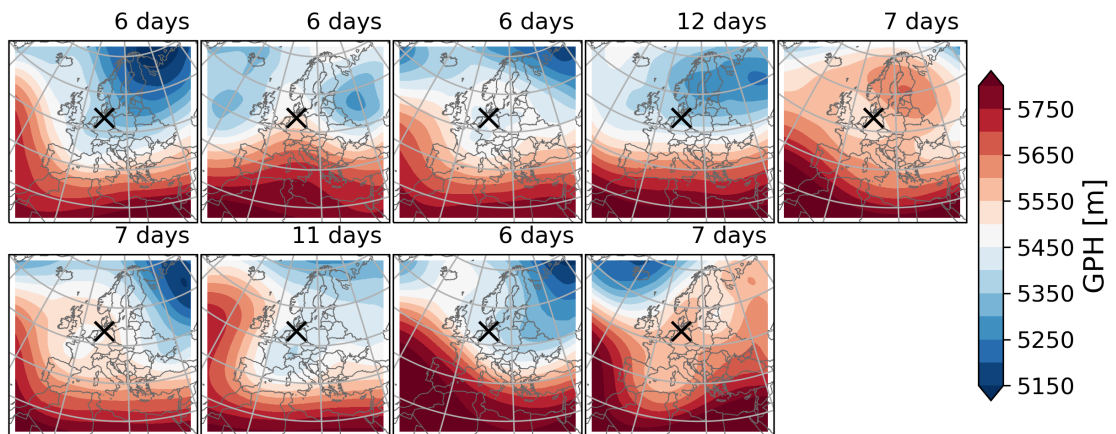

Figure C.19. Ice episode patterns (GPH$_{500}$) in far future based on RACMO22E with MPI-ESM-LR forcing. The number of days in the header indicates the duration of the respective ice episode.
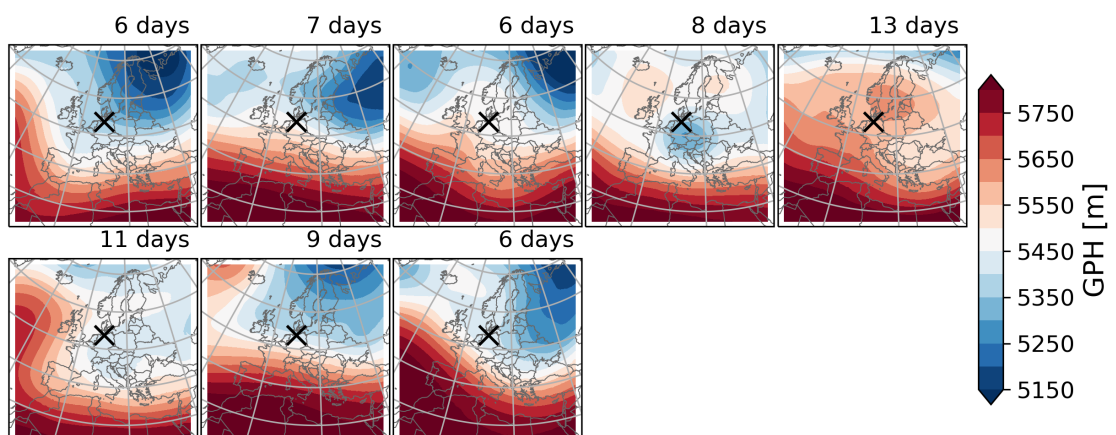

Figure C.20. Ice episode patterns (GPH$_{500}$) in far future based on REMO2009 with MPI-ESM-LR forcing. The number of days in the header indicates the duration of the respective ice episode.

Figure C.21. Ice episode patterns (GPH_{500}) in far future based on RCA4 with MPI-ESM-LR forcing. The number of days in the header indicates the duration of the respective ice episode.

Figure C.22. Ice episode patterns (GPH_{500}) in far future based on REMO2015 with NorESM1-M forcing. The number of days in the header indicates the duration of the respective ice episode.

Figure C.23. Ice episode patterns (GPH_{500}) in far future based on RACMO22E with NorESM1-M forcing. The number of days in the header indicates the duration of the respective ice episode.

C.2 Acknowledgements

D   Temperatures in Hamburg during ice episodes

This appendix is intended for informational purposes only. The temperatures referenced in the abstract and in the conclusions can be obtained from the following tables.

Table C.1. Average ice episode temperatures in Hamburg. The average is always calculated for the given time period. In addition to E-OBS and ERA5, average temperatures for four different climate periods of the GCM driven RCM projections (RCP8.5) are provided as ensemble mean values. All values are given in °C, changes (Δ) are given relative to 1971–2000.

data set	time period	T_{avg}		T_{max}		T_{min}	
E-OBS	1979–2005	-6.3		-3.5		-9.7	
ERA5	1979–2005	-6.1		-3.6		-8.9	
RCM ensemble	1979–2005	-6.1		-3.6		-8.7	
	1971–2000	-6.1		-3.7		-8.8	
	2030–2060	-5.2	$\Delta + 1.0$	-3.0	$\Delta + 0.6$	-7.5	$\Delta + 1.3$
	2065–2095	-4.6	$\Delta + 1.6$	-2.5	$\Delta + 1.2$	-6.7	$\Delta + 2.1$

Table C.2. Minimum ice episode temperatures in Hamburg. The minimum is always calculated for the given time period. In addition to E-OBS and ERA5, minimum temperatures for four different climate periods of the GCM driven RCM projections (RCP8.5) are provided as ensemble mean values. All values are given in °C, changes (Δ) are given relative to 1971–2000.

data set	time period	T_{avg}		T_{max}		T_{min}	
E-OBS	1979–2005	-15.2		-11.9		-20.2	
ERA5	1979–2005	-16.9		-13.3		-21.4	
RCM ensemble	1979–2005	-17.6		-14.6		-21.5	
	1971–2000	-18.6		-15.1		-22.7	
	2030–2060	-14.2	$\Delta + 4.4$	-11.5	$\Delta + 3.6$	-17.4	$\Delta + 5.2$
	2065–2095	-10.4	$\Delta + 8.2$	-7.6	$\Delta + 7.5$	-13.2	$\Delta + 9.4$

List of Acronyms

ACC	Anomaly Correlation Coefficient
AO	Arctic Oscillation
C3S	Copernicus Climate Change Service
CDOs	Climate Data Operators
CMIP	Coupled Model Intercomparison Project
DWD	Deutscher Wetterdienst (German Meteorological Service)
ECA&D	European Climate Assessment and Dataset project
ECMWF	European Centre for Medium-Range Weather Forecasts
ENSO	El Nino Southern Oscillation
(EURO-)CORDEX	Coordinated Regional Downscaling Experiment (for Europe)
GCM	Global Climate Model
GPH	Geopotential height
IPCC	Intergovernmental Panel on Climate Change
NAO	North Atlantic Oscillation
PDO	Pacific Decadal Oscillation
RCM	Regional Climate Model
RCP	representative concentration pathway
SOI	Southern Oscillation Index
SSIM	Structural Similarity (Index)
WCRP	World Climate Research Programme
WMO	World Meteorological Organization

References

BALDAUF, M., A. SEIFERT, J. FÖRSTNER, D. MAJEWSKI, M. RASCHENDORFER, T. REINHARDT, 2011: Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities. *Monthly Weather Review* **139**(12), 3887–3905, doi: 10.1175/MWR-D-10-05013.1, https://journals.ametsoc.org/view/journals/mwre/139/12/mwr-d-10-05013.1.xml.

BAUMGARTNER, E. A., M. BELSON, C. RUBIN, M. PATEL, 2008: Hypothermia and Other Cold-Related Morbidity Emergency Department Visits: United States, 1995–2004. *Wilderness & Environmental Medicine* **19**(4), 233–237, doi: 10.1580/07-WEME-OR-104.1, https://doi.org/10.1580/07-WEME-OR-104.1.

BAUR, F., P. HESS, H. NAGEL, 1944: Katalog der Grosswetterlagen Europas 1881–1939.

BECK, C., J. JACOBEIT, P. D. JONES, 2007: Frequency and within-type variations of large-scale circulation types and their effects on low-frequency climate variability in central europe since 1780. *International Journal of Climatology* **27**(4), 473–491, doi: 10.1002/joc.1410, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.1410.

BEDNAR-FIEDL, B., G. BERNDES, J. BIRKMANN, L. F. CABEZA, F. CREUTZIG, D. DERYNG, W. KIESSLING, S. KREIBIEHL, V. KREY, G. KRINNER, S. D. LANGHANS, F. LECOCQ, E. L. GUNN, J. MAROTZKE, V. MASSON-DELMOTTE, J. C. MINX, A. PIRANI, A. POPP, H.-O. PÖRTNER, R. RANASINGHE, K. RIAHI, J. ROGELJ, S. SENEVIRATNE, J. SKEA, T. STORELVMO, P. THORNE, M. V. ASLST, 2024: The Next Frontier for Climate Change Science: Insights from the authors of the IPCC 6th Assessment Report on knowledge gaps and priorities for research. doi: 10.2777/34601, https://op.europa.eu/en/publication-detail/-/publication/72f1cb35-cee7-11ee-b9d9-01aa75ed71a1/lang.

BELL, L. M., K. H. SCHLÜNZEN, K. SIECK, 2023: Influence of data uncertainty on cold season threshold based climate indices. *Meteorologische Zeitschrift* **32**(3), 195–206, doi: 10.1127/metz/2023/1158, http://dx.doi.org/10.1127/metz/2023/1158.

BELL, L. M., K. H. SCHLÜNZEN, K. SIECK, 2024: Reproducibility of local cold season characteristics, ice episodes and prevailing circulation patterns in RCMs. *Quarterly Journal of the Royal Meteorological Society* , in review, in review.

BLACKPORT, R., J. A. SCREEN, 2020: Weakened evidence for mid-latitude impacts of Arctic warming. *Nature Climate Change* **10**, 1065–1066, doi: 10.1038/s41558-020-00954-y, https://doi.org/10.1038/s41558-020-00954-y.

BLACKPORT, R., J. A. SCREEN, VAN DER K. WIEL, R. BINTANJA, 2019: Minimal influence of reduced Arctic sea ice on coincident cold winters in mid-latitudes. *Nature Climate Change* **9**, 697–704, doi: 10.1038/s41558-019-0551-4, https://doi.org/10.1038/s41558-019-0551-4.

BØSSING CHRISTENSEN, O., M. DREWS, J. HESSELBJERG CHRISTENSEN, K. DETHLOFF, K. KETELSEN, I. HEBESTADT, A. RINKE, 2007: The HIRHAM Regional Climate Model. Version 5 (beta).

BRANDS, S., 2022: A circulation-based performance atlas of the CMIP5 and 6 models for regional climate studies in the Northern Hemisphere mid-to-high latitudes. *Geoscientific Model Development* **15**(4), 1375–1411, doi: 10.5194/gmd-15-1375-2022, https://doi.org/10.5194/gmd-15-1375-2022.

BRÖDE, P., D. FIALA, K. BŁAŻEJCZYK, I. HOLMÉR, G. JENDRITZKY, B. KAMPMANN, B. TINZ, G. HAVENITH, 2012: Deriving the operational procedure for the Universal Thermal Climate Index (UTCI). *International Journal of Biometeorology* **56**, 481–494, doi: 10.1007/s00484-011-0454-1, https://doi.org/10.1007/s00484-011-0454-1.

BRUNNER, L., N. SCHALLER, J. ANSTEY, J. SILLMANN, A. K. STEINER, 2018: Dependence of Present and Future European Temperature Extremes on the Location of Atmospheric Blocking. *Geophysical Research Letters* **45**(12), 6311–6320, doi: 10.1029/2018GL077837, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL077837.

BUEHLER, T., C. C. RAIBLE, T. F. STOCKER, 2011: The relationship of winter season North Atlantic blocking frequencies to extreme cold or dry spells in the ERA-40. *Tellus A* **63**(2), 212–222, doi: 10.1111/j.1600-0870.2010.00492.x, https://doi.org/10.1111/j.1600-0870.2010.00492.x.

BUNGERT, U., 2008: Einfluss der Nestung auf die Ergebnisse meteorologischer Modelle Ph.D. thesis, Universität Hamburg [german].

BÖTTCHER, M., 2023: Difficulties of overcoming the jump in regional to microscales. personal communication.

CADULE, P., P. FRIEDLINGSTEIN, L. BOPP, S. A. SITCH, C. D. JONES, P. CIAIS, S. L. PIAO, P. PEYLIN, 2010: Benchmarking coupled climate-carbon models against long-term atmospheric $CO2$ measurements. *Global Biogeochemical Cycles* **24**(2), doi: 10.1029/2009GB003556, https://doi.org/10.1029/2009GB003556.

CHENG, A., D. CHEN, K. WOODSTOCK, N. H. OGDEN, X. WU, J. WU, 2017: Analyzing the Potential Risk of Climate Change on Lyme Disease in Eastern Ontario, Canada Using Time Series Remotely Sensed Temperature Data and Tick Population Modelling. *Remote Sensing* **9**(6), doi: 10.3390/rs9060609, https://doi.org/10.3390/rs9060609.

COHEN, J., X. ZHANG, J. FRANCIS, T. JUNG, R. KWOK, J. OVERLAND, T. J. BALLINGER, U. S. BHATT, H. W. CHEN, D. COUMOU, S. FELDSTEIN, H. GU, D. HANDORF, G. HENDERSON, M. IONITA, M. KRETSCHMER, F. LALIBERTE, S. LEE, H. W. LINDERHOLM, W. MASLOWSKI, Y. PEINGS, K. PFEIFFER, I. RIGOR, T. SEMMLER, J. STROEVE, P. C. TAYLOR, S. VAVRUS, T. VIHMA, S. WANG, M. WENDISCH, Y. WU, J. YOON, 2020: Divergent consensuses on Arctic amplification influence on midlatitude severe winter weather. *Nature Climate Change* **10**, 20–29, doi: 10.1038/s41558-019-0662-y, https://doi.org/10.1038/s41558-019-0662-y.

COLIN, J., M. DÉQUÉ, R. RADU, S. SOMOT, 2010: Sensitivity study of heavy precipitation in Limited Area Model climate simulations: influence of the size of the domain and the use of the spectral nudging technique. *Tellus A: Dynamic Meteorology and Oceanography* **62**(5), 591–604, doi: 10.1111/j.1600-0870.2010.00467.x, https://doi.org/10.1111/j.1600-0870.2010.00467.x.

COPERNICUS CLIMATE CHANGE SERVICE (C3S) Climate indicators — Temperature. https://climate.copernicus.eu/climate-indicators/temperature, last accessed 15 April 2024.

CORNES, R. C., VAN DER G. SCHRIER, VAN DEN E. J. M. BESSELAAR, P. D. JONES, 2018: An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research—Atmospheres* **123**(17), 9391–9409, doi: 10.1029/2017JD028200, https://doi.org/10.1029/2017JD028200.

DE FREITAS, C., E. GRIGORIEVA, 2017: A comparison and appraisal of a comprehensive range of human thermal climate indices. *International Journal of Biometeorology* **61**, 487–512, doi: 10.1007/s00484-016-1228-6, https://doi.org/10.1007/s00484-016-1228-6.

DEE, D. P., S. M. UPPALA, A. J. SIMMONS, P. BERRISFORD, P. POLI, S. KOBAYASHI, U. ANDRAE, M. A. BALMASEDA, G. BALSAMO, P. BAUER, P. BECHTOLD, A. C. M. BELJAARS, VAN DE L. BERG, J. BIDLOT, N. BORMANN, C. DELSOL, R. DRAGANI, M. FUENTES, A. J. GEER, L. HAIMBERGER, S. B. HEALY, H. HERSBACH, E. V. HÓLM, L. ISAKSEN, P. KÅLLBERG, M. KÖHLER, M. MATRICARDI, A. P. MCNALLY, B. M. MONGE-SANZ, J.-J. MORCRETTE, B.-K. PARK, C. PEUBEY, DE P. ROSNAY, C. TAVOLATO, J.-N. THÉPAUT, F. VITART, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **137**(656), 553–597, doi: 10.1002/qj.828, https://doi.org/10.1002/qj.828.

DÉQUÉ, M., D. ROWELL, D. LÜTHI, F. GIORGI, J. CHRISTENSEN, B. ROCKEL, D. JACOB, E. KJELLSTRÖM, DE M. CASTRO, VAN DEN B. HURK, 2007: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections.. *Climatic Change* **81**, 53–70, doi: 10.1007/s10584-006-9228-x, https://doi.org/10.1007/s10584-006-9228-x.

DOAN, Q.-V., H. KUSAKA, T. SATO, F. CHEN, 2021: S-SOM v1.0: a structural self-organizing map algorithm for weather typing. *Geoscientific Model Development* **14**(4), 2097–2111, doi: 10.5194/gmd-14-2097-2021, https://gmd.copernicus.org/articles/14/2097/2021/.

DOBRYNIN, M., D. I. V. DOMEISEN, W. A. MÜLLER, L. BELL, S. BRUNE, F. BUNZEL, A. DÜSTERHUS, K. FRÖHLICH, H. POHLMANN, J. BAEHR, 2018: Improved Teleconnection-Based Dynamical Seasonal Predictions of Boreal Winter. *Geophysical Research Letters* **45**(8), 3605–3614, doi: 10.1002/2018GL077209, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2018GL077209.

DODMAN, D., B. HAYWARD, M. PELLING, V. CASTAN BROTO, W. CHOW, E. CHU, R. DAWSON, L. KHIRFAN, T. MCPHEARSON, A. PRAKASH, Y. ZHENG, G. ZIERVOGEL, 2023: Cities, Settlements and Key Infrastructure. In: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, Chapter 6, 907–1040 doi: 10.1017/9781009325844.008.907, https://www.ipcc.ch/report/ar6/wg2/chapter/chapter-6/.

DUBOIS, G., J. P. CERON, C. DUBOIS, M. D. FRIAS, S. HERRERA, 2016: Reliability and usability of tourism climate indices. *Earth Perspectives* **3**, 2, doi: 10.1186/s40322-016-0034-y, https://doi.org/10.1186/s40322-016-0034-y.

DUNN-SIGOUIN, E., S.-W. SON, 2013: Northern Hemisphere blocking frequency and duration in the CMIP5 models. *Journal of Geophysical Research: Atmospheres* **118**(3), 1179–1188, doi: 10.1002/jgrd.50143, https://doi.org/10.1002/jgrd.50143.

ECMWF ERA5 documentation. https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation, last accessed 2022-08-22.

EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS (ECMWF) ECMWF Charts – 500 hPa geopotential height and 850 hPa temperature. https://charts.ecmwf.int/products/medium-z500-t850?base_time=202403070000&projection=opencharts_europe&valid_time=202403170000, last accessed 2024-03-07.

EUROPEAN CLIMATE ASSESSMENT AND DATASET Indices Dictionary. https://www.ecad.eu/indicesextremes/indicesdictionary.php, last accessed 2024-04-26.

EXPERT TEAM ON CLIMATE CHANGE DETECTION AND INDICES (ETCCDI) Climate Change Indices. http://etccdi.pacificclimate.org/index.shtml, last accessed 2024-04-30.

FARANDA, D., G. MESSORI, A. JEZEQUEL, M. VRAC, P. YIOU, 2023: Atmospheric circulation compounds anthropogenic warming and impacts of climate extremes in Europe. *Proceedings of the National Academy of Sciences* **120**(13), e2214525120, doi: 10.1073/pnas.2214525120, https://www.pnas.org/doi/abs/10.1073/pnas.2214525120.

FERNER, K., M. BOETTCHER, K. SCHLÜNZEN, 2023: Modelling the heterogeneity of rain in an urban neighbourhood with an obstacle-resolving model. *Meteorologische Zeitschrift* **32**(1), 67–81, doi: 10.1127/metz/2022/1149, http://dx.doi.org/10.1127/metz/2022/1149.

FISCHEREIT, J., K. H. SCHLÜNZEN, 2018: Evaluation of thermal indices for their applicability in obstacle-resolving meteorology models. *International Journal of Biometeorology* **62**, 1887–1900, doi: 10.1007/s00484-018-1591-6, https://doi.org/10.1007/s00484-018-1591-6.

FPS URB-RCC URBan environments and Regional Climate Change – A WCRP CORDEX Flagship Pilot Study. https://ms.hereon.de/cordex_fps_urban/index.php.en, last accessed 2024-04-24.

FRANCIS, J. A., S. J. VAVRUS, 2012: Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophysical Research Letters* **39**(6), doi: 10.1029/2012GL051000, https://doi.org/10.1029/2012GL051000.

FRANCIS, J. A., N. SKIFIC, S. J. VAVRUS, J. COHEN, 2022: Measuring "Weather Whiplash" Events in North America: A New Large-Scale Regime Approach. *Journal of Geophysical Research: Atmospheres* **127**(17), e2022JD036717, doi: 10.1029/2022JD036717, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022JD036717.

FRANCIS, J. A., N. SKIFIC, Z. ZOBEL, 2023: Weather whiplash events in Europe and North Atlantic assessed as continental-scale atmospheric regime shifts. *npj Climate and Atmospheric Science* **6**(1), 216, doi: 10.1038/s41612-023-00542-9, https://doi.org/10.1038/s41612-023-00542-9.

GASPARRINI, A., Y. GUO, M. HASHIZUME, E. LAVIGNE, A. ZANOBETTI, J. SCHWARTZ, A. TO-
BIAS, S. TONG, J. ROCKLÖV, B. FORSBERG, M. LEONE, M. DE SARIO, M. L. BELL, Y. L.
GUO, C. F. WU, H. KAN, S. M. YI, DE M. SOUSA ZANOTTI STAGLIORIO COELHO, P. H. SAL-
DIVA, Y. HONDA, H. KIM, B. ARMSTRONG, 2015: Mortality risk attributable to high and low am-
bient temperature: a multicountry observational study. *Lancet (London, England)* **369**(9991), 369–
–375, doi: 10.1016/S0140-6736(14)62114-0, https://doi.org/10.1016/S0140-6736(14)
62114-0.

GERMAN METEOROLOGICAL SERVICE DWD OpenData. https://opendata.dwd.de/
climate_environment/CDC/, last accessed 2024-01-08.

GERMAN METEOROLOGICAL SERVICE Weather Glossary. www.wetterlexikon.eu, last accessed
2024-04-26.

GERSTENGARBE, F.-W., P. C. WERNER, 2005: Katalog der Grosswetterlagen Europas (1881–
2004). https://www.pik-potsdam.de/en/output/publications/pikreports/
.files/pr100.pdf/view.

GIORGI, F., 2019: Thirty Years of Regional Climate Modeling: Where Are We and Where
Are We Going next?. *Journal of Geophysical Research: Atmospheres* **124**(11), 5696–5723,
doi: 10.1029/2018JD030094, https://agupubs.onlinelibrary.wiley.com/doi/abs/
10.1029/2018JD030094.

GIORGI, F., W. J. GUTOWSKI, 2015: Regional Dynamical Downscaling and the CORDEX Initiative.
Annual Review of Environment and Resources **40**(1), 467–490, doi: 10.1146/annurev-environ-102014-
021217, https://doi.org/10.1146/annurev-environ-102014-021217.

GLECKLER, P. J., K. E. TAYLOR, C. DOUTRIAUX, 2008: Performance metrics for climate models.
Journal of Geophysical Research **113**(D6), doi: 10.1029/2007JD008972, https://doi.org/10.
1029/2007JD008972.

HALLETT, T. B., T. COULSON, J. G. PILKINGTON, T. H. CLUTTON-BROCK, J. M. PEMBERTON,
B. T. GRENFELL, 2004: Why large-scale climate indices seem to predict ecological processes bet-
ter than local weather. *Nature* **430**, 71–75, doi: 10.1038/nature02708, https://doi.org/10.
1038/nature02708.

HANSEN, F., D. BELUŠIĆ, 2021: Tailoring circulation type classification outcomes. *International Journal
of Climatology* **41**(14), 6145–6161, doi: 10.1002/joc.7171, https://rmets.onlinelibrary.
wiley.com/doi/abs/10.1002/joc.7171.

HANSEN, F., D. BELUŠIĆ, K. WYSER, T. KOENIGK, 2023: Future changes of circulation types and
their effects on surface air temperature and precipitation in the SMHI large ensemble. *Climate Dynam-
ics* **61**, 2921–2936, doi: 10.1007/s00382-023-06704-y, https://doi.org/10.1007/s00382-
023-06704-y.

HAWKINS, E., R. SUTTON, 2011: The potential to narrow uncertainty in projections of regional
precipitation change. *Climate Dynamics* **37**, 407–418, doi: 10.1007/s00382-010-0810-6, https:
//doi.org/10.1007/s00382-010-0810-6.

HERRERA-LORMENDEZ, P., N. MASTRANTONAS, H. DOUVILLE, A. HOY, J. MATSCHULLAT, 2022: Synoptic circulation changes over Central Europe from 1900 to 2100: Reanalyses and Coupled Model Intercomparison Project phase 6. *International Journal of Climatology* **42**(7), 4062–4077, doi: 10.1002/joc.7481, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.7481.

HERSBACH, H., B. BELL, P. BERRISFORD, S. HIRAHARA, A. HORÁNYI, J. MUÑOZ SABATER, J. NICOLAS, C. PEUBEY, R. RADU, D. SCHEPERS, A. SIMMONS, C. SOCI, S. ABDALLA, X. ABELLAN, G. BALSAMO, P. BECHTOLD, G. BIAVATI, J. BIDLOT, M. BONAVITA, J.-N. THÉPAUT, 2018a: ERA5 hourly data on pressure levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). doi: 10.24381/cds.bd0915c6, https://doi.org/10.24381/cds.bd0915c6, accessed at DKRZ on 01-JUN-2022.

HERSBACH, H., B. BELL, P. BERRISFORD, S. HIRAHARA, A. HORÁNYI, J. MUÑOZ SABATER, J. NICOLAS, C. PEUBEY, R. RADU, D. SCHEPERS, A. SIMMONS, C. SOCI, S. ABDALLA, X. ABELLAN, G. BALSAMO, P. BECHTOLD, G. BIAVATI, J. BIDLOT, M. BONAVITA, J.-N. THÉPAUT, 2018b: ERA5 hourly data on single levels from 1979 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS). doi: 10.24381/cds.adbb2d47, https://doi.org/10.24381/cds.adbb2d47, accessed at DKRZ on 01-JUN-2022.

HERSBACH, H., B. BELL, P. BERRISFORD, S. HIRAHARA, A. HORÁNYI, J. MUÑOZ SABATER, J. NICOLAS, C. PEUBEY, R. RADU, D. SCHEPERS, A. SIMMONS, C. SOCI, S. ABDALLA, X. ABELLAN, G. BALSAMO, P. BECHTOLD, G. BIAVATI, J. BIDLOT, M. BONAVITA, J.-N. THÉPAUT, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* , doi: 10.1002/qj.3803, https://doi.org/10.1002/qj.3803.

HOFFMANN, P., C. MENZ, A. SPEKAT, 2018: Bias adjustment for threshold-based climate indicators. *Andvances in Science and Research* **15**, 107–116, doi: 10.5194/asr-15-107-2018, https://doi.org/10.5194/asr-15-107-2018.

HOFFMANN, P., J. LEHMANN, B. FALLAH, F. F. HATTERMANN, 2021: Atmosphere similarity patterns in boreal summer show an increase of persistent weather conditions connected to hydro-climatic risks. *Scientific Reports* **11**(22893), doi: 10.1038/s41598-021-01808-z, https://doi.org/10.1038/s41598-021-01808-z.

HOHENEGGER, C., P. KORN, L. LINARDAKIS, R. REDLER, R. SCHNUR, P. ADAMIDIS, J. BAO, S. BASTIN, M. BEHRAVESH, M. BERGEMANN, J. BIERCAMP, H. BOCKELMANN, R. BROKOPF, N. BRÜGGEMANN, L. CASAROLI, F. CHEGINI, G. DATSERIS, M. ESCH, G. GEORGE, M. GIORGETTA, O. GUTJAHR, H. HAAK, M. HANKE, T. ILYINA, T. JAHNS, J. JUNGCLAUS, M. KERN, D. KLOCKE, L. KLUFT, T. KÖLLING, L. KORNBLUEH, S. KOSUKHIN, C. KROLL, J. LEE, T. MAURITSEN, C. MEHLMANN, T. MIESLINGER, A. K. NAUMANN, L. PACCINI, A. PEINADO, D. S. PRATURI, D. PUTRASAHAN, S. RAST, T. RIDDICK, N. ROEBER, H. SCHMIDT, U. SCHULZWEIDA, F. SCHÜTTE, H. SEGURA, R. SHEVCHENKO, V. SINGH, M. SPECHT, C. C. STEPHAN, VON J.-S. STORCH, R. VOGEL, C. WENGEL, M. WINKLER, F. ZIEMEN, J. MAROTZKE, B. STEVENS, 2023: ICON-Sapphire: simulating the components of the Earth system and their interactions at kilometer and subkilometer scales. *Geoscientific Model Development* **16**(2), 779–811, doi: 10.5194/gmd-16-779-2023, https://gmd.copernicus.org/articles/16/779/2023/.

HORTON, D. E., N. C. JOHNSON, D. SINGH, D. L. SWAIN, B. RAJARATNAM, N. S. DIFFENBAUGH, 2015: Contribution of changes in atmospheric circulation patterns to extreme temperature trends. *Nature* **522**, 465–469, doi: 10.1038/nature14550, https://doi.org/10.1038/nature14550.

HUGUENIN, M. F., E. M. FISCHER, S. KOTLARSKI, S. C. SCHERRER, C. SCHWIERZ, R. KNUTTI, 2020: Lack of Change in the Projected Frequency and Persistence of Atmospheric Circulation Types Over Central Europe. *Geophysical Research Letters* **47**(9), e2019GL086132, doi: 10.1029/2019GL086132, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL086132.

HURRELL, J. W., Y. KUSHNIR, G. OTTERSEN, M. VISBECK, 2003: The North Atlantic Oscillation: Climatic Significance and Environmental Impact Geophysical Monograph Series. – American Geophysical Union (AGU) doi: 10.1029/GM134, https://agupubs.onlinelibrary.wiley.com/doi/book/10.1029/GM134.

HUTCHINSON, M., P. GESSLER, 1994: Splines — more than just a smooth interpolator. *Geoderma* **62**(1), 45–67, doi: 10.1016/0016-7061(94)90027-2, https://www.sciencedirect.com/science/article/pii/0016706194900272.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC), 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change – Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA doi: 10.1017/9781009157896, https://www.ipcc.ch/report/ar6/wg1/.

IPCC Scoping Meeting of the Special Report on Climate Change and Cities. https://www.ipcc.ch/event/scoping-meeting-of-the-special-report-on-climate-change-and-cities/, last accessed 2024-04-24.

JACOB, D., DEN B. J. J. M. V. HURK, G. E. U. ANDRÆ, C. FORTELIUS, L. P. GRAHAM, S. D. JACKSON, U. KARSTENS, C. KÖPKEN, R. LINDAU, R. PODZUN, B. ROCKEL, F. RUBEL, B. H. SASS, R. N. B. SMITH, X. YANG, 2001: A comprehensive model inter-comparison study investigating the water budget during the BALTEX-PIDCAP period. *Meteorology and Atmospheric Physics volume* **77**, 19–43, doi: 10.1007/s007030170015, https://www.sciencedirect.com/science/article/pii/0016706194900272.

JACOB, D., J. PETERSEN, B. EGGERT, A. ALIAS, O. B. CHRISTENSEN, L. M. BOUWER, A. BRAUN, A. COLETTE, M. DÉQUÉ, G. GEORGIEVSKI, E. GEORGOPOULOU, A. GOBIET, L. MENUT, G. NIKULIN, A. HAENSLER, N. HEMPELMANN, C. JONES, K. KEULER, S. KOVATS, N. KRÖNER, S. KOTLARSKI, A. KRIEGSMANN, E. MARTIN, VAN E. MEIJGAARD, C. MOSELEY, S. PFEIFER, S. PREUSCHMANN, C. RADERMACHER, K. RADTKE, D. RECHID, M. ROUNSEVELL, P. SAMUELSSON, S. SOMOT, J.-F. SOUSSANA, C. TEICHMANN, R. VALENTINI, R. VAUTARD, B. WEBER, P. YIOU, 2014: EURO-CORDEX: new high-resolution climate change projections for European impact research. *Regional Environmental Change* **14**(2), 563–578, doi: 10.1007/s10113-013-0499-2, https://doi.org/10.1007/s10113-013-0499-2.

JENDRITZKY, G., DE R. DEAR, G. HAVENITH, 2012: UTCI - Why another thermal index?. *International Journal of Biometeorology* **56**, 421–428, doi: 10.1007/s00484-011-0513-7, https://doi.org/10.1007/s00484-011-0513-7.

JUNGCLAUS, J. H., S. J. LORENZ, H. SCHMIDT, V. BROVKIN, N. BRÜGGEMANN, F. CHEGINI, T. CRÜGER, P. DE-VRESE, V. GAYLER, M. A. GIORGETTA, O. GUTJAHR, H. HAAK, S. HAGEMANN, M. HANKE, T. ILYINA, P. KORN, J. KRÖGER, L. LINARDAKIS, C. MEHLMANN, U. MIKOLAJEWICZ, W. A. MÜLLER, J. E. M. S. NABEL, D. NOTZ, H. POHLMANN, D. A. PUTRASAHAN, T. RADDATZ, L. RAMME, R. REDLER, C. H. REICK, T. RIDDICK, T. SAM, R. SCHNECK, R. SCHNUR, M. SCHUPFNER, VON J.-S. STORCH, F. WACHSMANN, K.-H. WIENERS, F. ZIEMEN, B. STEVENS, J. MAROTZKE, M. CLAUSSEN, 2022: The ICON Earth System Model Version 1.0. *Journal of Advances in Modeling Earth Systems* **14**(4), e2021MS002813, doi: 10.1029/2021MS002813, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021MS002813.

KATRAGKOU, E., M. GARCÍA-DÍEZ, R. VAUTARD, S. SOBOLOWSKI, P. ZANIS, G. ALEXANDRI, R. M. CARDOSO, A. COLETTE, J. FERNANDEZ, A. GOBIET, K. GOERGEN, T. KARACOSTAS, S. KNIST, S. MAYER, P. M. M. SOARES, I. PYTHAROULIS, I. TEGOULIAS, A. TSIKERDEKIS, D. JACOB, 2015: Regional climate hindcast simulations within EURO-CORDEX: evaluation of a WRF multi-physics ensemble. *Geoscientific Model Development* **8**(3), 603–618, doi: 10.5194/gmd-8-603-2015, https://doi.org/10.5194/gmd-8-603-2015.

KAUTZ, L.-A., O. MARTIUS, S. PFAHL, J. G. PINTO, A. M. RAMOS, P. M. SOUSA, T. WOOLLINGS, 2022: Atmospheric blocking and weather extremes over the Euro-Atlantic sector – a review. *Weather and Climate Dynamics* **3**(1), 305–336, doi: 10.5194/wcd-3-305-2022, https://doi.org/10.5194/wcd-3-305-2022.

KNUTTI, R., R. FURRER, C. TEBALDI, J. CERMAK, G. A. MEEHL, 2010: Challenges in Combining Projections from Multiple Climate Models. *Journal of Climate* **23**, 2739–2758, doi: 10.1175/2009JCLI3361.1, https://doi.org/10.1175/2009JCLI3361.1.

KOTLARSKI, S., K. KEULER, O. B. CHRISTENSEN, A. COLETTE, M. DÉQUÉ, A. GOBIET, K. GOERGEN, D. JACOB, D. LÜTHI, VAN E. MEIJGAARD, G. NIKULIN, C. SCHÄR, C. TEICHMANN, R. VAUTARD, K. WARRACH-SAGI, V. WULFMEYER, 2014: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geoscientific Model Development* **7**(4), 1297–1333, doi: 10.5194/gmd-7-1297-2014, https://doi.org/10.5194/gmd-7-1297-2014.

KOTTEK, M., J. GRIESER, C. BECK, B. RUDOLF, F. RUBEL, 2006: World Map of the Köppen?Geiger climate classification updated. *Meteorologische Zeitschrift* **15**(3), 259–263, doi: 10.1127/0941?2948/2006/0130, http://dx.doi.org/10.1127/0941?2948/2006/0130.

KUKAL, M., S. IRMAK, 2018: U.S. Agro-Climate in 20th Century: Growing Degree Days, First and Last Frost, Growing Season Length, and Impacts on Crop Yields. *Scientific Reports* **8**, 6977, doi: 10.1038/s41598-018-25212-2, https://doi.org/10.1038/s41598-018-25212-2.

KUČEROVÁ, M., C. BECK, A. PHILIPP, R. HUTH, 2017: Trends in frequency and persistence of atmospheric circulation types over Europe derived from a multitude of classifications. *International Journal of Climatology* **37**(5), 2502–2521, doi: 10.1002/joc.4861, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.4861.

LEI, J., R. CHEN, P. YIN, X. MENG, L. ZHANG, C. LIU, Y. QUI, J. S. JI, H. KAN, M. ZHOU, 2022: Association between Cold Spells and Mortality Risk and Burden: A Nationwide Study in China. *Environmental health perspectives* **130**(2), 27006, doi: 10.1289/EHP9284, https://doi.org/10.1289/EHP9284.

LHOTKA, O., J. KYSELÝ, 2018: Circulation-Conditioned Wintertime Temperature Bias in EURO-CORDEX Regional Climate Models Over Central Europe. *Journal of Geophysical Research: Atmospheres* **123**(16), 8661–8673, doi: 10.1029/2018JD028503, https://doi.org/10.1029/2018JD028503.

MARONGA, B., S. BANZHAF, C. BURMEISTER, T. ESCH, R. FORKEL, D. FRÖHLICH, V. FUKA, K. F. GEHRKE, J. GELETIČ, S. GIERSCH, T. GRONEMEIER, G. GROSS, W. HELDENS, A. HELLSTEN, F. HOFFMANN, A. INAGAKI, E. KADASCH, F. KANANI-SÜHRING, K. KETELSEN, B. A. KHAN, C. KNIGGE, H. KNOOP, P. KRČ, M. KURPPA, H. MAAMARI, A. MATZARAKIS, M. MAUDER, M. PALLASCH, D. PAVLIK, J. PFAFFEROTT, J. RESLER, S. RISSMANN, E. RUSSO, M. SALIM, M. SCHREMPF, J. SCHWENKEL, G. SECKMEYER, S. SCHUBERT, M. SÜHRING, VON R. TILS, L. VOLLMER, S. WARD, B. WITHA, H. WURPS, J. ZEIDLER, S. RAASCH, 2020: Overview of the PALM model system 6.0. *Geoscientific Model Development* **13**(3), 1335–1372, doi: 10.5194/gmd-13-1335-2020, https://gmd.copernicus.org/articles/13/1335/2020/.

MARSHALL, J., Y. KUSHNIR, D. BATTISTI, P. CHANG, A. CZAJA, R. DICKSON, J. HURRELL, M. MC-CARTNEY, R. SARAVANAN, M. VISBECK, 2001: North Atlantic climate variability: phenomena, impacts and mechanisms. *International Journal of Climatology* **21**(15), 1863–1898, doi: 10.1002/joc.693, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.693.

MCCUSKER, K. E., J. C. FYFE, M. SIGMOND, 2016: Twenty-five winters of unexpected Eurasian cooling unlikely due to Arctic sea-ice loss. *Nature Geoscience* **9**, 838–842, doi: 10.1038/ngeo2820, https://doi.org/10.1038/ngeo2820.

MENG, C., F. KE, Y. XIAO, S. HUANG, Y. DUAN, G. LIU, S. YU, Y. FU, J. PENG, J. CHENG, P. YIN, 2022: Effect of Cold Spells and Their Different Definitions on Mortality in Shenzhen, China. *Frontiers in Public Health* **9**, doi: 10.3389/fpubh.2021.817079, https://www.frontiersin.org/articles/10.3389/fpubh.2021.817079.

MESSORI, G., R. CABALLERO, D. FARANDA, 2017: A dynamical systems approach to studying midlatitude weather extremes. *Geophysical Research Letters* **44**(7), 3346–3356, doi: 10.1002/2017GL072879, https://doi.org/10.1002/2017GL072879.

MIECZKOWSKI, Z., 1985: The Tourism Climatic Index: A method of evaluating world climates for tourism. *Canadian Geographer* **29**, 220–233, doi: 10.1111/j.1541-0064.1985.tb00365.x, https://doi.org/10.1111/j.1541-0064.1985.tb00365.x.

MITTERMEIER, M., M. WEIGERT, D. RÜGAMER, H. KÜCHENHOFF, R. LUDWIG, 2022: A deep learning based classification of atmospheric circulation types over Europe: projection of future changes in a CMIP6 large ensemble. *Environmental Research Letters* **17**(8), 084021, doi: 10.1088/1748-9326/ac8068, https://dx.doi.org/10.1088/1748-9326/ac8068.

NOAA/NATIONAL WEATHER SERVICE, CLIMATE PREDICTION CENTER Teleconnections. https://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/teleconnections.shtml, last accessed 06.02.2024.

OZTURK, T., D. MATTE, J. H. CHRISTENSEN, 2022: Robustness of future atmospheric circulation changes over the EURO-CORDEX domain. *Climate Dynamics* **59**, 1799–1814, doi: 10.1007/s00382-021-06069-0, https://doi.org/10.1007/s00382-021-06069-0.

PFAHL, S., 2014: Characterising the relationship between weather extremes in Europe and synoptic circulation features. *Natural Hazards and Earth System Sciences* **14**(6), 1461–1475, doi: 10.5194/nhess-14-1461-2014, https://doi.org/10.5194/nhess-14-1461-2014.

PHILIPP, A., J. BARTHOLY, C. BECK, M. ERPICUM, P. ESTEBAN, X. FETTWEIS, R. HUTH, P. JAMES, S. JOURDAIN, F. KREIENKAMP, T. KRENNERT, S. LYKOUDIS, S. C. MICHALIDES, K. PIANKO-KLUCZYNSKA, P. POST, D. R. ÁLVAREZ, R. SCHIEMANN, A. SPEKAT, F. S. TYMVIOS, 2010: Cost733cat – A database of weather and circulation type classifications. *Physics and Chemistry of the Earth, Parts A/B/C* **35**(9), 360–373, doi: 10.1016/j.pce.2009.12.010, https://www.sciencedirect.com/science/article/pii/S1474706510000045.

RANASINGHE, R., A. RUANE, R. VAUTARD, N. ARNELL, E. COPPOLA, F. CRUZ, S. DESSAI, A. ISLAM, M. RAHIMI, D. RUIZ CARRASCAL, J. SILLMANN, M. SYLLA, C. TEBALDI, W. WANG, R. ZAABOUL, 2021: Climate Change Information for Regional Impact and for Risk Assessment. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, Chapter 12, 1767–1926 doi: 10.1017/9781009157896.014, https://www.ipcc.ch/report/ar6/wg1/chapter/chapter-12/.

REX, D. F., 1950: Blocking Action in the Middle Troposphere and its Effect upon Regional Climate. *Tellus* **2**(3), 196–211, doi: 10.1111/j.2153-3490.1950.tb00331.x, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2153-3490.1950.tb00331.x.

RIAHI, K., S. RAO, V. KREY, C. CHO, V. CHIRKOV, G. FISCHER, G. KINDERMANN, N. NAKICENOVIC, P. RAFAJ, 2011: RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change* **109**(1), 33–57, doi: 10.1007/s10584-011-0149-y, https://doi.org/10.1007/s10584-011-0149-y.

RICHARD, L., H. GOLDING, R. SASKIN, J. I. R. JENKINSON, K. F. PRIDHAM, E. GOGOIS, C. SNIDER, S. W. HWANG, 2023: Cold-related injuries among patients experiencing homelessness in Toronto: a descriptive analysis of emergency department visits. *Canadian Journal of Emergency Medicine* **25**, 695–703, doi: 10.1007/s43678-023-00546-7, https://doi.org/10.1007/s43678-023-00546-7.

ROCKEL, B., A. WILL, A. HENSE, 2008: The Regional Climate Model COSMO-CLM (CCLM). *Meteorologische Zeitschrift* **17**(4), 347–348, doi: 10.1127/0941-2948/2008/0309, http://dx.doi.org/10.1127/0941-2948/2008/0309.

RYTI, N. R., Y. GUO, J. J. JAAKKOLA, 2016: Global Association of Cold Spells and Adverse Health Effects: A Systematic Review and Meta-Analysis. *Environmental health perspectives* **124**(1), 12—22, doi: doi.org/10.1289/ehp.1408104, https://doi.org/10.1289/ehp.1408104.

SALIM, M. H., K. H. SCHLÜNZEN, D. GRAWE, M. BOETTCHER, A. M. U. GIERISCH, B. H. FOCK, 2018: The microscale obstacle-resolving meteorological model MITRAS v2.0: model theory. *Geoscientific Model Development* **11**(8), 3427–3445, doi: 10.5194/gmd-11-3427-2018, https://gmd.copernicus.org/articles/11/3427/2018/.

SAMUELSSON, P., C. G. JONES, U. WILLÉN, A. ULLERSTIG, S. GOLLVIK, U. HANSSON, E. JANSSON, C. KJELLSTRÖM, G. NIKULIN, K. WYSER, 2011: The Rossby Centre Regional Climate model RCA3:

model description and performance. *Tellus A: Dynamic Meteorology and Oceanography* **63**(1), 4–23, doi: 10.1111/j.1600-0870.2010.00478.x, https://doi.org/10.1111/j.1600-0870.2010.00478.x.

SCAIFE, A. A., T. WOOLLINGS, J. KNIGHT, G. MARTIN, T. HINTON, 2010: Atmospheric Blocking and Mean Biases in Climate Models. *Journal of Climate* **23**(23), 6143–6152, doi: 10.1175/2010JCLI3728.1, https://journals.ametsoc.org/view/journals/clim/23/23/2010jcli3728.1.xml.

SCAIFE, A. A., A. ARRIBAS, E. BLOCKLEY, A. BROOKSHAW, R. T. CLARK, N. DUNSTONE, R. EADE, D. FEREDAY, C. K. FOLLAND, M. GORDON, L. HERMANSON, J. R. KNIGHT, D. J. LEA, C. MACLACHLAN, A. MAIDENS, M. MARTIN, A. K. PETERSON, D. SMITH, M. VELLINGA, E. WALLACE, J. WATERS, A. WILLIAMS, 2014: Skillful long-range prediction of European and North American winters. *Geophysical Research Letters* **41**(7), 2514–2519, doi: 10.1002/2014GL059637, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014GL059637.

SCHLÜNZEN, K. H., P. HOFFMANN, G. ROSENHAGEN, W. RIECKE, 2010: Long-term changes and regional differences in temperature and precipitation in the metropolitan area of Hamburg. *International Journal of Climatology* **30**, 1121–1136, doi: 10.1002/joc.1968, https://doi.org/10.1002/joc.1968.

SCHLÜNZEN, K. H., W. RIECKE, B. BECHTEL, M. BOETTCHER, S. BUCHHOLZ, D. GRAWE, P. HOFFMANN, R. PETRIK, R. SCHOETTER, K. TRUSILOVA, S. WIESNER, 2018: Stadtklima in Hamburg. In: H. VON STORCH, I. MEINKE, and M. CLAUSSEN (Eds.), Hamburger Klimabericht - Wissen über Klima, Klimawandel und Auswirkungen in Hamburg und Norddeutschland, Springer Spektrum, 1 edition, Chapter 3, 37–53 doi: 10.1007/978-3-662-55379-4, https://doi.org/10.1007/978-3-662-55379-4.

SCHLÜNZEN, K. H., D. GRAWE, S. I. BOHNENSTENGEL, I. SCHLÜTER, R. KOPPMANN, 2011: Joint modelling of obstacle induced and mesoscale changes — Current limits and challenges. *Journal of Wind Engineering and Industrial Aerodynamics* **99**(4), 217–225, doi: 10.1016/j.jweia.2011.01.009, https://www.sciencedirect.com/science/article/pii/S0167610511000110.

SCHROEDER, G., K. H. SCHLÜNZEN, 2009: Numerical Dispersion of Gravity Waves. *Monthly Weather Review* **137**(12), 4344–4354, doi: 10.1175/2009MWR2890.1, https://journals.ametsoc.org/view/journals/mwre/137/12/2009mwr2890.1.xml.

SCHULZWEIDA, U., 2023: CDO User Guide, version 2.3.0. doi: 10.5281/zenodo.10020800, https://doi.org/10.5281/zenodo.10020800.

SENEVIRATNE, S., X. ZHANG, M. ADNAN, W. BADI, C. DERECZYNSKI, A. DI LUCA, S. GHOSH, I. ISKANDAR, J. KOSSIN, S. LEWIS, F. OTTO, I. PINTO, M. SATOH, S. VICENTE-SERRANO, M. WEHNER, B. ZHOU, 2021: Weather and Climate Extreme Events in a Changing Climate. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, Chapter 11, 1513–1766 doi: 10.1017/9781009157896.013, https://www.ipcc.ch/report/ar6/wg1/chapter/chapter-11/.

SINGH, M., N. ACHARYA, S. JAMSHIDI, J. JIAO, Z.-L. YANG, M. COUDERT, Z. BAUMER, D. NIYOGI, 2023: DownScaleBench for developing and applying a deep learning based urban climate downscaling-first results for high-resolution urban precipitation climatology over Austin, Texas. *Computational Urban Science* **3**, 22, doi: 10.1007/s43762-023-00096-9, https://doi.org/10.1007/s43762-023-00096-9.

SKAMAROCK, W. C., 2004: Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra. *Monthly Weather Review* **132**(12), 3019–3032, doi: 10.1175/MWR2830.1, https://journals.ametsoc.org/view/journals/mwre/132/12/mwr2830.1.xml.

SKAMAROCK, W. C., J. B. KLEMP, J. DUDHIA, D. O. GILL, D. BARKER, M. G. DUDA, YU X. HUANG, W. WANG, J. G. POWERS, 2008: A Description of the Advanced Research WRF Version 3 .

SMITH, D. M., R. EADE, M. B. ANDREWS, H. AYRES, A. CLARK, S. CHRIPKO, C. DESER, N. J. DUNSTONE, J. GARCÍA-SERRANO, G. GASTINEAU, L. S. GRAFF, S. C. HARDIMAN, B. HE, L. HERMANSON, T. JUNG, J. KNIGHT, X. LEVINE, G. MAGNUSDOTTIR, E. MANZINI, D. MATEI, M. MORI, R. MSADEK, P. ORTEGA, Y. PEINGS, A. A. SCAIFE, J. A. SCREEN, M. SEABROOK, T. SEMMLER, M. SIGMOND, J. STREFFING, L. SUN, A. WALSH, 2022: Robust but weak winter atmospheric circulation response to future Arctic sea ice loss. *Nature Communications* **13**, 727, doi: 10.1038/s41467-022-28283-y, https://doi.org/10.1038/s41467-022-28283-y.

STENSETH, N. C., G. OTTERSEN, J. W. HURRELL, A. MYSTERUD, M. LIMA, K.-S. CHAN, N. G. YOCCOZ, B. ÅDLANDSVIK, 2003: Studying climate effects on ecology through the use of climate indices: the North Atlantic Oscillation, El Niño Southern Oscillation and beyond. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 2087–2096, doi: 10.1098/rspb.2003.2415, https://doi.org/10.1098/rspb.2003.2415.

STUIVENVOLT-ALLEN, J., S. S. WANG, 2019: Data Mining Climate Variability as an Indicator of U.S. Natural Gas. *Frontiers in big data* **2**(20), doi: 10.3389/fdata.2019.00020, https://doi.org/10.3389/fdata.2019.00020.

SUN, Y., K. DENG, K. REN, J. LIU, C. DENG, Y. JIN, 2024: Deep learning in statistical downscaling for deriving high spatial resolution gridded meteorological data: A systematic review. *ISPRS Journal of Photogrammetry and Remote Sensing* **208**, 14–38, doi: 10.1016/j.isprsjprs.2023.12.011, https://www.sciencedirect.com/science/article/pii/S0924271623003489.

TAYLOR, K. E., R. J. STOUFFER, G. A. MEEHL, 2012: An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society* **93**(4), 485–498, doi: 10.1175/BAMS-D-11-00094.1, https://doi.org/10.1175/BAMS-D-11-00094.1.

TEBALDI, C., R. KNUTTI, 2007: The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **365**(1857), 2053–2075, doi: 10.1098/rsta.2007.2076, https://doi.org/10.1098/rsta.2007.2076.

UNDERWOOD, B. S., Z. GUIDO, P. GUDIPUDI, Y. FEINBERG, 2017: Increased costs to US pavement infrastructure from future temperature rise. *Nature Climate Change* **7**, 704–707, doi: 10.1038/nclimate3390, https://doi.org/10.1038/nclimate3390.

UNITED NATIONS (UN), DEPARTMENT OF ECONOMIC AND SOCIAL AFFAIRS, POPULATION DIVISION, 2019: World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420). New York: United Nations.

VAN DER WALT, S., J. L. SCHÖNBERGER, J. NUNEZ-IGLESIAS, F. BOULOGNE, J. D. WARNER, N. YAGER, E. GOUILLART, T. YU, THE SCIKIT-IMAGE CONTRIBUTORS, 2014: scikit-image: image processing in Python. *PeerJ* **2**, e453, doi: 10.7717/peerj.453, https://doi.org/10.7717/peerj.453.

VAN DER WIEL, K., H. C. BLOOMFIELD, R. W. LEE, L. P. STOOP, B. BLACKPORT, J. A. SCREEN, F. M. SELTEN, 2019: The influence of weather regimes on European renewable energy production and demand. *Environmental Research Letters* **14**(9), 094010, doi: 10.1088/1748-9326/ab38d3, https://doi.org/10.1088/1748-9326/ab38d3.

VAN MEIJGAARD, E., VAN L. H. ULFT, VAN DE W. J. BERG, F. C. BOSVELD, VAN DEN B. J. J. M. HURK, G. LENDERINK, A. P. SIEBESMA, 2008: The KNMI regional atmospheric climate model RACMO, version 2.1.

VAN ULDEN, A. P., VAN G. J. OLDENBORGH, 2006: Large-scale atmospheric circulation biases and changes in global climate model simulations and their importance for climate change in Central Europe. *Atmospheric Chemistry and Physics* **6**(4), 863–881, doi: 10.5194/acp-6-863-2006, https://doi.org/10.5194/acp-6-863-2006.

VAN VUUREN, D. P., E. STEHFEST, DEN M. G. J. ELZEN, T. KRAM, VAN J. VLIET, S. DEETMAN, M. ISAAC, K. KLEIN GOLDEWIJK, A. HOF, A. MENDOZA BELTRAN, R. OOSTENRIJK, VAN B. RUIJVEN, 2011: RCP2.6: exploring the possibility to keep global mean temperature increase below 2°C. *Climatic Change* **109**(1), 95–116, doi: 10.1007/s10584-011-0152-3, https://doi.org/10.1007/s10584-011-0152-3.

VIHMA, T., R. GRAVERSEN, L. CHEN, D. HANDORF, N. SKIFIC, J. A. FRANCIS, N. TYRRELL, R. HALL, E. HANNA, P. UOTILA, K. DETHLOFF, A. Y. KARPECHKO, H. BJÖRNSSON, J. E. OVERLAND, 2020: Effects of the tropospheric large-scale circulation on European winter temperatures during the period of amplified Arctic warming. *International Journal of Climatology* **40**(1), 509–529, doi: 10.1002/joc.6225, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.6225.

VON STORCH, H., I. MEINKE, M. CLAUSSEN, 2017: Hamburger Klimabericht – Springer Spektrum Berlin doi: https://doi.org/10.1007/978-3-662-55379-4, https://link.springer.com/book/10.1007/978-3-662-55379-4.

WANG, L., T. LIU, M. HU, W. ZENG, Y. ZHANG, S. RUTHERFORD, H. LIN, J. XIAO, P. YIN, J. LIU, C. CHU, S. TONG, W. MA, M. ZHOU, 2016: The impact of cold spells on mortality and effect modification by cold spell characteristics. *Scientific Reports* **6**, 38380, doi: 10.1038/srep38380, https://doi.org/10.1038/srep38380.

WANG, Z., A. C. BOVIK, H. R. SHEIKH, E. P. SIMONCELLI, 2004: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **13**(4), 600–612, doi: 10.1109/tip.2003.819861, https://doi.org/10.1109/tip.2003.819861.

WEILNHAMMER, V., J. SCHMID, I. MITTERMEIER, F. SCHREIBER, L. JIANG, V. PASTUHOVIC, C. HERR, S. HEINZE, 2021: Extreme weather events in europe and their health consequences – A systematic review. *International Journal of Hygiene and Environmental Health* **233**, 113688, doi: 10.1016/j.ijheh.2021.113688, https://www.sciencedirect.com/science/article/pii/S1438463921000018.

WILKS, D. S., 2019: Chapter 9 - Forecast Verification. In: Statistical Methods in the Atmospheric Sciences (Fourth Edition), Elsevier, 369–483 doi: 10.1016/B978-0-12-815823-4.00009-2, https://doi.org/10.1016/B978-0-12-815823-4.00009-2.

WORLD METEOROLOGICAL ORGANIZATION (WMO), 2022: State of the Climate in Europe 2022. In: Climate Reports, WMO-No. 1320 https://library.wmo.int/index.php?lvl=notice_display&id=22285.

YE, K., T. WOOLLINGS, S. N. SPARROW, P. A. G. WATSON, J. A. SCREEN, 2024: Response of winter climate and extreme weather to projected Arctic sea-ice loss in very large-ensemble climate model simulations. *npj Climate and Atmospheric Science* **7**, 20, doi: 10.1038/s41612-023-00562-5, https://doi.org/10.1038/s41612-023-00562-5.

ZUMWALD, M., B. KNÜSEL, C. BAUMBERGER, G. HIRSH HADORN, D. BRESCH, R. KNUTTI, 2020: Understanding and assessing uncertainty of observational climate datasets for model evaluation using ensembles. *WIREs Climate Change* **e654**, doi: 10.1002/wcc.654, https://doi.org/10.1002/wcc.654.
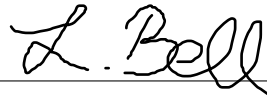
Versicherung an Eides statt

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, 02.12.2024

Datum, Ort

Unterschrift