

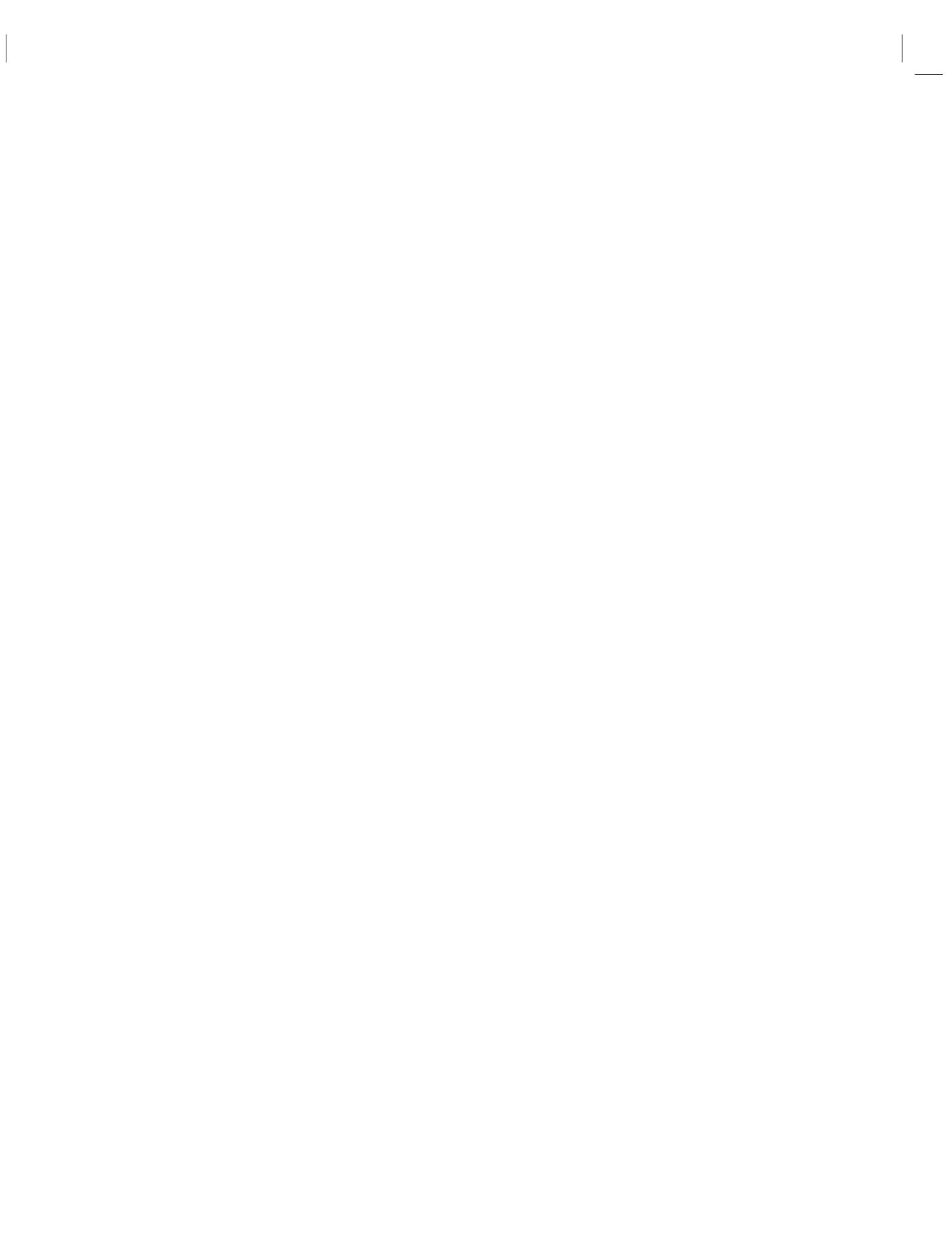
Helmut Küchenhoff, Henri Funk

Climate And Statistics



Contents

Preface	v
Foreword	1
1 Introduction	3
2 Natural Variability by internal variability	5
3 Standard Precipitation Evapotranspiration Index	17
4 Compound events	27
5 Asymmetric bivariate copulas for compounds	29
6 Teleconnections North Atlantic Oscillation	41
7 Low Flow Events	51
8 Statistical streamflow modelling	77
9 The Lancet Report 2023	91
10 Epidemiologic studies on the heat effects	101
11 Controversial issue : heat and humidity	117
12 Risk Projections	127
13 Climate Crisis and Mental Health	129
14 References	137
15 Open issue	139
16 Acknowledgements	141



Preface

Author: Henri Funk



As the world faces the reality of climate change, natural hazards and extreme weather events have become a major concern, with devastating consequences for nature and humans. The quantification and definition of climate change, extreme events and its implications for life and health on our planet is one of the major concerns in climate science.

This book explains current statistical methods in climate science and their application. The methods include compound events, low flow events and return periods, natural variability, teleconnections and causal discovery. All of those methods are used to quantify and anticipate the changing climate.

This book is the outcome of the seminar “Climate and Statistics” which took place in summer 2024 at the Department of Statistics, LMU Munich.



FIGURE 1: Creative Commons License

This book is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License¹.

¹<http://creativecommons.org/licenses/by-nc-sa/4.0/>



Foreword

Author: Christoph Molnar

This book is the result of an experiment in university teaching. Each semester, students of the Statistics Master can choose from a selection of seminar topics. Usually, every student in the seminar chooses a scientific paper, gives a talk about the paper and summarizes it in the form of a seminar paper. The supervisors help the students, they listen to the talks, read the seminar papers, grade the work and then ... hide the seminar papers away in (digital) drawers. This seemed wasteful to us, given the huge amount of effort the students usually invest in seminars. An idea was born: Why not create a book with a website as the outcome of the seminar? Something that will last at least a few years after the end of the semester. In the summer term 2019, some Statistics Master students signed up for our seminar entitled “Limitations of Interpretable Machine Learning”. When they came to the kick-off meeting, they had no idea that they would write a book by the end of the semester.

We were bound by the examination rules for conducting the seminar, but otherwise we could deviate from the traditional format. We deviated in several ways:

1. Each student project is part of a book, and not an isolated seminar paper.
 2. We gave challenges to the students, instead of papers. The challenge was to investigate a specific limitation of interpretable machine learning methods.
 3. We designed the work to live beyond the seminar.
 4. We emphasized collaboration. Students wrote some chapters in teams and reviewed each others texts.
-

Technical Setup

The book chapters are written in the Markdown language. The simulations, data examples and visualizations were created with R (?). To combine R-code and Markdown, we used rmarkdown. The book was compiled with the bookdown package. We collaborated using git and github. For details, head over to the book’s repository².

²https://github.com/henrifnk/Seminar_ClimateNStatistics



1

Introduction

Author:

Supervisor:

1.1 Intro About the Seminar Topic

1.2 Outline of the Booklet



2

Natural Variability by internal variability

Author: Senta Roßmayer

Supervisor: Henri Funk

Suggested degree: Bachelor

2.1 Abstract

Natural variability refers to the inherent fluctuations in the climate system that occur without external forcings, such as changes in solar radiation, volcanic eruptions, or human-induced alterations of the Earth's atmosphere and surface. This variability can be due to a variety of factors, including atmospheric processes, ocean currents, the El Niño-Southern Oscillation (ENSO), and other dynamic components of the Earth system. Natural variability occurs across all time scales, from short-term (daily, seasonal) to long-term (decadal, centennial) fluctuations.

2.2 Climate Model Ensemble

Climate models are sophisticated tools that simulate the interactions within the Earth's climate system. To understand and quantify natural variability, scientists use ensembles of climate model simulations. An ensemble consists of multiple runs of the same model, or different models, where each run has slightly different initial conditions or model parameters. This approach helps to capture the range of possible climate outcomes due to the inherent uncertainty and variability in the system.

2.3 Internal Variability

Within the context of climate model ensembles, internal variability refers to the variations in climate that arise from the system's internal processes, independent of external forcings. This variability is a fundamental part of the climate system's dynamics and can lead to different outcomes even if the external conditions are the same.

2.4 1. Introduction to climate modelling and climate variability

The Earth's climate is a complex system that is characterised by diverse and interrelated physical, chemical and biological processes. Climate models play a central role in analysing and predicting this system. They enable scientists to simulate and analyse the reactions of the Earth's atmosphere, oceans and land surfaces to various influences. By using them, past climate changes can be reconstructed and future climate trends and extremes can be predicted. The fundamentals of climate modelling include understanding the basic physical principles that regulate the exchange and distribution of energy and mass within the climate system. These models are based on mathematical representations of physical processes and use complex algorithms to calculate these representations. These calculations are performed repeatedly to simulate the dynamic changes of the climate system over time (?). A sound knowledge of the physical processes, such as radiative transfer, turbulent fluxes and the dynamics of the atmosphere and oceans, is essential for the development and interpretation of these models (?). Another important aspect of climate modelling is the consideration of climate variability. Climate variability refers to natural fluctuations in the climate system that can occur on different time scales, from a few years to centuries or even millennia. This variability can have various causes, including internal processes such as the North Atlantic Oscillation or global phenomena such as El Niño (?). It is crucial to distinguish between short-term climatic fluctuations and long-term trends in order to make more accurate climate projections. Various types of climate models are used in research, which differ in their complexity and scope of application. These range from simple energy-equilibrium models that consider basic climate variables to highly complex coupled atmosphere-ocean models that consider a variety of climate components and interactions (?). The Coupled Model Intercomparison Project Phase 6 (CMIP6) is a prominent example of such a coordinated modelling initiative that aims to overcome the challenges of climate modelling and improve the accuracy of climate predictions (?). Analysing and interpreting the results of these climate models requires a deep understanding of the uncertainties and accumulated sources of error in the projections. In particular, internal variability plays a crucial role in the overall uncertainty of climate predictions. These uncertainties need to be quantified through extensive ensemble simulations covering a wide range of scenarios and model parameters (?). To summarise, climate models and the study of climate variability are essential for understanding and predicting climate change. The continuous improvement of these models and their methods is of crucial importance in order to effectively meet the challenges of global climate change.

2.5 1.1 Basics of climate models

Climate models are complex mathematical representations of the Earth's climate system that are used to simulate the interactions between the atmosphere, ocean, land surface and ice. These models play a crucial role in predicting future climate changes and understanding historical climate patterns. A fundamental climate model is based on physical laws described by differential equations such as those found in radiative transfer, fluid dynamics and thermodynamics (?). The main components of a climate model include the atmosphere, the ocean, the cryosphere and the biosphere. The model takes into account processes such as energy exchange between the Earth's surface and atmosphere, ocean currents, ice cover and feedback effects. The models are divided into different horizontal and vertical grid cells, which represent the physical state variables such as temperature, pressure and humidity. Each cell interacts with neighbouring cells, resulting in very complex three-dimensional simulations. Validation is a key aspect in the development and utilisation of climate models. This is done by comparing the model results with historical observations. If past climate conditions are correctly reproduced, this strengthens confidence in the model's ability to predict future states of the climate system (?). Uncertainties play an important role here. A large part of this uncertainty stems from the internal variability of the climate system, i.e. the natural fluctuations that can occur even without

external influences. The accuracy of a climate model is also determined by the spatial resolution and the quality of the underlying physical parameterisations. Higher-resolution models can better capture smaller-scale phenomena, but are also more computationally intensive. One challenge is that not all climate-relevant processes can be simulated directly. For example, cloud formation or soil moisture must be approximated using parameterisation methods, which leads to further uncertainties (?). Another important approach in climate modelling is the use of ensemble simulations. By carrying out several simulations with slightly varying initial conditions, scientists can estimate the range of possible climate developments and better quantify the uncertainties resulting from internal variability. These approaches help to make more robust statements about future climate trends and to better predict possible extreme events. To summarise, it can be said that climate models are indispensable tools for investigating the complex and dynamic climate system. Through their continuous development and validation, they greatly contribute to deepening our understanding of climate processes and making well-founded predictions for the future.

2.6 1.2 Types of climate models: A comparison

Climate models are essential tools for analysing and predicting climatic changes. The models differ primarily in their complexity and applicability. The basic types of climate models can be divided into simple energy balance models, one-dimensional models, two-dimensional models and three-dimensional general circulation models (GCMs). Simple energy balance models aim to model the Earth's radiative imbalance and are relatively easy to use. They are particularly useful for understanding basic climate change mechanisms and providing rapid estimates of temperature changes. However, their low spatial resolution is a limitation that can affect the accuracy of the results. Some of the more advanced models are the so-called one- and two-dimensional climate models. A one-dimensional model can represent vertical temperature layers in the atmosphere or ocean, while a two-dimensional model also includes horizontal components such as latitude or altitude. These models offer improved spatial resolution and take more physical processes into account than simple energy balance models. However, they are usually not suitable for detailed regional forecasts either. General circulation models (GCMs) represent the most complex and detailed tools for climate modelling. They simulate atmospheric and oceanic circulation patterns in three dimensions, taking into account a variety of physical, chemical and biological processes. The GCMs are often used in multi-model ensembles to reduce uncertainties in predictions and produce more robust climate projections (?). These models have proven particularly useful for predicting long-term climate changes and their impacts, such as changes in sea ice cover (?). In addition, regional climate models (RCMs) are used, which can depict specific climatic conditions in smaller geographical areas in greater detail. Due to their finer spatial resolution, these models are better suited to analysing local climate changes and projecting more specific scenarios (?). By downscaling the results of the GCMs to regional models, more precise statements on local climate impacts can be made (?). Another important aspect in the evaluation of climate models is the uncertainty caused by internal variability and structural differences in the models. Multi-model ensembles help to quantify these uncertainties and provide a sound basis for the assessment of climate impacts (?). However, the challenge remains that different models and approaches often arrive at different results, which makes the interpretation of the data complex. Overall, the different types of climate models offer a wide range of tools for analysing the climate system, and their combination can help to paint a more comprehensive picture of future climatic conditions

2.7 2. Introduction to climate model ensembles

Climate model ensembles are an essential tool in modern climate research. These ensembles enable scientists to better understand and quantify the uncertainties and variability in the climate system. An ensemble typically consists of a series of climate model simulations performed using different models or different initial conditions of the same model (?). This multitude of simulations helps to capture the range of possible future climate scenarios and to identify and separate both internal and external influences on the climate system.

2.8 2.1 Definition and meaning of climate model ensembles

Climate model ensembles are an essential method for improving the accuracy of predictions and quantifying uncertainties in climate projections. A climate model ensemble consists of several climate models or several runs of a single model under different assumptions and input data (?). The fundamental advantage of ensembles is that they make it possible to better capture the variability and uncertainty in climate projections by taking into account a range of possible scenarios and their probabilities. This allows for a more robust prediction, as it is not sufficient to rely on a single model or simulation (?). Analysing a large number of models also helps to identify and address systematic errors and structural uncertainties within the modelling (?). In a climate model ensemble, various models are used that differ in their structural assumptions, parameterisations and input data. By combining these models, error tendencies of individual models can be compensated for and a more robust overall picture can be created (?). This is particularly important as individual climate models can have systematic errors and uncertainties that lead to inconsistencies in the predictions. Another important feature of climate model ensembles is their application in the investigation of extreme events and the estimation of local climate changes. In particular, the use of regionalised climate models (RCMs) within an ensemble enables a more precise representation of climatic conditions at smaller spatial scales (?). These regional models provide more detailed projections, which are crucial for climate change adaptation strategies. The Coupled Model Intercomparison Project Phase 6 (CMIP6) is one of the most comprehensive initiatives for the creation and use of climate model ensembles by coordinating global climate models from various research institutions worldwide. CMIP6 addresses central scientific questions such as the response of the Earth system to external influences and the causes of systematic model deviations (?). The importance of climate model ensembles lies not only in the reduction of uncertainty, but also in the improvement of the physical realism of the predictions. By combining different climate models, a broader range of physical processes and interactions can be considered, leading to more realistic and robust predictions (?). In addition, ensembles enable probabilistic analysis, which provides decision-makers with important information about the probability of different climate scenarios. Another key aspect of using climate model ensembles is the possibility of assessing uncertainty. Uncertainties in climate predictions result both from natural climate variability and from uncertainties in the models themselves and future emission scenarios (?). By using ensembles, the uncertainty can be quantified in a systematic way and integrated into the decision-making process. Finally, climate model ensembles play a central role in the assessment of regional climate changes and their impacts. Since regional climate changes can vary greatly, it is crucial that the uncertainties in these projections are taken into account by combining several models (?). The use of climate model ensembles thus enables the creation of detailed and reliable climate projections, which are essential for planning and adapting to climate change.

2.9 2.2 Methods for calculating climate model ensembles

The creation of climate model ensembles is a complex and methodologically diverse process that reduces uncertainties in climate projections and enables more robust predictions. A basic method is ensemble averaging, in which several climate models are combined to obtain an average result. This method aims to minimise the deviations caused by individual model errors and provide a more stable overall picture of the future climate (?). Another important method is Bayesian model averaging (BMA), which allows probabilistic statements to be made about future climate states. The different models are weighted according to their performance in order to obtain a combined posterior distribution that takes into account the uncertainties of the individual models. This technique enables a differentiated view of possible future climate changes and is considered a substantial progress over earlier approaches (?). In addition, the Reliability Ensemble Average (REA) approach was developed to assess and weight the reliability of the individual models. In this method, the projections are assessed on the basis of their agreement with observed climate data and their convergence with each other, which enables a reliable estimate of future climate conditions. REA is particularly useful for regionalised climate modelling, as different models may have different strengths and weaknesses in different regions (?). In addition, multivariate analysis methods play a decisive role in the creation of climate model ensembles. These approaches make it possible to consider several climate variables simultaneously and to analyse the interactions between them. Such a method improves the accuracy and reliability of the projections by better accounting for the inherent uncertainties of the models (?). The cross-validation method is another important tool for validating and fine-tuning climate model ensembles. This involves comparing the predictions of the models with new, independent data sets to check the accuracy and assumptions of the statistical models. This method ensures that the climate models not only accurately reproduce the historical data, but can also provide reliable forecasts for the future (?). Finally, the consideration of internal climate variability is of central importance for the creation of climate ensembles. This variability, which is characterised by both atmospheric and oceanic circulation patterns, contributes substantially to the uncertainty in climate projections. Methods such as ensemble simulations, which run a large number of simulations with a single model, make it possible to capture both the forced response of the climate system and the internal variability (?). Overall, these methods represent different approaches to reducing uncertainties and improving the reliability of climate model ensembles. They form the basis for sound climate science findings and enable decision-makers to take more effective measures to adapt to future climate change.

2.10 2.3 Example of a climate model ensemble

The climate model ensemble from (?) in more detail. A coupled ocean-atmosphere-land-cryosphere general circulation model (CCSM3) is used. It realistically simulates the main patterns of climate variability, but shows a higher regularity and frequency for ENSO than in nature. The 40-member CCSM3 ensemble uses the T42 version with specific resolutions for the different components and considers the A1B greenhouse gas scenario. The initial conditions for ocean, land and sea ice are identical for each ensemble member and are from 1 January 2000, while the atmospheric conditions vary. In addition, a 10,000-year control integration of CAM3, the atmospheric component of CCSM3, is used. This integration takes into account recurring seasonal cycles without annual variability based on observations from 1980-2000. A separate CMIP3 multi-model ensemble was formed for the study, comprising a single integration of 21 models with the SRES-A1B forcing scenario, excluding CCSM3. The climate response was calculated using epoch differences and linear trends for the period 2005-2060. Both methods provide almost identical results. The statistical significance of the results was assessed using a two-tailed Student's t-test. The left panels of Figure 1a show the ensemble mean epoch difference maps (2051-2060 minus 2005-2014) for sea level pressure (SLP), precipitation (Precip) and surface temperature (TS) during December-January-February (DJF) for the 40-member CCSM3 ensemble. Hatching indicates areas where the differences are significant at a 95% confidence level. After explaining the

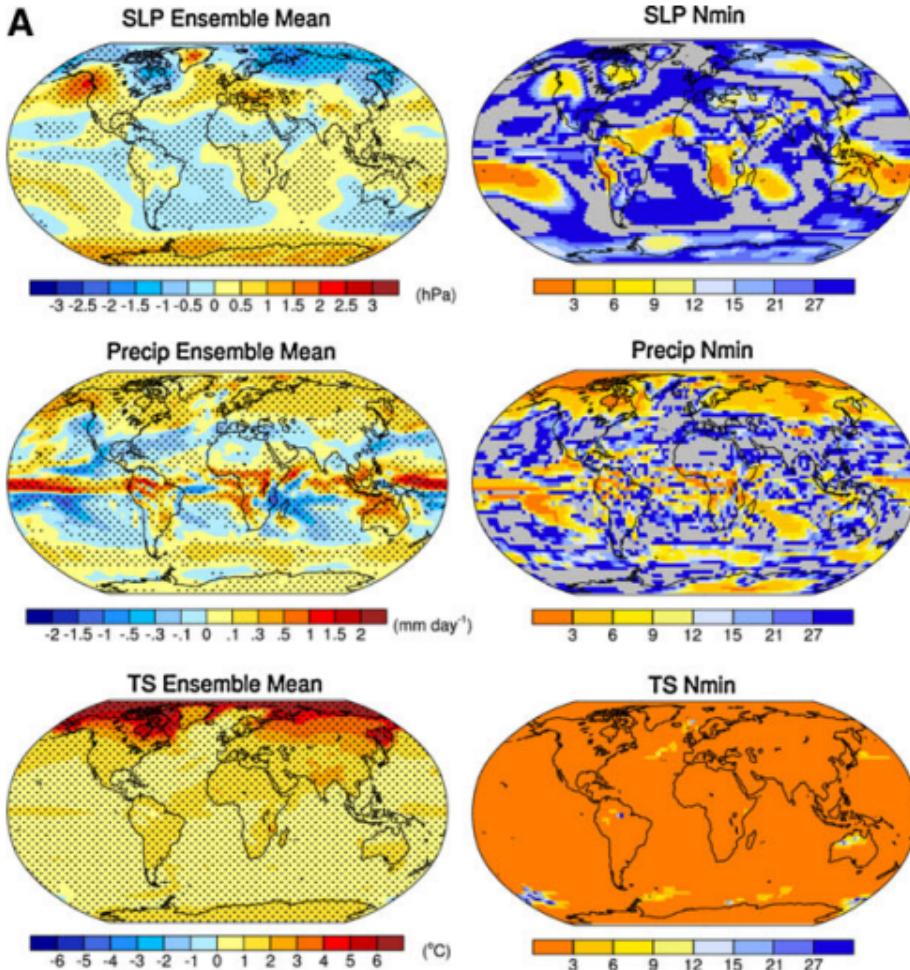


FIGURE 2.1: CCSM3 40-member ensemble (@deser)

model and the method, the results are now presented. The ensemble mean response is globally considerable for all three variables. Sea level pressure shows negative values at high latitudes in the northern hemisphere and positive values at mid-latitudes, similarly in the southern hemisphere, but with reversed polarity. These patterns correspond to the northern and southern annular modes (NAM and SAM). The global distribution of SLP changes is consistent with the results of 22 CMIP3 models. Precipitation shows positive values along the equator and negative values south of it. Subtropics generally have less precipitation, while extratropical regions have more precipitation. The surface temperature is rising everywhere, with stronger warming over land and maximum warming in the Arctic due to the loss of sea ice in autumn. The right panels of Figure 1a show the minimum number of ensemble members required to detect a significant response. Sea level pressure requires larger ensemble sizes than precip, and surface temperature requires the smallest ensemble sizes. Values for SLP range from less than 6 in the tropics to over 15 in other regions. Precip usually requires fewer members, while TS generally requires less than 3 members, except in some isolated regions. It can also be said that the grey areas indicate the locations where the mean response of the 40-member ensemble is not at the 95% confidence level.(?)

2.11 3. Internal variability in the climate system

The study of internal variability in the climate system is of central importance for understanding the complex interactions and fluctuations within the Earth's climate. Internal variability refers to naturally occurring fluctuations in the climate system that arise independently of external influences such as greenhouse gas emissions, volcanic eruptions or solar variations. This variability can manifest itself on different time scales, from monthly and seasonal fluctuations to decade and century scales. A prominent example of internal variability is the phenomenon of the North Atlantic Oscillation (NAO), which has a considerable impact on weather and climate conditions in Europe and North America. The NAO influences the temperature and precipitation patterns in these regions through its fluctuations, which in turn can be attributed to the internal dynamics of the circulation systems. The thermohaline circulation also plays a substantial role in internal variability, particularly in the North Atlantic region, where variations in the flow and the associated heat distribution have considerable climatic effects (?). The role of internal variability is particularly important for climate research and modelling, as it contributes to considerable uncertainties in climate projections. Models that do not adequately account for this internal variability tend to underestimate or overestimate the variability of the climate system. For example, studies show that the uncertainties in projections of ocean temperatures and ocean currents are influenced by internal variability (?). These uncertainties have a direct impact on the predictability of future climate trends and the planning of adaptation strategies. The fluctuations in surface temperature and ocean currents observed in the North Atlantic region illustrate the need for precise monitoring and modelling of internal variability processes. This is important in order to distinguish both natural and anthropogenic drivers of climate change and to develop appropriate climate adaptation measures. It is argued that ongoing observation programmes and high model resolutions are necessary to better understand and predict these complex processes (?). Internal variability can have both positive and negative effects, depending on the specific climate component and regional characteristics. For example, increased internal variability in the tropics can lead to changes in global atmospheric circulation, which in turn influence climate extremes such as droughts and floods (?). Understanding these internal processes is therefore crucial for improving the accuracy and reliability of climate models and thus also for developing effective climate change mitigation and adaptation strategies.

2.12 3.1 Definition and causes of internal variability

Internal variability in the climate system is a central element of climate research and describes fluctuations in the climate that are not caused by external influences such as volcanic eruptions or human activities, but arise from internal dynamics of the climate system itself. The main drivers of this internal variability are the complex interactions between the various components of the climate system, in particular between the atmosphere, oceans, cryosphere and biosphere. Internal variability is the natural variability of the climate system that occurs in the absence of external forcing and includes processes inherent to the atmosphere, ocean and coupled ocean-atmosphere system. A major cause of internal variability is the thermodynamic coupling between the atmosphere and the near-surface layers of the oceans. This coupling can cause long-lasting climate fluctuations that last for years to decades (?). A prominent example of such variability is the North Atlantic Oscillation (NAO), which is characterised by fluctuations in air pressure between the Icelandic low and the Azores high. These fluctuations have an influence on the weather and climate in the mid-latitudes of the North Atlantic (?). Another important factor is the wind-driven circulation systems of the oceans, such as the ocean gyre, which can contribute to the Pacific Decadal Oscillation and the Atlantic Multidecadal Oscillation. These processes can be influenced by changes in atmospheric circulation, such as different wind patterns and variations in air pressure (?). Within the North Atlantic, changes in the thermohaline circulation, also known as the Atlantic Meridional Overturning Circulation (AMOC), play a essential role. This circulation contributes to the distribution of heat within the ocean and is strongly

dependent on the salinity and temperature of the seawater. The natural fluctuations of the AMOC can have a considerable impact on the climate of the North Atlantic and lead to multidecadal variations (?). The internal variability of the climate system is therefore caused by a variety of mechanisms that operate on different spatial and temporal scales. These mechanisms are often stochastic in nature and can be described by internal dynamical processes such as interactions between ocean and atmospheric wind patterns (?). In contrast to internal variability, external variability is caused by external drivers; these external factors can influence the climate over longer periods of time. One example of this is changes in solar radiation or volcanic eruptions. In addition, human influences can also play a role, such as greenhouse gas emissions. To summarise, it can be said that the internal variability in the climate system represents a complex interplay of different factors that include both thermal and dynamic processes. This variability is fundamental to understanding the climate system in all its dynamics and results from the internal interactions of the various climate components. It forms an essential basis for better understanding the uncertainties in climate projections and thus also the planning of adaptation strategies for dealing with climate change.

2.13 3.2 Effects of internal variability on climate projections

The internal variability in the climate system represents a major uncertainty for climate projections. This variability results from natural processes within the climate system itself and is not caused by external forcing such as greenhouse gases or land-use changes. The internal variability manifests itself on different time scales, from seasonal to multi-decadal, which complicates the interpretation and prediction of future climate conditions. A key point that illustrates the uncertainties in climate projections is the role of atmospheric circulation. Studies have shown that variations in atmospheric circulation can contribute strongly to uncertainty in climate models (?). For example, differences in simulated atmospheric dynamics, such as the movement patterns of air masses, contribute essential to the variable amount of precipitation in the Sahel. These uncertainties remain in both older (CMIP5) and newer models (CMIP6) (?). Another example of the effects of internal variability is the Atlantic Meridional Overturning Current (AMOC). The AMOC is a large-scale ocean current that transports heat from the tropics to higher latitudes and thus has an influence on regional climate patterns. Studies have shown that the internal variability of the AMOC has been the dominant force since 1900 (?). Depending on its state, this internal variability can have short-term and long-term effects on the surface temperatures of the North Atlantic and the European climate. Internal variability is also increased by other factors such as the North Atlantic Oscillation (NAO) and the Atlantic Multidecadal Oscillation (AMO). Fluctuations in these systems can influence the European and North American climate and lead to seasonal to decadal variations in temperature and precipitation (?). Furthermore, internal variability complicates policymaking and decision-making as it leads to ranges in climate predictions. This makes accurate predictions difficult and emphasises the need for robust adaptation strategies that consider a variety of scenarios to minimise uncertainties (?). Finally, it can be stated that internal variability is an important, but often difficult to predict component in climate projections. It strongly influences the uncertainty range of model projections and therefore further studies are urgently needed to better understand the mechanisms of internal variability and to quantify its impact on climate projections more precisely.

2.14 4. Application and meaning of climate model ensembles in research

The use and importance of climate model ensembles in research has grown considerably in recent years. Climate model ensembles are a valuable method for assessing and minimising uncertainties in climate projections. These models make it possible to incorporate the internal variability of the climate system by performing multiple simulations with different initial conditions or model parameters. By using ensemble

techniques, the range of possible climate projections can be better understood and the uncertainties arising from internal variability and model specificities can be quantified (?). However, the role of climate model ensembles goes beyond the mere assessment of uncertainty. They also provide a robust basis for the assessment of climate dynamics under different scenarios and forcing conditions. Ensemble models facilitate the identification of systematic errors in models and enable the assessment of model performance by comparison with observations and other models (?). This comparability is particularly important to strengthen confidence in future climate projections and to check the reliability of the models. In addition, climate model ensembles offer the possibility of analysing the effects of climate change on different spatial and temporal scales. By considering ensembles on a regional scale, researchers can make more accurate predictions about local climate trends and extreme events. This is of particular importance for adaptation to climate change, as it enables decision-makers to develop well-founded and location-specific strategies. For example, modelling can be used to plan agricultural adaptation measures to changing climate conditions by assessing the uncertainties in the projected impacts and using statistical tools such as Monte Carlo simulations (?). Another key aspect is the promotion of cooperation within the scientific community. The Coupled Model Intercomparison Project Phase 6 (CMIP6) is an outstanding example that shows how coordinated modelling and data comparison between different research centres help to expand the scientific basis and the validity of climate models (?). These collective efforts not only strengthen the quality of the models, but also their ability to respond to the challenges and issues of climate change. To summarise, climate model ensembles are an indispensable tool in climate research. They help to reduce uncertainty in climate projections, support the development of robust climate predictions and foster deeper collaboration between research centres worldwide. The ongoing development and application of these models will be crucial to better understand the complexities of the climate system and to develop appropriate adaptation strategies.

2.15 4.1 Use of climate model ensembles for robust climate predictions

The benefits of climate model ensembles for robust climate predictions are manifold and crucial for the accuracy and reliability of long-term climate projections. These ensembles, which consist of a large number of independent model simulations, make it possible to evaluate and minimise the uncertainties in climate predictions due to internal variability and model-inherent uncertainties. By using climate model ensembles, the effects of internal variability on future climate projections can be better understood. Internal variability, which is due to natural fluctuations in the climate system, can cause considerable uncertainty in projections (?). Climate model ensembles offer the possibility to quantify these uncertainties by combining a large number of simulations under different initial conditions and with different models. This helps to better map the range of possible future climate states and to increase the robustness of the predictions. Another advantage of climate model ensembles is the improvement in the statistical representativeness of the climate model outputs. By superimposing many simulations, systematic errors of individual models can be compensated and a more reliable estimate of future climate conditions can be achieved. The probabilistic assessment of future climate trends, as made possible by Bayesian models, ensures that the uncertainties in the model forecasts are adequately taken into account (?). In addition, climate model ensembles make valuable contributions to the production of regional and seasonal climate predictions, which are of central importance for adaptation strategies at the local level. For example, high-resolution climate models within an ensemble can be used to project specific regional climate changes in more detail (?). This is particularly important for regions that are strongly affected by specific climate extremes, such as the Mediterranean region, where extreme temperature and precipitation events need to be intensively analysed (?). By considering different emission scenarios in the simulations, the range of possible future climate states can also be covered. This is essential in order to plan and implement adaptive measures that can react flexibly to different climatic developments. A multi-model ensemble thus enables a more robust and comprehensive consideration of climate risks and contributes to improved strategic decision-making (?; ?). To summarise, climate model ensembles provide an indispensable basis for making well-founded and reliable climate predictions. They not only enable a realistic

assessment of uncertainties, but also support the development of specialised and regionally adapted climate adaptation strategies. This makes them an essential tool in modern climate research and planning.

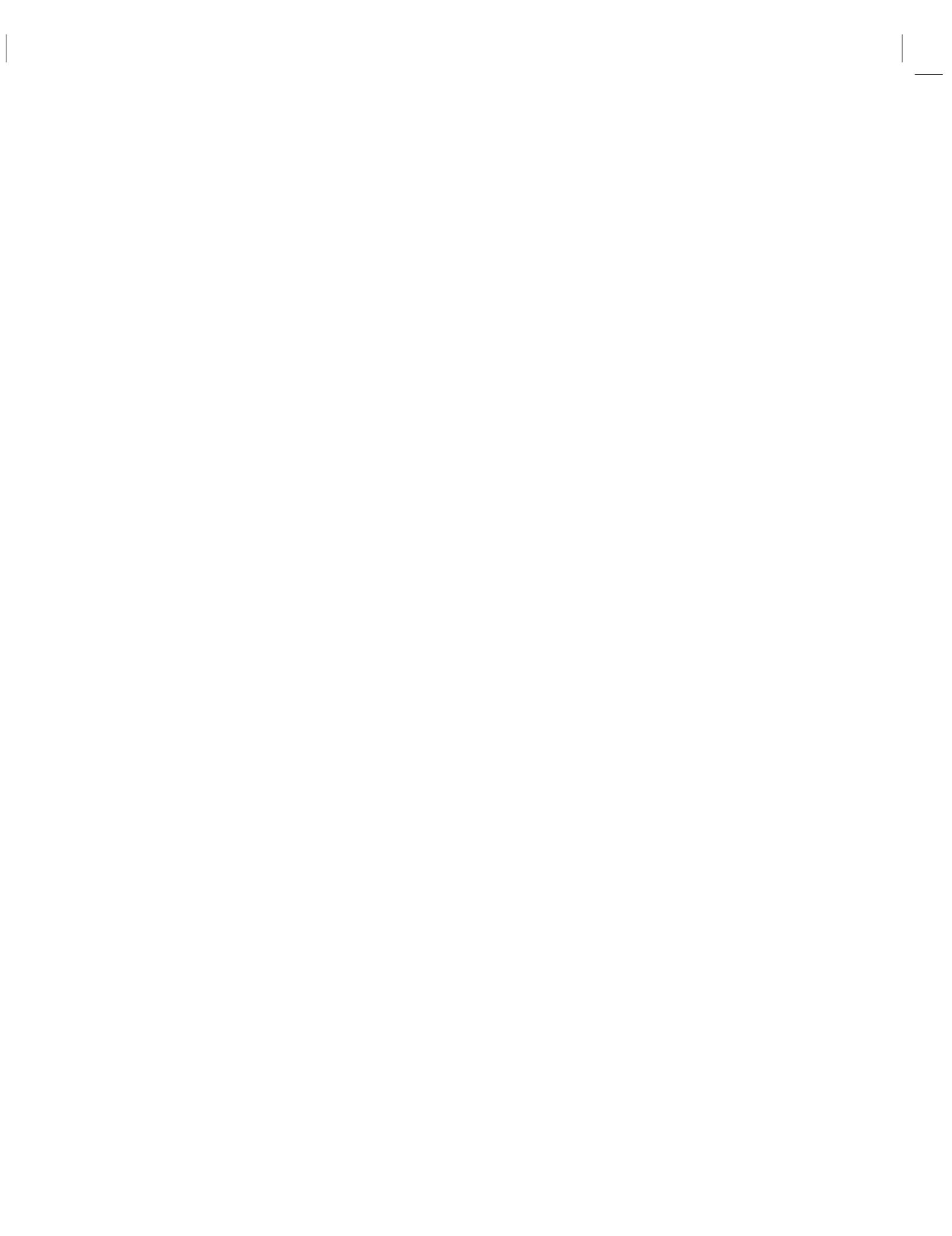
2.16 4.2 Application examples: regional and seasonal climate trends

Climate model ensembles play a crucial role in producing robust regional and seasonal climate predictions. By using a multi-model approach, uncertainties can be minimised and the accuracy of projected climate changes can be increased. An illustrative example of this is the study of long-term spatiotemporal variability and trends in extreme rainfall and temperatures in Bangladesh, which can have an impact on rice production (?). The analysis shows regional variations in climate trends, such as the decrease in precipitation in specific regions and the marked rise in temperature, which could substantially affect agricultural yields. Another example is the use of EURO-CORDEX RCMs and HighResMIP GCMs to analyse the daily precipitation distribution in Europe. These high-resolution models provide an improved representation of precipitation patterns compared to older CMIP5 models. In particular, the PRIMAVERA GCMs show precipitation distributions that are closer to observations in many European regions than the results of the EURO-CORDEX RCMs (?). Such regional climate models are particularly useful for simulating complex topographical and coastal areas more precisely, which is crucial for predicting extreme weather events, for example. Seasonal trends are of particular interest in climate research as they have a direct impact on agricultural productivity and water resources. Analyses of climate models show that intensified precipitation and longer dry periods have different regional effects. For example, some models show a tendency to overestimate winter rainfall and underestimate summer rainfall, which can be partially corrected by finer grid resolutions (?). Such seasonal discrepancies must be taken into account when planning and adapting agricultural and water management strategies in order to increase resilience to climate change. The importance of climate model ensembles is emphasised by the combined use of regional and global models. These models make it possible to understand and project the climate at different scales, which is particularly valuable for assessing potential climate risks in specific regions. In this way, the results of this modelling help to develop early warning systems and adaptation strategies that can counter the effects of extreme weather events and support long-term planning (?). Overall, these application examples illustrate that climate model ensembles are an indispensable tool for analysing and predicting regional and seasonal climate trends. They provide a sound basis for scientifically sound decision-making processes and adaptation planning in the face of advancing climate change.

2.17 5. Conclusion

The relationship between natural variability and climate modelling is a central aspect of climate science. Natural variability refers to the natural fluctuations in the climate system caused by internal processes such as El Niño/La Niña, volcanic activity or solar cycles. These fluctuations occur independently of human influences and can affect the climate in the short and medium term. Climate models attempt to take both these natural and anthropogenic (man-made) factors into account in order to provide a realistic simulation of the climate. They use historical data and physical laws to simulate and predict natural climate fluctuations. Natural variability can cause short-term fluctuations in temperature and precipitation that differ from the long-term trends due to climate change. Climate models show that natural variability can have a major impact on regional climate patterns in the coming decades, even if the long-term trend is dominated by anthropogenic climate change. Scientists compare the results of climate models with historical climate data to assess the accuracy of the models. Models that can correctly simulate natural variability are better able to predict future climate changes. Despite the progress made in modelling, there are still challenges. The complexity of the climate system and the large number of influencing factors make modelling natural variability a challenge.

Moreover, uncertainties regarding future natural events and their impact on the climate system pose an additional difficulty. Overall, natural variability is an integral part of climate modelling, and understanding its mechanisms and influences is crucial for the accuracy of climate projections.



3

Standard Precipitation Evapotranspiration Index

Author: Sophie Hopp

Supervisor: Henri Funk

Suggested degree: Bachelor

3.1 Introduction

Droughts are a common phenomenon that occurs worldwide and are characterized by prolonged periods of water shortage. These shortages can arise due to various factors such as a lack of precipitation or high temperatures leading to increased evaporation rates. These conditions can arise from natural climate variability and can be worsened by climate change. Droughts can have serious impacts, including agricultural effects such as crop failures and food shortages, disruptions to water supplies, and socio-economic consequences due to restrictions on shipping caused by low water levels (?). Therefore, it is important to monitor and assess droughts effectively.

Drought indices are used to quantify the severity, duration, and spatial extent of droughts. There are different types of droughts, such as meteorological, hydrologic, agricultural, and socio-economic droughts (?). The subjectivity in defining droughts has made it challenging to establish a universal drought index, leading to the development of various indices (?).

The Standardized Precipitation Index (SPI) is a commonly used index that is based on the probability of precipitation. It can be calculated for different time scales, which enables to distinguish between short and long-term droughts and is important to functionally separate different drought types. For example, agricultural droughts usually have a much shorter time scale than hydrological droughts (?). However, the SPI does not take into account the temperature, which can be a critical factor in drought development. It assumes that the variability of precipitation is much higher than that of other variables like temperature and potential evapotranspiration and considers these variables stationary, meaning they have no temporal trend (?). These assumptions are problematic, especially given the increases in temperature due to climate change, which can have an impact on drought conditions.

The Palmer Drought Severity Index (PDSI), an index primarily used for meteorological droughts, includes temperature by using a soil water balance equation that incorporates prior precipitation, moisture supply, runoff, and surface-level evaporation demand. However, it cannot be calculated for different time scales, has a strong influence of the calibration period, and faces issues with spatial comparability. The self-calibrated Palmer Drought Severity Index (scPDSI) was developed to address some of these issues, but it is still not multiscalar. Since droughts are multiscalar phenomena, they need to be assessed at different time scales (?).

To combine the advantages of both SPI and PDSI, the Standardized Precipitation Evapotranspiration Index (SPEI) was developed. The SPEI is based on the climatic water balance, which is the difference between precipitation and potential evapotranspiration. Like the SPI, the SPEI is based on a probabilistic approach, making the calculation easy and fast. It also shares the multitemporal nature of the SPI, allowing calculation for different time scales (?). The inclusion of evapotranspiration in the SPEI allows for a better capture of

water deficits or surpluses on the land surface, making it a more effective tool for monitoring and assessing droughts, especially under changing climatic conditions.

This report begins with an introduction to the data, followed by a description of the methods used to calculate the SPEI. The results of the SPEI calculation for Bavaria are then presented, including a comparison of different distributions for the water balance and the proportion of extreme droughts over time.

3.2 Data and Methods

3.2.1 Data

Meteorological data from January 1881 to March 2024 were provided by the Open Data Portal of the Climate Data Center of the German Weather Service (?). The dataset includes regional average monthly observations of the average air temperatures and precipitation in Bavaria. These data were aggregated from grid fields to represent the average conditions for the region. Bavaria is a federal state in the southeast of Germany with a latitude range of approximately 47° to 50° N and a longitude range of approximately 9° to 13° E.

3.2.2 Implementation of the SPEI

To implement the SPEI, we utilized the freely accessible SPEI package (?). This process requires the water balance, for which the potential evapotranspiration must first be determined.

Potential evapotranspiration (PET) represents the amount of water that would evaporate from the soil and transpire from plants if water supply were not limited. Although various methods exist for calculating PET , it has been shown that their differences have minimal impact on the results of the SPEI (?). For simplicity, we use the Thornthwaite method to calculate PET . This method relies on the monthly average temperature T_i , the latitude of the study area, and the month of the year. Latitude and time of year are used to estimate solar radiation.

The water balance D_i is the difference between precipitation P_i and potential evapotranspiration PET_i for month i . It is calculated for each month of the year as:

$$D_i = P_i - PET_i, \quad D_i \in \mathbb{R}, i \in \mathbb{Z}. \quad (3.1)$$

Negative values of the water balance indicate a water deficit, while positive values indicate a water surplus.

To calculate the SPEI for different time scales, the water balance is summed over the desired time scale. The aggregated water balance for a time scale of k months ($D_i^{(k)}$) is calculated as:

$$D_i^{(k)} = D_{i-k+1} + D_{i-k+2} + \dots + D_i. \quad (3.2)$$

$D_i^{(k)}$ is later used to compute the SPEI for the respective time scale of k months, denoted as $SPEI_i^{(k)}$.

To model the distribution of the aggregated water balance values $D_i^{(k)}$, a distribution function $F_D(D_i^{(k)})$ capable of handling negative values and skewed distributions is used.

For the SPI, the gamma distribution is typically employed to model precipitation since precipitation values are strictly non-negative. However, the gamma distribution is unsuitable for the SPEI because the water balance D_i can be negative. Instead, the three-parameter log-logistic distribution is used for the SPEI due to its flexibility in modeling skewed distributions, its ability to handle negative values and its relatively high kurtosis, which permits more gradual decrease in the curve for low values. This results in more coherent and

realistic probabilities for very low water balance values, unlike other distributions which suggest that these low values are extremely rare, especially at shorter time scales (?).

The probability density function of the three-parameter log-logistic distribution is given by:

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x-\gamma}{\alpha} \right)^{\beta-1} \left(1 + \left(\frac{x-\gamma}{\alpha} \right)^{\beta} \right)^{-2}, \quad \alpha > 0, \beta > 0, x > \gamma, \quad (3.3)$$

and its cumulative distribution function is defined as:

$$F(x) = \left(1 + \left(\frac{\alpha}{x-\gamma} \right)^{\beta} \right)^{-1}, \quad \alpha > 0, \beta > 0, x > \gamma, \quad (3.4)$$

where α is the scale parameter, β is the shape parameter, and γ is the origin parameter that can shift the distribution to model negative values.

These parameters are estimated using the method of L-moments. L-moments are analogous to conventional moments but are less sensitive to outliers, making them a robust and easy approach for parameter estimation (?).

The distribution for the SPEI is fitted to a reference period, which is typically 30 years. During this period, the climate is assumed to be stationary. The calibration using reference climate data allows for the intercomparison of the index among different stations or periods. Such reference periods are particularly important in climate change studies, as they offer a consistent baseline for assessing temporal changes in drought characteristics (?). In this report, the reference period used is from 1981 to 2010. This period serves as a baseline to better understand the extent and severity of droughts in the context of climate variability.

The SPEI values are obtained as the standardized values of $F_D(D_i^{(k)})$, meaning that the values are centered around zero and have a standard deviation of one. This transformation allows for easy comparison of the SPEI values across different time scales and regions (?). One approximation for the SPEI values is given by:

$$SPEI = W - \frac{C_0 + C_1 W + C_2 W^2 + C_3 W^3}{1 + d_1 W + d_2 W^2 + d_3 W^3} \quad (3.5)$$

with

$$W = \begin{cases} \sqrt{-2 \ln(P)} & \text{for } P \leq 0.5, \\ -\sqrt{-2 \ln(1-P)} & \text{for } P > 0.5 \end{cases} \quad (3.6)$$

and P is the probability that, for a given water balance value $D_i^{(k)}$, a random value from the fitted log-logistic distribution is greater than the observed value. This probability is calculated as:

$$P = 1 - F_D(D_i^{(k)}). \quad (3.7)$$

The constants are: $C_0 = 2.515517$, $C_1 = 0.802853$, $C_2 = 0.010328$, $C_3 = -0.000220$, $d_1 = 1.432788$, $d_2 = 0.189269$, $d_3 = 0.001308$. ?

3.3 Results

3.3.1 Comparison of Water Balance Distributions

To compare the water balance distributions in Bavaria for different time scales, histograms of the water balance values for the commonly used time scales of 1, 3, 6, 12, 18, and 24 months were created for the

reference period from 1981 to 2010 (see Figure 3.1). Both a normal and a log-logistic distribution were fitted to the empirical data. The additional normal distribution is fitted to provide a comparison to the log-logistic distribution. Moreover Figure 3.1 suggests that the normal distribution might be a good fit due to the bell-shaped, symmetric histograms.

For the normal distribution, parameters were estimated using Maximum Likelihood Estimation, which involves using the sample mean and standard deviation to determine μ and σ .

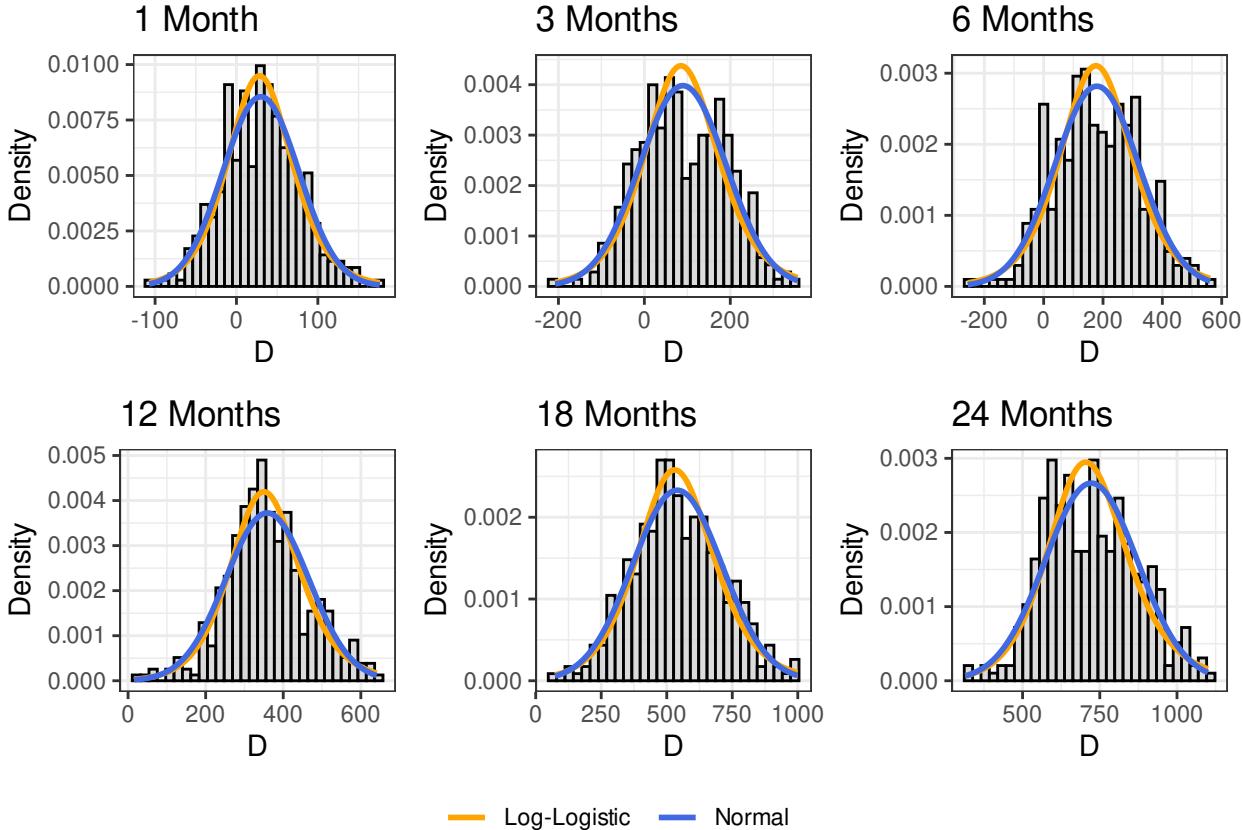


FIGURE 3.1: Histograms of the water balance for different time scales (1, 3, 6, 12, 18, and 24 months) with fitted log-logistic and normal distributions for the reference period 1981 - 2010.

To compare the goodness of fit of the normal and log-logistic distributions to the empirical water balance data, the Wasserstein distance was calculated.

The Wasserstein distance is a measure of the distance between two probability distributions. The Wasserstein distance is calculated by:

$$\mathcal{W} = \int_{-\infty}^{\infty} |F_{\text{empirical}}(x) - F_{\text{theoretical}}(x)| dx, \quad (3.8)$$

with $F_{\text{empirical}}(x)$ being the empirical cumulative distribution function and $F_{\text{theoretical}}(x)$ being the fitted theoretical cumulative distribution function. The smaller the Wasserstein distance, the better the fit of the theoretical distribution to the empirical distribution.

The results of the Wasserstein distance calculation between the empirical distribution and both the fitted log-logistic distribution and the fitted normal distribution for all time scales are shown in Table 3.1.

The Wasserstein distance between the empirical distribution and the fitted normal distribution is smaller

TABLE 3.1: Wasserstein distances between the empirical distribution of the water balance in the reference period from 1981 to 2010 and the fitted normal and log-logistic distributions for different time scales.

Time Scale	Normal	Log-Logistic
1	1.553	3.080
3	7.957	12.295
6	7.511	13.667
12	9.075	6.096
18	6.437	11.085
24	13.020	16.572

than that between the empirical distribution and the fitted log-logistic distribution for all time scales except the 12-month time scale. This suggests that the normal distribution fits the empirical distribution better than the log-logistic distribution for most of the given time scales in the reference period and the region of Bavaria. However, the log-logistic distribution with its three parameters offers more flexibility and can model skewness not captured by the normal distribution. Due to the reasons stated in the methods section, the log-logistic distribution is used for the calculation of the SPEI.

3.3.2 SPEI in Bavaria

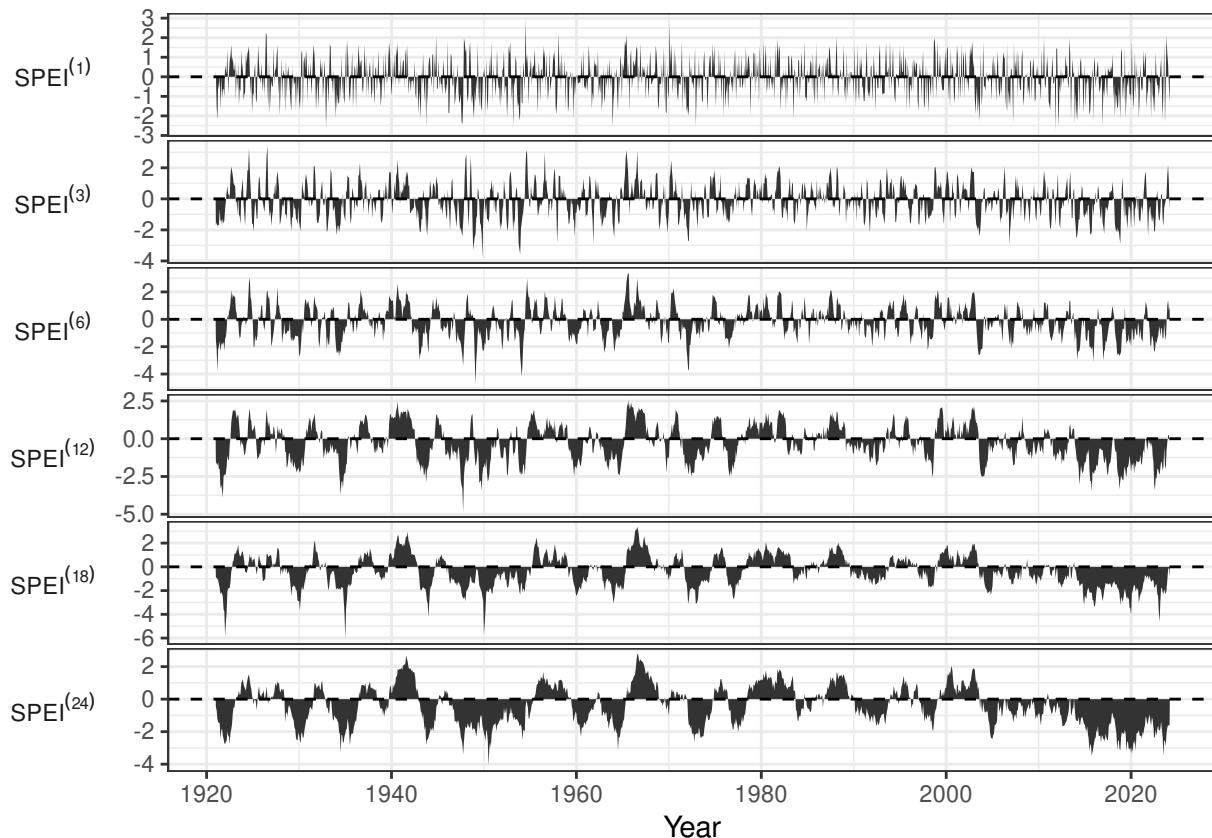


FIGURE 3.2: The 1-, 3-, 6-, 12-, 18-, and 24-month SPEI for Bavaria (1921 - 2024).

The SPEI for Bavaria for different time scales is shown in Figure 3.2. Values below 0 indicate that there is less

TABLE 3.2: Classification of Droughts Based on the SPEI (@mckee1993).

SPEI	Classification
0 to -0.99	Mild Drought
-1 to -1.49	Moderate Drought
-1.5 to -1.99	Severe Drought
≤ -2	Extreme Drought

water available than usual, while values above 0 indicate that water availability is higher than on average. All subplots show fluctuations around the zero line, indicating periods of both wet and dry conditions. Longer aggregation periods show smoother patterns compared to shorter periods. The 1-month SPEI shows the highest variability due to its sensitivity to short-term fluctuations in precipitation and temperature. There is a trend toward more negative SPEI values in the last years, indicating a decrease in water availability and increasing drought conditions. This trend is mostly visible in the longer time scales. Historically, Bavaria has experienced extended periods with both negative and positive SPEI values, but in recent years, there have been no prolonged periods with positive SPEI values visible in the 12-, 18-, and 24-month SPEI.

Table 3.2 shows the classification of droughts based on the SPEI. These drought categories were originally defined for the SPI but are commonly used for the SPEI due to the standardized nature of the indices. An index below -2 indicates extreme drought (?).

To visualize the frequency of extreme droughts in Bavaria over time, the proportion of months experiencing extreme drought within a given period is calculated using the following formula:

$$prop_{i,m}^{(k)} = \frac{\sum_{j=i-m+1}^i I(SPEI_j^{(k)} \leq -2)}{m} \quad (3.9)$$

where $prop_{i,m}^{(k)}$ represents the proportion of months experiencing extreme drought for a specific SPEI scale k within the period ending at month i and covering the last m months. I is an indicator function that equals 1 if the SPEI value is less than or equal -2 and 0 otherwise.

Figure 3.3 displays the proportion of 12-month SPEI values below -2 over consecutive 10-year periods. Historically, Bavaria has experienced periods of extreme drought, however the proportion of months classified under extreme drought conditions has continuously increased over the past decade. Currently, this proportion is at its highest level in the past century, with 32.5% of the months in the last 10 years recording an $SPEI^{(12)}$ value below -2.

When comparing the proportion of extreme droughts using different time scales for the SPEI, it is evident that the proportion of extreme droughts has increased across all time scales in recent years (see Figure 3.4). The data shows that the longer the time scale, the higher the proportion of extreme droughts. For the 24-month SPEI, the proportion of extreme droughts is at the moment above 50%. However, for the 3-month SPEI, the proportion of extreme droughts is not at its highest level in the last 100 years. Notably, from approximately 1952 to 1956, the proportion of extreme droughts was higher than in recent years when considering the 3-month SPEI. This mid-20th century period of increased drought conditions is also apparent across other time scales, but the proportion of extreme droughts in recent years surpasses that of the 1950s.

When using different periods for the calculation of the proportion of extreme droughts, the data reveals that the proportion of extreme droughts has a higher variability for shorter periods (see Figure 3.5).

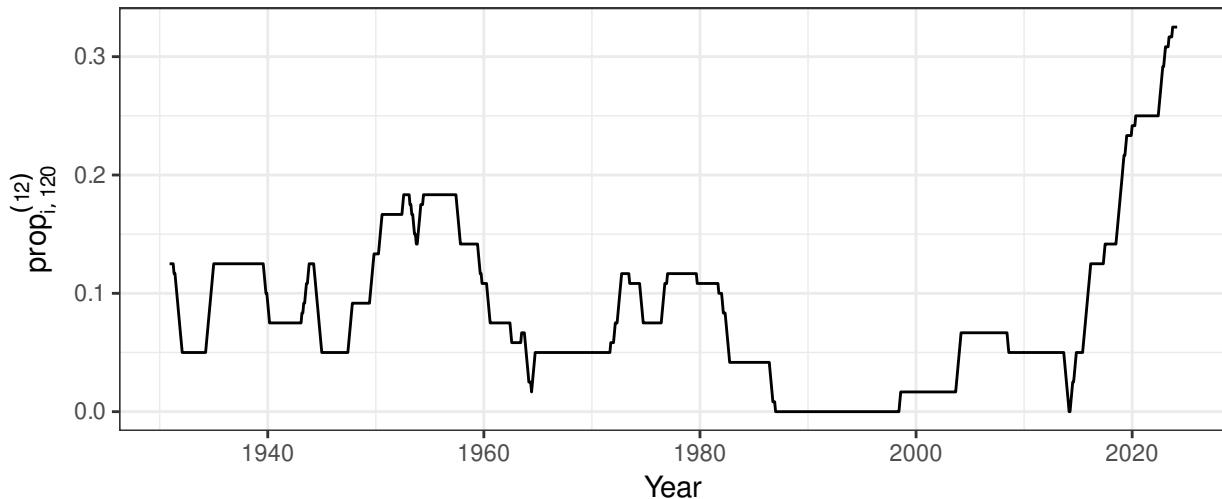


FIGURE 3.3: Proportion of 12-month SPEI values below -2 for periods of 10 years in Bavaria (1921 - 2024).

3.4 Discussion

The results demonstrate that the SPEI is a valuable and practical tool for monitoring drought conditions over different time periods. By considering both precipitation and potential evapotranspiration, the SPEI provides a comprehensive measure of water availability and accounts for climatic factors that affect the water balance beyond what precipitation-based indices, such as the SPI can offer. This sensitivity to climate change allows the SPEI to effectively identify the impacts of global warming on drought conditions.

The SPEI's minimal data requirements, ease of calculation, and multiscale nature contribute to its wide applicability. Additionally, it is robust across different climatic regions due to the log-logistic distribution used in its calculation.

However, a limitation of the SPEI is its dependence on the accuracy of potential evapotranspiration estimates, which can vary depending on the calculation method. For instance, the Thornthwaite method used in this report assumes that evapotranspiration does not occur when the temperature is below 0°C. This assumption might not be accurate in regions where evapotranspiration can still occur at low temperatures due to factors like wind speed and solar radiation. While other methods can provide more accurate PET estimates, they require additional data that may not be available in all regions.

The Global Precipitation Climatology Centre Drought Index (GPCC-DI) combines the SPEI with an adapted version of the SPI to account for the limitations of both indices. The SPEI works well in warm, dry regions

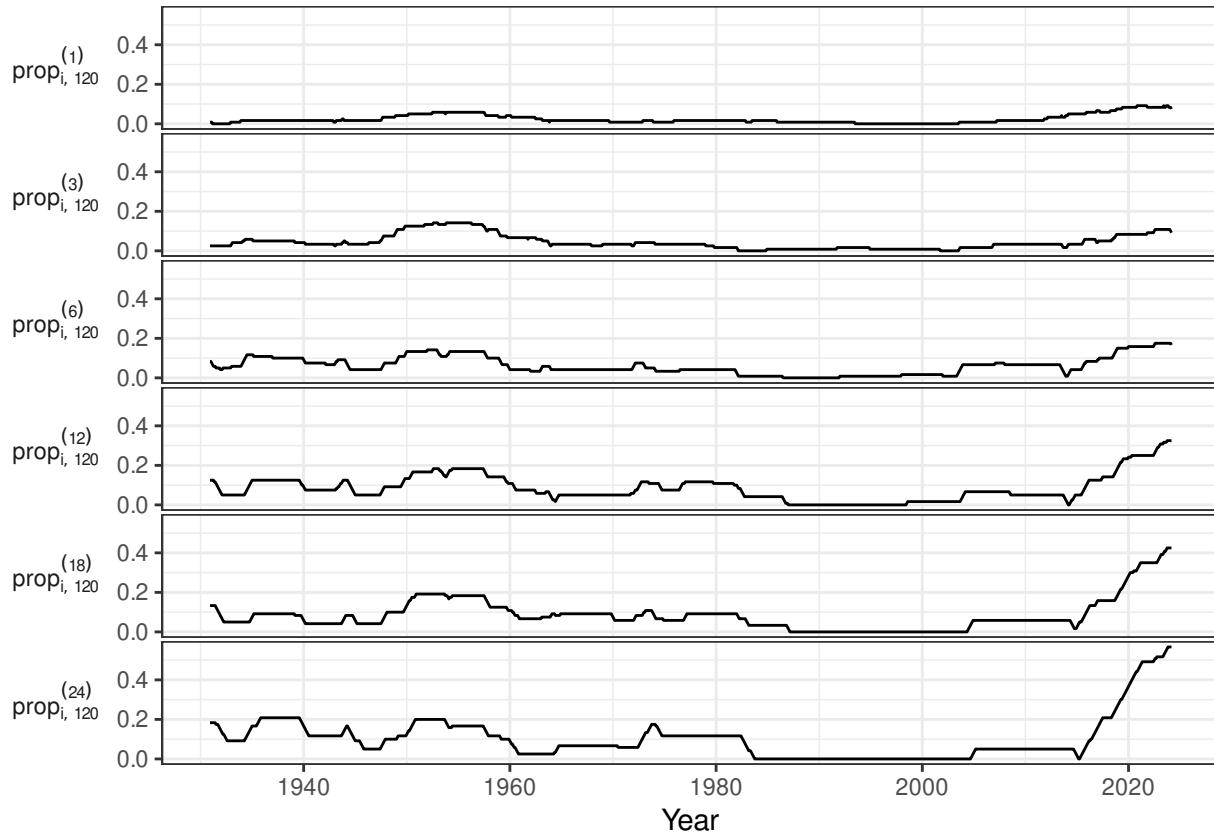


FIGURE 3.4: The proportion of extreme droughts in Bavaria in the last 10 years for different time scales of the SPEI (1, 3, 6, 12, 18, and 24 months).

where the SPI is not applicable, while the SPI is effective in cold regions where estimating potential evapotranspiration is challenging. This combination allows for almost global coverage (?).

An important consideration in using the SPEI is the choice of the reference period, as it influences the assessment of drought characteristics, particularly the severity and spatial extent, while its impact on frequency is relatively small. It is recommended that the reference period be clearly specified in drought assessments to improve understanding of regional drought characteristics and their temporal changes (?).

The observed trends in Bavaria indicate an increase in the frequency of extreme drought conditions, particularly in the last decade. These trends are especially pronounced at longer time scales.

One application of the SPEI is predicting future drought conditions. Accurate predictions can enable better preparedness and mitigation strategies, thereby reducing the negative impacts of droughts. ? investigated the use of generalized additive models to predict the SPEI, incorporating variables such as rain, minimum, maximum and average monthly temperature as predictors in addition to lagged SPEI values. Using observations such as temperature and precipitation, that can only be measured retrospectively to predict drought conditions of already passed months is not useful as they are not future predictions anymore. However, the use of lagged SPEI values can be a valuable tool for drought prediction. It would be beneficial to investigate the use of other models to predict future drought conditions based on lagged SPEI values.

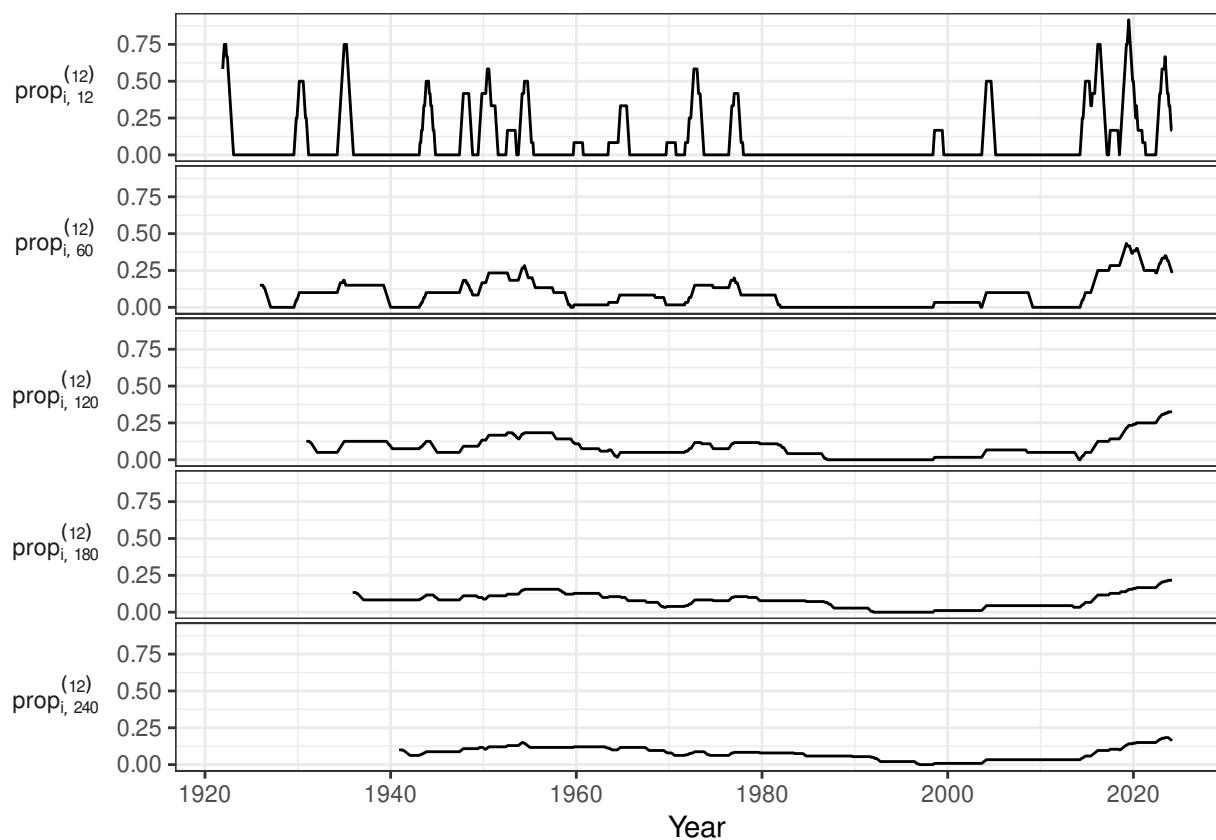


FIGURE 3.5: The proportion of extreme droughts in Bavaria by SPEI12 for different periods of the SPEI (1, 5, 10, 15, and 20 years).



4

Compound events

Author: Author

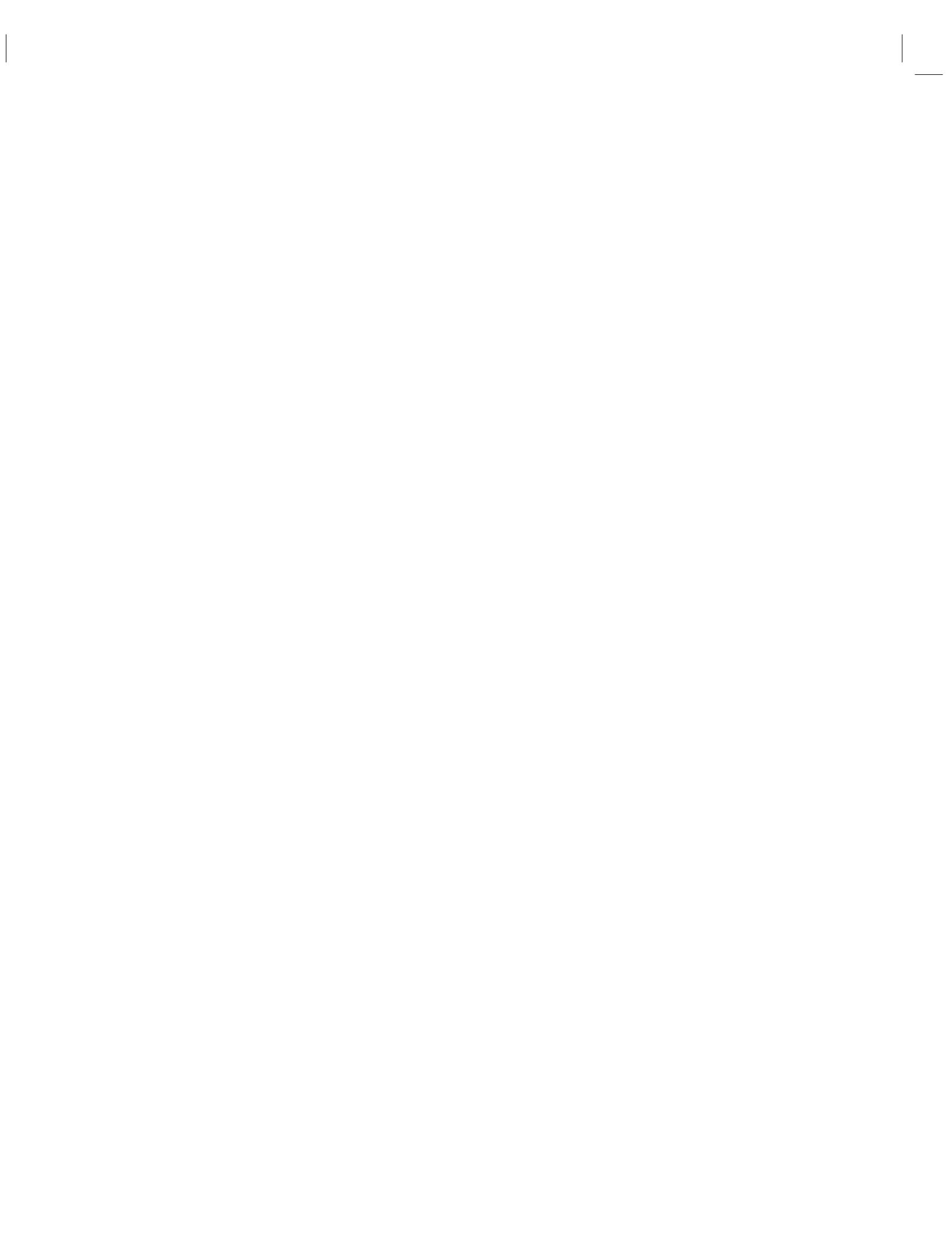
Supervisor: Henri Funk

Suggested degree: Master

In climate research, a compound event refers to the combination of multiple extreme weather or climate events occurring simultaneously or successively, leading to significant impacts on the environment, society, or economy. These events can be related either through meteorological factors (e.g., heatwaves and drought occurring together due to a prolonged high-pressure system) or through their impacts (e.g., heavy rainfall and storm surge combining to cause more severe flooding). The complexity of compound events lies in their interconnectedness and the way they can exacerbate each other's effects, often in a nonlinear manner, making them more challenging to predict and manage than individual extreme events.

4.1 Example: Hot and Dry Events

Consider a scenario where a region experiences both extreme heat and an extended period of low precipitation simultaneously. In this example, the compound nature of the hot and dry events creates a cascade of impacts that are more severe than what would be expected if these events occurred independently. The interplay between the heat and drought amplifies the consequences, underscoring the importance of understanding and managing compound events in the context of climate change and variability. ?



5

Asymmetric bivariate copulas for compounds

Author: Claudia Rettenbeck

Supervisor: Henri Funk

5.1 Introduction

There are different definitions of compound events. In this chapter (i.e. chapter 5), a compound event is defined as an extreme event, which is caused by at least two climate variables (e.g. hydrological variables such as river discharge). Thus, the question of how to model the interplay of several climate variables by means of their multivariate distribution arises. Answering this question lays the basis for modelling compound events, which can, but don't have to, arise from this interplay. After having briefly presented the definition of d -variate copulas in section 5.2.1, this chapter focuses on the interplay of *two* random variables. One approach for modelling multivariate cumulative distribution functions is based on copulas. According to Sklar's theorem, the joint cumulative distribution function of two univariate random variables can be modelled by combining its marginals with a copula (see section 5.2.2 for details). As ? state, an appropriate model for the dependence structure, which is determined by the copula, can therefore be selected independently from the choice of the marginals, which is advantageous in modelling. Since in this chapter, the focus is on modelling the interplay of two random variables by means of their multivariate distribution via the copula approach, it is possible to further narrow the focus to modelling the dependence structure of these random variables, which will be done in the following. In particular, this chapter concentrates on a specific type of dependence, namely asymmetric dependence. This is relevant for climate research since there are climate variables whose dependence structure is characterized by asymmetry, i.e. the corresponding random variables are asymmetrically dependent (see section 5.2.3 for details). For instance, the sections 5.2.3 and 5.3.3 show that the monthly average flow rates (flow rates measure river discharge) of the rivers Inn and Danube can be considered asymmetrically dependent (where the data set presented in the sections 5.2.3 and 5.3.3 is used to draw this conclusion). Thus, when modelling the dependence structure of two climate variables, the assumption of symmetry, which is made when using most classical copulas (see ?, p. 92), might be too restrictive. Therefore, the goal of this chapter is to lay the theoretical foundation for recognizing and modelling asymmetric dependence between two random variables. There are several approaches for modelling this kind of dependence structure by means of copulas. Due to its scope, this chapter focuses on Archimax copulas as modelling approach, but the interested reader may be referred to chapter 5.4 in ? for an overview of modelling approaches based on bivariate copulas. Moreover, following a large part of the relevant literature, from section 5.2.3 onwards, the focus lies on the asymmetric dependence of *continuous* random variables.

While theoretically, a compound event could arise from non-extreme events, climate research often defines compound events via the interplay of extreme events. Thus, while this chapter lays the theoretical foundation for modelling compound events defined according to a broader definition, chapter 4 presents a stricter version of this definition, which is often used in climate research.

The remainder of this chapter proceeds as follows: Section 5.2 provides theoretical background on copulas in general, where also the distinction between symmetric and asymmetric copulas is established. Section 5.3 covers theoretical background on measures of asymmetry in a copula and presents an application of a

selection of these asymmetry measures to hydrological data. Section 5.4 covers Archimax copulas. Specifically, it defines Archimax copulas, presents well-known special cases of these copulas and explains their asymmetry property. The asymmetry property is not the only relevant property of Archimax copulas (for instance, the fact that an Archimax copula can be constructed with a predetermined extreme value attractor (see ? for details) represents another property, which might be interesting in the context of the climate crisis). Due to the scope of this chapter however, the focus is on the property most relevant to it, i.e. the asymmetry property. Finally, section 5.5 concludes.

5.2 Theoretical background on copulas

5.2.1 Definition of d-variate copulas and the special case of bivariate copulas

Following ? (p. 10), copulas can be defined as follows: A d -dimensional copula C , i.e. a d -variate copula, is a d -dimensional cumulative distribution function concentrated on $[0, 1]^d$ with univariate marginals which are uniformly distributed on $[0, 1]$, where $d \geq 2$. Thus, as stated by ? (p. 10), C assumes values on \mathbb{R}^d . However, in the literature, C is frequently defined as a function with domain $[0, 1]^d$ and range $[0, 1]$, such that C can be described by $C : [0, 1]^d \rightarrow [0, 1]$ (see e.g. ? and ?). According to ? (p. 10), this can be attributed to the fact that C concentrates the probability distribution on $[0, 1]^d$ and therefore the values C assumes on $\mathbb{R}^d \setminus [0, 1]^d$ are usually not specified.

? (p.11) further remark that a random vector \mathbf{U} (defined on a suitable probability space) corresponds to each copula C such that the joint cumulative distribution function of \mathbf{U} is C .

The previous paragraph defines copulas based on their probabilistic interpretation. According to ? (p. 14), copulas can also be defined in terms of their analytical properties, where the probabilistic and the analytical definition are equivalent:

A function $C : [0, 1]^d \rightarrow [0, 1]$ is a d -variate copula if and only if the following holds:

- i) C is d -increasing,
- ii) $C(u_1, \dots, u_d) = 0$ if $u_k = 0$ for at least one $u_k \in \{u_1, \dots, u_d\}$,
- iii) $C(1, \dots, 1, u_k, 1, \dots, 1) = u_k$, where the left-hand side of the equation means that all arguments equal 1 except possibly for the k th one, which can assume any value in $[0, 1]$.

? (p. 14) point out that ii) and iii) are called boundary conditions, where ii) means that C is grounded (also named anchored) and iii) shows that C has univariate marginals which are uniform on $[0, 1]$.

Since this chapter focuses in the following on bivariate copulas, the expression d -increasing is not explained in full generality here. Instead an explanation will be provided for the special case of $d = 2$ in the subsequent paragraph. For a more general explanation, see ? (p.7).

For $d = 2$, the definition above can be translated to (see ?, p.15 as well as ?, p. 8 and p. 10): A function $C : [0, 1]^2 \rightarrow [0, 1]$ is a 2-variate copula if and only if the following holds:

- i) C is 2-increasing, meaning that for all $a_1, a_2, b_1, b_2 \in [0, 1]$, where $a_1 \leq a_2$ and $b_1 \leq b_2$, it holds that:

$$C(a_1, b_1) - C(a_1, b_2) - C(a_2, b_1) + C(a_2, b_2) \geq 0,$$
- ii) $C(0, v) = C(u, 0) = 0 \quad \forall u, v \in [0, 1]$,
- iii) $C(1, v) = v$ and $C(u, 1) = u \quad \forall u, v \in [0, 1]$.

Note that the names 2-variate copula and bivariate copula are used as synonyms in this chapter.

5.2.2 Sklar's theorem and the probability integral transformation

According to ? (p. 42), Sklar's theorem states that *any* multivariate cumulative distribution function can be expressed in terms of a copula and its univariate marginals. Specifically, let (X, Y) be a bivariate random vector on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, $H(x, y) := \mathbb{P}(X \leq x, Y \leq y)$ denotes the joint cumulative distribution function of (X, Y) , where $F(x) = \mathbb{P}(X \leq x)$ and $G(y) = \mathbb{P}(Y \leq y)$ represent the marginals of H . Then, there exists a bivariate copula $C = C_{X,Y}$ such that it holds that

$$H(x, y) = C(F(x), G(y)) \quad \forall (x, y) \in \mathbb{R}^2, \quad (5.1)$$

where the copula C is unique, if the marginals F and G are continuous.

Note that Sklar's theorem can be generalized to d -variate copulas easily, where $d > 2$ (see ?, p.42 for details), but since this chapter focuses on bivariate copulas, also the version of Sklar's theorem presented here is described for the bivariate case.

For the case of continuous marginals, on which this chapter focuses from section 5.2.3 onwards, the probability integral transformation plays an important role in understanding Sklar's theorem. Let \tilde{X} be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with continuous cumulative distribution function \tilde{F} . Then, $\tilde{F} \circ \tilde{X}$ is called the probability integral transformation and it can be shown that $\tilde{F} \circ \tilde{X}$ is uniformly distributed on $[0, 1]$ (see ?, pp. 5-6). As ? (p. 6) note, this means that transforming a random variable by its continuous cumulative distribution function always leads to a random variable that follows the standard uniform distribution, whereas transforming each component of a random vector by its continuous univariate cumulative distribution function can result in multivariate distributions that differ from the multivariate standard uniform distribution. This can be seen from the proof of equation (5.1) given in ? (p. 43) for continuous marginals F and G , which moreover explains the connection between Sklar's theorem and the probabilistic definition of copulas. The proof requires rewriting $H(x, y)$ for all $(x, y) \in \mathbb{R}^2$ as follows:

$$\begin{aligned} H(x, y) &= \mathbb{P}(X \leq x, Y \leq y) \\ &= \mathbb{P}(F(X) \leq F(x), G(Y) \leq G(y)) \\ &= \mathbb{P}(U \leq F(x), V \leq G(y)), \end{aligned}$$

where $U := F(X)$ and $V := G(Y)$. It can be seen from the description of the probability integral transformation above that U and V are uniformly distributed on $[0, 1]$. As a consequence, the cumulative distribution function of (U, V) has uniform univariate marginals on $[0, 1]$, from which it follows that this cumulative distribution function is a copula (see probabilistic definition in section 5.2.1). Therefore, $\mathbb{P}(U \leq F(x), V \leq G(y))$ can be written as $C(F(x), G(y))$ (recall that C denotes a copula). Thus, it has been shown that $H(x, y) = C(F(x), G(y))$.

If H is absolutely continuous, then the density h corresponding to H can be written in terms of the density c of the copula C . Specifically, h is given for almost all $(x, y) \in \mathbb{R}^2$ by

$$h(x, y) = c(F(x), G(y)) \cdot f(x) \cdot g(y), \quad (5.2)$$

where F and G are the univariate marginals (as above) with the corresponding densities being denoted by f and g , respectively (see ?, p. 42).

5.2.3 Distinction between symmetric and asymmetric copulas

There are different forms of symmetry. Among them are exchangeability, radial symmetry and joint symmetry, where radial symmetry is also called reflection symmetry (see ?, pp. 30-33). In line with that, ? (p. 92) note that in the copula literature, the term symmetry can refer to both exchangeability and radial symmetry. However, following the majority of the relevant literature on Archimax copulas reviewed, this chapter defines a copula as symmetric if and only if it is exchangeable.

According to ? (pp. 30-31), it can be shown that when the cumulative distribution function C of the random vector (U, V) on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a copula, also the cumulative distribution function \tilde{C} of (V, U) is a copula. Then, the bivariate copula C is defined as being exchangeable if

$$C = \tilde{C}. \quad (5.3)$$

Note that the definition of exchangeability presented above can be generalized to d -variate copulas. This generalization is based on the concept of a symmetry and in particular the group of symmetries under the operation of composition. This group includes permutations (see ?, pp. 30-31 for details), on which the definition above is based.

In the bivariate case, on which this chapter focuses, this translates to a copula being exchangeable if

$$C = C^T, \quad (5.4)$$

where $C^T(u, v) = C(v, u)$ for all $(u, v) \in [0, 1]^2$, i.e. C^T is the copula that results from permuting C 's arguments, and C^T is also called the transpose of C (see ?, p. 31).

Thus, a slightly different form of definition (5.4) one often finds in the literature is that a copula is exchangeable if (see e.g. ?, p. 92)

$$C(u, v) = C(v, u) \quad \forall (u, v) \in [0, 1]^2. \quad (5.5)$$

Asymmetry is then defined via non-symmetry, i.e. when equation (5.5) is not fulfilled for some $(u, v) \in [0, 1]^2$, the copula C is non-exchangeable, which, as explained above, is defined as being equivalent to asymmetric in this chapter (see ?, p. 92).

From now on, this chapter (i.e. chapter 5) focuses on continuous random variables, meaning that the random variables, whose univariate cumulative distribution functions enter the copula, are continuous. For instance, for the setting, which has been defined for equation (5.1), this means that X and Y are restricted to being continuous random variables for the rest of chapter 5. Recall from Sklar's theorem that any cumulative distribution function can be expressed in terms of a copula and its univariate marginals and that the copula C is unique if the marginals F and G are continuous. Thus, as described in equation (5.1), the cumulative distribution function $H(x, y)$ can be written as $H(x, y) = C(F(x), G(y))$ for all $(x, y) \in \mathbb{R}^2$ (for details and definitions see section 5.2.2). From equation (5.5), we know that the copula C is then, symmetric if $C(F(x), G(y)) = C(G(y), F(x))$ for all $(F(x), G(y)) \in [0, 1]^2$, i.e. $F(x)$ corresponds to u in equation (5.5), whereas $G(y)$ corresponds to v . Then, when this equation is not fulfilled for some $(F(x), G(y)) \in [0, 1]^2$, C is non-exchangeable, i.e. asymmetric, and the variables X and Y are asymmetrically dependent, which usually occurs, when there is a causal relationship between these variables (see ?, p. 92). As an example, ? (p. 92) give the monthly average flow rate (in m^3/s) of the rivers Inn (X) and Danube (Y), which are measured at the stations in Schärding and Nagymaros, respectively. The notation shown in brackets in the previous sentence was chosen since the monthly average flow rate of the rivers Inn and Danube are considered examples for the random variable X and Y , respectively, which have been introduced above. Note that the flow rate in m^3/s determined at a given point, measures river discharge, also called river flow, which is the volume of water moving through a river channel (see ? and ?). Moreover, river discharge is considered a climate variable by ?. Thus, ? investigate the dependence structure of two climate variables, namely (river) discharge of the Inn and (river) discharge of the Danube. The Inn flows into the Danube after it has passed Schärding and afterwards, the Danube passes Nagymaros. For illustration, this situation is depicted in the stylized map shown in figure 5.1:

Thus, high discharge of the Inn is likely to cause high discharge of the Danube, but not vice versa, where this circumstance describes a causal relationship. Consequently and as also pointed out by ? (p. 92), one would expect to observe high discharge of the Danube, when the discharge of the Inn is high, but not necessarily the other way round.

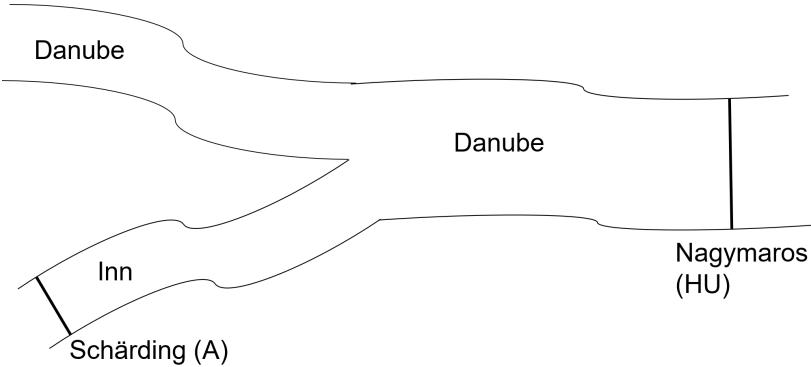


FIGURE 5.1: Stylized map, which shows the location of the Schärding and Nagymaros measurement stations. Adapted from ?, p. 92.

5.3 Theoretical background on measures of asymmetry in a copula and their application to hydrological data

5.3.1 Measures of asymmetry in a copula

The question arises how to measure the degree of asymmetry in a copula. ? introduce measures of asymmetry in an axiomatic way. Specifically, they define the term measure of asymmetry for the class of (bivariate) copulas \mathcal{C} as a function $\mu : \mathcal{C} \rightarrow \mathbb{R}_0^+$ that satisfies the following axioms:

- (A1) There exists $L \in \mathbb{R}_0^+$ such that: $\mu(C) \leq L \quad \forall C \in \mathcal{C}$,
- (A2) $\mu(C) = 0$ if and only if $C = C^T$, i.e. C is symmetric,
- (A3) $\mu(C) = \mu(C^T) \quad \forall C \in \mathcal{C}$,
- (A4) $\mu(C) = \mu(\check{C}) \quad \forall C \in \mathcal{C}$,
- (A5) $(\mu(C_n))_{n \in \mathbb{N}}$ converges to $\mu(C)$ if $(C_n)_{n \in \mathbb{N}} \in \mathcal{C}$ and $C \in \mathcal{C}$,

where \check{C} is the so-called survival copula of C with $\check{C}(u, v) = u + v - 1 + C(1-u, 1-v)$ for all $(u, v) \in [0, 1]^2$ (for further details on survival copulas, see ?, pp. 32-33).

Axiom (A1) ensures that the measure $\mu(C)$ is bounded which will help interpretation (see below), whereas axiom (A2) makes sure that when the measure $\mu(C)$ assumes the value zero, it can be concluded that the copula is symmetric (and vice versa). The remaining axioms (A3) to (A5) reflect certain invariance properties of the exchangeability of a random pair, which translates to the symmetry of the corresponding copula if exchangeability is given (see ? for details). ? then show that the function $\mu_p : \mathcal{C} \rightarrow \mathbb{R}_0^+$ is a measure of asymmetry for all $p \in [1, \infty]$, where for $p \in [1, \infty)$

$$\mu_p(C) = \left(\int_0^1 \int_0^1 |C(u, v) - C(v, u)|^p du dv \right)^{1/p} \quad (5.6)$$

and for $p = \infty$

$$\mu_\infty(C) = \max_{(u,v) \in [0,1]^2} |C(u, v) - C(v, u)|. \quad (5.7)$$

Moreover, they point out that equation (5.6) and equation (5.7) describe a large class of asymmetry measures.

Note that the class of asymmetry measures described above has been discussed by ? and ? before ?, but without basing it on axioms.

An alternative measure of asymmetry has been introduced by ? (pp. 96-97). Its main advantage is that the corresponding estimate can be calculated more easily than in case of μ_p for $p \neq 2$ and $p \neq \infty$. Since in this chapter, the focus is not on the estimation procedure, this is not explored further. Moreover, other alternative measures can be found in the literature, which were discussed shortly by ? (p. 94). Following their work however, these measures are not explored further here. More recently, additional measures of asymmetry have been proposed (see e.g. ?). Since the class of measures described above is relatively large, is based on a broad axiomatic approach, still appears to be relevant in practice (see ?) and has been studied relatively extensively in the context of Archimax copulas, this class of measures appears to be the most important one for the questions this chapter addresses, wherefore the focus is set accordingly.

5.3.2 Bounds for measures of asymmetry in a copula and the resulting possibility for normalization

? and ? show that $\frac{1}{3}$ is an upper bound for the asymmetry measure defined by (5.7), i.e. for every copula C the following inequality holds:

$$\mu_\infty(C) \leq \frac{1}{3} =: \eta_\infty. \quad (5.8)$$

Moreover, they show that inequality (5.8) holds with equality for some $C \in \mathcal{C}$.

Also for asymmetry measures defined by equation (5.6), an upper bound exists for every copula C (see ?). Specifically, for $p \in [1, \infty]$ it holds that

$$\mu_p(C) \leq \left(\frac{2 \cdot 3^{-p}}{(p+1)(p+2)} \right)^{1/p} =: \eta_p. \quad (5.9)$$

? conclude that thus for $p \in [1, \infty]$, an asymmetry measure $\mu_p(C)$ can be normalized as follows:

$$\frac{\mu_p(C)}{\eta_p} =: \nu_p(C) \quad (5.10)$$

and therefore $\nu_p : \mathcal{C} \rightarrow [0, 1]$, i.e. the range of ν_p is the interval $[0, 1]$.

Note however that it is unknown if for $p \in [1, \infty)$ the corresponding upper bound is attained. It is true that ? show that for every measure of asymmetry $\mu : \mathcal{C} \rightarrow \mathbb{R}_0^+$ there is a copula $\overset{\circ}{C}_\mu \in \mathcal{C}$ and a constant $L_\mu \in \mathbb{R}_0^+$ such that $\mu(C) \leq L_\mu$ for all $C \in \mathcal{C}$ and $\mu(\overset{\circ}{C}_\mu) = L_\mu$, but it is not clear if $\eta_p = L_\mu$ for $\mu = \mu_p$ and $p \in [1, \infty)$. Thus, ν_∞ takes values in the entire interval $[0, 1]$, but for $p \in [1, \infty)$ it can only be said that $\nu_p \in [0, 1]$ (see ?, p. 97).

5.3.3 Example of an application of a selection of the asymmetry measures described above to hydrological data

? (pp. 95-98) use two of the asymmetry measures described above to asses the asymmetry in the dependence of the flow rates of two rivers. Specifically, they use data on monthly average flow rates (in m^3/s) of the rivers Inn and Danube, which were determined at the Schärding and Nagymaros measurement stations in the period from 1936 to 1991, i.e. over a period of 55 years (see section 5.2.3 for an explanation why these two random variables are expected to be asymmetrically dependent). From this data, they obtain 660 values (u_i, v_i) , where for each $i \in \{1, \dots, n\}$, $(u_i, v_i) = (F_n(x_i), G_n(y_i))$, with n representing the sample size and x_i denoting the realization of the i th monthly average flow rate of the Inn X_i . Analogously, y_i represents the realization of the i th monthly average flow rate of the Danube Y_i . Furthermore, F_n and G_n denote the marginals of the (rescaled) empirical cumulative distribution function H_n corresponding to H , where H_n is calculated from the realizations of the sample $(X_1, Y_1), \dots, (X_n, Y_n)$ and H represents, as introduced above,

the theoretical joint cumulative distribution function of (X_i, Y_i) for all $i \in \{1, \dots, n\}$ with unique copula C . Therefore, for each $i \in \{1, \dots, n\}$, $(u_i, v_i) = (F_n(x_i), G_n(y_i))$ can be seen as a pseudo-observation from C . These pseudo-observations are shown in figure 5.2, which thus visualizes the dependence structure in the data (see ?, p. 43).

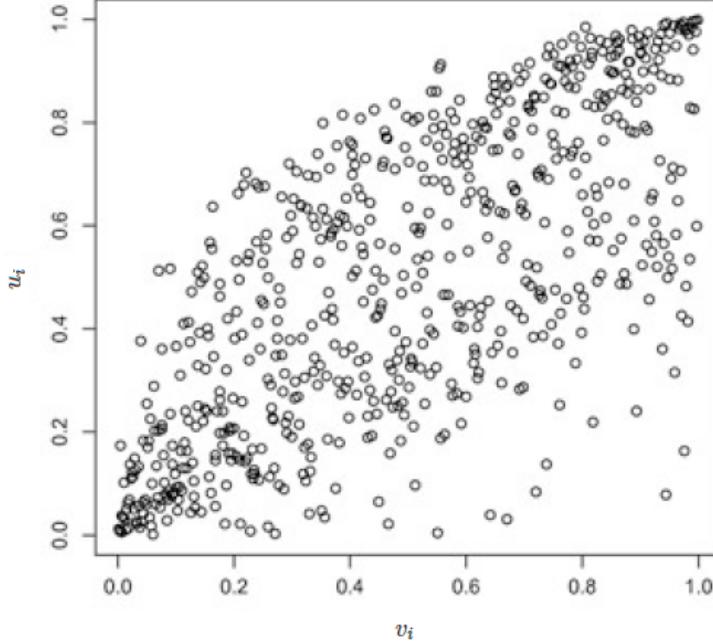


FIGURE 5.2: Scatter plot of the pseudo-observations $(u_1, v_1), \dots, (u_n, v_n)$ from C , where u_i is calculated from data on the Inn and v_i is obtained from data on the Danube ($\forall i \in \{1, \dots, n\}$). Adapted from ?, p. 93.

As can be seen from this figure, the visualization of the dependence structure supports the notion of asymmetry in the dependence between X and Y , where this conclusion has also been drawn by ? (p. 92).

Based on the insight that $(u_1, v_1), \dots, (u_n, v_n)$ can be seen as pseudo-observations from C , ? (p. 96) calculate the empirical copula \hat{C}_n as an estimate of C . Then, they use \hat{C}_n to calculate estimates of $\mu_2(C)$ and $\mu_\infty(C)$ (for details on their estimation procedure, see ?, pp. 95-96). These estimates are then standardized using the upper bounds for the corresponding measure, namely $\eta_2 = 1/\sqrt{54}$ and $\eta_\infty = 1/3$. Afterwards, these standardized estimates are used as estimates of $\nu_2(C)$ and $\nu_\infty(C)$ as defined by equation (5.10). This results in

$$\nu_\infty(\hat{C}_n) = 3\mu_\infty(\hat{C}_n) \approx 0.127$$

and

$$\nu_2(\hat{C}_n) = \sqrt{54}\mu_2(\hat{C}_n) \approx 0.083.$$

As explained in section 5.3.2, the values ν_∞ and ν_2 assume lie in the interval $[0, 1]$. For ν_∞ we know additionally that it takes values in the entire interval $[0, 1]$. This is not clear for ν_2 , since the bound η_2 might not be sharp (see section 5.3.2 as well as ?, p. 97), but following the assessment of ? (p. 97) that η_2 still is reasonable for standardization in practice, the estimates of both $\nu_\infty(C)$ as well as $\nu_2(C)$ appear to be relatively small. Since a value of 0 indicates symmetry (see section 5.3), this does not provide strong support for the notion of asymmetry. A similar result is obtained when the series is de-trended to avoid the time-invariant dependence measured between the two variables being distorted by serial dependence, which is possible since the data used shows time trends (see ?, p. 97). Still, the estimates ? (p. 98) obtain from this approach are relatively small. Specifically, they obtain

$$\nu_\infty(\hat{C}_n) = 3\mu_\infty(\hat{C}_n) \approx 0.091$$

and

$$\nu_2(\hat{C}_n) = \sqrt{54}\mu_2(\hat{C}_n) \approx 0.063.$$

Positive dependence supports symmetry in the sense that certain positive dependence properties can induce upper bounds on asymmetry measures which are relatively close to zero (see ?, p. 95 for details) and ? (p. 98) find that the two variables investigated here are clearly positively correlated. For instance, the sample value of Kendall's tau as reported by ? (p. 98) equals approximately 0.525 for the raw data and it assumes a similar value, namely 0.548, for the de-trended data. Thus, the small values of the asymmetry measures reported above might not be entirely surprising. In order to investigate if the estimates reported above are significantly greater than zero, and thus the dependence between the two variables can be considered asymmetric, ? (pp. 98-101) conduct formal statistical tests for both the raw and the de-trended data. Since the null hypothesis of symmetry was rejected in all cases, approaches to model asymmetry, such as Archimax copulas, should be considered for this example of hydrological data.

5.4 Archimax copulas

5.4.1 Definition

Since Archimax copulas are only symmetric under a condition that will be made precise in section 5.4.3, they can be used for modelling asymmetric dependence between random variables. Archimax copulas have been introduced by ?. Following ?, ?, as well as ?, the following bivariate function is a copula and is called Archimax copula:

$$C_{\phi,A}(u, v) = \phi^{[-1]} \left[(\phi(u) + \phi(v)) A \left(\frac{\phi(u)}{\phi(u) + \phi(v)} \right) \right] \quad \forall (u, v) \in [0, 1]^2 \quad (5.11)$$

with the conventions $0/0 = \infty/\infty = 0$ (see ?) and where

- i) $A : [0, 1] \rightarrow [1/2, 1]$ is a convex function such that $\max(t, 1-t) \leq A(t) \leq 1$ for all $0 \leq t \leq 1$,
- ii) $\phi : [0, 1] \rightarrow [0, \infty]$ is a convex, strictly decreasing and continuous function for which it holds that $\phi(1) = 0$ and with pseudo-inverse

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t), & 0 \leq t \leq \phi(0), \\ 0, & \text{otherwise.} \end{cases} \quad (5.12)$$

Functions A as defined in i) are called dependence functions (see ? and ?) or Pickands dependence functions (see ?, pp. 101-102 and p. 105) and functions ϕ as defined in ii) are called Archimedean generators (see ?) or just generators (see ?).

5.4.2 Archimedean and extreme value copulas as special cases of Archimax copulas

There are two special cases of Archimax copulas, namely Archimedean copulas and extreme value copulas (see e.g. ? and ?). In the following, it is described under which conditions these special cases arise:

When $A \equiv 1$, Archimax copulas can be written as

$$C_{\phi,A}(u, v) = C_{\phi}(u, v) = \phi^{[-1]} [\phi(u) + \phi(v)] \quad \forall (u, v) \in [0, 1]^2, \quad (5.13)$$

where the equation above corresponds to a bivariate Archimedean copula as described by ? (pp. 110-112). Note that Archimedean copulas are symmetric (see also ?, p. 103).

In contrast, for $\phi(t) = \log(1/t)$, an Archimax copula can be written in the general form of a bivariate extreme value copula (see ?), i.e. as

$$C_{\phi,A}(u,v) = C_A(u,v) = \exp \left[\log(uv) A \left(\frac{\log(u)}{\log(uv)} \right) \right] \quad \forall (u,v) \in (0,1)^2. \quad (5.14)$$

Extreme value copulas are only symmetric under a condition, which will be made precise in section 5.4.3. Note that this is in contrast to Archimedean copulas.

Well-known examples of extreme value copulas are the independence copula $\Pi(u,v) = u \cdot v$ as well as the comonotonicity copula $M(u,v) = \min\{u,v\}$ (see ?). The dependence structure the independence copula induces corresponds to the one its name suggests (see ?, p. 63 as well as p.11 for details). Also the comonotonicity copula obtains its name from the dependence structure it induces, where comonotonicity is the strongest concept of positive dependence (see ?, pp. 64-65 as well as p.11 for details).

5.4.3 Asymmetry

While the majority of classical copulas is symmetric (see ?, p. 92), according to ?, an Archimax copula $C_{\phi,A}$ as defined in (5.11) is symmetric if and only if

$$A(t) = A(1-t) \quad \forall t \in [0,1]. \quad (5.15)$$

Thus, when $A(t) \neq A(1-t)$ for some $t \in [0,1]$, Archimax copulas are asymmetric. Note that the same holds for the special case of extreme value copulas (see also ?, p. 102).

However, the degree of asymmetry that can be modelled by Archimax copulas is limited. This is shown by ?, who derive an upper bound for the degree of asymmetry in an Archimax copula (defined as in (5.11)) as measured by the normalized version of the asymmetry measure given in equation (5.7). Specifically, they show that for every Archimax copula $C_{\phi,A}$ it holds that

$$\nu_{\infty}(C_{\phi,A}) \leq 3 \left(\sup_{s \in (0,\infty)} \left| \phi^{[-1]}(s) - \phi^{[-1]} \left(\frac{5s}{4} \right) \right| \right). \quad (5.16)$$

Note that the upper bound in (5.16) depends on the generator ϕ . Thus, by choosing a different generator ϕ , the upper bound for the degree of asymmetry in an Archimax copula can be changed.

In addition, ? find that the upper bound in (5.16) is reached by $\nu_{\infty}(C_{\phi,A})$ if and only if $A \in \{A_1, A_2\}$, where A_1 and A_2 are defined as follows:

$$A_1(t) = \max \left(1-t, \frac{t+1}{2} \right), \quad (5.17)$$

$$A_2(t) = \max \left(t, \frac{2-t}{2} \right). \quad (5.18)$$

Since the upper bound presented in (5.16) depends on the generator ϕ , the interpretation of its size depends on ϕ as well. Thus, for getting an impression of the degree of asymmetry that is possible for Archimax copulas, it is useful to calculate the values the upper bound for $\nu_{\infty}(C_{\phi,A})$ assumes for different examples of Archimedean generators ϕ . This is done by ? (p. 615). The results are presented in table 5.1.

Recalling that the measure ν_{∞} takes values in the entire interval $[0,1]$ and that the upper bounds in table 5.1 are only reached if the dependence function A takes one of the two forms described by the equations

TABLE 5.1: Approximated sharp upper bounds for the measure of asymmetry $\nu_\infty(C_{\phi,A})$ of Archimax copulas $C_{\phi,A}$ with different Archimedean generators ϕ . Adapted from ?, p. 615.

Generator $\phi(t)$	$\max_A(\nu_\infty(C_{\phi,A}))$
$t^{-1} - 1$	0.167
$\log(\frac{3-t}{2t})$	0.274
$(-\log(t))^3$	0.082
$-\log(\frac{e^{-t}-1}{e^{-1}-1})$	0.214
$1 - t$	0.600

(5.17) and (5.18), the upper bounds appear to be relatively small for all of the generators shown in table 5.1 except for $\phi(t) = 1 - t$.

Since extreme value copulas form a special case of Archimax copulas, the question arises if Archimax copulas offer more flexibility in modelling asymmetric dependence than extreme value copulas. ? show that the upper bound for $\nu_\infty(C_A)$, i.e. the upper bound for ν_∞ applied to an extreme value copula as defined by equation (5.14), equals $3 \cdot \frac{4^4}{5^5} \approx 0.246$. Specifically, they show that for every extreme value copula C_A it holds that

$$\nu_\infty(C_A) \leq 3 \cdot \frac{4^4}{5^5}. \quad (5.19)$$

Moreover, Archimax copulas can also be symmetric, in which case $\nu_\infty(C_{\phi,A})$ equals 0 (see axiom A2 in section 5.3), which is the smallest value $\nu_\infty(C)$ can assume for any copula C . Combined with the fact that the maximum of the values presented in table 5.1 equals 0.600, which is larger than 0.246, this shows that the range of asymmetry of Archimax copulas is wider than that of extreme value copulas, which corresponds to the conclusion drawn by ?. Consequently, Archimax copulas in their general form are more flexible in modelling asymmetric dependence structures than one of their special cases, the extreme value copulas. This might be required in some applications.

For further illustration of the asymmetry property of Archimax copulas, scatter plots for examples of several asymmetric Archimax copulas as given by ? (p. 107) are presented in figure 5.3. Specifically, these scatter plots show samples of size 2,000 drawn from Archimax copulas generated based on Gumbel's asymmetric Pickand's dependence function and different Archimedean generators. According to ? (p. 102), the Gumbel family represents a typical example for symmetric extreme value copulas. It can be made asymmetric using Khoudraji's device, which relies on the parameters $\lambda, \kappa \in (0, 1)$ and produces an asymmetric extreme value copula if $\lambda \neq \kappa$ (see ?, p. 102 and references therein for further details). An Archimax copula with a dependence function corresponding to the one of such an asymmetric extreme value copula is also asymmetric and thus Khoudraji's device can be used for generating asymmetric Archimax copulas (see equation (5.15) and the accompanying explanations as well as ?, p. 105). In the examples, the parameters λ and κ are set to 0.5 and 0.7, respectively, while the parameter of the Gumbel's dependence function varies from left to right, i.e. it assumes three different arbitrary values (see ?, p. 102 and p. 107 and references therein for details). For ϕ , the Clayton (top), the Frank (middle) and the Joe (bottom) Archimedean generators are selected, where the corresponding parameter is chosen such that the Kendall's tau of the Archimedean copula with generator ϕ equals 0.5. The scatter plots corresponding to these asymmetric Archimax copulas, which are shown in figure 5.3, suggest that the degree of asymmetry is relatively small in the examples chosen.

5.5 Conclusion

Compound events are caused by at least two climate variables and the dependence between these climate variables can be asymmetric. This chapter laid the theoretical foundation for modelling compound events

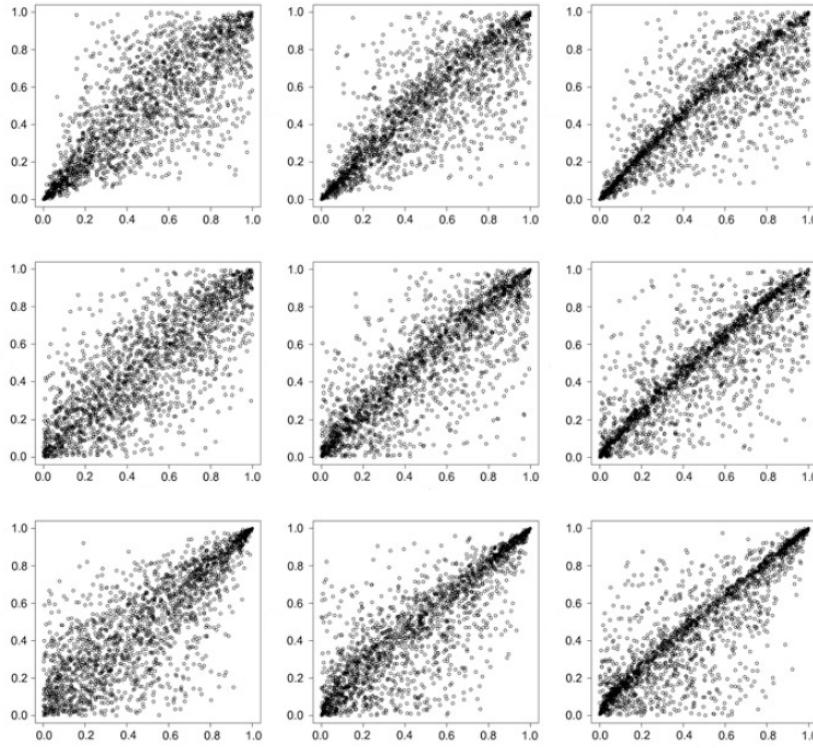


FIGURE 5.3: Samples of size 2,000 drawn from different Archimax copulas with Gumbel's asymmetric Pickands dependence function, where Khoudraji's device with $\lambda = 0.5$ and $\kappa = 0.7$ is used to induce asymmetry. The parameter of the Gumbel's dependence function varies from left to right, i.e. it assumes three different arbitrary values (see ?, p. 102 and p. 107 and references therein for details). For ϕ , the Clayton (top row), Frank (middle row), and Joe (bottom row) Archimedean generators are used, where the corresponding parameter is chosen such that Kendall's tau of the Archimedean copula with generator ϕ equals 0.5. In every of the scatter plots presented above, the x-axis shows the values of the first component of the realizations drawn from the respective Archimax copula, while the y-axis shows the value of the respective second component. Adapted from ?, p. 107.

characterized by asymmetric dependence. Specifically, by reviewing the relevant literature, this chapter investigated how to recognize asymmetric dependence between two random variables and if and to what extent this asymmetric dependence can be modelled via Archimax copulas. For doing so, first, theoretical background about copulas in general was given, which allowed for making the concept of asymmetry in the dependence structure of random variables precise. Then, it was explained how this asymmetry can be measured based on the axiomatic approach developed by ?. Thus, the foundation for estimating the degree of asymmetry in the dependence structure of two random variables from real-world data as well as for measuring the theoretical degree of asymmetry in Archimax copulas was laid. In the next step, bounds for these asymmetry measures were presented, which allowed for normalization and thus supports interpretation of the values these measures assume. For illustration, an application of a selection of the asymmetry measures to hydrological data was presented. Finally, the concept of Archimax copulas was explained, which form a relatively broad class of copulas in the sense that they include the well-known Archimedean and extreme value copulas as special cases. Since Archimax copulas are only symmetric under a condition which was stated in section 5.4.3, they can be used for modelling asymmetric dependence between random variables. However, the degree of asymmetry in Archimax copulas is limited by an upper bound presented in section 5.4.3. The specific value this upper bound assumes depends on the generator ϕ and the upper bound is reached if and only if the dependence function A takes one of two specific forms (see section 5.4.3 for details). However, Archimax copulas are more flexible in modelling asymmetric dependence structures than one of their special

cases, the extreme value copulas, which can also be asymmetric (see section 5.4.3 as well as ?, pp. 101-103 and ? for details).

Due to its scope, the focus of this chapter was on Archimax copulas. However, these copulas do not represent the only approach for modelling asymmetric dependence. In future work, one could compare different approaches for modelling this type of dependence structure, for which chapter 5.4, in ? might be a good starting point. Moreover, as already explained in section 5.1, this chapter focused on the asymmetry property of Archimax copulas. This however, is not the only property of these copulas. For example, one property, which might be interesting in the context of the climate crisis, is that Archimax copulas can be constructed with a predetermined extreme value attractor (see ? for details). Also this might be worth further investigation, for which ? as well as ? could serve as starting points. In terms of asymmetry measures, this chapter focused on a certain class of asymmetry measures, which is in line with the axiomatic approach established by ? (see section 5.3). Investigating alternative measures in more detail and comparing them to the ones explained above might be another interesting question one could address in the future.

6

Teleconnections North Atlantic Oscillation

Author: Pamela Sorelle Tueam

Supervisor: Henri Funk

Suggested degree: Bachelor

6.1 Abstract

The concept of teleconnections in climate science refers to climate anomalies or patterns that are related across large distances, often thousands of kilometers apart. Teleconnections represent atmospheric interactions that link weather and climate conditions in one region of the globe with those in another, often through the movement and behavior of large-scale atmospheric wave patterns (?). These connections are crucial for understanding regional climate variations, predicting weather patterns, and assessing climate change impacts. One prominent example of a teleconnection pattern in the Northern Hemisphere is the North Atlantic Oscillation (NAO). This oscillation plays a significant role in influencing the climate over the North Atlantic region. Its impacts are observed in various climate elements, including temperature, wind speed and direction, storm and precipitation, as well as ocean and sea ice dynamics.

This chapter explores various methods for defining and analyzing the North Atlantic Oscillation (NAO), following the framework established by ?. These methods include one-point correlation map, Empirical Orthogonal Function (EOF) analysis, and cluster analysis. Additionally, we present the long-term trends of the NAO using the NAO Index. Through a detailed examination of these approaches, we aim to provide a better understanding of the spatial and temporal characteristics of the NAO. Such understanding is important for improving climate models, enhancing weather prediction accuracy, and developing robust climate change adaptation and mitigation strategies. Furthermore, by assessing the NAO's influence on temperature, wind patterns, storm activities, and oceanographic conditions, this chapter contributes to a deeper comprehension of how large-scale atmospheric phenomena shape our climate system.

6.2 Introduction

The North Atlantic Oscillation (NAO), one of the most prominent and recurrent teleconnections over the Northern Hemisphere, is characterized by a redistribution of atmospheric mass at sea level between the Icelandic Low and the Azores High (?). It dictates climate variability from the eastern seaboard of the United States to Siberia and from the Arctic to the subtropical Atlantic, especially during boreal winter. Fluctuations from one phase of the NAO to another result in significant variations in the average wind speed and direction over the Atlantic, as well as in the transport of heat and moisture between the Atlantic and neighboring continents. This also affects the intensity and number of storms, their paths, and associated weather conditions (?). Especially during the months of December to March, the NAO is responsible for much variability of weather over the North Atlantic region. Thus, the NAO influences weather and climate

conditions across the North Atlantic region and beyond, affecting winter temperatures and storm tracks. In this chapter, our primary focus will be on exploring the spatial and temporal characteristics of the NAO. We will begin by discussing common methods used to define it, such as the one-point correlation map, which helps to identify the NAO by regions of positive and negative correlations. Additionally, we will employ the EOF analysis to discern the dominant patterns of variability associated with the NAO, and utilize the cluster analysis to identify distinct atmospheric patterns with similar pressure characteristics. Furthermore, we will analyze the NAO Index to observe its variations over long periods. Finally, we will investigate the impacts of the NAO on various climate factors including temperature, wind speed and direction, precipitation patterns, storm activities, oceanic conditions, and sea ice dynamics.

6.3 The spatial and temporal structure of the NAO

Understanding NAO requires a comprehensive understanding of its spatial and temporal characteristics. To analyze these aspects, we will review established methods commonly employed for defining or identifying NAO patterns. These methods include the one-point correlation map, the EOF analysis for its spatial structure, as well as cluster analysis. For its temporal structure, we will focus on the NAO index. These methods are used in the paper of (?) and the data used in his paper go up to 2006. The idea is to reproduce these on new data up to 2024 and see if the NAO patterns remain similar.

6.3.1 Data

Monthly geopotential height at 500 hPa (Z500) and mean sea level pressure (MSLP) data are used in this chapter. The monthly geopotential height at 500 hPa were obtained from the ERA5 reanalysis dataset for global climate and weather (?). Mean sea level pressure data used here were also sourced from the ERA5 reanalysis (?) dataset. For both monthly geopotential height and mean sea level pressure, only data from the Northern Hemisphere specifically within the coordinates 20°-70°N and 90°W-40°W, were extracted. Considering that the NAO is responsible for much variability of weather in boreal winter over the Northern Hemisphere, the period from December to March (DJFM) was selected for this analysis. The time series used consist of 84 DJFM seasons from 1940-2024. The data used to examine the NAO Index were collected from the National Center for Atmospheric Research (NCAR). This dataset includes station-based NAO index data of 159 DJFM seasons spanning from 1864 to 2023 (?). To examine the pattern of variability we need to compute the anomalies which are calculated by subtracting time means, computed for each calendar month from the data for individual year-months. These anomalies form the basis for all methods described in this chapter.

6.3.2 One-point correlation maps

One way to define the NAO is through conceptually simple one-point correlation map, identifying the NAO by regions of maximum negative correlation over the North Atlantic. For these one-point correlation we will focus on the methods described by ? and by ? in their respective papers.

One-point correlation correlates a variable at a specific location with the same variable at other locations. Given a field of data, we compute the correlation coefficient between the base point - referred to as the center of action - and each other points. The center of action is a specific location defined by latitude and longitude. The resulting correlation values generate a map that identifies the NAO by regions of positive and negative correlation. In this analysis, we use monthly geopotential heights data at 500 hPa for winter season December-March from 1940 to 2024. Anomalies are calculated by subtracting geopotential height time means, computed for each calendar month from the data for individual geopotential height year-months (? , 790).

The correlation coefficient for a given center of action is calculated using the Pearson's linear correlation coefficient at every point as follows:

$$r_i = \frac{\text{Cov}(X_i, Y_i)}{\sqrt{\text{Var}(X_i)\text{Var}(Y_i)}}$$

where X and Y are time series of contemporaneous observations, $[(x_i, y_i); i = 1, 2, \dots, N]$, where we fix one of the data points, say, X, at the center of action and compute the correlation over all Y. Covariance and variance are denoted as Cov and Var, respectively (?, 3721).

Figure 6.1 illustrates the corresponding one-point correlation map for the NAO pattern. The reference point that has been chosen is located in Iceland at 65°N and 18°W . The correlations are calculated for all points belonging to the Northern Hemisphere ($20\text{--}70^{\circ}\text{N}$, $90^{\circ}\text{W}\text{--}40^{\circ}\text{E}$). As we can see on the figure the map is characterized by more or less elliptical regions where correlation coefficients can be positive or negative, surrounding each base grid point. The size of these regions provides an indication of the horizontal scale of the anomalies, particularly near the base grid points (?, 790). The identifying dipole patterns located at the Icelandic Low and Azores High, reflect the NAO signature based on the chosen reference point. The color gradients on the figure indicate the strength of the correlation, with blue representing negative correlations and yellow to red representing positive correlations.

As implications, regions with strong positive correlations experience atmospheric conditions similar to Iceland (the center of action). These areas may share similar temperature anomalies. For instance, if the region of the center of action experiences warmer-than-average temperatures during high NAO phase, positive correlated regions might also experience similar warmth temperatures. We will also have similar precipitation patterns (if it is drier than the nearby regions may also be similar dry). Conversely, the region with negative correlations such as the anticorrelation center at Azores, will exhibit opposite conditions, including temperature contrasts (warmer vs. cooler) and differing precipitation patterns (wetter vs. drier).

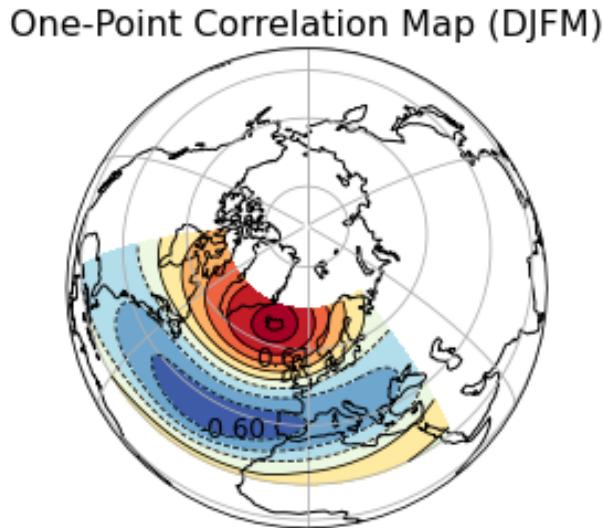


FIGURE 6.1: One-point correlation map of 500 hPa geopotential height for boreal winter (December–March) over 1940–2024. The reference point is 65°N , 18°W . Negative correlation coefficients are in blue and the lines are dashed.

6.3.3 EOF Analysis

The second method is the Empirical Orthogonal Function (EOF) analysis, also known as Principal Component Analysis (PCA). EOF analysis is commonly used to identify and analyze dominant patterns of variability in spatial and temporal datasets. The EOFs of a dataset are simply the eigenvectors of the covariance matrix of the dataset, which provide maximal data compression (?). The method reduces a dataset with many

variables to a smaller set of new variables, which are linear combinations of the original ones, chosen to capture the maximum possible variability of the original data (?). Given multiple observations of a ($K \times 1$) data vector \mathbf{x} , PCA finds ($M \times 1$) vector \mathbf{u} , which are linear combinations of the elements of \mathbf{x} and capture most of the information from the original data (?). The steps by performing the EOF are as follows:

- Collection of the dataset: We used the mean sea level pressure (MSLP) data for boreal winter DJFM over 84 years from 1940 to 2024.
- Calculation of the anomalies: Anomalies are calculated by subtracting the MSLP time means, computed for each calendar month from the data for individual MSLP year-months.
- Formation of the Covariance Matrix as we want to analyze the variance and patterns in the data.
- Performance of the eigenvalue decomposition: The covariance matrix is decomposed using eigenvalue decomposition to obtain the eigenvalues and eigenvectors. The eigenvectors represent the spatial patterns, and the eigenvalues indicate the amount of variance explained by each EOF.
- The first principal component, u_1 is obtained as the projection of the anomalies vector onto the first eigenvector \mathbf{e}_1 $\mathbf{u}_1 = \mathbf{e}_1^T = \sum_{k=1}^K e_{k,1} x'_k$. This first EOF/PC explains the most variance, followed by the second, and so on (?).

Computing the EOFs allows for comparison with the results of the one-point correlation analysis. While correlation analysis focuses on the variability relative to the primary center of action of specific teleconnections, EOF analysis is not centered on specific points. The loading at any two points are not solely determined by the temporal correlation between the time series at those points (? S 3724). Furthermore, EOF analysis is not explicitly designed to highlight regional patterns of strong correlation. Instead, EOF analysis sequentially identifies the leading patterns that explain the most variance within the analyzed dataset (? S 3724). Therefore, values of the same sign at two different spatial points in an EOF do not necessarily imply a significant correlation between those two points.

Figure 6.2 illustrates the leading eigenvectors of the cross-covariance matrix calculated from seasonal (4-month average) MSLP anomalies in the North Atlantic sector (20° - 70° N; 90° W- 40° E). Since the eigenvectors are, by definition, structured to explain maximum variance, it is expected that the center of actions of the leading EOFs will coincide with the regions of strongest variability. Although EOF patterns are not exactly identical to the one-point correlation map, there is a noticeable similarity between the two. Both correlation map and EOF analysis consistently illustrate the distinct behavior of variability across different time scales. During the winter season (December-March), the NAO accounts for more than one-third (35.8%) of the total variance in SLP over the North Atlantic and appears with a slight northwest-to-southeast orientation.

6.3.4 Cluster Analysis

Another method to examine the dynamical signature of interannual variability in the North Atlantic is the cluster analysis. This non-linear approach groups similar data points into clusters based on their similarities, and is used here to identify distinct atmospheric patterns with similar pressure characteristics. It is applied to 84 years of monthly MSLP data from December to March using procedures based on a clustering algorithm. The algorithm groups in this case the monthly MSLP data into a small number of representative states (or regimes) according to an objective criterion of similarity. Given a prescribed number of clusters k , the goal of the algorithm is to find a partition P of the data points into k clusters C_1, C_2, \dots, C_k that minimizes the sum of the variances within the clusters, expressed as (?):

$$W(P) = \sum_{j=1}^k \sum_{x \in C_j} d^2(X, Y_j)$$

where Y_j is the centroid cluster C_j and $d(X, Y)$ represents the Euclidean distance between two points X and Y as a measure of similarity. The global minimum of the function $W(P)$ therefore corresponds to a partition that achieves the best separation of data points. The dynamic cluster algorithm defines iterative partitions $P^{(n)}$ for which $W(P^{(n)})$ decreases with n and eventually converges to a local minimum which generally differs from the global one. This algorithm thus provides the optimal k number of clusters to be retained (? S 1241). The steps by performing the cluster Analysis are:

EOF1 MSLP DJFM 1940-2024: 35.8

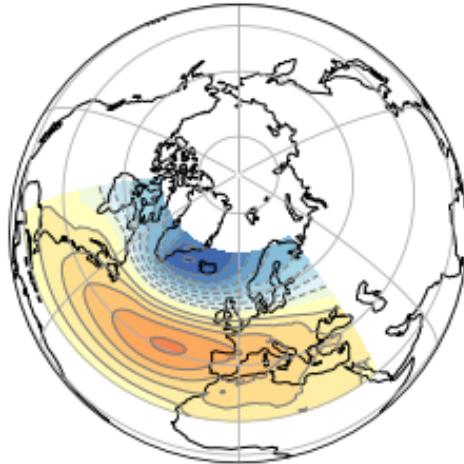


FIGURE 6.2: Leading empirical orthogonal function (EOF 1) of the seasonal mean sea level pressure anomalies in the North Atlantic sector (20° – 70° N, 90° W– 40° E), and the percentage of the total variance it explains. The data cover the period of 1940–2024.

- Data collection and anomalies calculation.
- Selection of a clustering algorithm: In this case the k-means clustering is used.
- Determination of the optimal number of clusters.
- Application of the algorithm: Assign each monthly SLP data to a cluster with each cluster representing a distinct atmospheric regime or pattern.

The clustering algorithm applied over the Atlantic domain (20° – 70° N; 90° W– 40° E) identifies four winter climate regimes Figure 6.3. The clustering partition yields the optimal $k = 4$ number of significant winter climate regimes subsequently represented by the cluster centroid. Two of these clusters correspond to the negative and positive phases of the NAO and are characterized by a zonally elongated meridional pressure dipole between the Icelandic Low and the Azores High. The third regime exhibits a zonal pressure dipole between Greenland and Scandinavia (the “Blocking” regime) with a clear southeastward extension of low-pressure anomalies toward the Iberian Peninsula. The fourth regime displays a strong anticyclonic ridge (RDG) over western Europe (the “Atlantic Ridge” regime) (?).

6.3.5 NAO positive and negative Index

The most modern NAO indices are derived either from the simple difference in surface pressure anomalies between selected northern and southern locations, or from the PC time series of the leading Empirical Orthogonal Function (EOF) of the sea level pressure (SLP) field (?). Although the NAO is observable throughout the year, it is most pronounced during winter, accounting for over one-third of the total variance in the SLP field over the North Atlantic (?; 305). (?) defined a simple index of the NAO as the difference between the normalized mean winter (December–March) SLP anomalies at Lisbon, Portugal and Stykkisholmur, Iceland.

Figure 6.4 illustrates the positive and negative index of the NAO from 1864 to 2023. The average winter sea level pressure data at each station were normalized by division of each seasonal pressure by the long-term mean (1864–2023) standard deviation. Positive values of the index are in red and negative values in blue. The NAO index, as depicted in Figure 6.4, exhibits a well-defined oscillation. However, several periods exist where the NAO index remained in one phase over multiple winters, such as from 1903 to 1914 (positive phase), from 1915 to 1919 (negative phase), from 1958 to 1971 (negavite phase, excluding 1961 and 1967),

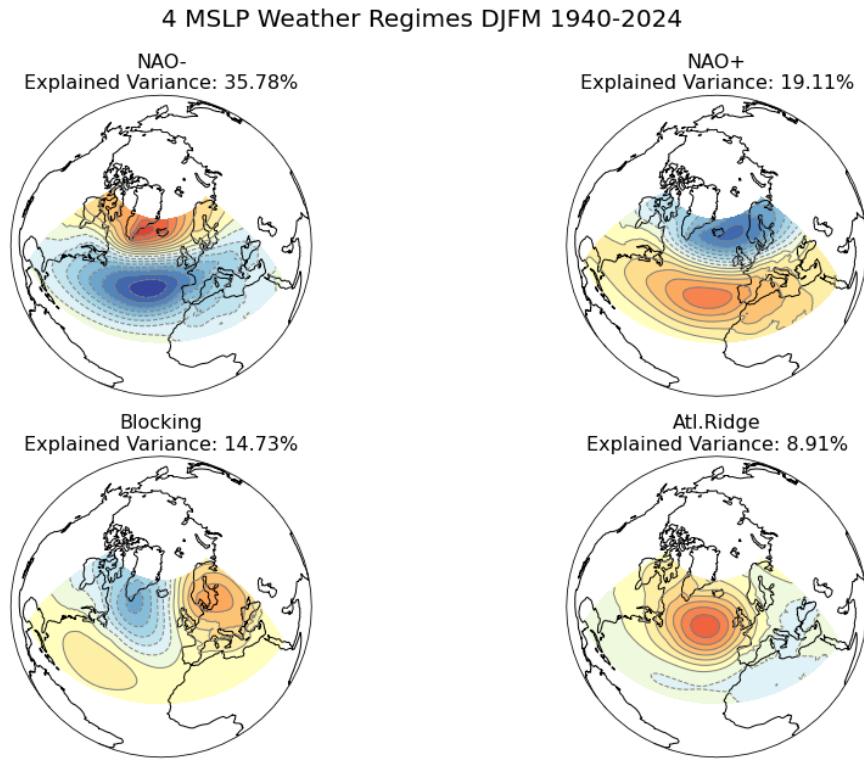


FIGURE 6.3: Boreal winter (December–March) climate regimes in sea level pressure (hPa) over the North Atlantic domain (20° – 70° N, 90° W– 40° E) using monthly data over the period from 1940 to 2024. The percentage on each panel expresses the explained variance of each cluster.

and from 2014 to 2020 (positive phase). On this figure of the NAO index 6.4 we can notice that there is little evidence for the NAO to vary on any preferred time scale. Significant changes can occur from one winter to the next, as well as from one decade to another.

The NAO is in a positive phase when both Icelandic low-pressure center and the high-pressure center at the Azores are stronger than the average (figure 6.5). During the positive NAO phase, the increased difference in pressure between these two regions results in a stronger Atlantic jet stream and a northward shift of the storm track. Consequently, northern Europe experiences increased storminess, higher precipitation, and warmer-than-average temperatures due to the air masses that arrive from lower latitudes. At the same time, southern Europe experiences decreased storminess and below-average precipitation. In eastern North America, the positive phase of the NAO generally brings higher air pressure, a condition associated with fewer cold-air outbreaks and decreased storminess (?).

Conversely, the NAO is in a negative phase when both the Icelandic low and Azores high are weaker than average. During the negative NAO phase, the Atlantic jet stream and storm track have a more west-to-east orientation. This results in decreased storminess, below-average precipitation, and lower-than-average temperatures in northern Europe. Conversely, southern Europe experiences increased storminess, above-average precipitation, and warmer-than-average temperatures. In eastern North America, the negative phase of NAO generally brings lower air pressure, a condition associated with stronger cold-air outbreaks and increased storminess (?).

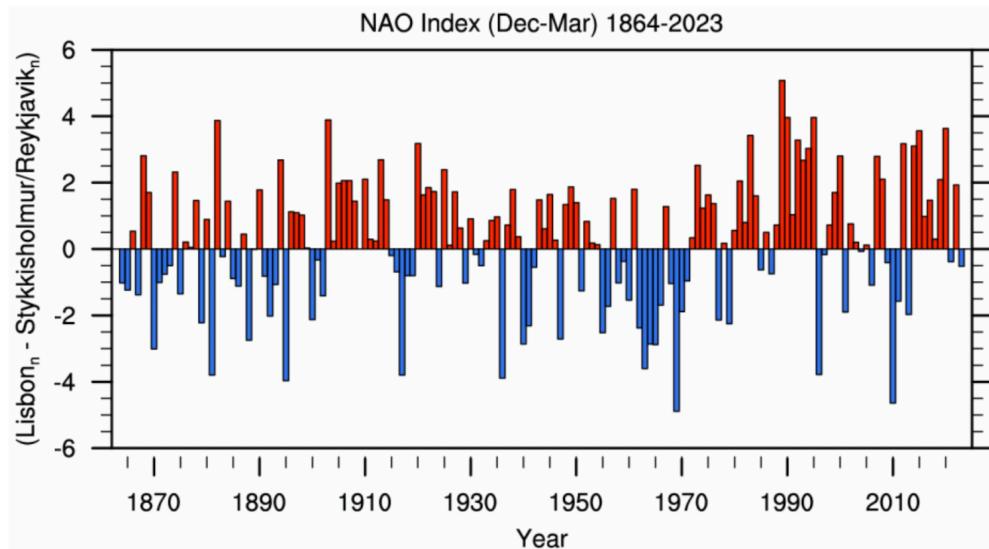


FIGURE 6.4: Normalized indices of the mean winter (December–March) NAO constructed from sea level pressure data. The index is based on the difference of normalized sea level pressure between Lisbon, Portugal and Stykkisholmur/Reykjavik, Iceland. The average winter sea level pressure data at each station were normalized by division of each seasonal pressure by the long-term mean (1864–2023) standard deviation. The indicated year corresponds to the January of the winter season (e.g., 1990 is the winter of 1989/1990)

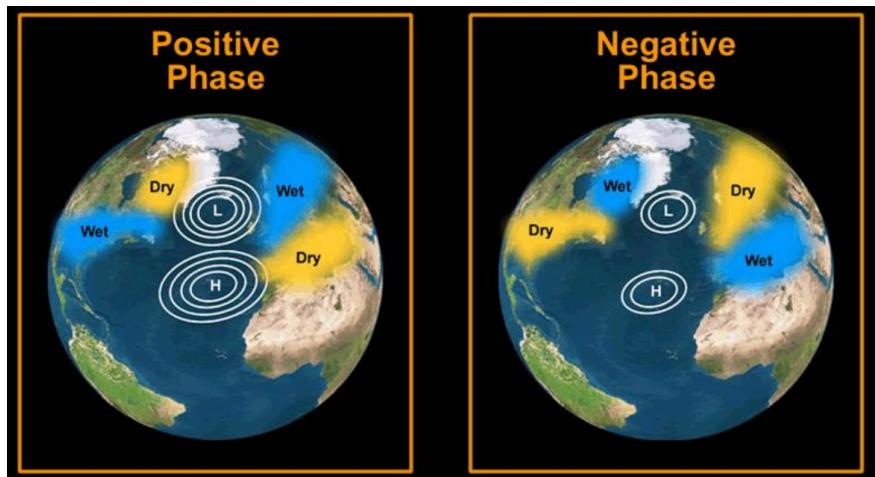


FIGURE 6.5: Positive and negative phase of the North Atlantic Oscillation. Image from @NAO_phase2013

6.4 The Implications of the North Atlantic Oscillation

The NAO plays a critical role in shaping climate patterns across the North Atlantic and surrounding areas. Changes in its spatial and temporal structures can lead to significant variations in temperature, precipitation, and weather extremes, influencing ecosystems, human activities, and natural resources.

6.4.1 Temperature, wind speed and direction

During the positive phase of the North Atlantic Oscillation (NAO), the atmospheric pressure difference between the Icelandic low and the Azores high intensifies. This results in a stronger-than-usual westerly flow across the North Atlantic during the winter months. Consequently, relatively warm, and moist maritime air is transported over much of Europe and extends across Asia (?). This influx of warm air can lead to milder winter conditions across the continent, often resulting in higher temperatures (?). Meanwhile, while stronger northerly winds over Greenland and northeastern Canada transport cold air southward leading to a decrease in land temperatures in the northwest Atlantic (?). The cold air intrusions can lead to harsher winter conditions in parts of eastern Canada and the northeastern United States.

Additionally, the stronger clockwise circulation around the subtropical Atlantic high-pressure center during this phase of the NAO significantly impacts other regions. In North Africa and the Middle East, this results in cooler-than-average temperatures due to the altered wind patterns. In contrast, parts of North America experience warmer conditions as the strengthened high-pressure system influences the weather patterns across the continent (?).

6.4.2 Storm and precipitation

NAO signals can also be found in winter precipitation and storm over the North Atlantic. During the positive phase of the NAO, storm activities in the Atlantic shift northeastward. This means there are more intense and frequent storm activities from southern Greenland across Iceland into northern Europe, and slightly less storm activities from the Azores across the Iberian Peninsula and the Mediterranean (?). Furthermore, the positive phase is typified by more intense and frequent storms in the vicinity of Iceland and the Norwegian Sea. During the NAO-, a weak subtropical high and a weak Icelandic Low dominate, resulting in a reduced pressure gradient and fewer, weaker winter storms with a more west-east trajectory, while moist air is brought into the Mediterranean and cold air into northern Europe (?).

Changes in the mean flow and storminess associated with oscillations in the NAO index lead to significant changes in the transport and convergence of atmospheric moisture, impacting the distribution of evaporation and precipitation (?). These results in drier conditions of this magnitude are observed over central and southern Europe, the Mediterranean, and parts of the Middle East, while more precipitation than normal occurs from Iceland through Scandinavia (?).

6.4.3 Ocean and Sea Ice

Another implication of the NAO mentioned by ? is its impact on ocean and sea ice. SST and NAO strength fluctuate together, with the main pattern of SST variability in boreal winter forming a tri-polar structure: a cold anomaly in the subpolar North Atlantic, a warm anomaly in mid-latitudes, and a cold anomaly in subtropical regions (?). These SST anomalies are driven by changes in surface wind and air-sea heat exchanges related to NAO variations. Persistent SST anomalies over longer periods also relate to consistent anomalous SLP patterns, including the NAO, reflecting the ocean's response to long term atmospheric variability. The reduction of the Arctic Sea ice is one of the most prominent indicators of climate change. less compact ice in high NAO winters plays an important role in the reduction of summer sea ice extent in high NAO years. In winter, the stronger wind in high NAO years drives more ice away from the Eurasian coastal region, suppressing sea ice buildup there (?).

6.5 Conclusion

In this chapter we have presented the NAO as the dominant mode of regional climate variability over the Northern Hemisphere, focusing particularly on common methods used to identify its spatial and temporal structure. These methods, although they are different, each have their strengths and limitations. One point correlation map is easy to implement and interpret providing clear insight into how the chosen center of action is related to the variability at other locations. However, the result heavily depends on the choice of the center of action. Selecting a different center may not well capture the full spatial pattern of the NAO. EOF analysis captures the dominant patterns of variability in a dataset and provides a quantitative measure of how much variance a pattern explains. But it assumes linear relationships and might not capture non-linear features of the NAO. Cluster Analysis is effective at identifying distinct patterns and regimes including non-linear and non-stationary features, but it is computationally intensive and requires careful selection of clustering parameters. Additionally, interpreting the results can be challenging if the patterns are not well separated.

As we have seen, the NAO is a significant climatic driver in the North Atlantic region, influencing temperature, wind speed and direction, precipitation, storm tracks, ocean and sea ice, particularly over Europe and North America. Its variability poses challenges and opportunities for understanding and predicting weather and climate impacts in affected regions. Understanding the implications of the NAO for climate change is essential for developing adaptive strategies to deal with the resulting implications on weather patterns and sea ice.

By improving climate models, and implementing robust adaptation plans, we can better address the challenges posed by NAO-induced climate variability. Predicting the future behavior of the NAO in a changing climate is challenging due to the complexity of the interactions between the atmosphere, oceans, and sea ice. Enhanced climate models that incorporate these interactions are crucial for reducing uncertainties and improving forecasts. Despite the strong influence of the NAO, many open issues remain about which climate processes govern NAO variability, how the phenomenon has varied in the past or will vary in the future, and to what extent it is predictable. External factors, such as the rise in greenhouse gas emissions, have a significant impact on the recent patterns of the NAO (?). Research indicates that higher greenhouse gas concentrations are associated with alterations in atmospheric circulation, which may result in more frequent occurrences of positive NAO phases (?). Understanding these shifts is essential for developing strategies to mitigate and adapt to the climate change effects influenced by the NAO (?).



7

Low Flow Events

Author: Gamze Gizem Kasman

Supervisor: Henri Funk

Degree: Master

7.1 Abstract

This chapter introduces some of the most used advanced statistical methods for predicting and analyzing low-flow events, characterized by reduced streamflow and water availability. Initially, key univariate concepts such as nonstationary models with time-varying parameters and meteorological covariates are presented, along with the concept of return periods to capture the dynamic nature of these events accurately. Following this, bivariate concepts, including the use of copulas to model dependencies between hydrological variables and the integration of reservoir indices to account for human impacts on flow regimes, are discussed. The concept of return periods is also introduced in a bivariate context. After these theoretical introductions, the chapter demonstrates practical applications, including the Statistical Downscaling Model (SDSM), which connects large-scale climate projections with localized hydrological impacts. These methodologies enhance the prediction and management of low-flow events, offering more robust strategies for water resource management in a changing climate.

7.2 Introduction

Low-flow events or periods are hydrological conditions marked by significantly reduced streamflow and water availability in rivers and streams. These events often result from prolonged dry weather conditions, reduced precipitation, high evaporation rates, or significant water withdrawals (?). Their extremeness distinguishes them from normal variations in water flow, as they can profoundly and lastingly impact water availability and quality.

To effectively manage and mitigate the consequences of these extreme low-flow periods, it is essential to predict their occurrence and severity. This is where the concept of return periods becomes crucial. The return period, also known as the recurrence interval, is a statistical measure that estimates the frequency at which a particular hydrological event, such as a low-flow period, is expected to occur (?). It provides a probabilistic assessment of how often extreme low-flow events of a certain magnitude are likely to happen over a given time frame. The concept of return periods is mostly used in the context of univariate analysis, where it pertains to a single hydrological variable, such as streamflow. However, return periods are also utilized in bivariate analysis in the context of low-flow events. Bivariate return periods consider the joint occurrence of two hydrological variables, such as streamflow and precipitation, or the dependence between two low-flow series, providing a more comprehensive assessment of extreme events. The bivariate approach

is particularly useful when the interaction between variables significantly influences the event's occurrence and impact (??).

Closely associated with the return period is the concept of risk. In the context of hydrology, risk refers to the probability that a specific event, such as a low-flow period, will occur within a given timeframe. Understanding the risk associated with different return periods allows water resource managers to better prepare for and mitigate the impacts of these extreme low-flow events, ensuring more resilient and adaptable water management practices.

Climate change adds a significant layer of complexity to the management and prediction of low-flow events. Changes in climate patterns, such as alterations in precipitation regimes and increased temperatures, can exacerbate the frequency and severity of low-flow periods. For instance, climate change may lead to longer dry seasons, reduced snowpack, and increased evapotranspiration, all of which contribute to more severe low-flow conditions (?). Furthermore, shifts in weather extremes mean that historical data may no longer be a reliable predictor of future conditions, complicating the estimation of return periods and associated risks. Therefore, understanding how flood and low-flow events are evolving and how likely they are to unfold beyond their historical distributions under the influence of multiple drivers, such as climate variability, human activities like reservoir operations etc., is one of the major challenges facing the hydrology field (??; ??; ??; ??; ??). In particular, predicting the likelihood of low-flow events occurring beyond their historical distributions due to factors like climate variability and human activities remains a key challenge in hydrology (??; ??; ??; ??; ??; ??).

This necessitates the use of advanced statistical models that can incorporate climate projections, climate indices such as oscillation data, and reservoir indices to better predict low-flow events in a changing climate (?). Understanding how climate change impacts low-flow events and their return periods is crucial for adapting water management practices to future conditions.

This chapter aims to provide a focused study on low-flow events and their return periods through a statistical lens. We will introduce and discuss the most widely used models in hydrological frequency analysis that take climate change and human activities into account. These modern methods incorporate a variety of data sources, providing a more accurate prediction of low-flow events under changing climate conditions. For comparison purposes, we will also present traditional approaches that do not account for climate change, which assume that the statistical properties of low-flow events remain constant over time. By exploring both modern and traditional methods, we will highlight their respective advantages and limitations, and demonstrate their applications in the context of hydrological frequency analysis of low-flow events and return periods.

7.3 Data Analysis

In the context of hydrological frequency analysis, understanding the most commonly used datasets that underpin our insights into low-flow events and return periods is essential. These datasets vary in their sources, methodologies, and applications, reflecting advancements in the field and the growing acknowledgment of climate change.

7.3.1 Observed Hydrological Data

One of the foundational datasets in hydrological analysis is observed hydrological data, which encompasses both temporal and spatial dimensions. This data, typically gathered from gauging stations along rivers and streams, captures key measurements such as river discharge, water levels, and flow rates. The temporal aspect is reflected in the continuous time-series records, while the spatial aspect is evident in the geographical distribution of gauging stations across a watershed. The value of this data lies in its direct representation of

the hydrological processes within a given watershed, providing a reliable basis for understanding past and present flow conditions.

Long-term records from these stations are crucial for identifying trends, variability, and the statistical characteristics of hydrological events, including low-flow and high-flow occurrences. This type of data can be directly utilized in statistical analyses for hydrological frequency analysis, enabling researchers to assess event probabilities and predict future hydrological behaviors (?).

7.3.2 Observed Meteorological Data

In the context of models where climate change is acknowledged, observed meteorological data plays a significant role. This dataset encompasses measurements of various atmospheric parameters, including precipitation, temperature, and humidity, collected from meteorological stations. The linkage between meteorological conditions and hydrological responses is vital, as weather patterns significantly influence water availability and flow dynamics. By analyzing this data, researchers can better understand the drivers of hydrological extremes and their temporal patterns, thus enhancing the prediction and management of low-flow events in the context of a changing climate (?). Initially, this data is point data as it is collected from individual meteorological stations. It is then spatially averaged to form areal data representing the broader catchment areas.

7.3.3 Reanalysis Data

Reanalysis data adds another critical dimension to modern hydrological studies. These datasets synthesize historical weather observations with modern forecasting models to produce comprehensive and consistent records of atmospheric conditions over extended periods. An example of this is the NOAA National Centers for Environmental Prediction (NCEP) reanalysis data. This data offers high temporal and spatial resolution, filling gaps where direct observations may be sparse or unavailable. Reanalysis data uses a variety of atmospheric predictors such as sea level pressure, air temperature at various altitudes, geopotential heights, and humidity, integrating them with observed data to create a continuous and coherent dataset (?). This linkage to observed meteorological data ensures that reanalysis products accurately reflect past atmospheric conditions, making them a critical tool for studying long-term trends and variability in low-flow events under changing climatic conditions. This data is considered gridded spatial data due to its continuous spatial coverage over a global grid.

7.3.4 Global Circulation Model Outputs

Another dataset for hydrological frequency analysis comes from climate modeling efforts. General Circulation Models (GCMs) simulate the Earth's climate system, incorporating the interactions between the atmosphere, oceans, land surface, and ice. These models are integral to projecting future climate scenarios and assessing potential impacts on hydrological systems. The Coupled Model Intercomparison Project (CMIP), organized by the World Climate Research Programme (WCRP), is a key initiative that brings together various GCM outputs to facilitate a comparative analysis of these projections. The IPCC uses results from these GCMs to provide comprehensive assessments of future climate change in its reports. For hydrological frequency analysis, GCM outputs are crucial for assessing future scenarios of low-flow events under different greenhouse gas emission pathways. However, the coarseness of GCM outputs often requires them to be downscaled to a finer resolution to be useful for regional or local studies. GCM outputs are also gridded spatial data, providing large-scale projections over a global grid. Downscaling techniques, which can be statistical or dynamical, help bridge the gap between the broad, global scale of GCMs and the finer scales needed for practical hydrological applications. In this regard, NCEP reanalysis data is frequently used as a reference to improve the downscaling of GCM outputs, ensuring that they more accurately reflect observed climatic conditions at finer scales (?).

7.3.5 Reservoir Indices

Reservoir indices are critical datasets in hydrological studies, especially for understanding and managing low-flow events. These indices provide information on reservoir levels, inflows, outflows, storage capacities, and water releases. Reservoir operations significantly influence streamflow regimes, particularly during periods of low flow, by regulating the amount of water released downstream. The management of reservoir water levels can mitigate the impacts of droughts and ensure a stable water supply for various uses, including agriculture, drinking water, and hydroelectric power generation.

López and Francés (2013) proposed the reservoir index (RI) as an indicator of the impact of reservoirs on flow regimes. It is defined as follows:

$$RI = \sum_{i=1}^N \left(\frac{A_i}{A_T} \right) \cdot \left(\frac{V_i}{C_T} \right) \quad (1)$$

where N is the number of reservoirs upstream of the gauge station, A_i is the catchment area of each reservoir, A_T is the catchment area of the gauge station, V_i is the total capacity of each reservoir, and C_T is the mean annual streamflow of the gauge station.

This definition may be modified in specific studies to better suit their research needs. The table below provides detailed information about the datasets used in these studies, including any modifications to the reservoir index.

Data Used in Key Studies

Throughout this chapter, information and results from the following studies will be referenced. The table below provides detailed information about the datasets used in these studies.

Table: Detailed Information of Data Used in Studies

Study	Data Type	Details	Period	Source
Du et al. (2015)	Observed Hydrological Data	Mean daily streamflow data	1954-2009	Huaxian and Xianyang gauging stations
	Observed Meteorological Data	Daily average temperature and daily total precipitation	1954-2009	22 meteorological stations in the Wei River Basin
	Reanalysis Data	NCEP reanalysis data	1954-2009	NOAA (National Oceanic and Atmospheric Administration)
	GCM Outputs	Downscaled GCM outputs	2010-2099	CMIP5 (Coupled Model Intercomparison Project Phase 5) under the RCP 8.5 scenario
Wang et al. (2022)	Observed Hydrological Data	Mean daily streamflow data	1955-2016	Multiple gauging stations in the Yellow River Basin
	Observed Meteorological Data	Daily temperature and precipitation data	1955-2016	Meteorological stations in the Yellow River Basin
	Reservoir Data	Reservoir inflow, outflow, and storage capacity data	1955-2016	Reservoirs in the Yellow River Basin
Jiang et al. (2015)	Observed Hydrological Data	Mean daily streamflow data	1961-2010	Ankang (Qa) and Huangzhuang (Qh) gauges
	Observed Meteorological Data	Daily temperature and precipitation data	1961-2010	Meteorological stations in the Weihe River Basin

Study	Data Type	Details	Period	Source
Reservoir Data	Reservoir inflow, outflow, and storage capacity data	1961-2010		Reservoirs in the Weihe River Basin

7.4 Stationarity vs. Nonstationarity in Low-flow Analysis

In hydrological frequency analysis, understanding the concept of stationarity is crucial for predicting and managing low-flow events. Traditionally, hydrological analyses have assumed stationarity, meaning that the statistical properties of hydrological processes are constant over time. Under this assumption, the historical data on low-flow events can be used to predict future occurrences with the expectation that the governing processes and factors remain unchanged. This approach simplifies the analysis and provides a straightforward basis for calculating return periods and risks (??).

Stationary analysis operates under the premise that factors such as climate and land use, which influence hydrological events, do not change significantly over time. This allows for the use of historical data to estimate future low-flow events with a reasonable degree of confidence. The traditional frequency analysis methods, such as fitting probability distributions to historical low-flow data, rely on this assumption. These methods are useful for long-term planning and water resource management as they provide stable and predictable estimates of low-flow probabilities and return periods (?). However, the assumption of stationarity is increasingly being questioned due to evident and ongoing changes in climate and land use patterns. Studies have shown that these changes can significantly alter hydrological processes, rendering stationary models less reliable(?).

Nonstationary analysis acknowledges that the statistical properties of hydrological processes can change over time due to factors like climate change, urbanization, and land-use changes. In a nonstationary framework, the parameters of probability distributions used to model hydrological events are allowed to vary over time or with respect to external covariates, such as meteorological variables. This approach provides a more realistic representation of the evolving nature of hydrological processes under the influence of changing climatic and environmental conditions (??).

The most common method for addressing nonstationarity in hydrological time series is the time-varying moment method. This technique assumes that while the type of distribution function for the hydrological variable remains constant, its statistical parameters can vary over time (??; ??; ??; ??). The time-varying moment method is typically used in a univariate context, where it focuses on a single hydrological variable, adjusting its statistical parameters to reflect temporal changes. Furthermore, the time-varying moment method can be extended to perform covariate analysis by replacing time with physical factors known to influence the hydrological variable of interest, such as temperature and precipitation. This provides a dynamic model that can adapt to projected changes in climate. This approach is critical for modern water resource management as it helps anticipate the impacts of climate variability and change, thus supporting more resilient and adaptable planning and infrastructure development (??).

In addition to the time-varying moment method, copula-based models have become increasingly important in nonstationary hydrological analysis. Copulas allow for the modeling of complex interdependencies between hydroclimatic variables, such as streamflow, precipitation, and temperature, as well as the dependence structure between low-flow events. Copula models are particularly useful in a bivariate context, as they can capture the evolving dependencies between two or more variables over time or under changing climatic conditions. This capability is crucial for understanding and predicting compound extreme events, which occur due to the concurrence or succession of multiple hydroclimatic extremes. By using copula-based probability distributions, researchers can more accurately assess the risk and impact of such compound events, providing a more comprehensive approach to hydrological frequency analysis (?).

Comparing stationary and nonstationary analyses highlights the limitations of traditional methods in the

face of a changing climate. Stationary models, while simpler and easier to apply, may lead to underestimation or overestimation of future low-flow events if the underlying conditions change significantly. Nonstationary models, though more complex, offer a nuanced understanding by integrating evolving environmental variables into the analysis, thus providing more accurate and reliable predictions (?).

The choice between stationary and nonstationary analysis has significant implications for water resource management. For instance, infrastructure designed based on stationary assumptions might not withstand future conditions, leading to potential failures or inefficiencies. Conversely, incorporating nonstationarity into planning and design processes ensures that water management strategies are robust and flexible enough to cope with future uncertainties (?; ?).

This chapter aims to explore these two approaches through a detailed comparison, providing insights into their respective methodologies, advantages, and limitations. The discussion will be grounded in recent research and applied examples, emphasizing the practical implications for water resource management under changing climatic conditions.

7.5 Detecting Nonstationarity

To identify nonstationarity in hydroclimatological extremes, it is crucial to thoroughly examine the data before applying statistical tests. These tests aim to detect two primary types of nonstationarity: step changes (change points) and monotonic changes (trends). Such variations may be considered indicative of nonstationarity if they significantly deviate from normality in long-term records. Nonstationarity can be detected either in individual time series (point-based analysis) or across larger groups of stations (spatially coherent trends; (?; ?)).

7.5.1 Change Point Analysis

Change point analysis is a statistical technique used to detect points in a time series where the properties of the data change. In the univariate context, this analysis is relatively straightforward as it deals with a single variable. A widely used method for identifying change points in univariate time series is the Pettitt test (?). This approach considers a time series as two samples represented by x_1, \dots, x_t and x_{t+1}, \dots, x_N . Then, the Pettitt indices $U_{t,N}$ can be calculated as:

$$U_{t,N} = \sum_{j=1}^t \sum_{i=1}^N \text{sgn}(x_j - x_i) \quad (t = 1, \dots, N) \quad (2)$$

where

$$\text{sgn}(\theta) = \begin{cases} +1 & \theta > 0 \\ 0 & \theta = 0 \\ -1 & \theta < 0 \end{cases}$$

The p value of the $U_{t,N}$ can be computed as follows:

$$p \cong 2 \exp \left[\frac{-6(U_{t,N})^2}{N^3 + N^2} \right]$$

While the Pettitt test effectively identifies change points in univariate time series, real-world data often involve multiple variables that can change simultaneously. In the multivariate context, detecting change points becomes more complex due to the higher dimensionality of the data. To address this complexity,

Lavielle and Teyssi  re (2006) developed a method for identifying multiple change points in multivariate time series. Their approach involves segmenting the time series into several homogeneous segments, each assumed to follow a different multivariate distribution with distinct parameters.

The time series is divided into K segments, where K is the number of change points plus one. For each segment, the likelihood of the observed data is calculated based on the assumed multivariate distribution (e.g., multivariate normal). The method then optimizes a penalized likelihood criterion to balance the goodness of fit and model complexity, often using a penalty term similar to the Bayesian Information Criterion (BIC). Dynamic programming algorithms are employed to efficiently find the optimal segmentation that maximizes the penalized likelihood.

7.5.2 Temporal Trend Analysis

The Mann-Kendall (MK) trend test method, endorsed by the World Meteorological Organization, is a widely used nonparametric approach for trend analysis. For a series $X = \{x_1, x_2, \dots, x_n\}$, the standard normal statistic Z is estimated as follows:

$$Z = \begin{cases} (S - 1)/\sqrt{\text{var}(S)} & S > 0 \\ 0 & S = 0 \\ (S + 1)/\sqrt{\text{var}(S)} & S < 0 \end{cases} \quad (3)$$

where

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(x_j - x_i)$$

and

$$\text{var}(S) = \frac{n(n-1)(2n+5)}{18}$$

Applications of Change Point Analysis in Detecting Nonstationarity

Du et al. (2015): This study applies the Mann-Kendall test to detect trends in low-flow series under climate change. Significant decreasing trends were detected at two stations, Huaxian and Xianyang, with Mann-Kendall statistics Z_{MK} values of -2.71 and -3.63, respectively. The results highlight the impact of climate change on hydrological processes and the necessity of incorporating nonstationarity in trend analyses.

Wang et al. (2022): In this study, change point analysis is applied using the Pettitt test and the Mann-Kendall trend test to detect nonstationarity in univariate time series. Specifically, the Pettitt test is utilized to identify significant change points in annual low-flow series. For instance, a significant change point in the low-flow series of the ZGP station in 1980 ($p < 0.1$) was detected, revealing modes of nonstationarity in the hydrological series. This analysis is illustrated in the table below, showing the results of the Pettitt test for low-flow series.

Table: Results of Pettitt and MK tests of low-flow series

Stations	Pettitt	MK
DPL	0.156	0.156
CTG	0.360	0.360
ZGP	0.001*	0.001*
XX	0.279	0.279

Notes: The numerical values correspond to the p values of the Pettitt and MK tests. The asterisks indicate significance at the 10% level. A value of 0.293 / 0.573 indicates p values before and after a change point.

Jiang et al. (2015): This study examines nonstationarities in low-flow series at Ankang and Huangzhuang Gauges and their dependence structure using both univariate and bivariate aspects. For the univariate analysis, the nonparametric Pettitt test identifies an abrupt change in the variance of Q_a in 1991, though its mean remains stationary. For Q_h , changes in mean and variance are detected in 1971 and 1972, respectively. For the bivariate analysis, using Lavielle and Teyssi  re's (2006) method, a significant change in the dependence structure between Q_a and Q_h is found in 1970, indicating nonstationarity in both individual series and their dependence structure.

7.6 Univariate Statistical Analysis

To effectively analyze and manage low-flow events, it is essential to understand the concepts of exceedance probability and design low-flow quantile. These concepts are foundational for conducting return-period and risk analysis, which are critical in hydrological studies. Return-period analysis helps in estimating the frequency of low-flow events, while risk analysis assesses the probability of these events occurring within a specific timeframe.

7.6.1 Exceedance Probability of a Low-flow Event

The low-flow characteristics of the flow regime are represented by the random variable Z . Focusing on water scarcity, we define the design low-flow quantile z_{p_0} , which has a probability p_0 in any given year that the streamflow falls below this quantile. The likelihood of a flood event exceeding a design flood quantile is commonly referred to as the exceedance probability or the exceeding probability (e.g., (?; ?)). For a low-flow event, exceedance or exceeding indicates that the severity of lack of precipitation is surpassed, meaning the flow statistic value is lower than the design quantile (?).

7.6.1.1 Stationary Case

Under stationary conditions, the cumulative distribution function of Z is denoted by $F_Z(z, \theta)$, where θ represents the constant statistical parameter set. The aim is to analyze the future evolution of the return period and hydrological risk for a specific design low-flow determined at a particular year based on historical observed data. This particular year is typically defined as the base year or initial year, denoted by $t = 0$. Therefore, the given design low-flow at $t = 0$, corresponding to an initial return period T_0 , can be derived using $z_{p_0} = F_Z^{-1}(p_0, \theta)$, where $p_0 = 1/T_0$ and F_Z^{-1} is the inverse function of F_Z (similarly hereinafter). The period of years following the initial year is referred to as the “future” and initial year $t = 0$ corresponds to the last observation year (?). In the stationary case, meaning the controlling environmental factors for future years remain the same as in the initial year $t = 0$, θ is constant for every year, and the exceedance probability corresponding to the design quantile z_{p_0} is p_0 for each future year (Fig. 1a), which can be expressed as:

$$p_t = F_Z(z_{p_0}, \theta) = p_0, \quad t = 1, 2, \dots, \infty \quad (4)$$

7.6.1.2 Nonstationary Case

Under nonstationary conditions, the cumulative distribution function of Z is denoted by $F_Z(z, \theta_t)$, where θ_t varies with time or, more specifically, with meteorological variables. In the nonstationary case, the design low-flow quantile corresponding to the initial exceedance probability $p_0 = 1/T_0$ can be derived from $z_{p_0} = F_Z^{-1}(p_0, \theta_0)$, where θ_0 is the statistical parameter set for the initial year $t = 0$. Since the statistical parameters change over time, the future exceedance probability corresponding to z_{p_0} is no longer constant. The temporal variation in the exceedance probability corresponding to z_{p_0} can be characterized by the way the low-flow

distribution or, more specifically, the statistical parameters change over time (Fig. 1b) The exceedance probability for each future year can be obtained by:

$$p_t = F_Z(z_{p_0}, \theta_t), \quad t = 1, 2, \dots, \infty \quad (5)$$

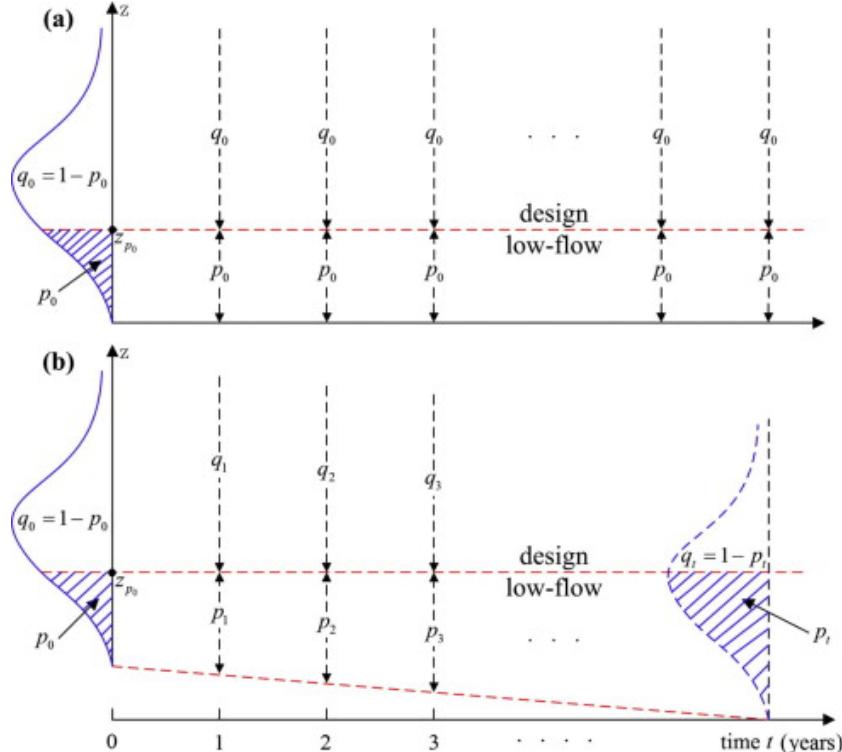


FIGURE 7.1: Schematic depicting the design low-flow quantile z_{p_0} with (a) constant exceedance probability p_0 , and (b) time-varying exceedance probabilities $p_t, t = 1, 2, \dots, \infty$. (#fig:fig1_gamze)

Source: (?)

7.6.2 Return Period

With the method of time-varying moment, deriving a value of a hydrological variable for a return period with a specific design quantile is straightforward (?; ?). Specifically, $T_t = 1/p_t = 1/(1 - F_z(z_{p_0}, \theta_t))$, where T_t and p_t are the annual return period and exceedance probability, respectively, of the given design quantile z_{p_0} with fitted annual statistical parameters θ_t , and F_Z is the cumulative distribution function of the hydrological variable of interest. However, for many planning and design applications, a return period measure that varies annually is impractical. To address this issue, various studies have been conducted on return period estimation and risk analysis under nonstationary conditions (?; ?; ?; ?, 2010; ?; ?; ?). These studies have proposed two different interpretations of return period. The first interpretation is that the expected waiting time (EWT) until the next exceedance event is T years (?; ?; ?; ?, ?, ?). The second interpretation is that the expected number of exceedances (ENE) of the event in T years is 1 (?; ?, ?).

7.6.2.1 Return Period Using Expected Waiting Time (EWT) Interpretation

Under stationary conditions, if X is the random variable representing the year of the first occurrence of a low-flow that exceeds (i.e., is lower than) the design quantile, then the occurrence of a low-flow Z exceeding

the design value z_{p_0} for the first time in year $X = x$, $x = 1, 2, \dots, \infty$, follows the geometric probability law (?, ?, ?):

$$f(x) = P(X = x) = (1 - p_0)^{x-1} p_0, \quad x = 1, 2, \dots, \infty \quad (6)$$

Given that Eq. (3) is based on the assumptions of independence and stationarity, the expected value of X —representing the return period (expected waiting time interpretation) of the low-flow exceeding the design quantile z_{p_0} under stationary conditions, is:

$$T = E(X) = \sum_{x=1}^{\infty} x f(x) = 1/p_0 \quad (7)$$

Under nonstationary conditions, the exceedance probability associated with z_{p_0} is no longer constant (Fig. 1b). Consequently, the geometric probability law, which takes into account time-varying exceedance probabilities p_t , is expressed as (?, ?, ?):

$$\begin{aligned} f(x) &= P(X = x) = (1 - p_1)(1 - p_2) \dots (1 - p_{x-1}) p_x \\ &= p_x \prod_{t=1}^{x-1} (1 - p_t), \quad x = 1, 2, \dots, \infty \end{aligned} \quad (8)$$

Therefore, the EWT-return period T for a low-flow event exceeding the design quantile z_{p_0} under nonstationary conditions is:

$$T = E(X) = \sum_{x=1}^{\infty} x f(x) = \sum_{x=1}^{\infty} x p_x \prod_{t=1}^{x-1} (1 - p_t) \quad (9)$$

7.6.2.2 Return Period Using Expected Number of Exceedances (ENE) Interpretation

Under stationary conditions, if M is the random variable representing the number of exceedances in T years, then $M = \sum_{t=1}^T I(Z_t < z_{p_0})$, where $I(\cdot)$ is the indicator function. In this scenario, M follows a binomial distribution (?):

$$f(m) = P(M = m) = \binom{T}{m} p_0^m (1 - p_0)^{T-m} \quad (10)$$

It follows that the expected value of M is 1 :

$$E(M) = \sum_{t=1}^T p_0 = T p_0 = 1 \quad (11)$$

Therefore, the return period (expected number of exceedances interpretation) for a low-flow event exceeding the design quantile z_{p_0} under stationary conditions is $T = 1/p_0$.

Under nonstationary conditions, the exceedance probability is not constant, and M does not follow a binomial distribution. In this context, the expected number of exceedances is expressed as (?):

$$\begin{aligned} E(M) &= \sum_{t=1}^T E[(Z_t < z_{p_0})] = \sum_{t=1}^T P(Z_t < z_{p_0}) = \sum_{t=1}^T F_Z(z_{p_0}, \theta_t) \\ &= \sum_{t=1}^T p_t \end{aligned} \quad (12)$$

The ENE-return period T for a low-flow event exceeding the design quantile z_{p_0} under nonstationary conditions can be derived by setting Eq. (12) equal to 1 and solving:

$$\sum_{t=1}^T p_t = 1 \quad (13)$$

7.6.3 Hydrological Risk

In practical applications of hydrological frequency analysis, the management question often revolves around risk. Specifically, for a design life of n years, the hydrological risk R is defined as the probability of a low-flow event exceeding the design value z_{p_0} before or at year n . This risk can be derived from the complement perspective, meaning there is no exceedance during the design life of n years.

As an example, consider a low-flow event with a return period of 100 years. The risk for $t = 1$ year is 1% (i.e., $R = 0.01$). For $t = 5$ years, the risk increases to approximately 5%. By $t = 69$ years, the risk reaches 50%; for $t = 100$ years (the return period), the risk is 63%; and for $t = 300$ years, the risk climbs to 95%, and so on (?).

Under the assumptions of independence and stationarity, the probability of the complement is $(1 - p_0)^n$. Consequently, the hydrological risk under stationary conditions is given by (?):

$$R = 1 - (1 - p_0)^n \quad (14)$$

In recent years, hydrological **risk analysis under nonstationary conditions** has gained popularity (?; ?; ?). This approach offers designers a different perspective from traditional return period and design level methods by incorporating the basic information of design life n and time-varying exceedance probabilities p_t . Similar to the stationary case, for a design life of n years, the probability of a low-flow event exceeding the design quantile z_{p_0} before or at year n under time-varying exceedance probabilities p_t is:

$$R = 1 - [(1 - p_1)(1 - p_2) \dots (1 - p_n)] = 1 - \prod_{t=1}^n (1 - p_t) \quad (15)$$

Return Period and Risk Results from Du et al. (2015)

Nonstationary Return Period Using Meteorological Covariates

Figure 2 presents the return period and risk results using meteorological covariates (explained in the model selection part) for Huaxian and Xianyang stations.

- **Panel (a) and (b):** The black diagonal line represents the stationary case where $T = T_0$, meaning the return period remains constant over time as expected under stationary assumptions. The other lines depict the nonstationary return periods calculated using the meteorological covariates (temperature and precipitation). These lines show the variability in return periods under nonstationary conditions. The deviations from the diagonal line indicate the impact of changing climatic conditions on the frequency of low-flow events. In nonstationary cases, return periods are generally shorter than the stationary assumption, indicating more frequent low-flow events.
- **Panel (c) and (d):** These panels illustrate the nonstationary hydrological risk R over different design lives (n). The black solid line shows the stationary risk, providing a baseline for comparison. The nonstationary risk curves, calculated using the covariates, are generally higher than the stationary risk, especially for longer design lives. This indicates a greater likelihood of experiencing low-flow events in the future under nonstationary conditions. The comparison between stationary and nonstationary risks highlights the importance of incorporating nonstationary models in hydrological risk assessments to account for changing climatic conditions.

- **Huaxian Station:**

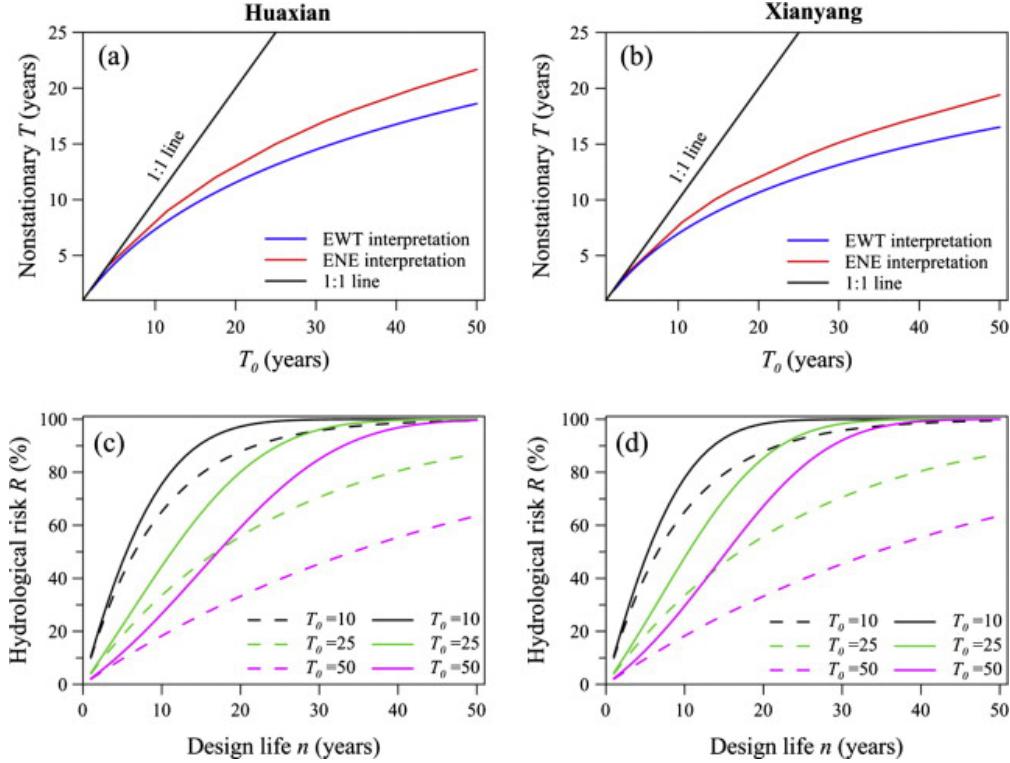


FIGURE 7.2: Nonstationary return period T and hydrological risk R of the Wei River (#fig:fig2_gamze)

- **Expected Waiting Time (EWT):**
 - * For a specified initial return period $T_0 = 50$ years, the nonstationary return period T was calculated to be 18.6 years.
- **Expected Number of Exceedances (ENE):**
 - * For $T_0 = 50$ years, the nonstationary return period T was calculated to be 22 years.
- **Implications:**
 - * Under nonstationary conditions, the return periods are significantly shorter than the specified T_0 , indicating more frequent low-flow events than expected under stationary assumptions.

- **Xianyang Station:**

- **Expected Waiting Time (EWT):**
 - * Similar to Huaxian, the nonstationary return period T for $T_0 = 50$ years was shorter than the specified T_0 .
- **Expected Number of Exceedances (ENE):**
 - * Similar shorter return periods were observed under the ENE interpretation.
- **Implications:**
 - * The results underscore the necessity of accounting for nonstationarity in return period analysis for accurate water resource planning.

Nonstationary Hydrological Risk Using Meteorological Covariates

- **Huaxian Station:**

- **Risk Calculations:**

- * For a design life n of 10 years and $T_0 = 50$ years, the hydrological risk R was 14.1% under nonstationary conditions, compared to 18.3% under stationary conditions.
- * For $n = 40$ years and $T_0 = 50$ years, the risk R was 79.0% under nonstationary conditions, significantly higher than the stationary risk of 55.4%.

– **Implications:**

- * The nonstationary risk is higher, especially for longer design lives, indicating an increased likelihood of low-flow events over time.

• **Xianyang Station:**

– **Risk Calculations:**

- * For $n = 10$ years and $T_0 = 50$ years, the risk R was 5.8% under nonstationary conditions.
- * For $n = 40$ years and $T_0 = 50$ years, the risk R was 68.4% under nonstationary conditions.

– **Implications:**

- * Similar to Huaxian, the nonstationary risk for Xianyang is significantly higher, emphasizing the impact of nonstationarity on hydrological risk assessments.

7.6.4 Nonstationary frequency analysis of low-flow series

The standard method for calculating nonstationary return periods under the EWT and ENE interpretations, as well as the risk of a design quantile z_{p_0} corresponding to an initial return period T_0 , is outlined by Eqs. (9), (13), (15), respectively. A crucial aspect of this procedure is the derivation of time-varying exceedance probabilities p_t (Eq. (5)) for future years. This process involves determining the relationships between the statistical parameters of the low-flow distribution and the explanatory variables, which is typically referred to as nonstationary frequency analysis.

7.6.4.1 GAMLSS

The nonstationary low-flow series is modeled using the Generalized Additive Models in Location, Scale, and Shape (GAMLSS) framework (?). The GAMLSS framework assumes that independent observations z_i for $i = 1, 2, 3, \dots, n$ have a distribution function $f(z_i | \theta^i)$, where $\theta^i = (\theta_1^i, \theta_2^i, \dots, \theta_p^i)$ is a vector of p distribution parameters accounting for the location, scale, and shape of the random variable characteristics. Usually, p is less than or equal to four, since the 1-, 2-, 3-, and 4-parameter families ensure sufficient flexibility for most applications in hydrology.

Given an n -length vector of the response variable $z^T = (z_1, \dots, z_n)$, let $g_k(\cdot)$, for $k = 1, 2, \dots, p$, be known monotonic link functions relating the distribution parameters to explanatory variables and random effects through an additive model given by:

$$g_k(\theta_k) = \eta_k = X_k \beta_k + \sum_{j=1}^{J_k} Z_{jk} \gamma_{jk} \quad (16)$$

where η_k and θ_k are vectors of length n (e.g., $\theta_k^T = (\theta_k^1, \theta_k^2, \dots, \theta_k^n)$), $\beta_k^T = (\beta_{1k}, \beta_{2k}, \dots, \beta_{J_k k})$ is a parameter vector of length J_k , X_k is a matrix of explanatory variables (i.e., covariates) of order $n \times J_k$, Z_{jk} is a fixed known $n \times q_{jk}$ design matrix, and γ_{jk} is a q_{jk} -dimensional random variable. The linear predictors η_k for $k = 1, \dots, p$ comprise a parametric component $X_k \beta_k$ (linear functions of explanatory variables) and the additive component $Z_{jk} \gamma_{jk}$ (linear functions of stochastic variables). If $Z_{jk} = I_n$, where I_n is an $n \times n$ identity matrix and $\gamma_{jk} = h_{jk} = h_{jk}(x_{jk})$ for all combinations of j and k , we obtain the semi-parametric additive formulation of GAMLSS:

$$g_k(\theta_k) = \eta_k = X_k \beta_k + \sum_{j=1}^{J_k} h_{jk}(x_{jk}) \quad (17)$$

where θ_k is the parameter vector of length n , \mathbf{X}_{jk} is a matrix of explanatory variables of order $n \times m$, β_k is a parameter vector of length m , and h_{jk} denotes the functional dependence of the distribution parameters on explanatory variables x_{jk} , which is a column of matrix \mathbf{X}_{jk} . This dependence can be linear or smoothed using smoothing terms, and the smooth dependence is based on a cubic spline function (?).

Various probability distributions have been suggested for modeling low-flow events (?; ?; ?; ?). Matalas (1963) investigated four distributions in modeling low-flow data from 34 streams and found that the Gumbel and Pearson-III distributions fit the data well and were more representative than the Lognormal and Pearson-V distributions. Eratakulan (1970) discovered that the Gamma and Weibull distributions were the most suitable for modeling the low-flow series from 37 stations in the Missouri River basin. Hewa et al. (2007) introduced the GEV distribution into the frequency analysis of low-flow data from 97 catchments in Victoria, Australia. Liu et al. (2015) tested six distributions for modeling the annual low flows at the Yichang station, China, under nonstationary conditions and found that the GEV distribution provided the best fit.

Based on these studies, five two-parameter distributions (Gamma (GA), Weibull (WEI), Gumbel (GU), Logistic (LO), and Lognormal (LOGNO)) and two three-parameter distributions (Pearson-III (P-III) and GEV) that are widely used in modeling low-flow data are listed in Table 1. Considering that the shape parameter κ of the P-III and GEV distributions is quite sensitive and difficult to estimate, it is mostly assumed to be constant (?; ?; ?; ?). Nonstationarities in both the location μ and scale σ parameters are examined through monotonic link functions $g(\cdot)$.

Distribution	Probability Density Function	Moments	Link Functions
Gamf _Z (z μ, σ)	$\frac{1}{(\mu\sigma^2)^{1/\sigma^2}} \frac{z^{(1/\sigma^2-1)} e^{-z(\mu\sigma^2)}}{\Gamma(1/\sigma^2)} z > 0, \mu > 0, \sigma > 0$	$E(Z) = \mu$ $SD(Z) = \mu\sigma$	$g_1(\mu) = \ln(\mu)$ $g_2(\sigma) = \ln(\sigma)$
Weibf _Z (z μ, σ)	$\frac{\sigma z^{\sigma-1}}{\mu^\sigma} \exp\left[-\left(\frac{z}{\mu}\right)^\sigma\right] z > 0, \mu > 0, \sigma > 0$	$E(Z) = \mu\Gamma\left(\frac{1}{\sigma} + 1\right)$ $SD(Z) = \mu\sqrt{\Gamma\left(\frac{2}{\sigma} + 1\right) - [\Gamma\left(\frac{1}{\sigma} + 1\right)]^2}$	$g_1(\mu) = \ln(\mu)$ $g_2(\sigma) = \ln(\sigma)$
Gumbf _Z (z μ, σ)	$\frac{1}{\sigma} \exp\left\{-\left(\frac{z-\mu}{\sigma}\right) - \exp\left[-\left(\frac{z-\mu}{\sigma}\right)\right]\right\}$ $-\infty < z < \infty, -\infty < \mu < \infty, \sigma > 0$	$E(Z) = \mu + \gamma\sigma \approx \mu + 0.57722\sigma$ $SD(Z) = \frac{\pi}{\sqrt{6}}\sigma \approx 1.28255\sigma$	$g_1(\mu) = \mu$ $g_2(\sigma) = \ln(\sigma)$
Logistf _Z (z μ, σ)	$\frac{1}{\sigma} \left\{ \exp\left[-\left(\frac{z-\mu}{\sigma}\right)\right] \left\{ 1 + \exp\left[-\left(\frac{z-\mu}{\sigma}\right)\right] \right\}^{-2} \right\}$ $-\infty < z < \infty, -\infty < \mu < \infty, \sigma > 0$	$E(Z) = \mu$ $SD(Z) = \frac{\pi}{\sqrt{3}}\sigma \approx 1.81380\sigma$	$g_1(\mu) = \mu$ $g_2(\sigma) = \ln(\sigma)$
Lognf _Z (z μ, σ)	$\frac{1}{\sqrt{2\pi}} \frac{1}{z} \exp\left\{-\frac{(\log(z)-\mu)^2}{2\sigma^2}\right\}$ $z > 0, \mu > 0, \sigma > 0$	$E(Z) = w^{1/2}e^\mu$ $SD(Z) = \sqrt{w(w-1)}e^\mu$ $w = \exp(\sigma^2)$	$g_1(\mu) = \mu$ $g_2(\sigma) = \ln(\sigma)$
Pearsf _Z (z μ, σ, κ)		$E(Z) = \mu$ $Cv = \sigma$ $Cs = 2\kappa$	$g_1(\mu) = \ln(\mu)$ $g_2(\sigma) = \ln(\sigma)$ $g_3(\kappa) = \kappa$
III	$\frac{1}{\sigma\mu\kappa\Gamma(1/\kappa^2)} \left(\frac{z-\mu}{\mu\sigma\kappa} + \frac{1}{\kappa^2} \right)^{\frac{1}{\kappa^2}-1} \exp\left[-\left(\frac{z-\mu}{\mu\sigma\kappa} + \frac{1}{\kappa^2}\right)\right]$ $\sigma > 0, \kappa \neq 0, \frac{z-\mu}{\mu\sigma\kappa} + \frac{1}{\kappa^2} \geq 0$		
GEVf _Z (z μ, σ, κ)	$\frac{1}{\sigma} [1 + \kappa\left(\frac{z-\mu}{\sigma}\right)]^{(-1/\kappa)-1} \exp\left\{-[1 + \kappa\left(\frac{z-\mu}{\sigma}\right)]^{-1/\kappa}\right\}$ $-\infty < \mu < \infty, \sigma > 0, -\infty < \kappa < \infty$	$E(Z) = \mu - \frac{\sigma}{\kappa} + \frac{\sigma}{\kappa}\eta_1$ $SD(Z) = \sigma\sqrt{\eta_2 - \eta_1^2}/\kappa$ $\eta_m = \Gamma(1-m\kappa)$	$g_1(\mu) = \mu$ $g_2(\sigma) = \ln(\sigma)$ $g_3(\kappa) = \kappa$

Table 1: Summary of the distributions used to model the low-flow series (?)

7.6.4.2 Model Selection

The optimal nonstationary model was selected by penalizing more complex models using the Akaike Information Criterion (AIC) (?) and the Bayesian Information Criterion (BIC) (?). The AIC is calculated as:

$$\text{AIC} = -2 \ln(\text{ML}) + 2k \quad (18)$$

where ML is the maximum likelihood function of the models and k is the number of independently adjusted parameters within the model. Theoretically, $-\infty < \text{AIC} < \infty$. Similarly, BIC is calculated as:

$$\text{BIC} = -2 \ln(\text{ML}) + k \ln(n) \quad (19)$$

where n is the number of observations. The model with the smallest AIC and BIC values is considered the optimal one.

7.6.4.3 Goodness-of-fit

While the information criterion value identifies the optimal model, it is not a measure of model performance. The goodness-of-fit of the selected optimal model was assessed qualitatively using the worm plots (Buuren and Fredriks, 2001) and centile curves diagnostic plots, and quantitatively using the statistics of the Filliben correlation coefficient (denoted by F_r) (?) and the Kolmogorov-Smirnov (KS) test (denoted by D_{KS}) (?). F_r and D_{KS} are calculated by Eqs. (20) and (22), respectively:

$$F_r = \text{Cor}(S, B) = \frac{\sum_{i=1}^{\tau} (S_{(i)} - \bar{S})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^{\tau} (S_{(i)} - \bar{S})^2 \sum_{i=1}^{\tau} (B_i - \bar{B})^2}} \quad (20)$$

where $S_{(i)}$ are the ordered residuals obtained by sorting $\Phi^{-1}[F_Z(z_i, \theta_i)]$ in ascending order for $1 \leq i \leq \tau$. Here, Φ^{-1} denotes the inverse function of the standard normal distribution, and τ represents the length of the observation period (similarly hereinafter). B_i are the standard normal order statistic medians calculated from $\Phi^{-1}(b_i)$, where b_i are derived by:

$$b_i = \begin{cases} 1 - b_{\tau} & i = 1 \\ (i - 0.3175)/(\tau + 0.365) & i = 2, 3, \dots, \tau - 1 \\ 0.5^{(1/\tau)} & i = \tau \end{cases} \quad (21)$$

F_r ranges from $(0, 1]$, and an F_r value greater than the critical value F_{α} indicates that the nonstationary model passes the goodness-of-fit test.

$$D_{KS} = \max_{1 \leq i \leq \tau} |\hat{G}_i - G_{(i)}| \quad (22)$$

where \hat{G}_i are the empirical cumulative probabilities calculated as $i/(\tau + 1)$, and $G(i)$ are the ordered theoretical cumulative probabilities obtained by sorting $F_Z(z_i, \theta_i)$ in ascending order for $1 \leq i \leq \tau$. D_{KS} ranges from $[0, 1]$, and a D_{KS} value smaller than the critical value D_{α} indicates that the nonstationary model passes the goodness-of-fit test.

To summarize, the main steps in deriving time-varying exceedance probabilities p_t for future years are:

- (i) Nonstationary modeling of the observed low-flow series using either time alone or meteorological variables as covariates.
- (ii) Calculating the design low-flow quantile z_{p_0} corresponding to an initial return period T_0 from the quantile function $z_{p_0} = F_Z^{-1}(p_0, \theta_0)$, where $p_0 = 1/T_0$ and θ_0 is the fitted statistical parameter set of the initial year $t = 0$.
- (iii) Deriving time-varying exceedance probabilities p_t corresponding to z_{p_0} for future years $t = 1, 2, \dots, \infty$ (when using time as a covariate) or $t = 1, 2, \dots, t_{\max}$ (when using meteorological covariates) from Eq. (5) where θ_t is calculated by extending the optimal nonstationary model from step (i) into the future under the respective case of covariates.

Applications of Model Selection and Goodness of Fit

This table summarizes the model selection and goodness of fit for low flow series as analyzed in three different studies. Each study applied various statistical criteria to select the best-fit models for different hydrological stations. The goodness of fit measures include diagnostic tools like Worm plots, Kolmogorov-Smirnov test, and others. The selected distributions and their parameters are highlighted for each station.

Study	Selected Model for Each Station	Goodness of Fit		Selected Distribution with Parameters
		Model	teria	
Du et al. (2015)	Huaxian: AIC et Xianyang: Weibull Gamma	Worm plots, Centile curves, Filliben Correlation Coefficient, Kolmogorov- Smirnov test		Huaxian: $\ln(\mu) \sim \text{Temp}$, $\ln(\sigma) \sim \text{Pr}$ Xianyang: $\ln(\mu) \sim \text{Temp}$, $\ln(\sigma) \sim \text{constant}$
Wang et al. (2022)	DPL: AIC et Weibull TG: Weibull ZGP: Weibull XX: Weibull	Worm plots, Kolmogorov- Smirnov test		DPL: $\mu \sim \text{cs}(\text{Pr}, 2)$, $\sigma \sim \text{constant}$ CTG: $\mu \sim \text{cs}(\text{Pr}, 3)$, $\sigma \sim \text{constant}$ ZGP: $\mu \sim \text{cs}(\text{Pr}, 2) + \text{RI}$, $\sigma \sim \text{RI}$ XX: $\mu \sim \text{cs}(\text{Pr}, 2)$, $\sigma \sim \text{constant}$
Jiang et al. (2015)	Qa: Gamma AICc	Worm plots, Kolmogorov- Smirnov test Gamma		Qa: $\mu \sim \text{constant}$, $\sigma \sim \ln(\text{RI}_a)$ Qh: $\mu \sim \ln(\text{RI}_h)$, $\sigma \sim \text{constant}$

- **Du et al. (2015):**
 - **Huaxian Station:** Selected Weibull distribution with temperature (Temp) and precipitation (Pr) as covariates.
 - **Xianyang Station:** Selected Gamma distribution with temperature and a constant scale parameter.
- **Wang et al. (2022):**
 - Selected Weibull distribution for **DPL**, **CTG**, **ZGP**, and **XX** stations with various covariates. Precipitation was modeled using cubic splines (cs) and reservoir index (RI) was included for some stations.
- **Jiang et al. (2015):**
 - Selected Gamma distribution for **Qa** and **Qh** low-flow series of the respective stations with reservoir index as covariate.

These results underscore the importance of incorporating climatic variables (e.g., temperature and precipitation) and anthropogenic factors (e.g., reservoir index) into nonstationary models to better capture the variability and trends in low flow series. The use of cubic splines to model non-linear relationships and the selection of appropriate distributions based on AIC and other diagnostic tools (e.g., worm plots, Filliben correlation coefficient) ensure that the models are well-fitted to the observed data, thereby providing robust tools for hydrological analysis and water resource management.

7.7 Bivariate Statistical Analysis

Bivariate statistical analysis is an essential tool in hydrology, providing insights into the interactions between different hydrological variables, especially during low-flow events. Unlike univariate analysis, which examines a single variable in isolation, bivariate analysis allows us to explore the joint behavior of two variables, offering a more comprehensive understanding of hydrological phenomena.

One of the advanced methodologies in this domain is the use of copulas, which enable the modeling of dependencies between variables without the restrictive assumptions of linearity or normality. Copulas provide a flexible framework to construct joint distributions, capturing the dependence structure between variables more accurately.

Given the dynamic nature of hydrological processes influenced by climate change and human activities such as reservoir operations, it is crucial to account for nonstationarity in these analyses. This is where time-varying copulas come into play. Unlike static copulas, time-varying copulas allow the parameters defining the dependence structure to evolve over time, reflecting the changing environmental conditions.

In this context, the content that will be presented here is heavily based on the methodologies and findings from the study “Bivariate frequency analysis of nonstationary low-flow series based on the time-varying copula” by Jiang et al (2015). This study exemplifies how time-varying copulas can effectively model the joint distribution of low-flow events under nonstationary conditions, taking into account both the marginal distributions and the evolving dependence structure.

7.7.1 Joint Return Period Under Nonstationary Framework

Three methods have been used to calculate the joint return period (JPR) of low-flow events in stationary bivariate frequency analysis: the AND method, corresponding to the probability $P(Y_1 \leq y_1 \wedge Y_2 \leq y_2)$; the OR method, corresponding to the probability $P(Y_1 \leq y_1 \vee Y_2 \leq y_2)$; and the Kendall (KEN) return period method. The Kendall return period, a multivariate return period first defined by Salvadori and De Michele (2004b), has been widely applied in analyzing the joint return periods of floods or droughts (? , ? , ? , ? , ?). The JPR using the KEN method for low-flow events is given by:

$$\begin{aligned} T_{KEN} &= \frac{\lambda}{P[C(u_1, u_2 | \theta_c) \leq p_{KEN}]} \\ &= \frac{\lambda}{P\{C[F_1(y_1 | \theta_1), F_2(y_2 | \theta_2) | \theta_c] < p_{KEN}\}} \\ &= \frac{\lambda}{K_C(p_{KEN})} \end{aligned} \quad (23)$$

where λ is the average interarrival time between low-flow event occurrences. In this paper, the annual minimum low-flow series is investigated, so λ is set to 1 (i.e., $\lambda = 1$). $K_C(\cdot)$ is the Kendall distribution function (? , ? , ?), which provides a univariate representation of multivariate information, and p_{KEN} is a critical probability level corresponding to the value of $K_C(p_{KEN})$ (? , ?).

Similar to calculating the JPR of low-flow events in stationary bivariate frequency analysis, the JPRs for AND, OR, and KEN methods in nonstationary circumstances are defined as follows:

$$\begin{aligned}
T_{AND}^t &= \frac{1}{P(Y_1 \leq y_1 \wedge Y_2 \leq y_2)} \\
&= \frac{1}{C[F_1(y_1 | \theta_1^t), F_2(y_2 | \theta_2^t) | \theta_c^t]} \\
T_{OR}^t &= \frac{1}{P(Y_1 \leq y_1 \vee Y_2 \leq y_2)} \\
&= \frac{1}{P(Y_1 \leq y_1) + P(Y_2 \leq y_2) - P(Y_1 \leq y_1 \wedge Y_2 \leq y_2)} \\
&= \frac{1}{F_1(y_1 | \theta_1^t) + F_2(y_2 | \theta_2^t) - C[F_1(y_1 | \theta_1^t), F_2(y_2 | \theta_2^t) | \theta_c^t]} \\
T_{KEN}^t &= \frac{1}{P[C(u_1^t, u_2^t | \theta_c^t) \leq p_{KEN}]} \\
&= \frac{1}{P\{C[F_1(y_1 | \theta_1^t), F_2(y_2 | \theta_2^t) | \theta_c^t] < p_{KEN}\}} \\
&= \frac{1}{K_C^t(p_{KEN})}
\end{aligned} \tag{24}$$

Joint Return Period in Jiang et al. (2015)

Jiang et al. (2015) analyzed the joint return period of low-flow events at the Ankang (Qa) and Huangzhuang (Qh) gauges on the Hanjiang River using a time-varying copula model. The study used three methods to calculate the JRP: AND, OR, and Kendall (KEN) methods.

Time Variation in the Joint Return Periods:

- **1954–1967:** Before major reservoirs were built.
- **1968–1991:** After the Danjiangkou Reservoir was operational.
- **1992–2011:** After the Ankang Reservoir was operational.

The JRP isolines of the design low-flow pairs for a given JRP of 50 years were calculated for each time segment. The study found that reservoir operations significantly impacted the joint distribution and return periods of low flows at both gauges.

Results:

- **AND Method:** The JRP isoline moved horizontally to the right from 1954–1967 to 1968–1991 due to the increased mean of Huangzhuang low flow. It then moved downward from 1968–1991 to 1992–2011 due to the increased coefficient of variation of Ankang low flow.
- **OR and KEN Methods:** Showed similar variations to the AND method.

Implications:

- The movement of JRP isolines indicates the significant influence of reservoir operations on the hydrological regime. Specifically, the changes in JRP isolines reflect how human activities, such as the construction and operation of reservoirs, alter the dependence structure between low-flow events at different gauges.
- The results highlight the necessity of considering nonstationarity in hydrological frequency analysis to accurately assess the impacts of climate variability and human interventions on water resources.

These findings emphasize the importance of integrating nonstationary models and anthropogenic factors in hydrological studies to improve the reliability of low-flow frequency analysis and water resource management.

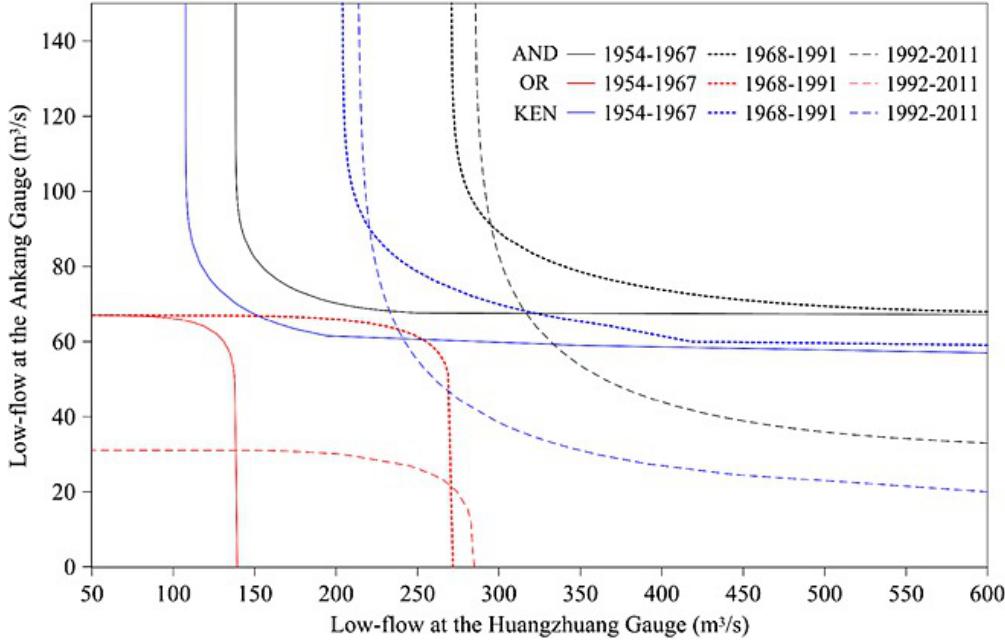


FIGURE 7.3: The isolines of design low-flow events with JRP=50 years for three different time periods of 1954-1967, 1968-1991 and 1992-2011
(#fig:fig3_gamze)

7.7.2 Time-Varying Copula Model

Based on the definition of the copula (?, ?) the time-varying copula can be expressed as:

$$\begin{aligned} H_{Y_1, Y_2}(y_1^t, y_2^t) &= C[F_1(y_1^t | \theta_1^t), F_2(y_2^t | \theta_2^t) | \theta_c^t] \\ &= C(u_1^t, u_2^t | \theta_c^t) \end{aligned} \quad (26)$$

where $C(\cdot)$ denotes the copula function, $F(\cdot)$ is the cumulative distribution function, θ_1^t and θ_2^t represent the time-varying marginal distribution parameters, and θ_c^t is the time-varying copula parameter. Additionally, the marginal probabilities u_1^t and u_2^t within the time-varying copula must both follow a uniform distribution over the interval $[0, 1]$.

According to Equation (27), three scenarios for the time-varying copula can be derived:

- all the marginal distribution parameters are constant while the copula parameter is time-varying,
- at least one marginal distribution parameter is time-varying while the copula parameter remains constant, and
- both at least one marginal distribution parameter and the copula parameter are time-varying.

The implementation of the time-varying copula model in Equation (27) involves two major steps: first, determining the time variation in the marginal distributions, and second, modeling the evolution of the copula parameter. The first step is done under the GAMLSS framework we have discussed before. As the second step the copula parameter θ_c can be expressed as a linear function of the time-varying explanatory variables $x_i^t (i = 1, 2, \dots, m)$ through a suitable link function $g_c(\cdot)$, as follows:

$$g_c(\theta_c^t) = \beta_0 + \sum_{i=1}^m \beta_i x_i^t \quad (27)$$

where $\beta_0, \beta_1, \dots, \beta_m$ are the parameters represented by the vector $\beta = (\beta_0, \beta_1, \dots, \beta_m)^T$. The link function $g_c(\cdot)$ depends on the domain of the copula parameter. For example, if $\theta_c \in \mathbb{R}$, then $g_c(\theta_c) = \theta_c$ (for the Frank copula), or if $\theta_c > 0$, then $g_c(\theta_c) = \log(\theta_c)$ (for the GH and Clayton copulas). The time-varying copula parameters β_i are estimated by Inference Function for Margins (IFM) method (?).

7.7.3 Model Selection and Goodness-of-fit

Model selection in terms of copula functions involves evaluating different copula models to determine which best fits the data. Similar to bivariate analysis, criteria such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are commonly used for this purpose. Additionally, variations like the corrected Akaike Information Criterion (AICc) can be employed, especially when dealing with smaller sample sizes. By comparing the AIC, BIC, or AICc values of various copula models, the model with the lowest criterion value is chosen, indicating the best trade-off between model complexity and goodness of fit.

The goodness-of-fit of the time-varying copula model is evaluated by assessing the fit of both the two marginal distributions and the copula function. The goodness-of-fit test for the copula is conducted using Rosenblatt's probability integral transform (? , ?).

Chosen Copula Models and AICc Values in Jiang et al. (2015)

The study by Jiang et al. (2015) focused on analyzing the time-varying dependence structure between the low-flow series at the Ankang (Qa) and Huangzhuang (Qh) gauges. The authors selected three Archimedean copulas (Frank, Gumbel-Hougaard (GH), and Clayton) due to their flexibility and ability to capture different dependence structures. The models were evaluated using two different explanatory variables: time and reservoir index (RI). The selection criteria include the corrected Akaike Information Criterion (AICc), which is preferred over the regular AIC when dealing with small sample sizes as it adjusts for the number of parameters in the model, providing a more accurate measure of model fit.

With Time as the Explanatory Variable

Copula Type	Copula Parameter Model	AICc
Frank	$\theta(t) = 7.580 + 0.091t$	-29.41
GH	$\theta(t) = \exp(0.772 + 0.007t)$	-27.50
Clayton	$\theta(t) = \exp(0.442 + 0.022t)$	-14.33

With Reservoir Index as the Explanatory Variable

Copula Type	Copula Parameter Model	AICc
Frank	$\theta(RI_h) = 9.096 - 9.753RI_h$	-31.10
GH	$\theta(RI_h) = \exp(0.970 - 0.946RI_h)$	-30.16
Clayton	$\theta(RI_h) = \exp(0.461 - 1.681RI_h)$	-14.61

The tables shows that the Frank copula with the reservoir index (RI) as the explanatory variable had the lowest AICc value of -31.10, indicating the best fit among the evaluated models. This is followed by the GH copula with an AICc of -30.16, and the Clayton copula with an AICc of -14.61. In comparison, when time was used as the explanatory variable, the Frank copula still performed best with an AICc of -29.41, but its fit was not as good as with the reservoir index.

The results demonstrate that the Frank copula model, particularly when using the reservoir index as the explanatory variable, provides the most accurate representation of the time-varying dependence structure between the low-flow series at the two gauges. This highlights the significant impact of reservoir operations on the hydrological behavior of the Hanjiang River, emphasizing the importance of considering such anthropogenic factors in hydrological modeling.

7.8 Hybrid Dynamical–Statistical Approaches

Hybrid dynamical–statistical (or statistical–dynamical) approaches offer several advantages over purely statistical or purely dynamical methods for predicting and projecting nonstationary extremes. In nonstationary hydrological analysis, a time or climate covariate can be used to detect changes in the distribution parameters. The benefit of using climate covariates is that climate model predictions or projections can then be utilized as covariates to estimate future changes.

Hydrologists are interested in catchment-scale projections, but climate models operate at much larger resolutions, necessitating downscaling for use in process-based models. Due to the coarse spatial and temporal resolution of climate forecasts, downscaling is required to provide regional information at finer temporal and spatial scales (?, ?). Various statistical downscaling and bias correction methods are employed. In the following section, we will present the Statistical Downscaling Model employed by Du et al. (2015) in the context of low-flow events.

7.8.1 Statistical Downscaling Model (SDSM)

GCMs are tools used for predicting future time series of meteorological variables, thereby extending the time-varying exceedance probabilities described in Eqs. (13) and (15). The usefulness of GCMs for local impact studies is limited by their coarse spatial resolution, typically around 50,000 km², and their inability to resolve important sub-grid scale features such as clouds and topography which can be seen from the schematic illustration of general approach of SDSM Figure ?? below. Therefore, the coarse spatial resolution of GCM data limits its direct application to local impact studies (?, ?).

Source: SDSM Project Team, 2024

This issue can be addressed through a technique known as downscaling. The statistical downscaling model (SDSM) developed by Wilby et al. (2002) serves as a decision support tool for assessing local climate change impacts. SDSM can be best described as a hybrid of stochastic weather generator and regression-based methods. It employs a robust statistical downscaling technique that combines a weather generator and multiple linear regression. SDSM uses large-scale daily circulation patterns and atmospheric moisture variables ($j = 1, 2, \dots, n$) to linearly condition local-scale weather generator parameters (e.g., precipitation occurrence and intensity) at individual sites. It generates local weather scenarios by combining weather generator outputs with future climate scenarios projected by GCMs.

The diagram ?? illustrates the general approach of SDSM, showing how it integrates both NCEP predictors and GCM predictors through a downscale predictand process. First, the NCEP predictors feed into a weather generator. Then, the simulated historical data coming from weather generator is assessed using the Nash–Sutcliffe efficiency (NSE) between the simulated and observed meteorological variables of interest. With the input of the GCMs large-scale predictors and outputs of the weather generator, future values for the meteorological variable of interest are generated by the scenario generator. Both the weather generator and the scenario generator contribute to the model output, which represents the localized weather patterns and events based on projected climate changes.

Source: SDSM Project Team, 2024

It has been widely used in climate change research (Wilby and Dawson, 2013) and is comprehensively described in Wilby et al. (2002) and Wilby and Dawson (2007).

Theoretically, the downscaling process with SDSM can be either unconditional or conditional on an event. For the unconditional process, a direct multiple linear regression equation is constructed between the unconditional predictand y_i^{UC} and the chosen normalized large-scale predictors \hat{u}_i^j on day i (?, ?):

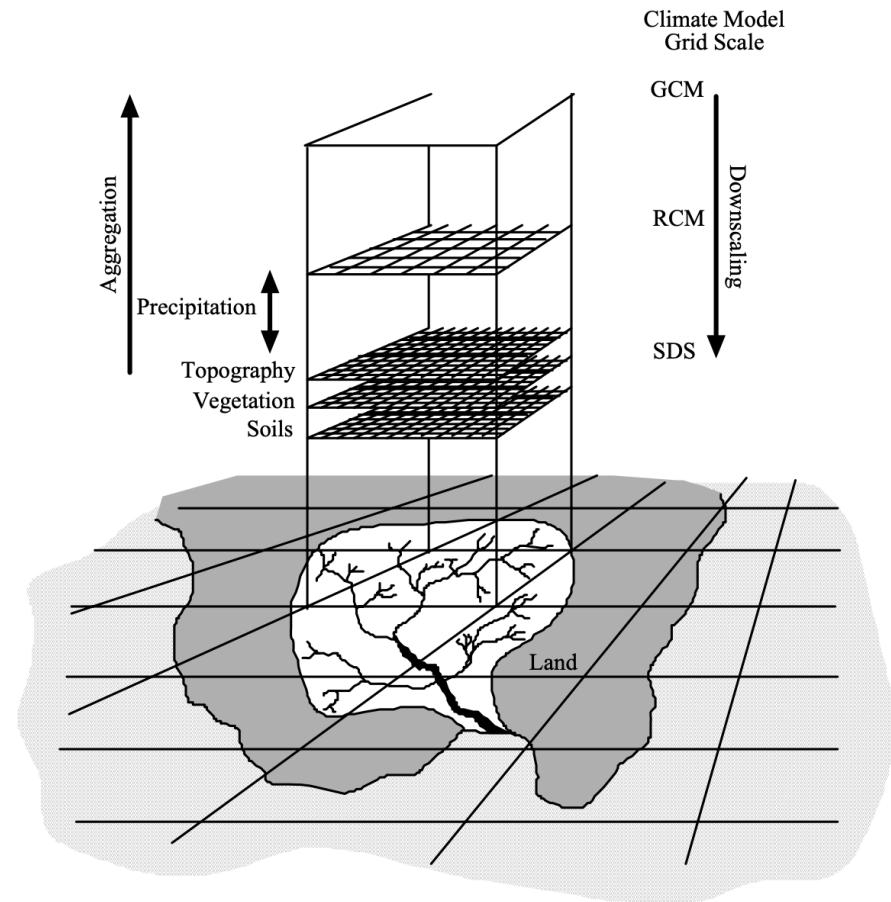


FIGURE 7.4: Schematic illustration of SDSM
(#fig:fig4_gamze)

$$y_i^{UC} = \gamma_0 + \sum_{j=1}^l \gamma_j \hat{u}_i^j + \varepsilon \quad (28)$$

where γ_j are the estimated regression coefficients determined by the least squares method, l is the number of selected predictors, and ε is a normally distributed stochastic error term.

For the conditional process, the conditional probability of predictand occurrence ω_i on day i is directly represented as a multiple linear regression equation of \hat{u}_i^j :

$$\omega_i = \eta_0 + \sum_{j=1}^l \eta_j \hat{u}_i^j \quad (29)$$

where η_j are the estimated regression coefficients. If $\omega_i \leq r_i$, where r_i is a uniformly distributed random number ($0 \leq r_i \leq 1$), the conditional predictand y_i^c occurs with an amount of:

$$y_i^C = F^{-1} [\Phi(Z_i)] \quad (30)$$

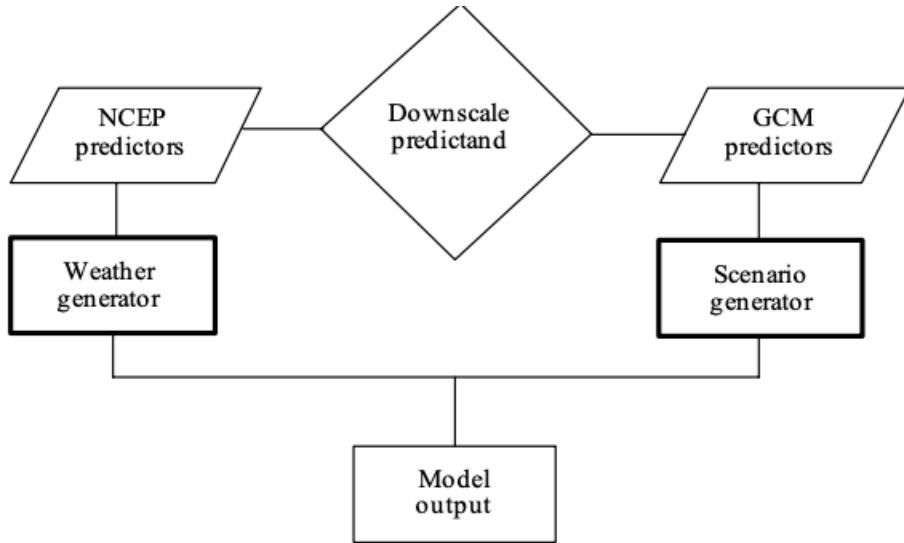


FIGURE 7.5: SDSM Climate Scenario Generation
(#fig:fig5_gamze)

where F is the empirical distribution function of y_i^c , and Φ is the normal cumulative distribution function. Z_i is the z -score for day i , expressed as $Z_i = \lambda_0 + \sum_{j=1}^l \lambda_j \hat{u}_j i^j + \varepsilon$, where λ_j are the estimated regression coefficients.

The simulation results were assessed by the Nash-Sutcliffe efficiency (NSE) between the simulated and observed Temp and Prep. NSE is calculated as (?):

$$NSE = 1 - \frac{\sum_{i=1}^{\tau} (Y_i^{\text{obs}} - Y_i^{\text{sim}})^2}{\sum_{i=1}^{\tau} (Y_i^{\text{obs}} - Y^{\text{mean}})^2} \quad (31)$$

where Y_i^{obs} are the observed meteorological variables, Y_i^{sim} are the simulated values corresponding to Y_i^{obs} , Y^{mean} is the mean value of Y_i^{obs} . NSE ranges from $-\infty$ to 1, with $NSE = 1$ being the perfect simulation.

Typically, the predictand-predictor relationship in the SDSM, represented by the calibrated multiple linear regression equation, is assumed to be transferable to future projection periods (? , ? , ? , ? , ?). Using the large-scale predictors from GCMs as input, future scenarios of the meteorological variables of interest can then be derived, providing localized weather patterns and events based on projected climate changes.

Statistical Downscaling Model (SDSM) in Du et al. (2015)

Du et al. (2015) employed the Statistical Down Scaling Model developed by Wilby et al. (2002) as a decision support tool for assessing local climate change impacts.

Key Features of SDSM:

- **Unconditional Process** is used for the simulation of temperature.
- **Conditional Process** is used for the simulation of precipitation.

Downscaling Procedure:

- **Selection of Predictors:** Mean sea level pressure, 500 hPa geopotential height, 500 hPa eastward wind, 850 hPa air temperature, and near-surface air temperature were selected for downscaling daily average temperature. For daily total precipitation, 850 hPa specific humidity was included.

- **Model Calibration:** SDSM was calibrated using National Centers for Environmental Prediction (NCEP) reanalysis data to establish the relationship between predictands (daily average temperature and daily total precipitation) and large-scale predictors.
- **Simulation:** Daily average temperature and daily total precipitation for the period 1954–2009 were simulated using SDSM driven by NCEP reanalysis predictors. The Nash-Sutcliffe Efficiency (NSE) between the simulated and observed data was used to assess model performance.
- **Future Projections:** The calibrated SDSM model was used to generate future scenarios of temperature and precipitation for 2010–2099 based on predictors from multiple GCMs under the RCP8.5 scenario.

Results:

- The simulated data for 1954–2009 showed adequate agreement with observed data, particularly for temperature. - Future projections indicated strong increasing trends in temperature for both Huaxian and Xianyang stations, while precipitation remained relatively stable for Huaxian.
- The use of SDSM allowed Du et al. (2015) to effectively incorporate climate change projections into their nonstationary frequency analysis of low-flow events, providing valuable insights for water resource management under changing climatic conditions.
- It should be noted that there are uncertainties in the downscaled future scenarios of meteorological variables used in this study. These uncertainties could be partially mitigated by exploring additional downscaling methods and different General Circulation Models (GCMs).

7.9 Conclusion

In this chapter, we delved into the complex nature of low-flow events, emphasizing their prediction and analysis through advanced hydrological frequency models. Low-flow events, characterized by significantly reduced streamflow and water availability, pose serious challenges to water resource management, especially under changing climatic conditions. Our exploration highlighted the critical role of return periods and risk assessments in understanding and mitigating the impacts of these events.

Traditional hydrological models, which assume stationarity, rely on historical data to predict future occurrences of low-flow events. However, this assumption is increasingly challenged by the evident impacts of climate change and human activities, which alter hydrological processes over time. Nonstationary models, which account for these dynamic factors, provide a more accurate and realistic approach to predicting low-flow events.

One of the key advancements in nonstationary hydrological analysis is the incorporation of time-varying parameters and meteorological covariates. By integrating these factors, nonstationary models can better capture the evolving nature of low-flow events. Meteorological covariates, such as temperature and precipitation, have been shown to significantly enhance the accuracy of low-flow predictions by linking hydrological responses to climatic conditions. This approach allows for a more nuanced understanding of how climate change influences hydrological extremes.

The use of copulas in bivariate analysis further enriches our understanding of low-flow events by modeling the dependencies between multiple hydrological variables. Time-varying copulas, which allow the dependence structure between variables to evolve over time, offer a sophisticated method for analyzing the joint behavior of hydrological extremes. This is particularly useful in assessing the risk and return periods of compound extreme events, where multiple variables interact to influence the occurrence and severity of low-flow conditions.

Reservoir indices, which account for the impact of reservoir operations on flow regimes, also play a crucial

role in low-flow predictions. These indices provide a more comprehensive view of how human interventions, such as water withdrawals and reservoir management, affect streamflow patterns. By integrating reservoir indices into nonstationary models, we can better predict low-flow events in regulated river systems.

The hybrid dynamical–statistical approach, exemplified by the Statistical Downscaling Model (SDSM), represents another significant advancement in hydrological modeling. This approach combines the strengths of both dynamical and statistical methods, using climate model outputs to predict future hydrological conditions at a local scale. SDSM effectively bridges the gap between large-scale climate projections and localized hydrological impacts, providing valuable insights for water resource management under future climate scenarios.

The integration of nonstationary models, meteorological covariates, and advanced statistical methods offers a more robust framework for understanding and predicting low-flow events. These approaches provide a more accurate and comprehensive view of hydrological extremes, enabling better preparation and mitigation strategies in water resource management.

Looking forward, the ongoing development and refinement of nonstationary models are crucial for enhancing our ability to predict and manage low-flow events. As future research addresses the expansion of explanatory variables, improvements in model accuracy, and the integration of more comprehensive datasets, these efforts will enable the creation of more resilient and adaptive water management strategies, effectively tackling the challenges posed by climate change and other dynamic factors.



8

Statistical streamflow modelling

Author: Lennart Marx

Supervisor: Henri Funk

Degree: Master

8.1 Abstract

This study evaluates and compares the performance of Long Short-Term Memory (LSTM) networks and Temporal Fusion Transformers (TFT) in forecasting streamflows for up to seven day ahead forecasts, using historical streamflow data alongside precipitation and temperature as covariates. Freely available data published by the bavarian hydrological authority from the Regen river in Bavaria, Germany, was used in conjunction with meteorological data from the ERA5 reanalysis dataset. River basins are defined via the HydrorIVERS and HydroBASINS dataset to obtain the area over which precipitation is aggregated. Results indicate that while both models face challenges in predicting extreme values, TFT maintains more consistent accuracy over longer forecast horizons while pure LSTM model's predictions decline sharply in performance with increasing lead times.

8.2 Introduction

Streamflow forecasting plays a crucial role in water resource management, particularly in flood prediction. As climate change intensifies extreme weather events and alters precipitation patterns, accurate streamflow forecasts have become increasingly important for mitigating flood risks by providing information about the timing and magnitude of potential flood events. This knowledge enables authorities to implement timely flood control measures, operate reservoirs effectively, and issue early warnings to vulnerable communities. As flooding remains one of the most common and destructive natural disasters, especially in developing countries, improved forecasting capabilities can significantly reduce the human and economic toll of these events. (?)

Hydropower is the dominant source of renewable energy, accounting for 16% of the world's electricity output and 62% of renewable electricity generation (?). Precise streamflow predictions can be beneficial in optimizing reservoir operations. ? show that this is the case for 94% of all analyzed dams dependent on forecasting ability and reservoir characteristics.

A streamflow refers to the flow of water in rivers, streams or other water bodies, originating from precipitation, melting snow, and groundwater discharge. It is typically measured in cubic meters per second (m^3/s).

In recent years, machine learning techniques have emerged as powerful tools for forecasting in various domains, including hydrology. Traditional hydrological models often rely on extensive datasets and intricate

physical parameters, which can be challenging to obtain and process. In contrast, machine learning models such as Long Short-Term Memory (LSTM) networks offer an alternative approach by learning patterns directly from historical data.

This study aims to evaluate and compare the performance of LSTM and TFT models in forecasting streamflows up to seven days ahead. By incorporating precipitation and temperature as future known covariates alongside historical streamflow data, this work seeks to determine the effectiveness of these models in predicting streamflows.

8.3 Data

8.3.1 Preparation

The streamflow data is gathered from the bavarian hydrology authority (GKD). They provide freely accessible data on most water bodies in Bavaria [?](#). Focusing on rivers of medium size where the entire length of the river is located inside of the state of Bavaria, the river Regen positioned north/east of the city of Regensburg was chosen as a bases for this study. The GKD has data on 21 streamflow gauging stations located at the Regen or any of its tributary rivers. For the Regen river, data between 2014 and 2024 is available with daily measurements on the streamflows including the coordinates of the the gauging station. This study focused on the 15207507 Marienthal gauging station that was located closest to the final outflow towards the Danube river within the city of Regensburg. Utilizing the HydroRIVERS dataset, which contains the shapes of rivers all over the world, it was possible the acquire the shape of the Regen along with its geolocation as shown in figure 8.1 [?](#).

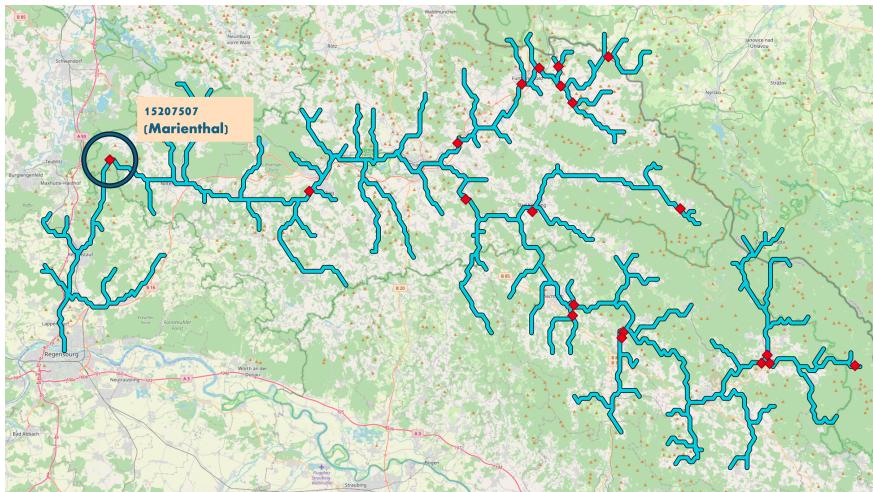


FIGURE 8.1: Streamflow gauging stations that provide there measurements at the GKD along the Regen river (adapted from @OpenStreetMap2017)

A catchment also known as a drainage basin or watershed, is an area of land where all precipitation collects and drains off into a common outlet, such as a river, bay, or other body of water. The boundaries of a catchment are defined by the highest points of land around the drainage area, often referred to as the drainage divide. These catchment areas will later be used to determine the variables for precipitation that are used to forecast the streamflows of the river. Taking advantage of the HydroBASINS dataset, that contains the shapes of basins all over the world in different resolutions [?](#). With the software QGIS a geoinformation system (GIS) all basins were chosen that completely contain the river Regen shape, which led to the 19 defined catchments that can be seen figure 8.2.

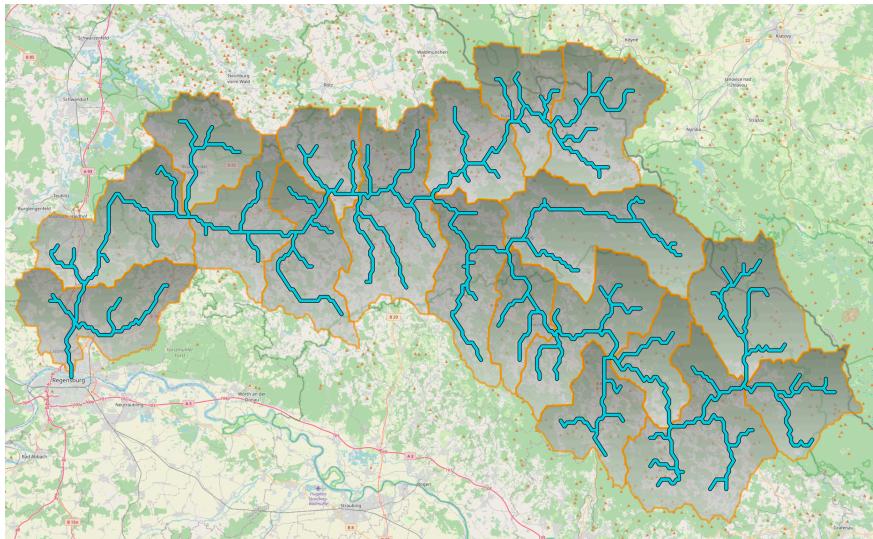


FIGURE 8.2: Catchments defined for the Regen river based on the HydroBASINS shapefile dataset (adapted from @OpenStreetMap2017)

The ERA5 reanalysis dataset, contains a plethora of meteorological data in a rasterized form. Each data cell contains information of when it occurred, where it occurred in the form of longitude and latitude coordinates as well as the information on a meteorological variables ?. As proposed in ?, for this study the variables 2 meter temperature and total precipitation were selected around the area of the Regen and for the past 10 years.

8.3.2 Preprocessing

The measuring station 15207507 contained 1 missing value which was imputed by linear interpolation. As suggested in ? non centered moving average smoothing with a window size of 3 was applied to the streamflow data as can be seen in figure 8.3.

To combine the rasterized meteorological data from ERA5 with the streamflows of the Regen, it is necessary to only take precipitation into account that occurs within the defined catchments. To achieve this, a weighted average is calculated, where the weights are determined by the area size of the intersection between the raster cell of the meteorological data and the catchment. This can be seen in a schematic visualization in figure 8.4.

$$tp(\text{catch}_i) = \frac{1}{A(\text{catch}_i)} \sum_{\text{cell}_j \in \text{ERA5 raster}} (tp(\text{cell}_j) \cdot A(\text{catch}_i \cap \text{cell}_j)) \quad (1)$$

A : area

catch : catchment

cell : one cell in the raster data set of ERA5

tp : total precipitation

Since the catchments are geospatially very close to each other, the different catchments are highly correlated to each other and therefore provide little variance (see Figure 8.5). To reduce the feature space for both variables temperature and precipitation, the mean is taken over all catchments. Finally to reduce the noise in the features a moving average of window size 3 was applied.

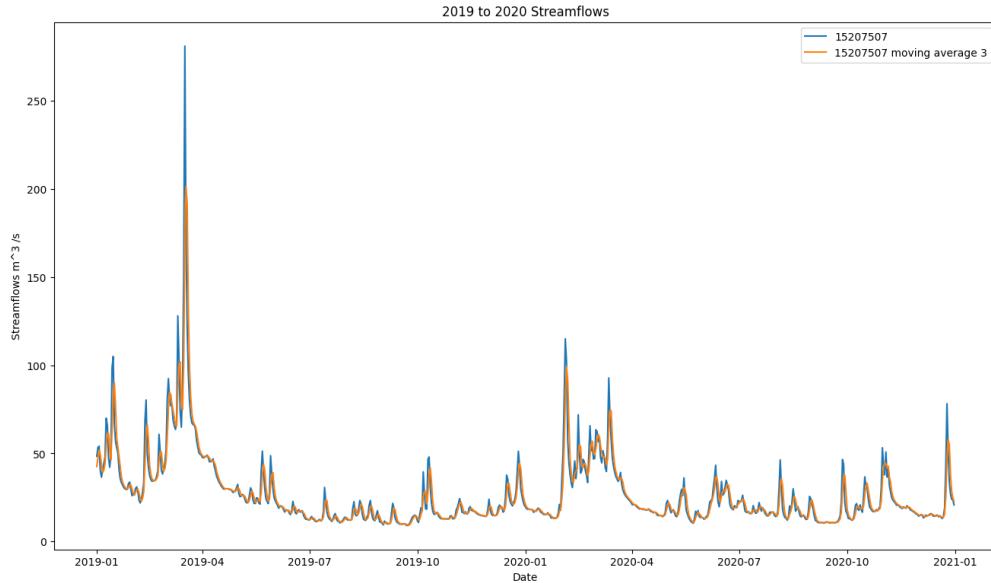


FIGURE 8.3: streamflows of the 15207507 Marienthal gauging stations before (blue) and after (orange) applying moving average smoothing

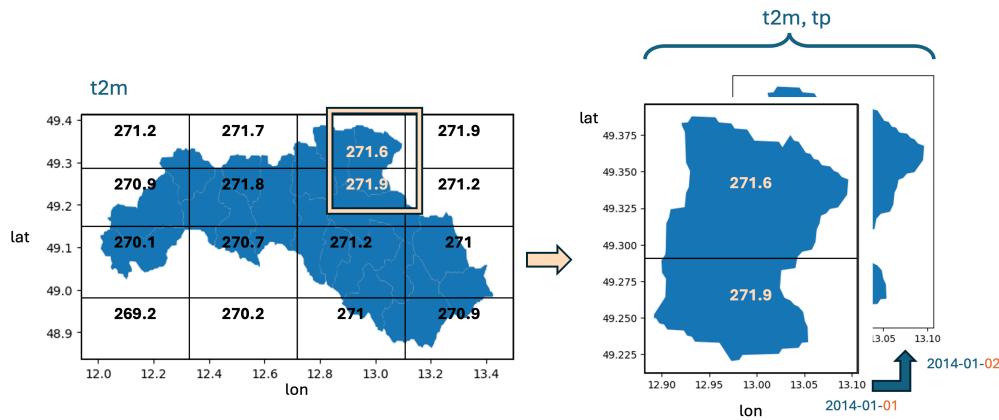


FIGURE 8.4: schematic visualization of the spatial averaging performed for temperature and total precipitation during data preprocessing

8.4 Models

8.4.1 LSTM

The Long Short-Term Memory (LSTM) cell (figure 8.6) is a type of recurrent neural network (RNN) architecture designed to model temporal sequences and their long-range dependencies more accurately than traditional RNNs. LSTMs were introduced by Sepp Hochreiter and Jürgen Schmidhuber in 1997 to address the issue of vanishing and exploding gradients encountered in traditional RNNs ?.

1. **Cell State (C_t):** The internal memory of the cell, which can carry information across many time steps.

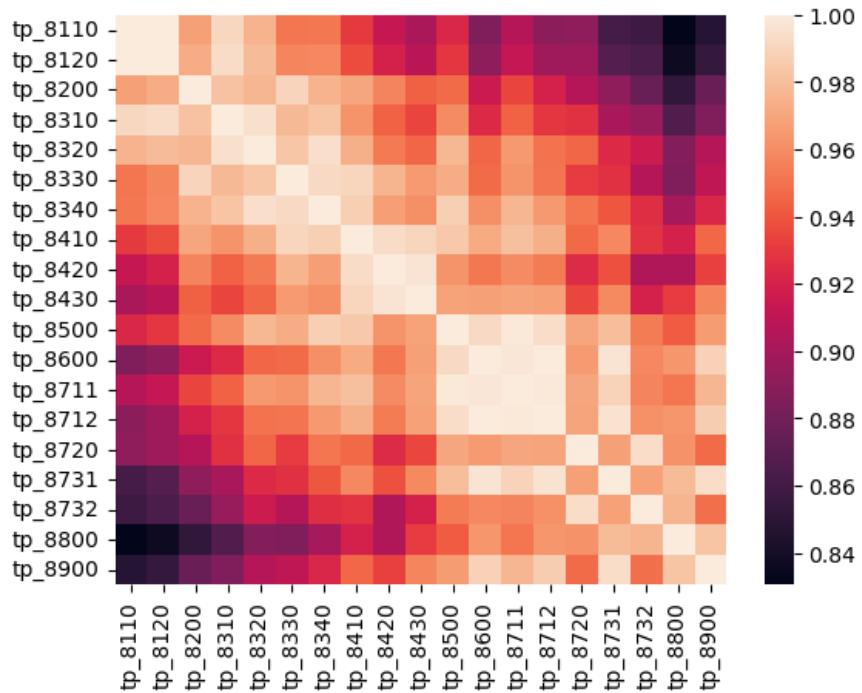
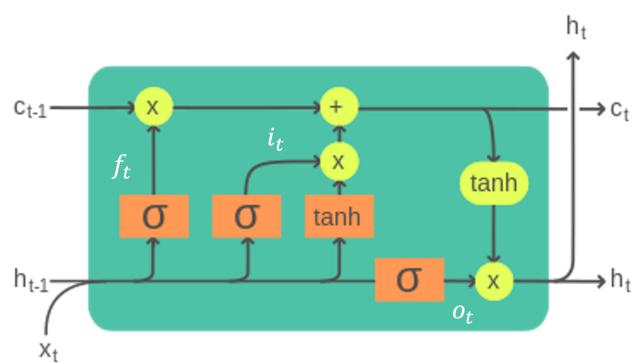


FIGURE 8.5: correlation table for total precipitation over all catchments



Legend: Layer Componentwise Copy Concatenate

FIGURE 8.6: visualization of an LSTM cell at time step t (image adapted from @chevalier2023)

2. **Hidden State (h_t):** The output of the LSTM cell at a given time step, also serving as the input to the next cell.
3. **Input Gate (i_t):** Controls how much of the new information from the current input is used to update the cell state.
4. **Forget Gate (f_t):** Decides how much of the past cell state should be forgotten.
5. **Output Gate (o_t):** Determines the output of the LSTM cell based on the cell state.

LSTMs are particularly well-suited for tasks that involve sequential data and temporal dependencies, such as:

1. Natural Language Processing (NLP):

- Language modeling
- Machine translation
- Speech recognition

2. Time Series Forecasting:

- Stock price prediction
- Weather forecasting
- Anomaly detection

Streamflow forecasting involves predicting the flow of water in rivers and streams over time, which is inherently a time-series problem with temporal dependencies influenced by various factors such as rainfall, snowmelt, and upstream water management. When used in an encoder decoder architecture LSTM-cells can also incorporate future known covariates such as weather forecasts. These specifications make LSTM-based architectures beneficial for modeling and forecasting streamflow data.

8.4.2 Temporal Fusion Transformer

The Temporal Fusion Transformer (TFT) (figure 8.7) is a neural network architecture specifically designed for multi-horizon time series forecasting. It combines the strengths of both recurrent and attention-based models, offering an advanced approach to handling complex time series data. The TFT was introduced by Bryan Lim et al. in 2019, aiming to provide interpretability and accuracy for forecasting tasks ?.

A TFT consists of several key components:

1. Temporal Processing:

- **Local Processing with LSTMs:** LSTMs are used to process local temporal dependencies within the time series.
- **Multi-Head Attention:** An attention mechanism to capture long-range dependencies across different time steps.

2. Variable Selection:

- **Static Covariate Encoders:** Handle static features that do not change over time (e.g., location-specific data).
- **Temporal Covariate Encoders:** Manage time-varying features (e.g., weather data, past values of the time series).

3. Gating Mechanisms:

- **Gated Residual Network (GRN):** Ensures that only relevant information is passed through layers, improving the network's efficiency and interpretability.
- **Variable Selection Networks:** Dynamically select relevant variables at each time step to enhance model performance and interpretability.

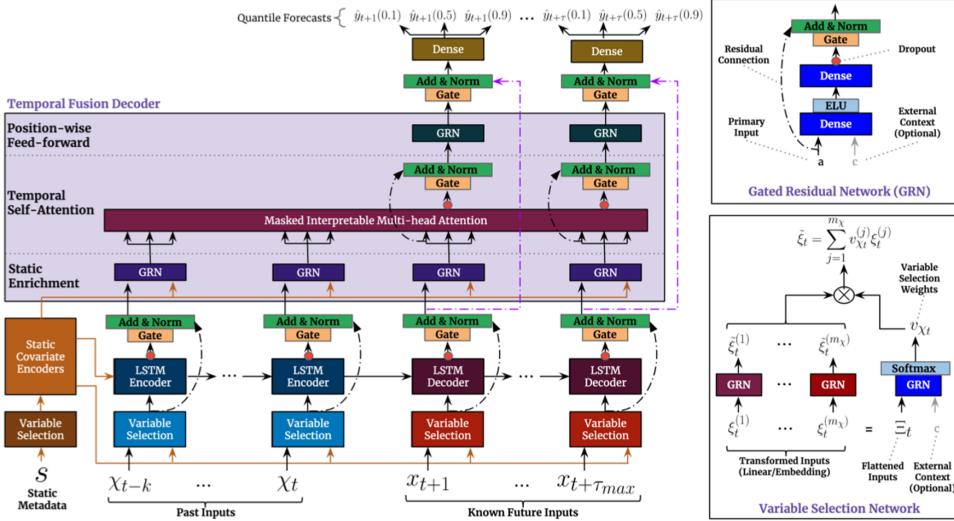


FIGURE 8.7: model architecture for the Temporal Fusion Transformer (image from @lim2021)

4. Multi-Horizon Forecasting:

- **Sequence-to-Sequence Framework:** Allows the TFT to generate forecasts for multiple future time steps simultaneously.

5. Interpretable Outputs:

- **Attention Weights:** Provide insights into which time steps and variables the model is focusing on, aiding interpretability.

The Temporal Fusion Transformer represents an advancement in time series forecasting, offering both high accuracy and interpretability. Its ability to capture complex dependencies, dynamically select relevant features, and provide insights into the decision-making process makes it a useful tool for streamflow forecasting.

8.4.3 Kling Gupta Efficiency

The Kling-Gupta Efficiency (KGE) is a statistical metric used to evaluate the performance of hydrological models by comparing simulated data to observed data. Developed by Gupta et al. in 2009, the KGE addresses limitations found in traditional metrics such as the Nash-Sutcliffe Efficiency (NSE). The KGE decomposes the evaluation of model performance into three distinct components, providing a more comprehensive assessment. These components are correlation, bias, and variability, which help in understanding different aspects of the model's accuracy.

The KGE is calculated using the following formula:

$$\text{KGE} = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (2)$$

- r is the Pearson correlation coefficient between the simulated and observed values.
- α is the bias ratio, defined as the ratio of the mean of the simulated values to the mean of the observed values.
- β is the standard deviation ratio, defined as the ratio of the standard deviation of the simulated values to the standard deviation of the observed values.

TABLE 8.1: Hyperparameter comparison between LSTM and TFT.

Hyperparameter	LSTM	TFT
Batch Size	128	128
Epochs	100	80
Hidden Size	128	128
Attention Head Size	-	2
Learning Rate	0.001	0.003
Dropout	0.2	0.1
Weight Decay	0.001	1e-04
Gradient Clipping	0.1	0.1
Loss Function	Mean Absolute Error	Mean Absolute Error
Optimizer	Adam	Adam
Reduce on Plateau Patience	7	7
Time Varying Known Features	t2m, tp	t2m, tp
Time Varying Unknown Features	streamflow 15207507	streamflow 15207507

8.5 Results

8.5.1 Training Setup

The 2 different model architectures were trained using the historical streamflows as well as temperature and precipitation as covariates. Using an input sequence length of 364 days and an output lead time of up to 7 days. Temperature and precipitation can be used as future known values when considering weather forecasts. For example when trying to predict one step ahead forecast for the streamflow in addition to the past 364 days of precipitation values one can consider the precipitation forecast for the next day to get the best predictions possible.

The LSTM Model is run in an encoder decoder architecture, were the past 364 days are the input for an LSTM cell which returns a hidden state and an output as the encoder step. During the decoder step, the encoder hidden state is fed into the decoder LSTM together with the future known inputs. The model predicts incrementally in the sense that for example to predict a 3 step ahead forecast it firsts predicts 1 and 2 step forecast and uses both forecasts to then predict the 3 step prediction. Both model architectures were used from the pytorch-forecasting library. The models were retrained for the different lead times. The used hyperparameters for both models are shown in Table 8.1.

The dataset is split into 70% train set, 15% validation set and 15% test set respectivell.

8.5.2 Results

The models evaluated on the holdout test set show good performance for lead times of 1 and 2 days especially considering since the training and validation loss that is used is the MAE and not the KGE. The performance declines sharply for the LSTM model across the lead times while the decline for the TFT is more gradual as can be seen in Table 8.2.

TABLE 8.2: Performance comparison between TFT and LSTM models across different lead times on a holdout test set. Better performing model for each lead time in bold

Lead.Time	TFT.KGE	LSTM.KGE
1	0.8352	0.9696

Lead.Time	TFT.KGE	LSTM.KGE
2	0.7103	0.8821
3	0.6410	0.6716
4	0.6096	0.4943
5	0.5901	0.4302
6	0.5778	0.3312
7	0.5717	0.3185

Forecasting the peaks of the streamflows is challenging and neither model performs particularly well on this task. Especially when considering that the peaks were already drastically reduced due to the moving average smoothing of the target variable. The model routinely undershoots the observed true streamflows here shown for a lead time of 5 days for the LSTM Model and the TFT Model in figure 8.8 and in figure 8.9 respectively. This behavior can be observed for all lead times except 1 and 2 days where the performance is reasonable even in the peaks.

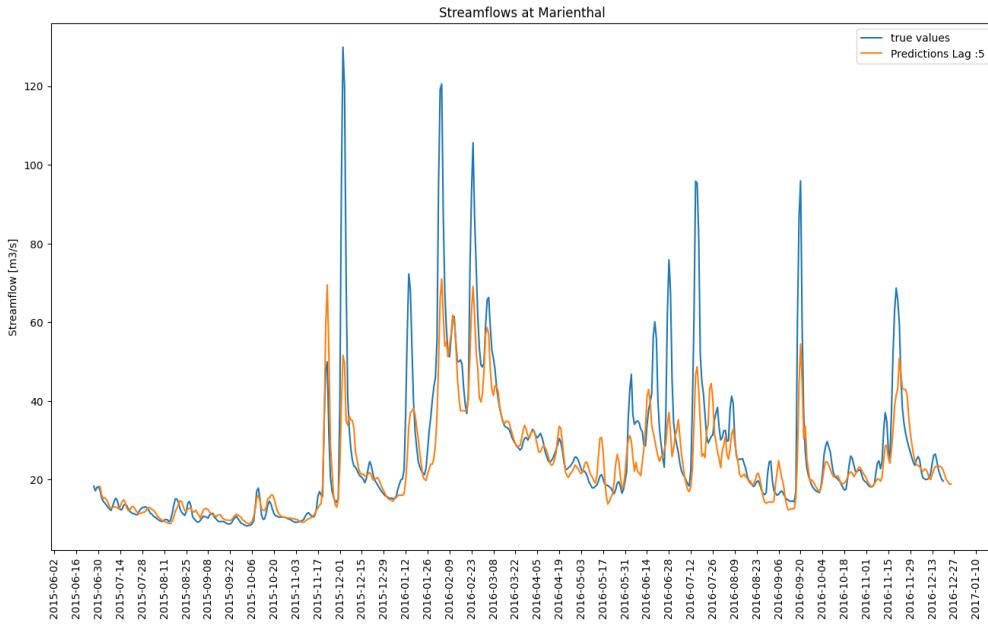


FIGURE 8.8: LSTM predicted streamflows for a lead time of 5 days (orange) compared to the observed streamflows (blue)

8.5.3 Feature Importance

For an inspection of feature dependence it is possible to make use of the inbuilt function that the pytorch-forecasting library provides. The models show very similar behavior for both used covariates. As can be seen in 8.11 the TFT model observes a slightly stronger relationship between high total precipitation and large streamflows, but the plot for both models has the expected shape, that high precipitation leads to high streamflows and vice versa.

In analyzing the relationship between t2m and streamflow, both graphs demonstrate a positive relationship. Specifically, the TFT dependence plot in Figure 8.13 shows an increase up to approximately 280 K, beyond which the effect plateaus. Similarly, the LSTM model depicted in Figure 8.12 shows a more gradual increase, with the effect leveling off at higher mean temperatures of around 295 K.

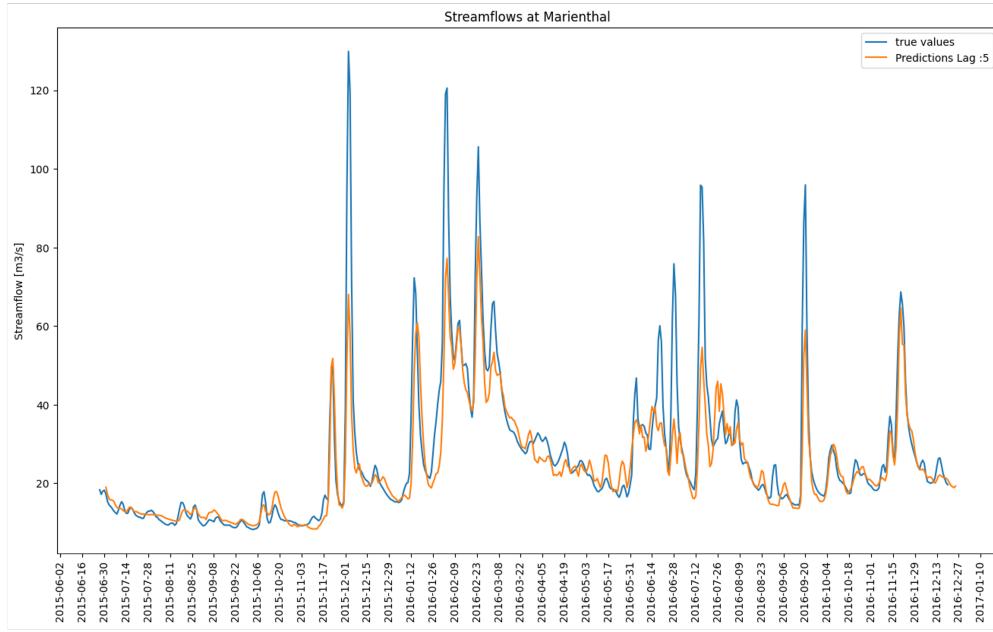


FIGURE 8.9: TFT predicted streamflows for a lead time of 5 days (orange) compared to the observed streamflows (blue)

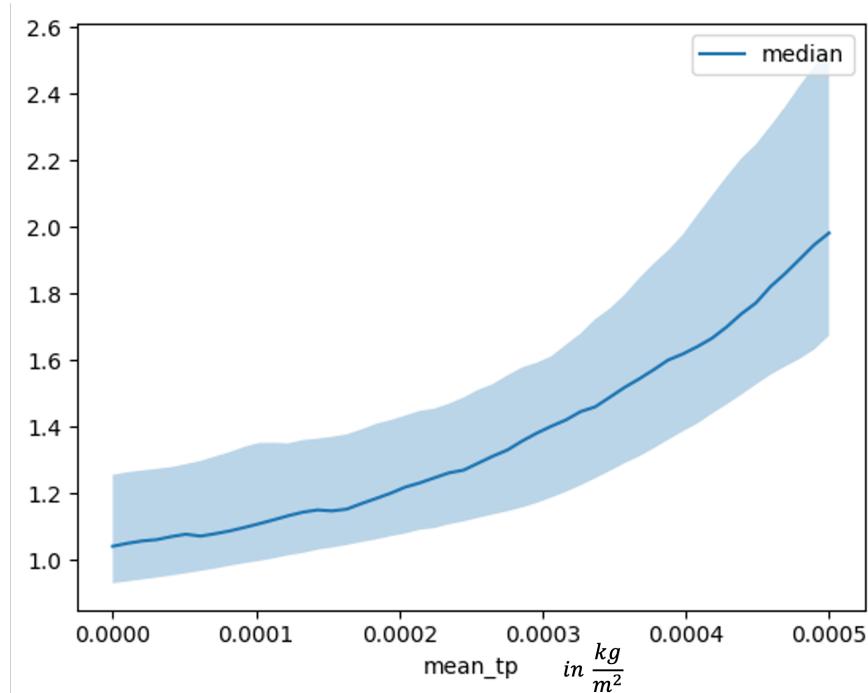


FIGURE 8.10: Feature dependence plot for mean total precipitation for a lead time 5 trained LSTM

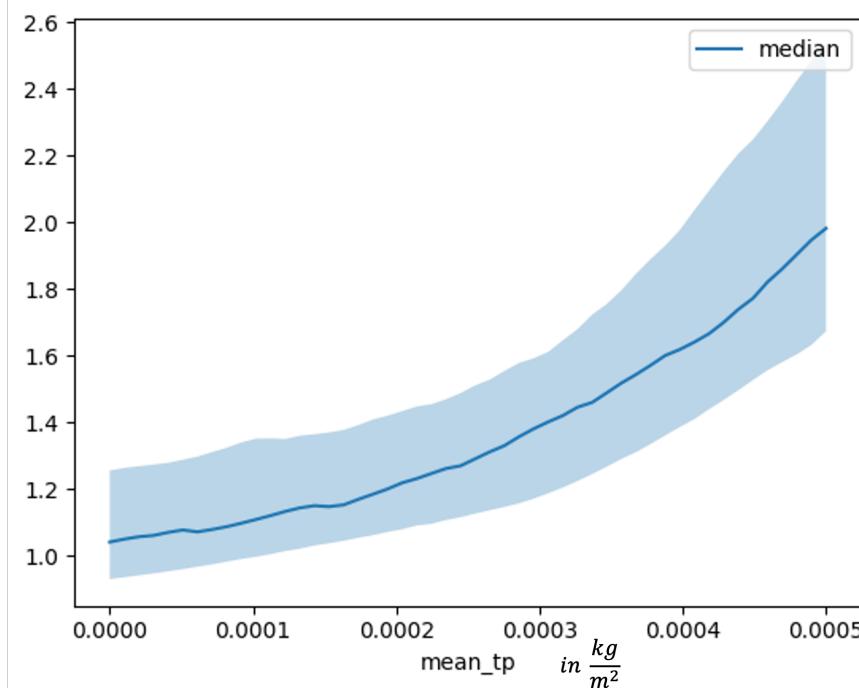


FIGURE 8.11: Feature dependence plot for mean total precipitation for a lead time 5 trained TFT

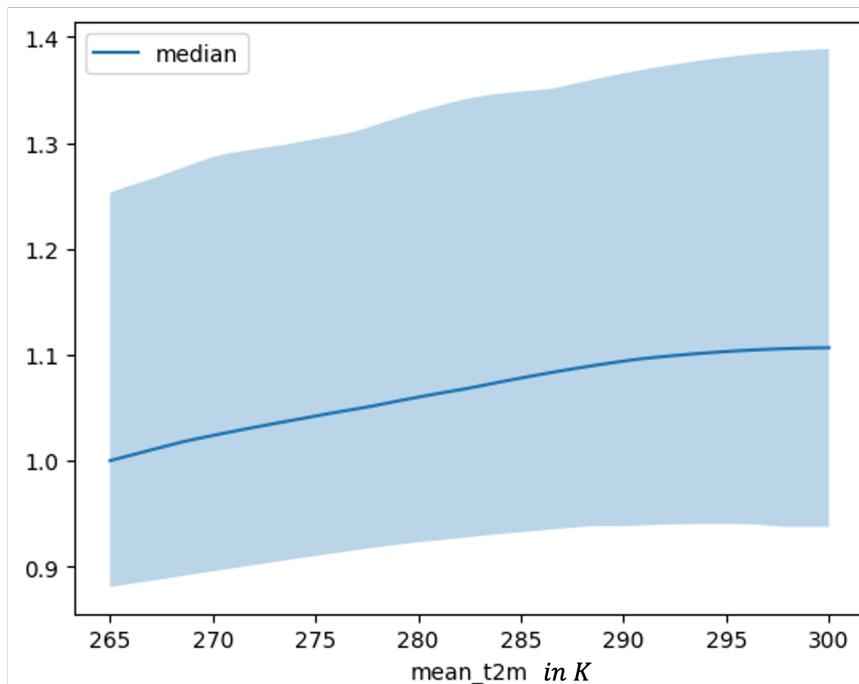


FIGURE 8.12: Feature dependence plot for mean two meter temperature for a lead time 5 trained LSTM

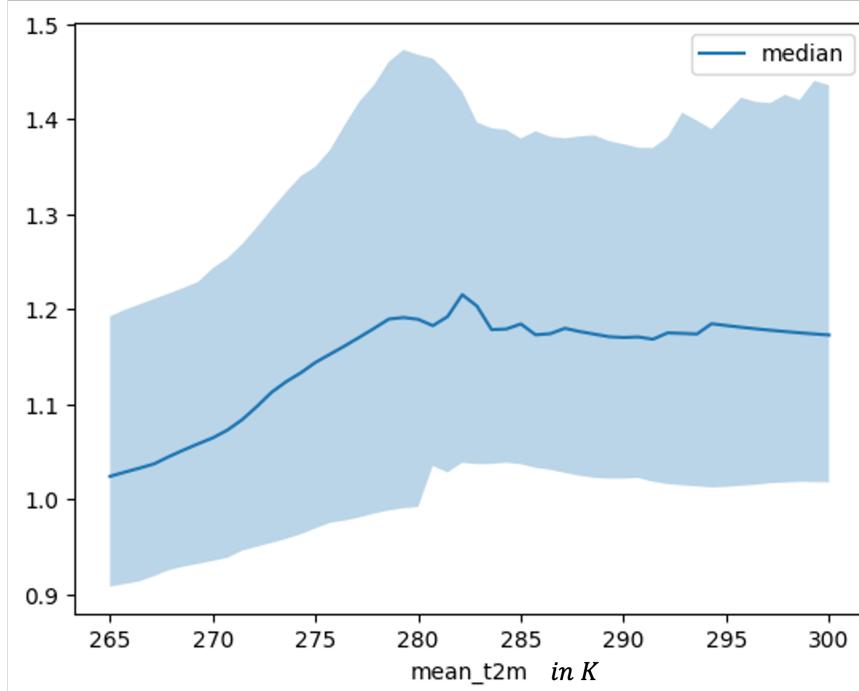


FIGURE 8.13: Feature dependence plot for mean two meter temperature for a lead time 5 trained TFT

8.6 Conclusion

This study compares the performance of LSTM (Long Short-Term Memory) and Temporal Fusion Transformer (TFT) models in forecasting streamflows of the bavarian Regen river for up to seven days ahead, using precipitation and temperature as future known covariates alongside historical streamflow data. The data used was obtained from freely available data provided by the Bavarian hydrology authority and Copernicus Climate Project.

The findings indicate that both models exhibit limitations in predicting extreme values such as floods, with KGE (Kling-Gupta Efficiency) scores significantly lower than those reported in similar studies like ?, likely due to limited data amounts and the challenges inherent in modeling a river system instead of a reservoir. The results also demonstrate a clear difference in performance trends between the two models across different lead times.

Although the KGE was not used as the loss function in training the models. The LSTM's observed KGE scores are high starting at 0.9696 for a one-day lead time, before dropping sharply to 0.3185 for a seven-day lead time. Conversely, the TFT model shows a more gradual decline, from 0.8352 at one day to 0.5717 at seven days, suggesting it maintains more consistent accuracy over longer forecast horizons.

Despite the sharper decline in performance for longer lead times, the LSTM model is notably less resource-dependent, making it a viable option for scenarios where computational resources are limited. However, attempts to forecast streamflows without future known meteorological variables were unsuccessful, underscoring the importance of these covariates in achieving accurate predictions.

While the moving average was applied here, it is not advised to use this tool for streamflow forecasting. By reducing the peaks in the target variable it artificially boosts the predictive ability of the model and forces the model to miss the peaks by an even larger margin. In a flood scenario even with a high KGE, the model would miss the exact streamflow value as it was trained on lower peaks and would underestimate the posed danger in the situation.

8.7 Outlook

Implementing a robust hyper-parameter tuning routine is essential to optimize model performance. This process will require additional computational resources due to the complexity and extensive search space. Given the high dependency of hyper-parameters on lag, it might be necessary to tune hyper-parameters for each lead time separately.

To make the models more sensitive to extreme events such as floods, specialized loss functions could be employed or training a consecutive model that specifically forecasts the peaks taking the first models predictions as inputs.

The ERA5 dataset used for obtaining meterological data in this study only provides a very coarse representation of the underlying variables. The use of down scaling techniques to obtain a finer grid than the one used in this study might be able to boost the accuracy of the model.

Another next step could be to test the models ability to generalize by training on multiple datasets from different rivers. Including static variables such as river basin properties and land use information can help in creating a more comprehensive model that can adapt to various river systems.



9

The Lancet Report 2023

Author: Michael Strobl

Supervisor: Prof. Dr. Helmut Küchenhoff

Suggested degree: Bachelor

9.1 Introduction

The 2023 report of the Lancet Countdown on health and climate change (?) aims to monitor the evolving impact of climate change on health and the emerging health opportunities of climate action. It is in its eighth iteration and was created by 114 scientists and health practitioners from 52 research institutions and UN agencies from all continents (but Antarctica). While this chapter is based on the 2023 report, there are also regional reports available, such as the 2024 report for Europe (?).

The current report focuses on 47 indicators, tracking changes based on observed data as well as projections, in the following fields:

1. Health hazards, exposures, and impacts
2. Adaptation, planning, and resilience for health
3. Mitigation actions and health co-benefits
4. Economics and finance
5. Public and political engagement with health and climate change

The remainder of this chapter provides some background information on climate models, detailed information on three indicators, and a discussion.

9.2 Background

Shared Socioeconomic pathways (SSPs) (?) were established by the climate research community in order to make the analysis of future climate impacts, vulnerabilities, adaptation, and mitigation possible. They describe alternative socioeconomic developments regarding their energy, land use, and emissions implications, e.g. more or less renewable energy vs. fossil fuels. The following SSPs are of interest for the rest of this chapter:

1. SSP1: Sustainability (low challenges to mitigation and adaptation), corresponding to a 2°C emission scenario
2. SSP3: Regional Rivalry (high challenges to mitigation and adaptation), corresponding to a 3.7°C emission scenario

These pathways are considered by climate models, e.g. CMIP6 (?), in order to describe multiple scenarios for climate projections. Figure 9.1 shows CMIP6 simulations for multiple models and all SSPs. In this chapter, we consider SSP1, representing a sustainable path, and SSP3, representing a non-sustainable path, which humanity is following right now.

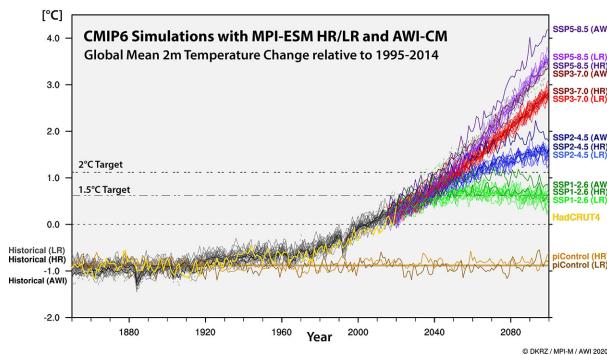


FIGURE 9.1: CMIP6 climate simulations for SSPs with projections until 2100 (https://www.dkrz.de/en/communication/climate-simulations/cmip6-en/cmip6-activities-at-dkrz-overview?set_language=en; accessed on July 10, 2024)

9.3 Selected Indicators

We selected three indicators from the first field (Health hazards, exposures, and impacts) with detailed descriptions and outcomes.

9.3.1 Indicator 1.1.2: Exposure of vulnerable populations to heatwaves

Heatwaves have a severe or even life-threatening impact on human health, e.g. high temperatures can cause heat stroke, heat exhaustion, heat syncope, and heat cramps (see, for example, ? or ?). The following variables (among others) can influence mortality: increased risk for older adults (>65 years of age), low-income countries due to a reduced number of health workers, or pre-existing cardiovascular and chronic respiratory conditions.

This indicator tracks the number of heatwave days and the exposure of vulnerable populations (older adults and infants <1 years of age) to heatwaves. A heatwave day is a period of 2 or more days where both the minimum and maximum temperatures are above the 95th percentile of temperatures in 1986-2005.

The following datasets were used for this indicator:

- ERA5 monthly averaged data on single levels (?): World-wide weather data on a lat-lon grid of 0.25 degrees from 1940 onwards.
- ISIMIP3b Bias Adjustment dataset (?): A bias adjusted and downscaled version of the output of CMIP6 climate models with projections up to the year 2100.
- Hybrid gridded demographic data for the world (?): Demographics data from 1950-2020 with 5-year population bands and a 0.5 degree grid.
- 2020 Revision of World Population Prospects (<https://population.un.org/wpp/>): This dataset contains population estimates as well as projections for 237 countries or areas between 1950 and 2100.
- A global downscaled age structure dataset for assessing human impacts of climate change (?): This

dataset contains historic population estimates starting in 1970 and projections considering SSP1 and SSP3 in a 0.5 degree grid.

Headline Finding:

“In 2013-2022, infants and people over 65 experienced, on average, 108% more days of heatwave per year than in 1986-2005.”

Figures 9.2 and 9.3 show the increase in the number of heatwave days, comparing the time periods 1986-2005 and 2013-2022, for infants and older adults, respectively. This increase (or decrease for Portugal) per country can be calculated, for example, through ERA5 monthly averaged weather data and the hybrid gridded demographics dataset.

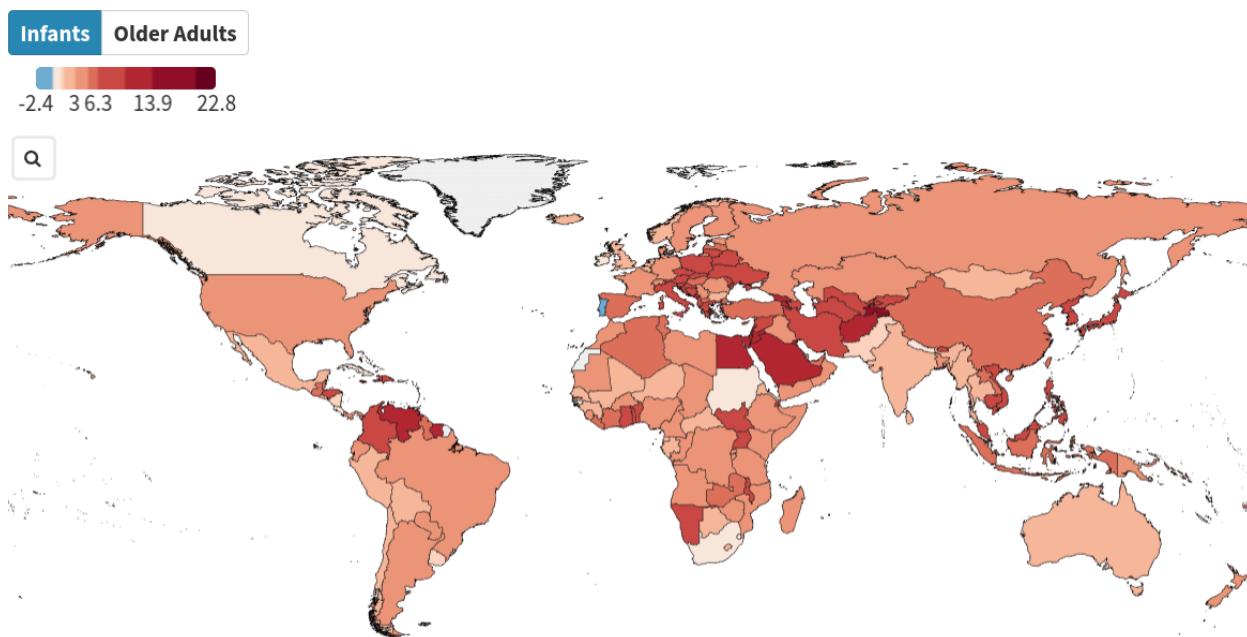


FIGURE 9.2: Change in number of heatwave days per country for infants, comparing the time periods 1986-2005 and 2013-2022.

The change in the number of heatwave days ranges from approximately -1 in Portugal to 16 in Tajikistan, with only minimal differences between infants and older adults.

Similarly, the ISIMP3b Bias Adjustment dataset combined with the global downscaled age structure dataset can be used to project the change in heatwave days up until the year 2100. Figures 9.4 and 9.5 show the projected change in heatwave days for older adults at mid-century (2041-2060) with baseline period 1995-2014. For example, the increase for Venezuela ranges from approximately 88 (SSP1) to 153 (SSP3) heatwave days per year.

9.3.2 Indicator 1.1.5: Heat-related mortality

This indicator is tracking days with temperatures exceeding safe levels for humans and heat-related mortality. Only older adults are considered.

Similar datasets are used here as for the previous indicator. However, the difficulty arises due to the fact that safe temperatures may vary across regions (see ?).

Headline Finding:

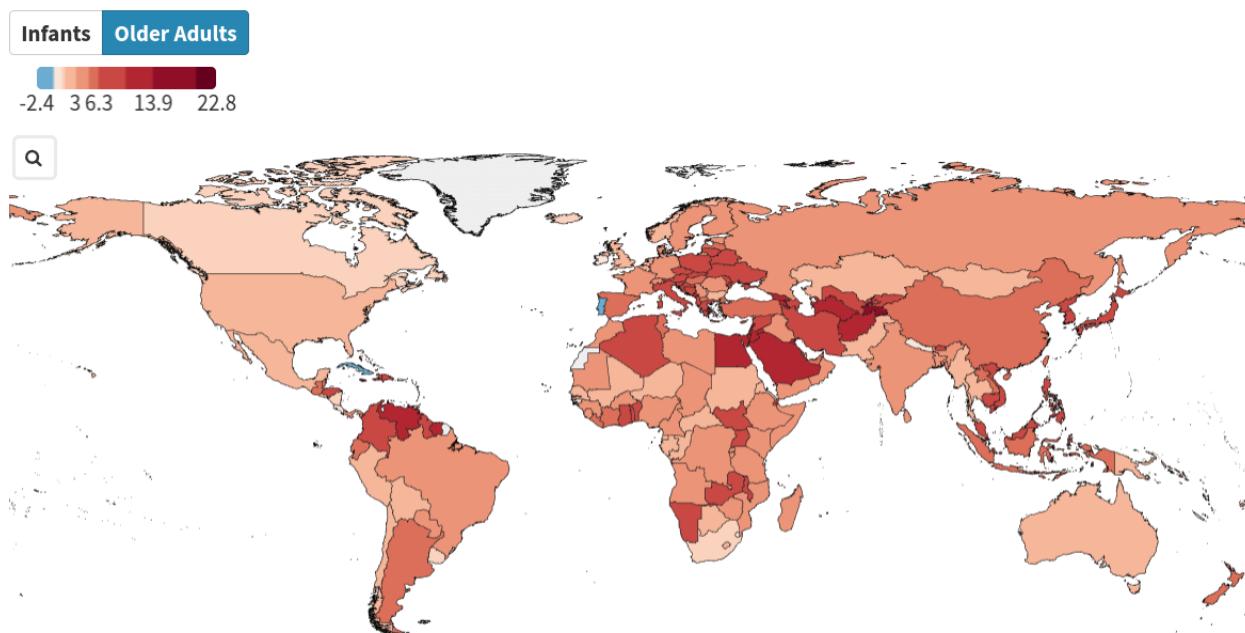


FIGURE 9.3: Change in number of heatwave days per country for older adults, comparing the time periods 1986-2005 and 2013-2022.

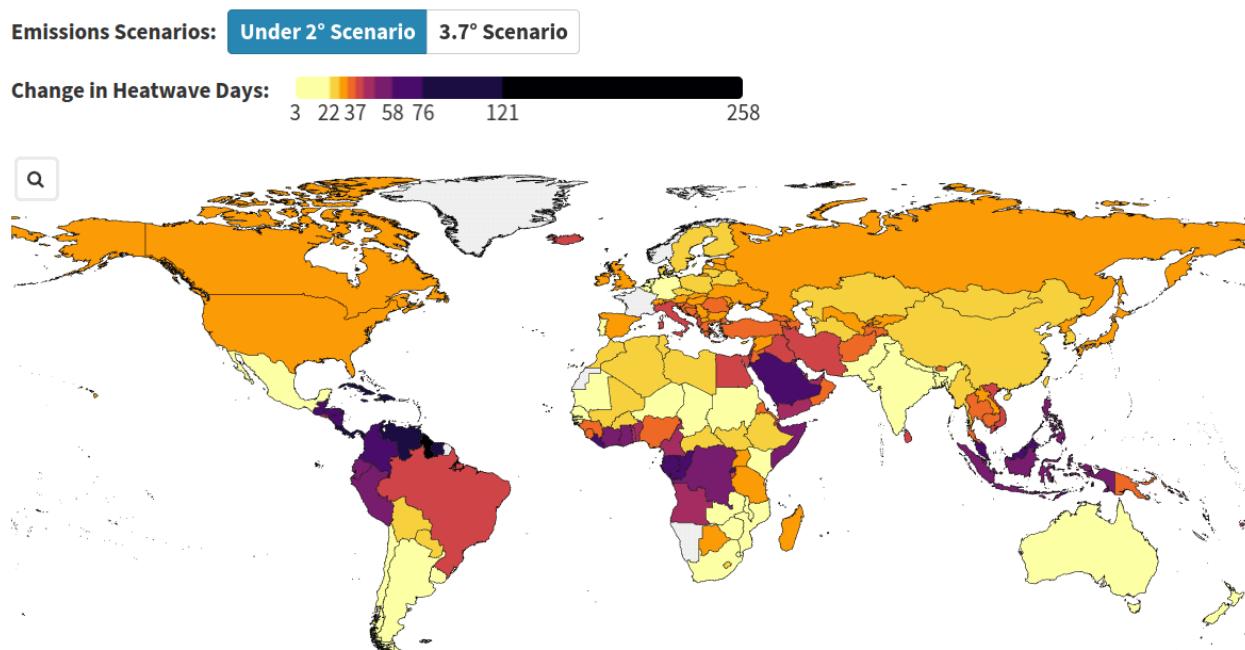


FIGURE 9.4: Projections of the change in number of heatwave days for older adults per country for SSP1 (under 2 degree scenario) for mid-century (2041-2060) with baseline period 1995-2014.

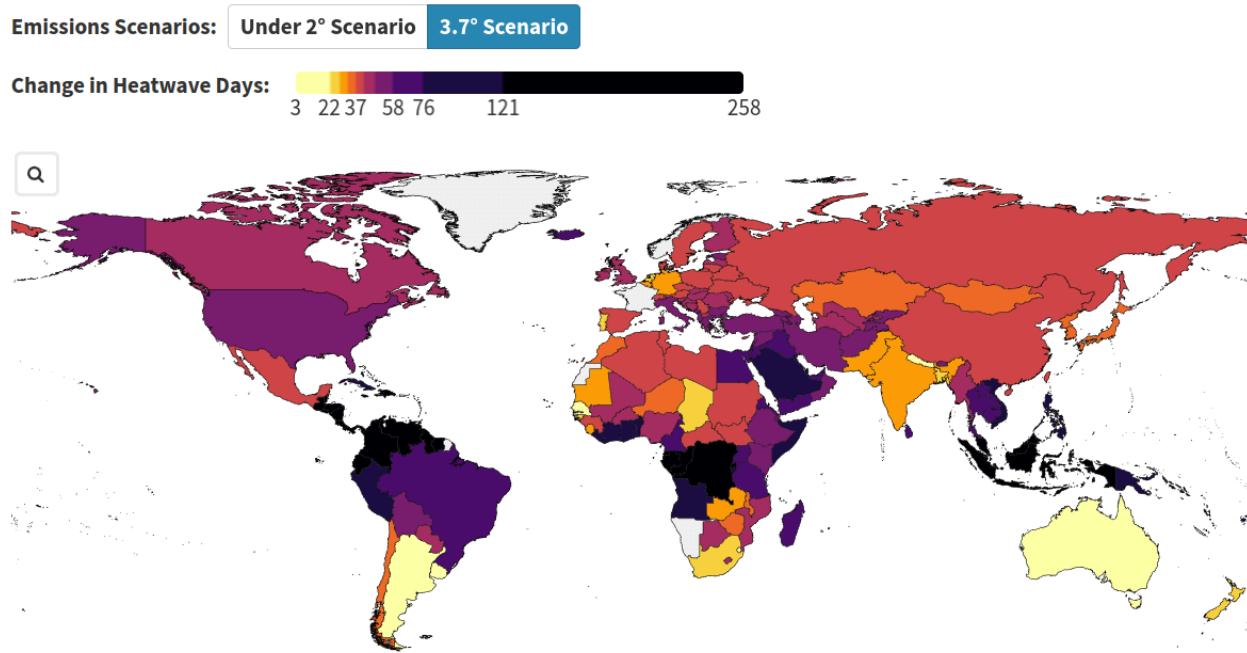


FIGURE 9.5: Projections of the change in number of heatwave days for older adults per country for SSP3 (3.7 degree scenario) for mid-century (2041-2060) with baseline period 1995-2014.

"In 2018-2022, people experienced on average 86 days of health-threatening high temperatures annually. 60% of such temperatures were made more than twice as likely to occur by human-caused climate change."

Therefore, [] created the notion of an Optimum Temperature (OT), at which the mortality is the lowest. Figure 9.6 shows an example for the Tokyo Prefecture and the period 1972-2008. A smoothing spline with six degrees of freedom was used to model the Relative Risk (RR) for mortality and observed daily maximum temperature data. OT represents the daily maximum temperature where RR is the lowest. RR was used here in order to account for different populations of regions and an RR=1.0 is the reference mortality at around OT, i.e. the average number of deaths observed when the daily maximum temperature was within the range of the 75th to the 85th percentile for each year.

The shaded area to the right of OT represents temperatures with increased mortality, i.e. daily maximum temperatures above OT are considered as risky for older adults. The average OT percentile of 47 Japanese prefectures was 83.6, which was used to decide which daily maximum temperatures exceeded safe levels for this indicator. In addition, a distributed lag nonlinear model [] was used to model the increase in mortality (represented by RR) depending on temperature. This model is considered as a response function, which can be used to calculate the RR and therefore the absolute increase of mortality depending on the observed or projected temperatures.

Figure 9.7 shows the increase in the number of days with unsafe temperatures (daily maximum temperatures exceeding OT). This includes the increasing number of days that were made at least twice as likely due to human induced climate change (it is not described how the red line was created).

In addition, Figures 9.8 and 9.9 show the absolute projected increase in heat-related mortality per country for older adults at mid-century (2041-2060) compared to 1995-2014 as baseline, for SSP1 and SSP3, respectively. In both cases there is an increase, although it seems that the two scenarios, SSP1 and SSP3, have been flipped accidentally. In addition, only the absolute increase is shown, i.e. it is obvious that highly populated countries, such as India and China, have a very high increase compared to smaller countries.

The response function is kept constant for these projections, i.e. adaptation is ignored for this indicator.

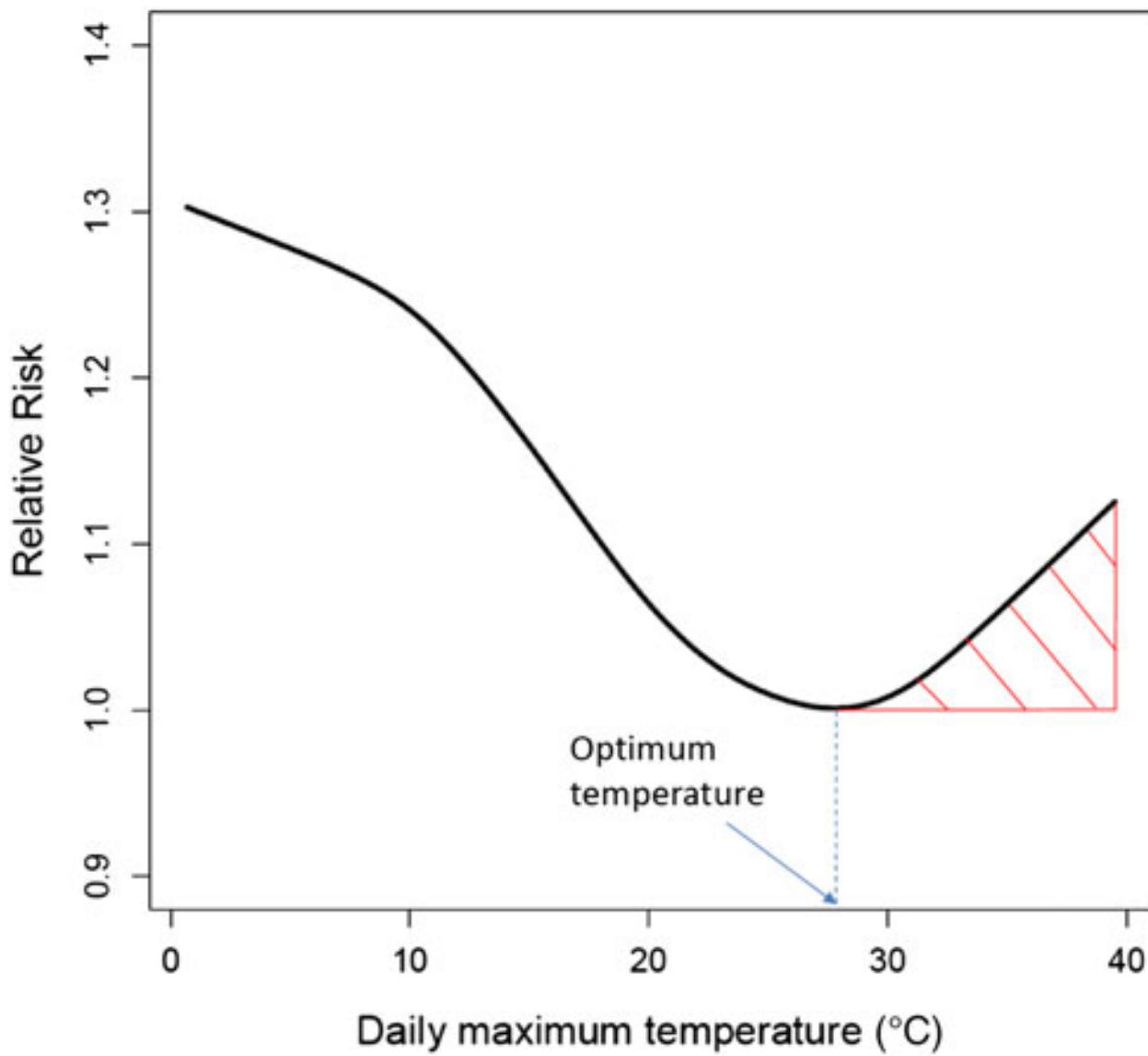


FIGURE 9.6: A smoothing spline with six degrees of freedom was used for temperature data from the Tokyo Prefecture and the period 1972–2008, showing the daily maximum temperatures and the Relative Risk (RR) for mortality.

Also, absolute numbers of mortality are used here, i.e. it was not distinguished between heat-related deaths and other causes.

9.3.3 Indicator 1.4: Food insecurity and undernutrition

In 2022, 725 million people faced hunger and in 2021, 3.1 billion people were unable to afford a healthy diet (?). For example, climate change is undermining crop yields, affecting labour capacity, and threatening food security of marine resources.

In addition to the previously mentioned datasets, specifically ERA5 and ISIMP3b, data from the Food and Agriculture Organization Food Insecurity Experience Scale (FIES) (?) was used. This dataset contains survey data from 153 countries and territories during the years 2014, 2015, and 2016. Among others, people

Days of health-threatening temperatures ■ Days made at least twice as likely due to human-induced climate change

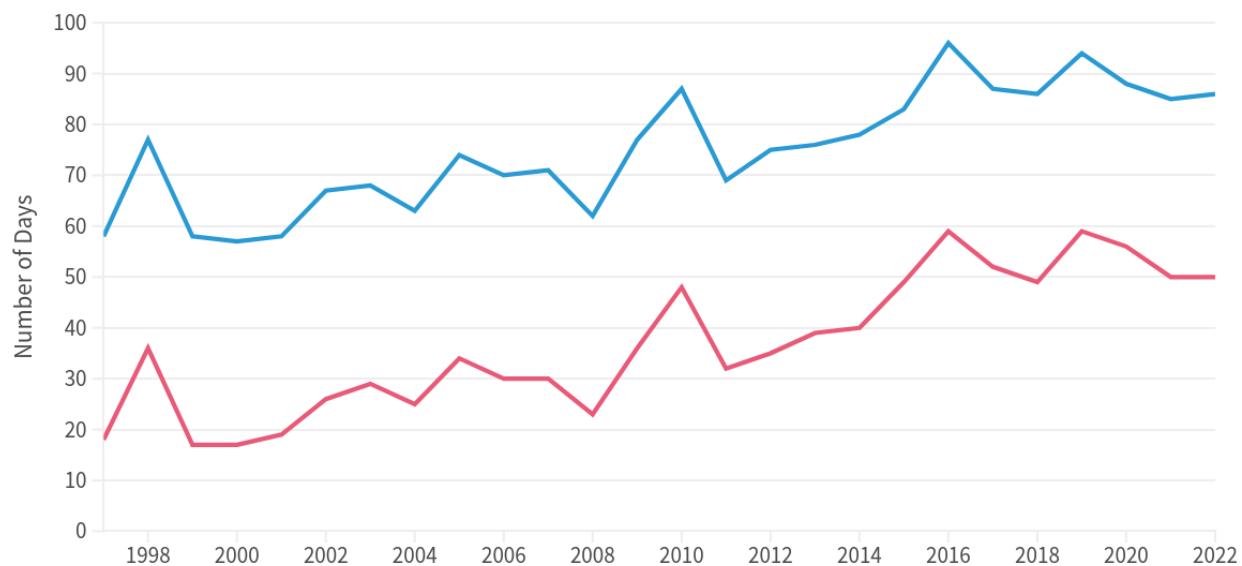


FIGURE 9.7: Average number of days with unsafe temperatures for older adults from 1997 to 2022, including days that are twice as likely due to climate change.

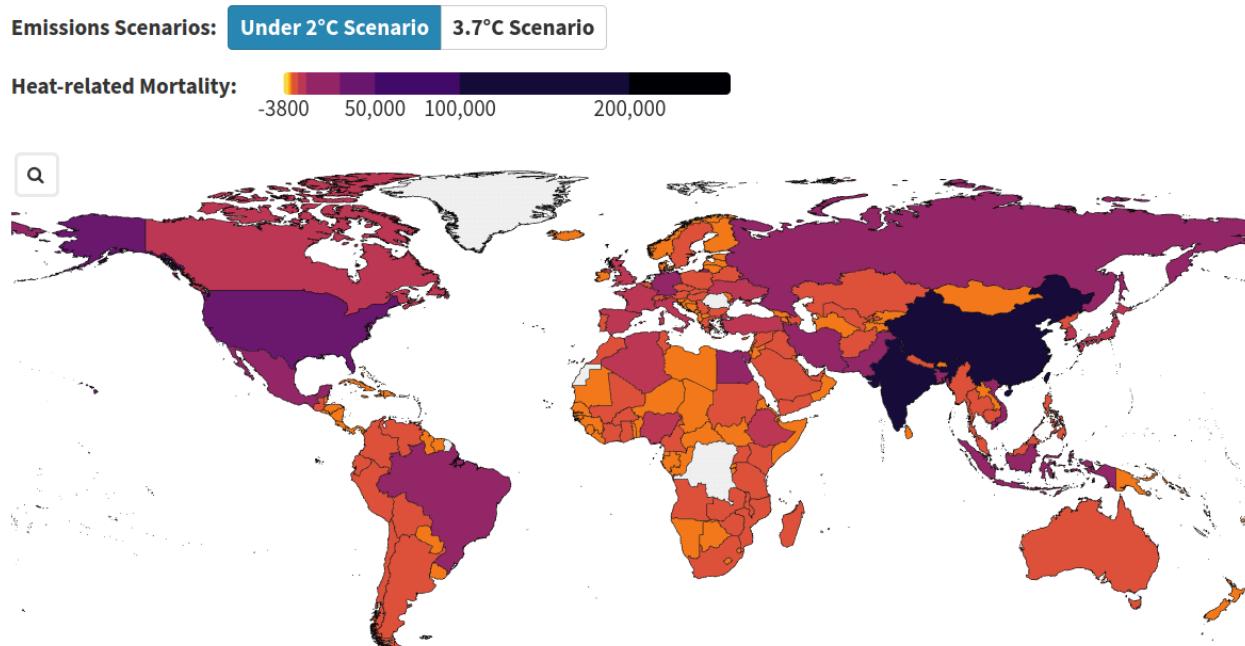


FIGURE 9.8: Projections of the absolute change in heat-related mortality for older adults per country for SSP1 (under 2 degree scenario) for mid-century (2041-2060) with baseline period 1995-2014.

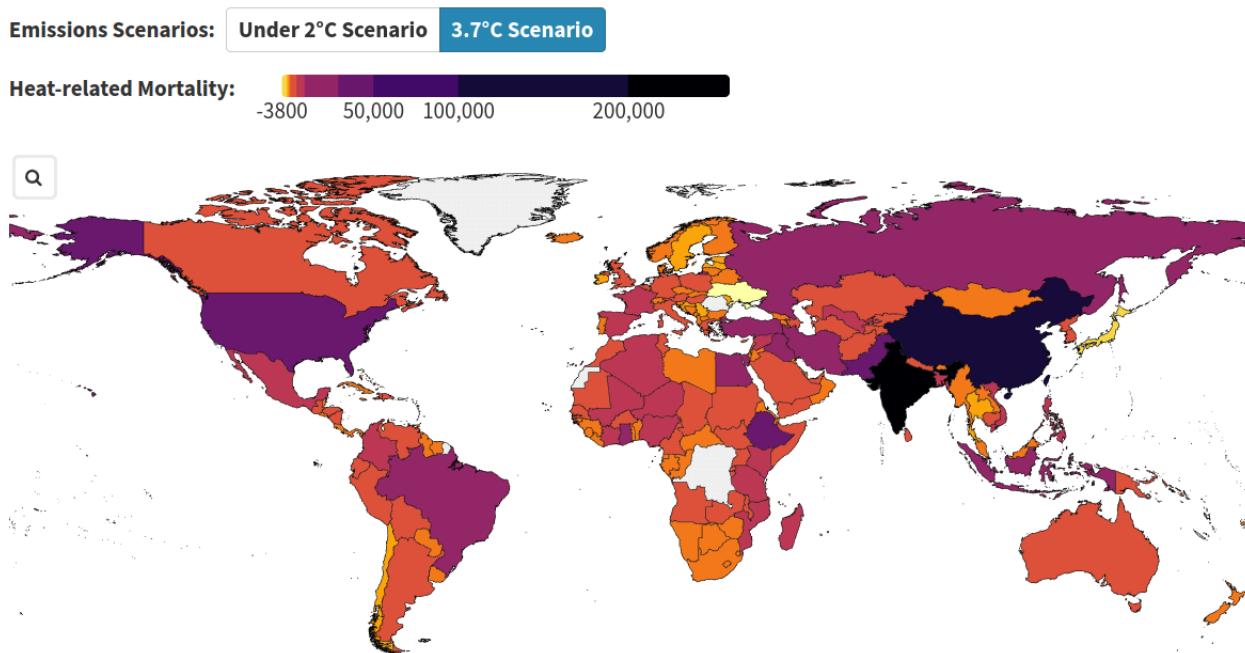


FIGURE 9.9: Projections of the absolute change in heat-related mortality for older adults per country for SSP3 (3.7 degree scenario) for mid-century (2041–2060) with baseline period 1995–2014.

were asked for whether they experienced moderate or severe food insecurity within the past 12 months. This was matched with weather data in these people's specific regions in order to find out whether they experienced heatwaves (as previously defined) or drought months (12-monthly SPEI; only during the growth season of specific crops). In order to make predictions for all other years, including the future, a time-varying panel regression model was used (without further detail). The goal was to predict the percentage of people reporting food insecurity based on heatwaves and drought months experienced within a specific year.

Headline Finding:

"The higher frequency of heatwave days and drought months in 2021 compared to 1981–2010, is associated with 127 million more people experiencing moderate or severe food insecurity."

Figure 9.10 shows world-wide predictions of the change in percentage points (using the previously mentioned model) for heatwaves and drought months for the period 2014–2021. As temperatures, number of heatwave days as well as number of drought months increase, more people are expected to experience severe or moderate food insecurity.

Figures 9.11 and 9.12 show projections (SSP1 and SSP3, respectively) for the percentage point changes for both kinds of extreme weather combined for mid-century (2014–2060), compared to the baseline 1995–2014. For example for Somalia, this change ranges from 4.02 (SSP1) to 10.32 (SSP3).

It is unclear why for many countries there is missing data.

9.4 Discussion

The 2023 report of the Lancet Countdown on health and climate change provides a comprehensive overview of the current state of the impact of climate change on health and opportunities of climate action. 47 indicators are used to keep track of where humanity is at right now.

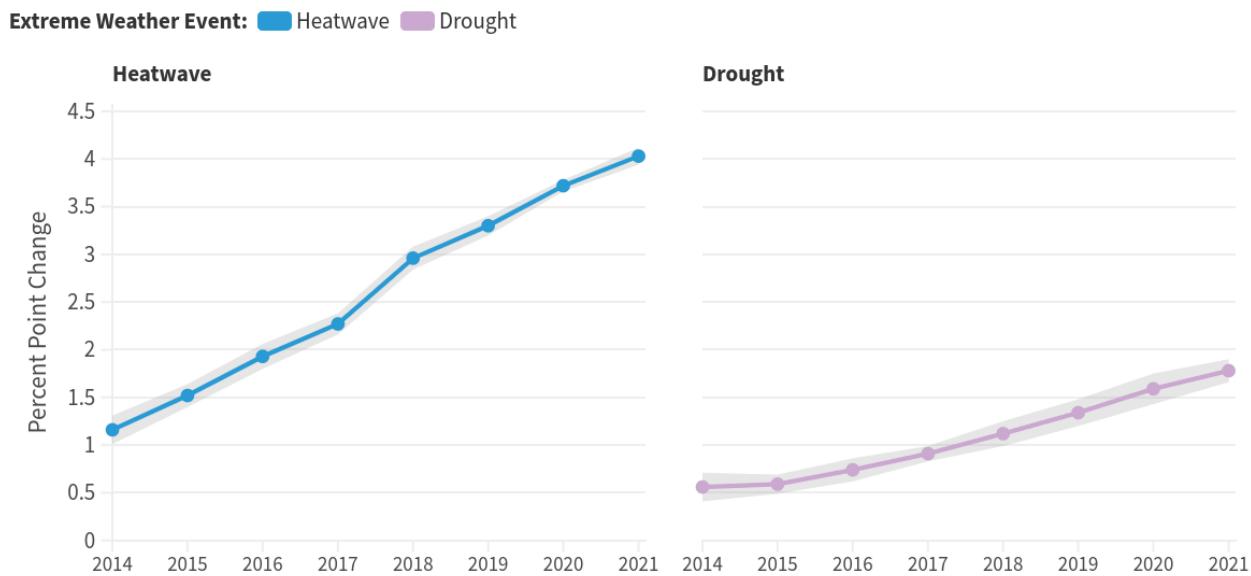


FIGURE 9.10: Impact of extreme weather (heatwaves and droughts) on food insecurity from 2014-2021, based on survey data.

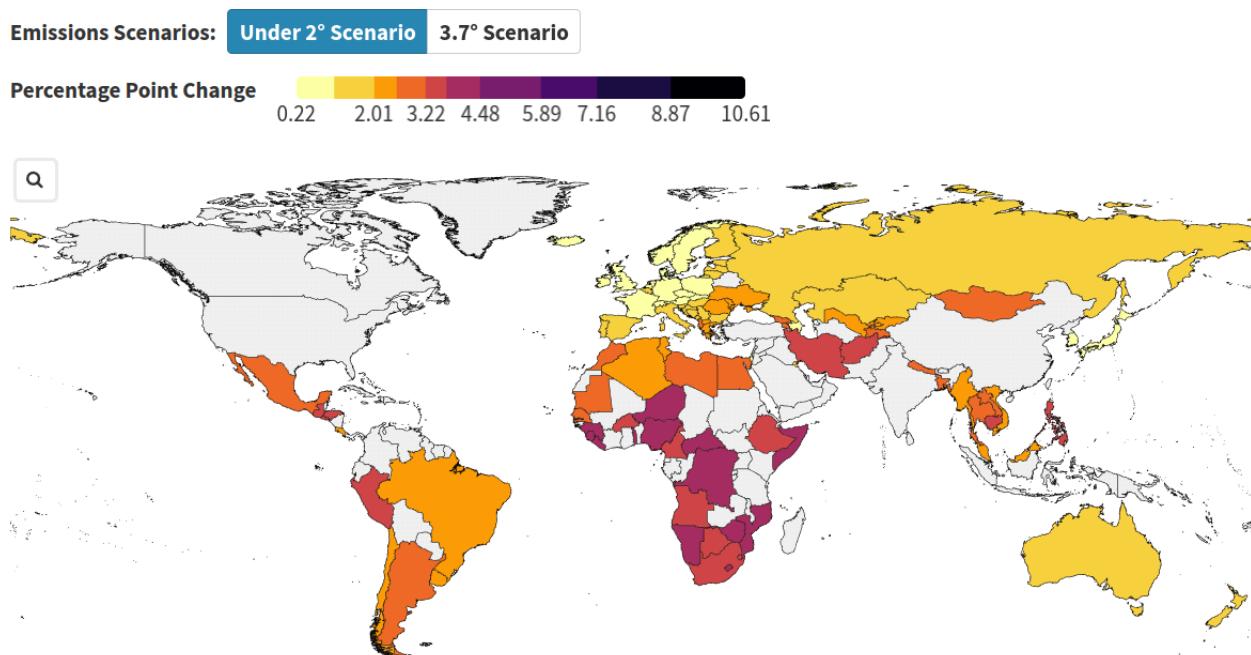


FIGURE 9.11: Projections of the impact of extreme weather on food insecurity per country for SSP1 (under 2 degree scenario) for mid-century (2041-2060) with baseline period 1995-2014.

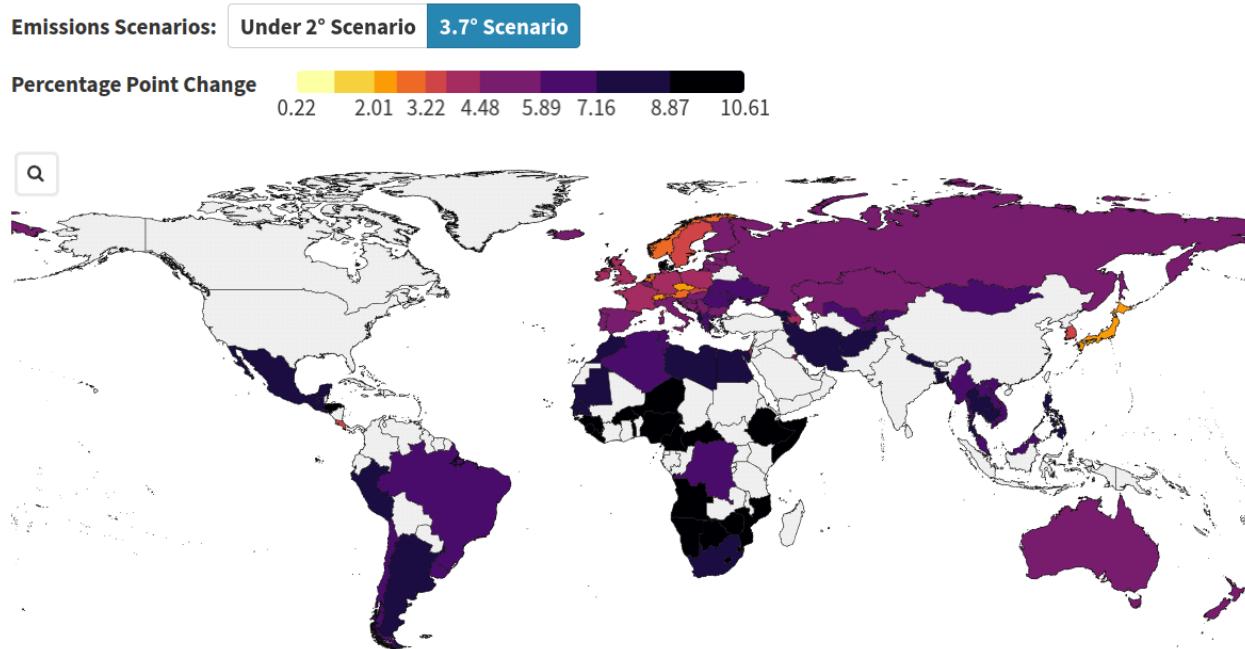


FIGURE 9.12: Projections of the impact of extreme weather on food insecurity per country for SSP3 (3.7 degree scenario) for mid-century (2041-2060) with baseline period 1995-2014.

We provided detailed descriptions of three main indicators and the modelling techniques used, if mentioned. However, there are a few issues discovered:

- These modelling techniques are typically not described in much detail. Sometimes, as in the case for ?, some more information is provided in cited articles, but still on a relatively high level.
- The baselines for comparisons or projections vary. It is not clear, nor described, why different baselines are used. For projections, it is obvious that the baseline for CMIP6 projections (1995-2014) was used, without information on why not the one that was previously used for the indicator at hand.
- In general, the main findings are reported, without going much into detail on how the authors came up with these. In some cases, e.g. indicator 1.1.5, cited articles provide more information, but in other cases, e.g. indicator 1.4, it seems to be mainly the authors own work, without providing details.

10

Epidemiologic studies on the heat effects

Author: Oussama Mesbah

Supervisor: Helmut Kuechenhoff

Degree: Bachelor

10.1 Introduction

This chapter covers epidemiologic studies on the effects of heat on health. The aim is to discuss the excess mortality attributed to heat and cold based on recent studies by Masselot et al. (2023) (?) and the joint effects of air pollution and heat on mortality as examined by Stafoggia et al. (2023) (?).

10.2 Excess Mortality Attributed to Heat and Cold

10.2.1 Background for Study

Masselot et al. (2023) conducted a comprehensive study on the mortality impacts of non-optimal ambient temperatures. Both heat and cold are established risk factors, and the study aims to provide a robust assessment of these impacts. Previous research has predominantly focused on country-specific data or omitted significant factors.

10.2.2 Added Value

This study offers several key advancements:

First, it has unprecedented city coverage, examining 854 urban areas across Europe. This extensive coverage offers a more comprehensive assessment compared to prior studies that focused on specific countries or regions. Furthermore, the study balances its focus on both heat and cold impacts, providing a complete picture of temperature-related health risks. It explicitly models age-specific risks and provides age-standardized mortality rates, which allow for fair comparisons across different locations with varying demographic compositions. This is particularly important as it highlights the heightened vulnerability of the elderly population.

Compared to Zhao et al. (2021) (?) , Masselot et al. (2023) focuses on a more detailed regional scope by examining 854 cities in Europe, while Zhao's study covered 750 locations globally. Both studies cover the same period from 2000 to 2019, allowing for direct comparison. Masselot's study utilizes higher resolution ERA5-Land reanalysis data, which may provide a more accurate exposure assessment, especially for smaller cities. Additionally, the inclusion of age-specific mortality data in Masselot's study allows for a more nuanced

analysis of age-related vulnerability to temperature extremes. Here is a table summarizing the key differences between the two studies:

Feature	Zhao et al. (2021)	Masselot et al. (2023)	Added Benefit of Masselot et al. (2023)
Geographical Scope	Global (750 locations in 43 countries)	Europe (854 cities in 30 countries)	More detailed focus on a specific region.
Temporal Scope	2000-2019	2000-2019	Allows direct comparison of results.
Exposure Data	Gridded temperature data ($0.5^\circ \times 0.5^\circ$)	ERA5-Land reanalysis data (9 km resolution)	Higher resolution data provides more accurate exposure assessment.
Outcome Data	Daily all-cause or non-external mortality counts	Daily all-cause or non-accidental mortality counts, with age-specific data for a subset of cities	Inclusion of age-specific mortality data allows for nuanced analysis.
Modeling Framework	Three-stage: time-series regression, meta-regression, prediction	Three-stage: quasi-Poisson regression with DLNM, meta-regression with PLS, kriging interpolation	Incorporation of PLS for dimensionality reduction and kriging for spatial interpolation.
Vulnerability Assessment	Limited to 5 meta-predictors	Includes 22 city-specific characteristics	Comprehensive assessment of vulnerability factors.
Age-Specific Analysis	Not explicitly mentioned	Provides age-specific relative risks and age-standardized mortality rates	More targeted understanding of vulnerable age groups.

10.2.3 Limitations

The study also has limitations. The extrapolation of exposure-response functions (ERFs) across European cities may introduce bias, especially in less-covered areas like Eastern Europe. Additionally, reliable ERFs could not be estimated for people younger than 20 years due to low death counts, limiting the analysis for younger populations.

10.2.4 Findings

The study by Masselot et al. (2023) provides a detailed analysis of the impact of non-optimal temperatures on mortality across Europe. The study found an annual excess of 203,620 deaths due to cold and 20,173 deaths due to heat. Age-standardized excess death rates were calculated at 129 per 100,000 person-years for cold and 13 per 100,000 person-years for heat.

10.2.4.1 Regional Variations

The results indicated significant variations across different regions of Europe. The highest mortality effects were observed in eastern European cities, where the infrastructure and healthcare systems might be less equipped to handle extreme temperatures compared to their western counterparts.

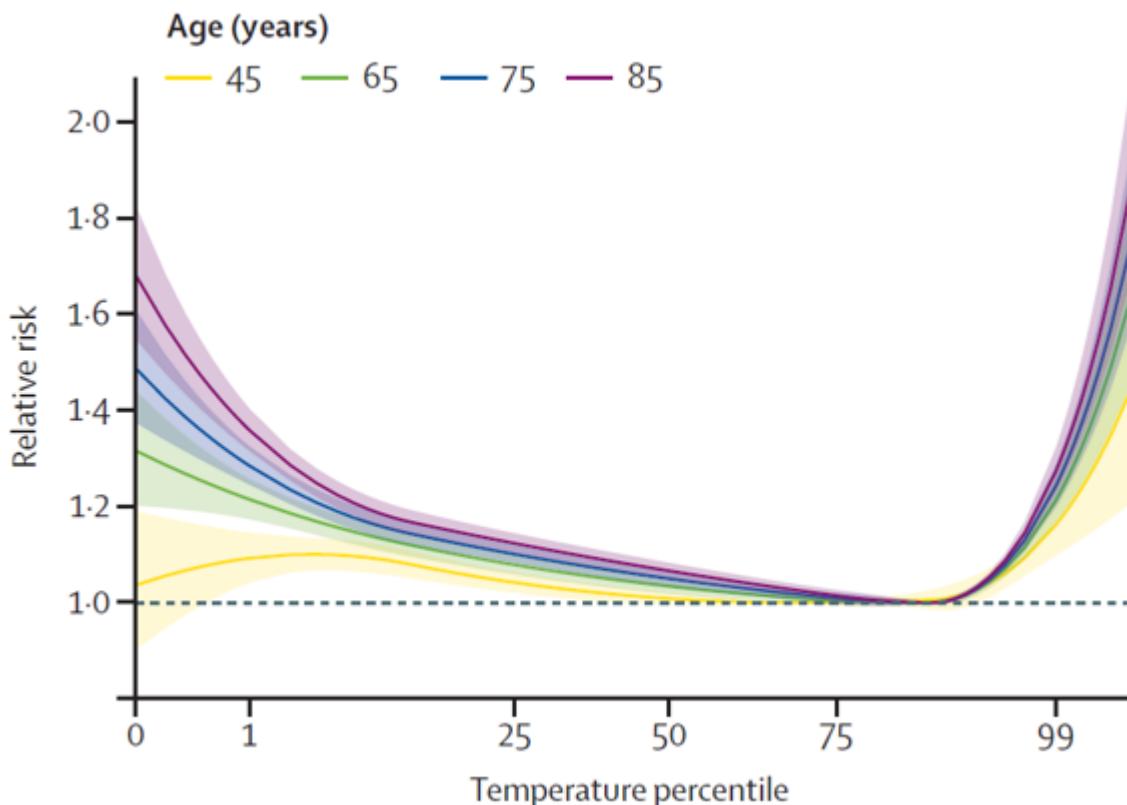


FIGURE 10.1: Relative Risk at Different Temperatures and Ages

10.2.4.2 Age-Related Vulnerability

The study highlighted that vulnerability increases with age, particularly for individuals aged 85 and above. The pooled overall cumulative exposure-response relationship was predicted at several ages: 20-44 years, 45-64 years, 65-74 years, 75-84 years, and 85+ years. This data provides crucial insights into how different age groups respond to temperature extremes.

Several key terms are defined in the study to better understand the results:

- **Exposure-Response Function (ERF):** This describes the relationship between exposure levels (temperature) and mortality risk.
- **Minimum Mortality Temperature (MMT):** The specific temperature at which mortality risk is lowest.
- **Minimum Mortality Percentile (MMP):** The temperature range where the population is generally healthiest.
- **Relative Risk (RR):** The ratio of mortality risk at a specific temperature compared to the MMT.

The study also analyzed the relative risks for cold and heat in capital cities across five age groups. The findings showed that excess mortality is generally lower in Western Europe compared to Northern and Southern Europe.

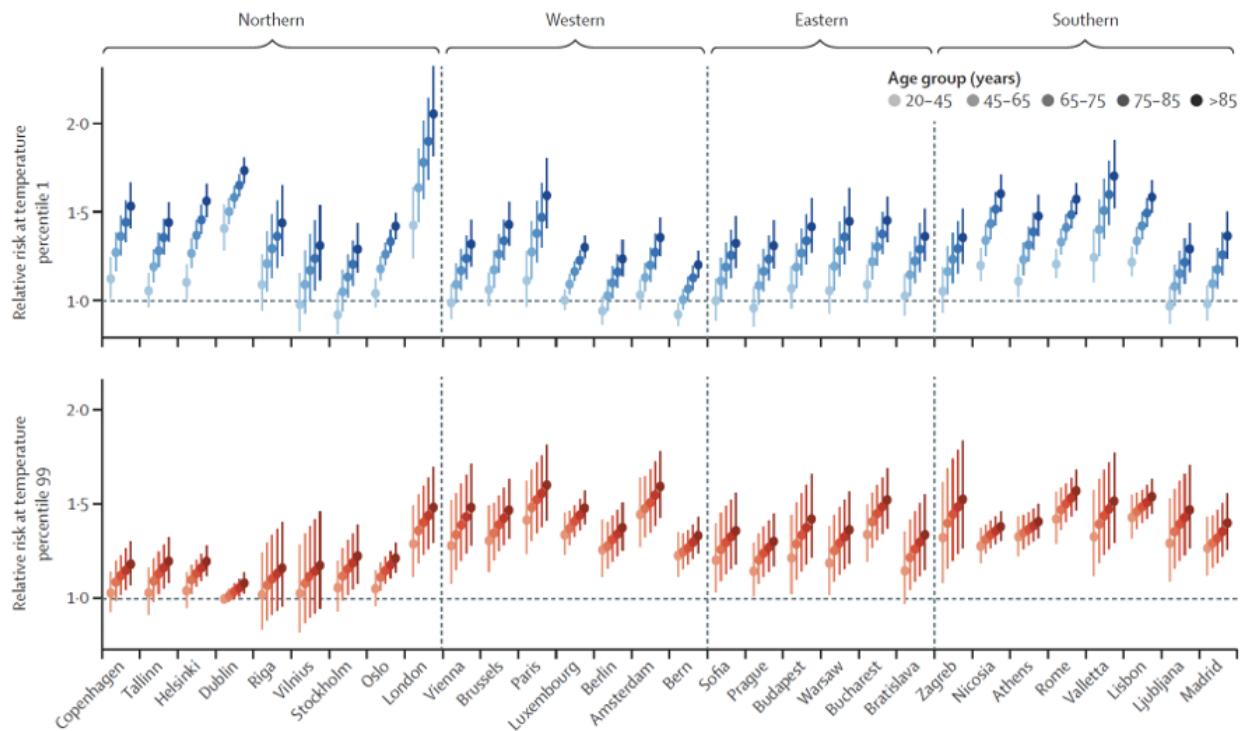
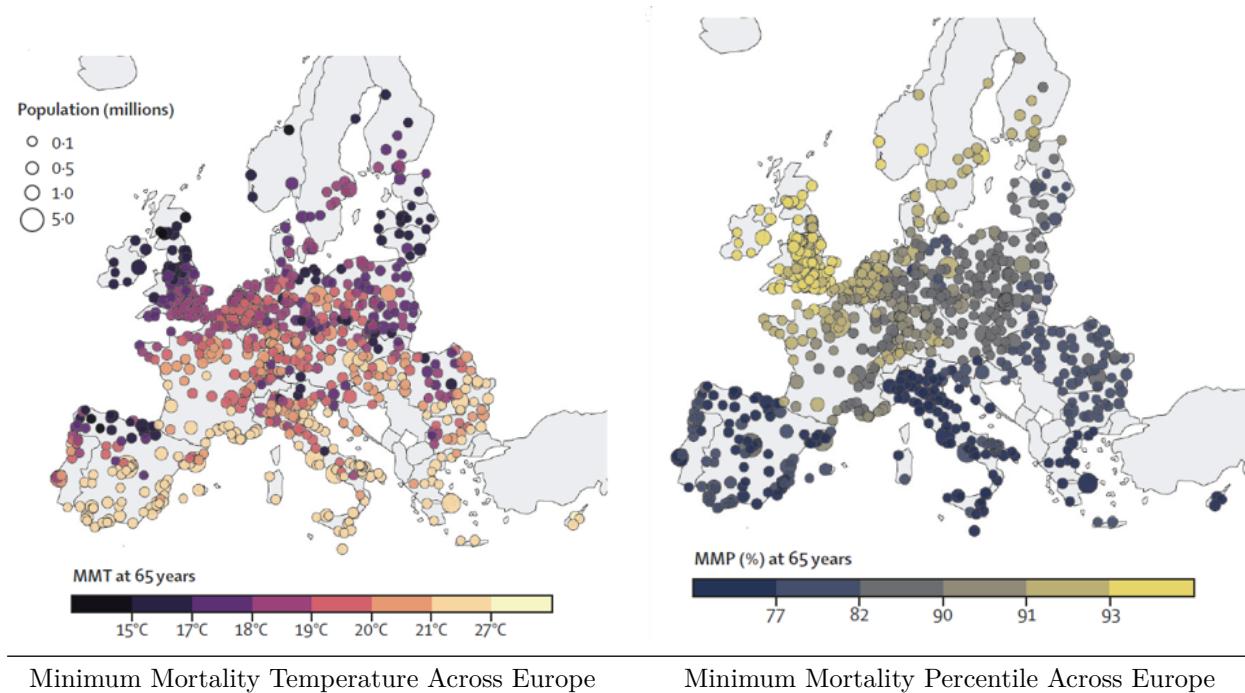


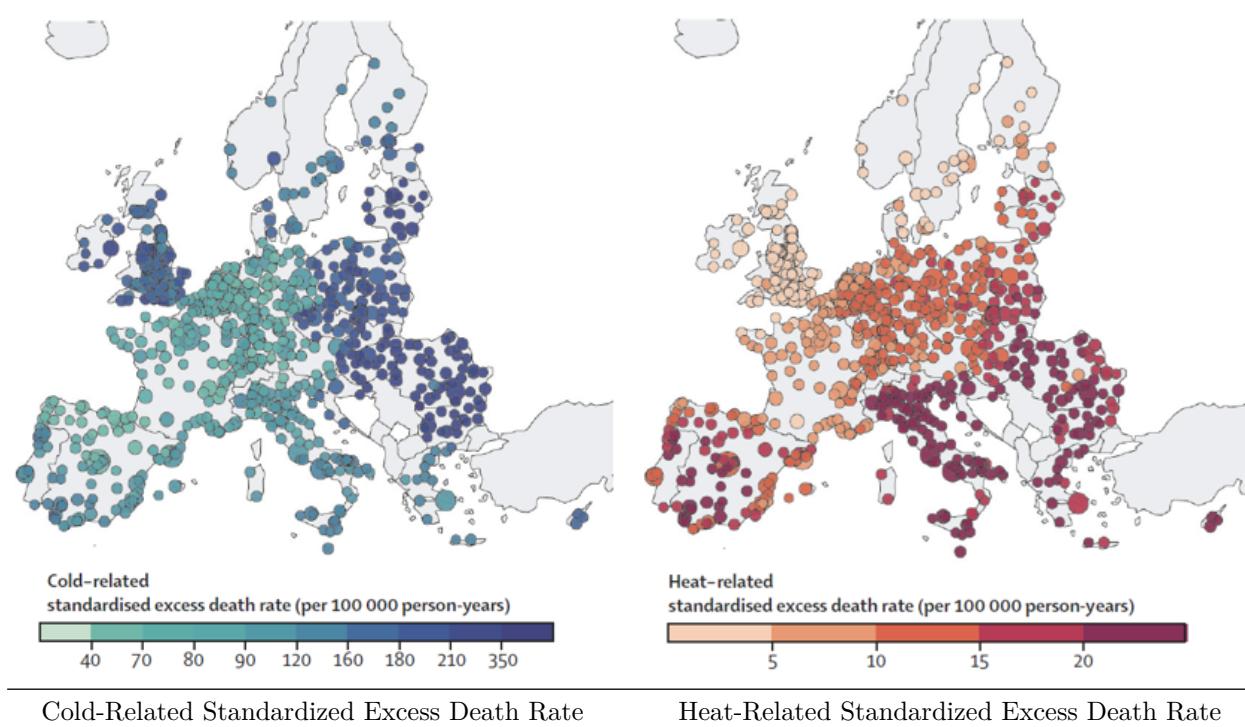
FIGURE 10.2: Relative Risk for Cold and Heat in Capital Cities

10.2.4.3 Geographical Gradient

One notable finding is the north-south gradient in the Minimum Mortality Percentile (MMP) and Minimum Mortality Temperature (MMT). As temperatures get hotter, the MMP decreases and the MMT increases, indicating that populations in hotter climates have adapted to higher temperatures as their baseline for optimal health.



The study's maps illustrate these gradients and the spatial distribution of temperature-related mortality risks. For example, cold-related mortality shows higher excess death rates in Eastern Europe, while heat-related mortality is more prominent in Southern Europe.



These visualizations provide a comprehensive overview of the temperature-mortality relationship across Europe, highlighting regions and populations that are most vulnerable to extreme temperatures.

10.2.5 Conclusion

The detailed findings from this study underscore the importance of tailored public health strategies that consider regional and age-specific vulnerabilities. By understanding the nuanced impacts of heat and cold across different European cities and demographic groups, policymakers can better prepare for and mitigate the adverse health effects of extreme temperatures.

10.2.6 Data and Framework

The study utilized data from 854 urban audit cities from Eurostat and 232 cities from the Multicountry Multi-city (MCC) Collaborative Research Network. City-specific meta-predictors included total population, proportion of population above 65 years, population density, GDP per capita, and environmental factors like PM2.5 concentration and annual temperature range.

Dimensionality reduction is achieved using PLS regression to create composite indices of vulnerability. This method helps in capturing geographical variability in risks while avoiding overfitting. PLS regression is used for dimensionality reduction and predictive modeling, particularly useful when dealing with a large number of highly correlated predictors. The Akaike Information Criterion (AIC) helps select the optimal number of PLS components by balancing goodness of fit with model complexity.

10.2.7 Modelling Framework

The modelling framework includes several stages: 1. **City-Specific Risk Estimation:** Using a quasi-Poisson regression model with a distributed lag non-linear model (DLNM). 2. **Meta-Regression and Spatial Interpolation:** Pooling city-specific estimates into a multivariate multilevel meta-regression model. 3. **Prediction and Impact Assessment:** Predicting ERFs for all 854 cities and estimating excess deaths.

Here is a Flowchart of the Modelling Framework:

10.2.8 Partial Least Squares (PLS) Regression

Partial Least Squares (PLS) regression is a statistical method used for dimensionality reduction and predictive modeling. It is particularly useful when dealing with a large number of highly correlated predictors. The goal of PLS is to find the fundamental relations between two matrices (predictors and responses) by projecting the predictors to a new space.

In the context of epidemiologic studies on the heat effects, PLS regression helps in creating composite indices of vulnerability from city-specific characteristics. This method captures geographical variability in risks while avoiding overfitting.

10.2.8.1 PLS Algorithm

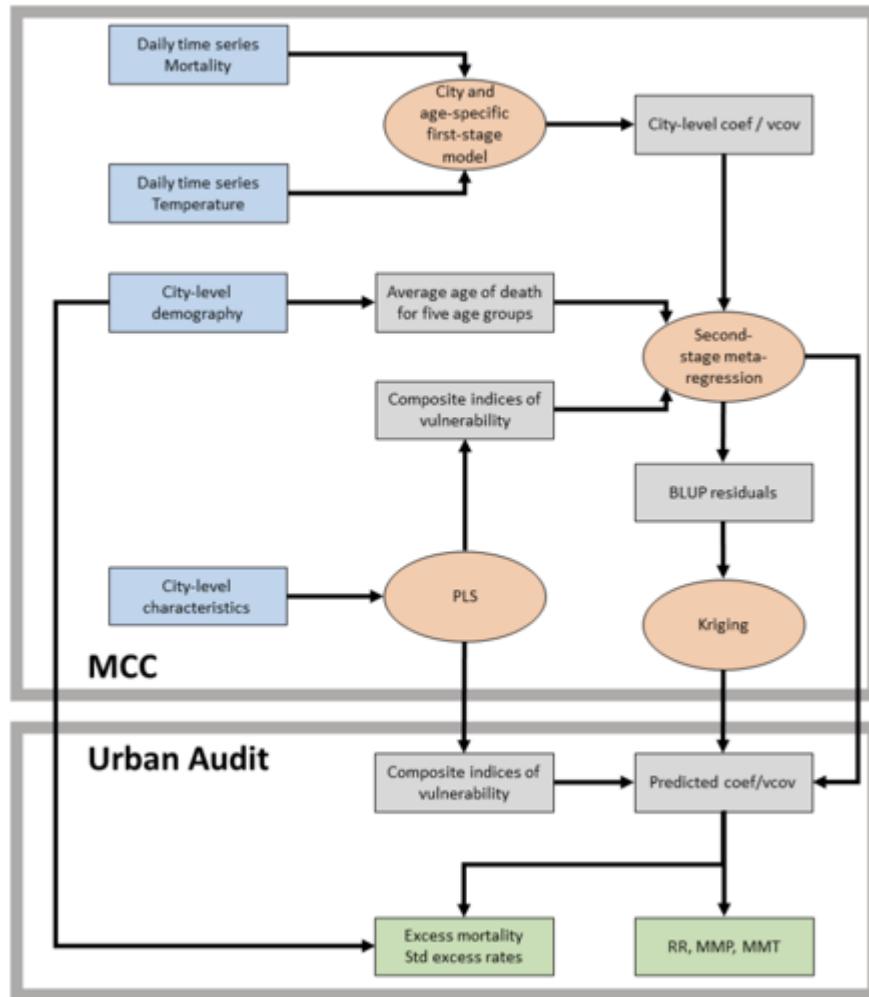
The PLS algorithm iteratively constructs new predictor variables (latent variables) that are linear combinations of the original predictor variables. These new variables are uncorrelated and explain the maximum possible variance in the response variables.

The steps involved in PLS are as follows:

1. Standardize the predictors and responses:

$$\mathbf{X} = \frac{\mathbf{X} - \bar{\mathbf{X}}}{\mathbf{S}_{\mathbf{X}}}$$

$$\mathbf{Y} = \frac{\mathbf{Y} - \bar{\mathbf{Y}}}{\mathbf{S}_{\mathbf{Y}}}$$

**FIGURE 10.3:** Modelling Framework Flowchart

2. Calculate the covariance matrix:

$$\mathbf{C} = \mathbf{X}^T \mathbf{Y}$$

3. Extract the first latent variable:

$$\mathbf{w}_1 = \arg \max (\mathbf{w}^T \mathbf{C} \mathbf{w})$$

Subject to $\|\mathbf{w}\| = 1$.

4. Compute the scores for the first latent variable:

$$\mathbf{t}_1 = \mathbf{X} \mathbf{w}_1$$

$$\mathbf{u}_1 = \mathbf{Y} \mathbf{q}_1$$

Where $\mathbf{q}_1 = \mathbf{C} \mathbf{w}_1$.

5. Deflate the matrices:

$$\begin{aligned} \mathbf{X}_1 &= \mathbf{X} - \mathbf{t}_1 \mathbf{p}_1^T \\ \mathbf{Y}_1 &= \mathbf{Y} - \mathbf{t}_1 \mathbf{c}_1^T \end{aligned}$$

Where $\mathbf{p}_1 = \mathbf{X}^T \mathbf{t}_1 / (\mathbf{t}_1^T \mathbf{t}_1)$ and $\mathbf{c}_1 = \mathbf{Y}^T \mathbf{t}_1 / (\mathbf{t}_1^T \mathbf{t}_1)$.

6. Repeat steps 2-5 for the remaining components.

10.2.8.2 Geometric View

Geometrically, PLS regression can be viewed as projecting the original predictor space \mathbf{X} into a lower-dimensional latent space while maximizing the covariance between the predictors and the responses \mathbf{Y} .

The projections are chosen to ensure that each successive pair of latent variables captures the maximum possible covariance between \mathbf{X} and \mathbf{Y} .

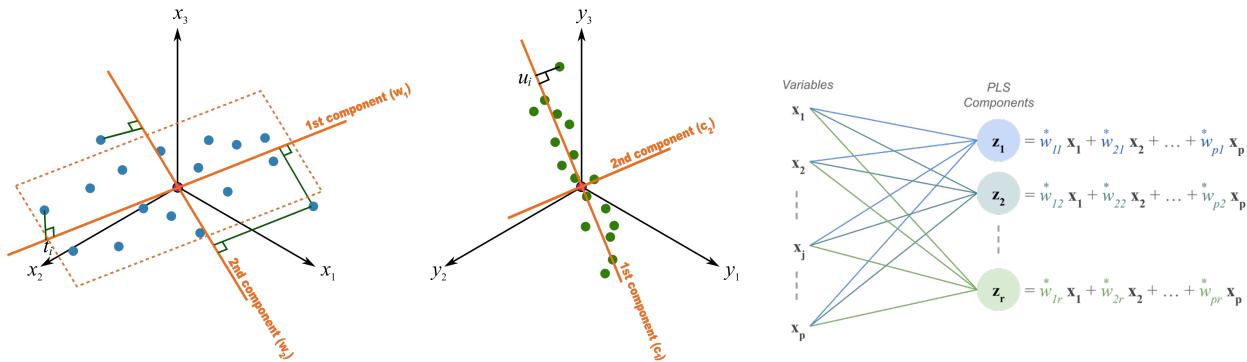


FIGURE 10.4: Geometric View of PLS

10.2.9 Akaike Information Criterion (AIC) and Model Selection

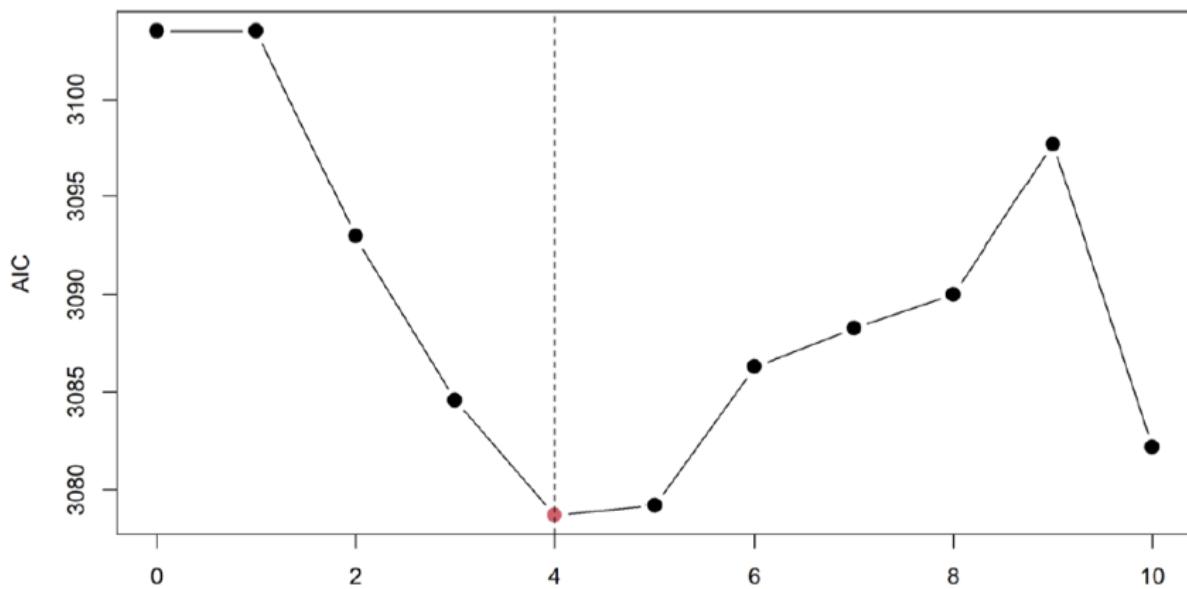
AIC is used to select the optimal number of PLS components. It balances the goodness of fit with the complexity of the model. Lower AIC values indicate a better model fit, accounting for the number of parameters used.

The formula for AIC is:

$$\text{AIC} = 2k - 2 \ln(\hat{L})$$

where k is the number of parameters in the model and \hat{L} is the maximized value of the likelihood function for the estimated model.

The plot below shows the AIC values for different numbers of PLS components. The optimal number of components is indicated by the lowest AIC value, which in this case is four components. This balance ensures that the model is neither overfitted nor underfitted, providing a parsimonious and explanatory model for predicting temperature-related mortality risks.

¹

10.2.10 Application in the Study

In the epidemiologic study on the heat effects, PLS regression was performed by reducing the dimensionality of city-specific characteristics into a few uncorrelated components. These components were then used in the meta-regression model to predict temperature-related mortality risks across different cities and age groups.

The PLS regression allowed the study to incorporate a large number of variables, such as demographic and environmental factors, without overfitting the model. This approach provided a robust framework for assessing the vulnerability of different urban areas to heat and cold effects.

By selecting the optimal number of components through AIC, the study ensured that the models were both parsimonious and explanatory, leading to reliable predictions of excess mortality due to temperature extremes.

10.2.11 Application in the Study

In the epidemiologic study on the heat effects, PLS regression was performed by reducing the dimensionality of city-specific characteristics into a few uncorrelated components. These components were then used in the meta-regression model to predict temperature-related mortality risks across different cities and age groups.

The PLS regression allowed the study to incorporate a large number of variables, such as demographic and environmental factors, without overfitting the model. This approach provided a robust framework for assessing the vulnerability of different urban areas to heat and cold effects.

By selecting the optimal number of components through AIC, the study ensured that the models were both parsimonious and explanatory, leading to reliable predictions of excess mortality due to temperature extremes.

¹Source: Excess mortality attributed to heat and cold

10.2.12 Model

The model used to estimate the main effect of air temperature on mortality employs a quasi-Poisson regression framework. This framework is suitable for count data, such as daily death counts, and can handle overdispersion commonly observed in such data.(?)

10.2.12.1 Quasi-Poisson Regression Model

The primary model is formulated as follows:

$$\log(\mu_i) = \alpha + \sum_{j=1}^p \beta_j x_{ij} + \gamma z_i$$

where: - μ_i is the expected daily death count for day i , - α is the intercept, - x_{ij} represents the predictor variables (e.g., temperature, humidity, etc.) for day i and predictor j , - β_j are the regression coefficients for the predictor variables, - z_i includes other confounders such as air pollution levels, day of the week, and long-term trends, - γ represents the coefficients for the confounders.

The quasi-Poisson model is an extension of the Poisson model, which allows for overdispersion by adding a dispersion parameter θ :

$$\text{Var}(Y_i) = \theta \mu_i$$

10.2.13 Assessing Overdispersion

In epidemiological studies involving count data, such as daily mortality counts, it is essential to assess for overdispersion. Overdispersion occurs when the observed variance in the count data is greater than what is expected under a standard Poisson model. This can lead to underestimated standard errors and consequently overconfident statistical inferences.

10.2.13.1 Overdispersion in Poisson Models

The Poisson regression model assumes that the mean and variance of the count data are equal ($\mu = \sigma^2$). However, in real-world data, this assumption often does not hold, especially in the presence of unobserved heterogeneity or clustering effects.

To assess for overdispersion, we can use the following diagnostic measures:

- Dispersion Parameter (θ):** The dispersion parameter is defined as the ratio of the Pearson chi-square statistic to the degrees of freedom:

$$\theta = \frac{\sum_{i=1}^n (y_i - \hat{\mu}_i)^2 / \hat{\mu}_i}{n - p}$$

where y_i is the observed count, $\hat{\mu}_i$ is the predicted count, n is the number of observations, and p is the number of parameters estimated in the model. A θ value greater than 1 indicates overdispersion.

- Pearson Chi-Square Statistic:** The Pearson chi-square statistic is calculated as:

$$X^2 = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}$$

This statistic is compared against the chi-square distribution with $n-p$ degrees of freedom. A significantly large value indicates overdispersion.

- Residual Deviance:** The residual deviance should be approximately equal to the degrees of freedom under the Poisson assumption. A much larger residual deviance suggests overdispersion:

$$\text{Deviance} = 2 \sum_{i=1}^n \left[y_i \log \left(\frac{y_i}{\hat{\mu}_i} \right) - (y_i - \hat{\mu}_i) \right]$$

10.2.13.2 Addressing Overdispersion

When overdispersion is detected, alternative modeling approaches should be considered to obtain valid statistical inferences:

1. **Quasi-Poisson Model:** The quasi-Poisson model adjusts for overdispersion by introducing a dispersion parameter:

$$\text{Var}(Y_i) = \theta\mu_i$$

This model provides robust standard errors that account for overdispersion.

2. **Negative Binomial Model:** The negative binomial model is another approach that introduces an additional parameter to model the overdispersion. The variance is modeled as:

$$\text{Var}(Y_i) = \mu_i + \alpha\mu_i^2$$

where α is the overdispersion parameter. This model is particularly useful when the overdispersion increases with the mean.

3. **Generalized Estimating Equations (GEE):** GEEs provide a flexible approach to handle correlated and overdispersed data by specifying a working correlation structure. They produce robust standard errors that are valid even if the specified correlation structure is incorrect.

10.2.13.3 Application in the Study

In the context of the study on the effects of temperature on mortality, assessing overdispersion is crucial for ensuring the validity of the model estimates. The quasi-Poisson regression model was chosen for this study due to its ability to handle overdispersed data effectively. The dispersion parameter was estimated, and diagnostic checks were performed to confirm the presence of overdispersion.

By using the quasi-Poisson model, the study obtained robust standard errors and more reliable estimates of the temperature-mortality relationship. This approach ensured that the findings were not biased by under-estimated variability, leading to more accurate and confident public health recommendations.

10.2.13.4 Quasi-Poisson Regression Model

The Quasi-Poisson regression model adjusts for overdispersion in count data like mortality. The model is represented as:

$$\log(\mu_i) = \alpha + \sum_{j=1}^p \beta_j x_{ij} + \gamma z_i$$

where μ_i is the expected daily death count, α is the intercept, x_{ij} represents the predictor variables (e.g., temperature), β_j are the regression coefficients, and z_i includes other confounders like air pollution levels.

10.2.13.5 Distributed Lag Non-Linear Model (DLNM)

To capture the delayed and non-linear effects of temperature on mortality, a Distributed Lag Non-Linear Model (DLNM) is used (?). The DLNM incorporates the lagged effects of temperature, recognizing that exposure to extreme temperatures can affect mortality over several days. The model can be expressed as:

$$\log(\mu_i) = \alpha + f(\text{temperature}_{i-l}) + \sum_k \beta_k z_{ik}$$

where $f(\text{temperature}_{i-l})$ represents a smooth function of temperature over lag l days, and the other terms are as previously defined.

The smooth function f is typically modeled using splines, allowing for flexible non-linear relationships:

$$f(\text{temperature}_{i-l}) = \sum_{m=1}^M s_m(\text{temperature}_{i-l})$$

where s_m are the spline basis functions and M is the number of basis functions.

10.2.13.6 Meta-Analysis

In the second stage, the results from individual city-specific models are pooled using a multilevel random-effects meta-analysis. This approach combines the estimates from different cities while accounting for between-city variability. The random-effects model is given by:

$$\hat{\theta}_j = \theta + u_j + \epsilon_j$$

where: - $\hat{\theta}_j$ is the estimated effect size for city j , - θ is the overall effect size, - u_j is the random effect for city j , assumed to follow a normal distribution with mean 0 and variance τ^2 , - ϵ_j is the within-city error term.

The overall effect size θ is estimated by weighting the city-specific estimates based on their variances:

$$\hat{\theta} = \frac{\sum_{j=1}^J w_j \hat{\theta}_j}{\sum_{j=1}^J w_j}$$

where $w_j = \frac{1}{\text{Var}(\hat{\theta}_j) + \tau^2}$ are the weights.

10.2.13.7 Incorporating Age-Specific Effects

The model also considers age-specific effects by stratifying the analysis by different age groups. Age-stratified models help in understanding the differential impact of temperature on various age cohorts. The age-specific model is given by:

$$\log(\mu_{i,a}) = \alpha_a + \sum_{j=1}^p \beta_{j,a} x_{ij} + \gamma_a z_i$$

where $\mu_{i,a}$ is the expected daily death count for age group a , and the other terms are as previously defined but specific to age group a .

10.2.14 Findings and Interpretation

The combination of quasi-Poisson regression, DLNM, and multilevel meta-analysis provides a robust framework for estimating the temperature-mortality relationship. The pooled estimates across cities indicate the overall effect of temperature on mortality while accounting for local variations and confounders.

The findings from the model underscore the significant impact of temperature extremes on mortality, with cold temperatures generally showing a larger effect than heat in many European cities. The results highlight the importance of considering both immediate and lagged effects of temperature, as well as the need for city-specific and age-specific analyses to inform targeted public health interventions.

By employing these advanced statistical techniques, the study provides comprehensive and nuanced insights into the public health implications of temperature variability and extreme weather events.

10.3 Joint Effect of Heat and Air Pollution on Mortality

10.3.1 Findings

The study by Stafoggia et al. (2023) examined the combined effects of high temperatures and air pollution on mortality. The analysis involved data from 620 cities in 36 countries, providing a comprehensive overview of the joint impact of these environmental stressors. The findings indicated that both high temperatures and air pollution independently increase mortality rates. However, their combined effect is significantly more pronounced. Specifically, a temperature increase from the 75th to the 99th percentile was associated with an 8.9% rise in mortality.

10.3.2 Model

To estimate the joint effect of air temperature and air pollution on mortality, a two-stage modeling approach was employed.

10.3.2.1 First Stage: City-Specific Models

In the first stage, city-specific models were developed using a quasi-Poisson regression framework. The model is expressed as:

$$\log(\mu_i) = \alpha + \beta_1 \text{Temp}_i + \beta_2 \text{PM}_i + \beta_3 (\text{Temp}_i \times \text{PM}_i) + \sum_{j=1}^p \gamma_j z_{ij}$$

where: - μ_i is the expected daily death count for day i , - α is the intercept, - Temp_i represents the temperature on day i , - PM_i represents the air pollution level (e.g., PM2.5 concentration) on day i , - $\text{Temp}_i \times \text{PM}_i$ is the interaction term between temperature and air pollution, - β_1, β_2 , and β_3 are the coefficients for temperature, air pollution, and their interaction, respectively, - z_{ij} represents other covariates (e.g., day of the week, long-term trends), - γ_j are the coefficients for the covariates.

The interaction term (β_3) captures the joint effect of temperature and air pollution on mortality, indicating how the combined exposure amplifies the risk.

10.3.2.2 Distributed Lag Non-Linear Model (DLNM)

The DLNM was used to account for the delayed effects of temperature and air pollution. The model is formulated as:

$$\log(\mu_i) = \alpha + f(\text{Temp}_{i-l}, \text{PM}_{i-l}) + \sum_{k=1}^p \gamma_k z_{ik}$$

where: - $f(\text{Temp}_{i-l}, \text{PM}_{i-l})$ is a bivariate smooth function of temperature and air pollution over lag l days, - z_{ik} are other covariates.

This model allows for flexible modeling of the non-linear and lagged relationships between temperature, air pollution, and mortality.

10.3.2.3 Second Stage: Meta-Analysis

In the second stage, the city-specific estimates were pooled using a multilevel random-effects meta-analysis. This approach combines the results from different cities, accounting for between-city heterogeneity.

The random-effects model is given by:

$$\hat{\theta}_j = \theta + u_j + \epsilon_j$$

where: - $\hat{\theta}_j$ is the estimated effect size for city j , - θ is the overall effect size, - u_j is the random effect for city j , assumed to follow a normal distribution with mean 0 and variance τ^2 , - ϵ_j is the within-city error term.

The overall effect size θ is estimated by weighting the city-specific estimates based on their variances:

$$\hat{\theta} = \frac{\sum_{j=1}^J w_j \hat{\theta}_j}{\sum_{j=1}^J w_j}$$

where $w_j = \frac{1}{\text{Var}(\hat{\theta}_j) + \tau^2}$ are the weights.

10.3.3 Comparison of Findings: Masselot et al. (2023) vs. Stafoggia et al. (2023)

Both studies highlight the significant impact of temperature on mortality, but they approach the analysis from different angles and include different additional variables.

10.3.3.1 Masselot et al. (2023) (?)

- **Focus:** Assessed the impact of non-optimal temperatures (both heat and cold) on mortality across Europe.
- **Model:** Used a quasi-Poisson regression with DLNM for temperature effects, and meta-regression for pooling city-specific estimates.
- **Key Findings:** Higher mortality rates associated with cold temperatures compared to heat. Significant regional variations with Eastern Europe showing higher vulnerability. Age-specific analysis revealed increased vulnerability in older age groups.
- **Data:** Included 854 cities across Europe.

10.3.3.2 Stafoggia et al. (2023) (?)

- **Focus:** Examined the joint effect of high temperatures and air pollution on mortality.
- **Model:** Used a quasi-Poisson regression with an interaction term for temperature and air pollution, combined with DLNM for lagged effects. Pooled city-specific estimates using multilevel random-effects meta-analysis.
- **Key Findings:** Both high temperatures and air pollution independently increase mortality, but their combined effect is more pronounced. The interaction between temperature and air pollution significantly amplifies mortality risk.
- **Data:** Included 620 cities in 36 countries.

10.3.3.3 Comparison of Temperature Effects

- **Temperature Alone:**
 - Both studies confirm that temperature extremes (both high and low) significantly affect mortality rates.
 - Masselot et al. focused on a broad spectrum of temperature effects, including both heat and cold.
 - Stafoggia et al. specifically addressed the interaction between heat and air pollution.
- **Combined Effects:**
 - The interaction term in Stafoggia et al.'s model revealed a synergistic effect, where the combined exposure to heat and air pollution resulted in a higher mortality risk than the sum of their individual effects.
 - This aspect was not covered in Masselot et al.'s study, which did not include air pollution as a variable.
- **Methodological Differences:**
 - Masselot et al. used a wider range of temperatures and focused on a comprehensive assessment of cold and heat effects, including age-specific vulnerabilities.
 - Stafoggia et al. provided insights into the compounded effects of environmental stressors, emphasizing the need to consider multiple factors in public health policies.

10.3.4 Conclusion

Both studies contribute valuable insights into the public health impacts of temperature extremes. Masselot et al. (2023) (?) highlighted the broader spectrum of temperature effects across Europe, while Stafoggia et

al. (2023) (?) provided a nuanced understanding of the combined effects of heat and air pollution. Together, these studies underscore the importance of considering both direct and interactive effects of environmental stressors in epidemiological research and public health planning.



11

Controversial issue : heat and humidity

Author: Mona Niethammer

Supervisor: Helmut Kuechenhoff

Degree: Master

11.1 Abstract

As climate change progresses, heat-related mortality and adverse health outcomes are also increasing. In the chapters before it was mainly talked about temperature related outcomes. This chapter focuses on heat events and how humidity plays a crucial role in high ambient temperatures and the effect on health outcomes. Physiological research indicates that both heat and humidity significantly impact the body's response to heat. Thus, influencing mortality and adverse health outcomes. Several epidemiological studies aimed to show this relationship, found however only minimal or even no effect of humidity on negative health outcomes. This discrepancy between physiological knowledge and epidemiological findings creates a, as ([Baldwin et al. \(2023\)](#)) describes it, controversial issue. This will be further explored in this chapter.

11.2 Introduction

11.2.1 Background

Climate change is a broad and pressing topic as the number of heat-related deaths and the air-temperature break records. Extreme heat waves have become a consistent feature of summer seasons, leading to an increase in heat-related fatalities ([K.L.Ebi \(2021\)](#)). Rising greenhouse gas concentrations further contribute to rising temperatures and humidity at the same time. Physiological knowledge indicates that heat and humidity both contribute to human heat stress, adversely affecting the human body ([Baldwin et al. \(2023\)](#)). Heat stress occurs when environmental conditions overwhelm the body's cooling mechanism ([Buzan and Huber \(2020\)](#)). When the body is not able to sufficiently lower its core temperature, the human body suffers health problems, potentially leading to death. From a physiological perspective, the body's cooling mechanisms are clearly influenced both by heat and humidity.

In contrast, a wide range of epidemiological studies have found either no or only a weak correlation between humidity and human heat stress or adverse health outcomes. Those studies all infer that a rise in air-temperature (heat) is associated with negative health outcomes. These two contradictions lead, according to ([Baldwin et al. \(2023\)](#)), to a controversial issue as physiologically it is known that there is an effect of humidity on health outcomes, whereas in epidemiological studies they did not find evidence for such an effect. Figure 1 describes this contradiction between physiological knowledge and epidemiological studies. Heat in this chapter is used as synonym for high ambient temperature.

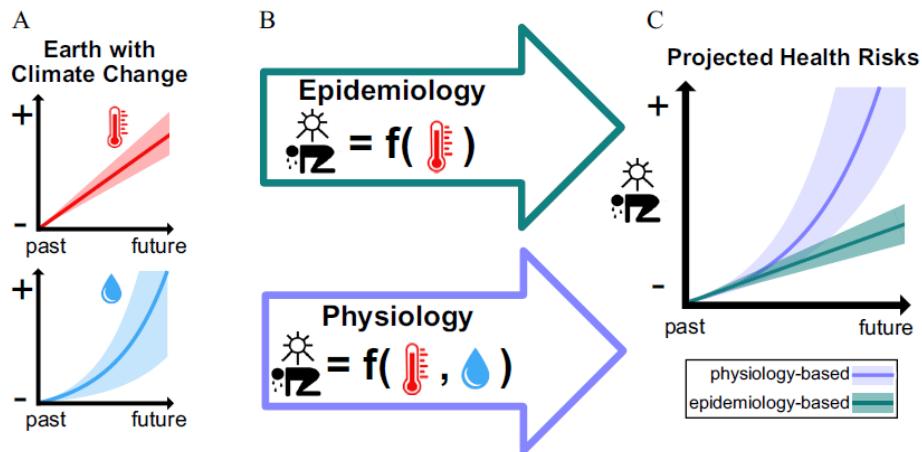


Figure 1: Flowchart illustrating the controversial issue between epidemiological studies and physiological knowledge.

A: illustrates how heat and humidity changes with climate change. B: illustrates that epidemiological studies only take heat as driver into account, the physiological view heat and humidity. C: shows the predictions from an epidemiological and physiological view.

11.2.2 Epidemiological Studies

Numerous epidemiological studies investigated the influence of heat and humidity on negative health outcomes such as death or cardiovascular diseases. Results of these studies are mixed regarding the association of humidity and negative health outcomes. As an example, a study in New York, conducted by Barreca found a positive correlation between high levels of relative humidity (RH) and cardio-respiratory hospitalizations. However, this study is one of few that identified an effect of humidity on negative health outcomes. Baldwin et al. reviewed several epidemiological studies and conducted further research. They concluded six key reasons why epidemiological studies found less or even no effect of humidity on health outcomes:

- 1) At high temperatures, humidity has minimal or even no influence on negative health outcomes
- 2) Study results may be skewed by limited data sets
- 3) Analyses often focus on vulnerable populations, like elderly, infants or individuals with impaired sweating functions
- 4) Necessary extreme values of heat and humidity to show a significant impact on heat strain are rare
- 5) The relationship between heat and humidity is often not adequately considered. This may lead to inappropriate model results
- 6) Sub-daily meteorological phenomena, such as rain, which occur during high heat and humidity, may bias studies based on daily data (Baldwin et al. (2023)).

11.2.3 Physiological knowledge

Physiological evidence indicates that heat and humidity increase health risks, leading to higher rates of morbidity, mortality, reduced physical work capacity, and other adverse health outcomes as the climate crisis worsens. ([Buzan and Huber \(2020\)](#))

The human body has two primary cooling mechanism to control the body temperature. One mechanism redistributes the blood flow to the skin (vasodilation), and therefore allows metabolically generated heat from inside of the body to reach the skin's surface and subsequently the environment. This process requires an increased blood flow and elevates cardiac demand. This may trigger cardiovascular events such as myocardial infarction in subjects with heart conditions. The second mechanism involves secreting sweat onto the skin. Subsequently, the sweat evaporates to the environment and heat from the body is removed. These two mentioned cooling mechanisms are crucial for the human body to maintain a sustainable core temperature. ([K.L.Ebi \(2021\)](#))

The effectiveness of these cooling mechanisms decline with the increase of heat and humidity. A higher ambient humidity reduces the proportion of sweat that evaporates from the skin into the air. Hence, the effectiveness of sweating decreases and sweating does not lead to heat loss anymore, causing a rise in body temperature. The human body responds to such ineffective sweating by an even higher rate of sweating. This can lead to heat stroke, dehydration, or the body's physical limits may simply be exceeded. ([Baldwin et al. \(2023\)](#))

Certain populations are at higher risk due to less efficient cooling mechanisms. For instance, people older than 65 years may have a reduced sweating ability same as individuals with cardiovascular diseases who cannot effectively increase blood flow to the skin. These factors, among others, exacerbate the difficulty in cooling the body. ([K.L.Ebi \(2021\)](#))

When the body can no longer cool itself, heat stress occurs. However, these physiological insights are not consistently reflected in epidemiological studies. Several methodological and physiological factors might explain this discrepancy.

Firstly, heat-vulnerable individuals often have impaired sweating capacities, especially with aging this capacity decreases. Thermoregulatory sweating can decrease by up to 25% in individuals older than 60. Consequently, they will be less sensitive to high humidity. Since most studies only include individuals over 65 years, this might explain the lack of association between humidity and heat stress.

Secondly, it is possible that the threshold at which differences in humidity impact the body's regulatory mechanisms are rarely attained. This means, that only at a specific threshold of humidity the body's cooling mechanisms have trouble and this threshold may be reached rarely in especially dry areas.

Thirdly, heat-related outcomes may result from other causes unrelated to humidity. For example, cardiovascular diseases are a leading cause of death, and high temperatures can exacerbate these conditions independently of humidity.

Lastly, dehydration, a major cause of heat-related deaths, reduces the amount of sweat, thereby minimizing the impact of high humidity on evaporation and cooling. ([Baldwin et al. \(2023\)](#))

11.3 Heat and Humidity in Studies

11.3.1 Humidity definitions

Humidity has multiple definitions. This makes it crucial to select an appropriate measure to answer specific research questions. ([Baldwin et al. \(2023\)](#)) categorize humidity variables into the four following groups:

1. Simple relative variables including Relative Humidity (RH), Dew-point depression (DPD) and Saturation

deficit.

2. Mass-based variables including specific humidity, absolute humidity, dew point temperature, vapor pressure and mixing ratio.
3. Composite indicators such as the Heat index, Humidex, Wet-bulb-temperature, UTCI and further more.
4. Physiologic-based humidity indicators including maximum evaporation, required evaporation, total evaporation.

The simple relative variable RH (%) is the ratio of actual vapor pressure (e) in the air to the saturation vapor pressure (e_s), indicating proximity to saturation ($RH = e/e_s$). The dew-point depression (DPD) ($^{\circ}\text{C}$) indicates the relative dryness of the air and is the difference between the near-surface air temperature (T_a) and the dew point temperature (T_d) ($DPD = T_a - T_d$). Saturation deficit refers to the amount by which the water vapor must be increased in a given environment to reach saturation, assuming unchanged temperature and pressure.

The mass-based variable specific humidity (q) (g/kg) is the ratio of the mass of water vapor (g) to the total mass of moist air (kg). Absolute humidity is the density of water vapor, defined as the ratio of the mass of water vapor (g) to the volume of the air mixture (m^3). The dew point temperature (T_d) ($^{\circ}\text{C}$) is the temperature at which saturation occurs. The mixing ratio (w) refers to the ratio of the mass of water vapor (g) to the mass of dry air (kg).

The Heat Index (HI) describes a “feels like” temperature, while the Humidex is a discomfort index. The wet-bulb temperature is the temperature at which air becomes saturated through evaporation at constant pressure. The Universal Thermal Climate Index (UTCI) ($^{\circ}\text{C}$) is a physiologically based thermal stress scale designed to reflect the average human physiological response to the outdoor thermal environment.

The physiological-based humidity indicator maximum evaporation (E_{max}) represents the maximal evaporation rate in a given environment and clothing conditions. The required evaporation (E_{req}) measures the evaporation needed to achieve thermal balance, and the total evaporation (E) is the total amount of heat lost via evaporation of sweat from the body.

Davis et al. reviewed the use of humidity variables in epidemiological studies. They found that 62% of the studies used RH, while absolute and specific humidity were used in only 5.3% and 1.5% of the studies, respectively. They discussed the appropriate use of different humidity variables, especially the use of RH due to its strong inverse correlation with temperature and its diurnal and seasonal variability was mentioned. It is recommended to use RH sparingly. Instead, they recommend using mass-based variables such as absolute or specific humidity, as they provide a more accurate measure of exposure to humidity levels.

The selection of humidity variables must be done carefully, they can be highly correlated with other variables, for example with heat. Davis et al. concluded that using RH as humidity variable is often inappropriate and should be done with caution if necessary. ([Davis \(2016\)](#))

The widespread use of RH in epidemiological studies is largely caused by the available data and the measurements which are performed. Mostly, if humidity is measured, then RH is measured. Deriving mass water-vapor based variables such as absolute and specific humidity can introduce uncertainty and bias in estimates.

11.3.2 Composite Indicators

Composite indicators are frequently used in studies. There are over 120 heat stress metrics available ([Buzan and Huber \(2020\)](#)), most of which were originally designed for heat warning systems. Composite indicators consider heat, humidity, and additional variables simultaneously, making them practical for certain applications. Often they are just a weighted sum of different variables. Using such composite indices is very practical, but it comes at the cost of interpretability. Changes in the composite index can result from variations in heat, humidity, or other included variables. Therefore, information about individual effects is lost

and the effects can not be separated. Furthermore, the use of such indices may lead to biased or misleading estimates. Hence, the effects of heat or humidity may be underestimated. Baldwin et al. recommend the use of alternative methods such as confounding or effect modification, to separately assess the effects of heat and humidity ([Baldwin et al. \(2023\)](#)).

11.3.3 Confounder

A confounder is a variable that affects both the exposure of interest and the outcome. Figure 2 illustrates how humidity acts as a confounder in the relationship between heat and health outcomes.

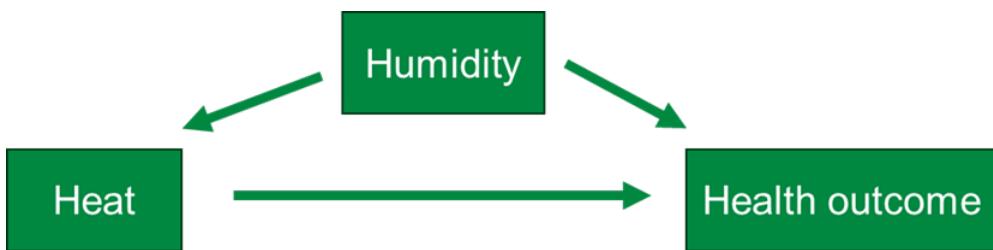


Figure 2: Humidity as confounding effect.

Humidity has been used in many epidemiological studies as a confounder when the effect of heat on health outcomes was analyzed. Researchers aim to remove the confounding influence of humidity on the relationship between heat and health outcomes by adjusting for humidity. This means the indirect effect of humidity through heat will be eliminated. Not adjusting for confounder variables, the causal effect of heat may be overestimated. This adjustment leads to a different interpretation of the estimates, as shown in the following linear model:

$$E[Y | t, h] = \beta_0 + \beta_1 * t + \beta_2 * h$$

Y represents the health outcome of interest, t is heat, and h humidity. In this model, β_1 represents the conditional total effect of heat on the health outcome at any given level of humidity. β_2 represents the direct effect of humidity on the outcome. That is the causal effect of humidity on the outcome when heat is held fixed and thus blocking the humidity effect on heat.

Interpretation of estimates can however get very complex and sometimes even misleading. Consequently, the causal relation between the variables needs to be considered carefully ([Baldwin et al. \(2023\)](#), [D. \(2013\)](#))

One significant problem of adjusting for confounding is to potentially run into multicollinearity issues. If heat (the exposure of interest) and humidity (the confounder) are highly correlated and both variables are included in the same regression model multicollinearity is present. Consequently, the estimates capture similar information and therefore making it difficult to interpret individual variable effects. Additionally, the precision of the estimated coefficients may be reduced. Including further variables such as calendar variables to adjust for long-term or seasonal trends can exacerbate multicollinearity if the correlation between heat and humidity varies significantly over time. Furthermore, including other variables which are highly correlated with heat and humidity will lead to further multicollinearity issues.

Baldwin et al. suggest using alternative approaches to significance testing that do not rely solely on p-values. Furthermore, it is crucial to carefully consider the relationship between heat and humidity and their individual effects on health outcomes to avoid misleading results. ([Baldwin et al. \(2023\)](#))

11.3.4 Effect Modification and Interaction

Including humidity as an effect modifier or interaction with heat is another method to test an effect of humidity in the framework of negative health outcomes and was used in some epidemiological studies ([Baldwin](#)

et al. (2023)). Unlike confounding, which distorts relationships and should therefore be eliminated, effect modification and interaction are desirable biases. They provide valuable insights into how variables interact and influence outcomes, revealing inherent aspects of causal relationships that need clarification.

Often, the two terms *Effect Modification* and *Interaction* are used interchangeably. However, from a causal view these two concepts differ in a counterfactual perspective and slightly differ in their definition. The counterfactual perspective involves hypothetical scenarios, asking “What if?” questions such as, “What if heat remained the same but humidity was lower? Would the health outcome change?”. (Bours (2021)) Counterfactual outcomes are typically written with a superscript such as Y^t and answers the question: “What would the outcome Y be if the exposure was t?”. Or as another example Y^{th} answers the question: “What would the outcome Y be if both exposures t and h occurred together?”. As in real-world data typically not all combinations of exposures are observed in each individual we are talking about hypothetical scenarios. Both concepts are scale dependent and their presence depends on the scale used. Here, we use the risk difference scale for formal definitions.

Interaction refers to the combined causal effect of two exposures, such as heat and humidity. *Effect modification* pertains to how the causal effect of one exposure changes across different levels of a second exposure. For example, “How does the effect of heat change across different levels of humidity?”. Let T refer to heat, H to humidity and Y represent the outcome of interest such as heat-related death. Y^t is the counterfactual outcome under exposure t and Y^{th} is the counterfactual outcome under t and h.

Humidity (H) is **effect modifier** on the risk difference scale for the effect of heat (T) on the outcome (Y) if H is not affected by T and there are two levels of T (t_0, t_1) and two levels of H (h_0, h_1), such that:

$$E[Y^{t_1} | H = h_1] - E[Y^{t_0} | H = h_1] \neq E[Y^{t_1} | H = h_0] - E[Y^{t_0} | H = h_0]$$

Effect modification indicates that the effect of heat on the outcome differs across different levels of humidity. By recognizing effect modification, public health interventions can better address how heat exposure interacts with humidity. This enables more targeted strategies to mitigate and adapt health risks associated with specific combinations of heat and humidity.

There is an **interaction** on the risk difference scale between heat (T) and humidity (H) on the outcome (Y) if there are two levels of T (t_0, t_1) and two levels of H (h_0, h_1), such that: $E[Y^{t_1, h_1}] - E[Y^{t_0, h_1}] \neq E[Y^{t_1, h_0}] - E[Y^{t_0, h_0}]$

Interaction requires that the joint effect of heat and humidity is different from the sum of their individual effects. If interaction is present, studying each exposure separately may not fully capture the complexity of their joint influence on the outcome. (VanderWeele (2009))

According to Baldwin and colleagues, analytical approaches to estimate interaction of effect modification are similar. They suggest that by including interaction terms or effect modifiers the main question of interest needs to be considered same as the scale of interest and the potential policy implications. (Baldwin et al. (2023))

Unfortunately, the interpretation of interaction terms and effect modifiers often leads to misinterpretation and often the scale (multiplicative or additive) also is not considered correctly. Therefore, it is crucial to understand the concepts of interaction and effect modification and how to interpret them correctly.

11.3.5 Data Limitations

The availability and quality of weather data vary across different locations all over the world. Low-income countries, like for example most african regions, often have limited access to weather data compared to high-income countries due to limited resources for data collection. Figure 3 illustrates global sub-daily weather data from the HadISD data source. This highlights the inconsistency in data coverage. Many regions in Africa, South America, and Australia have few measurements compared to Europe or the US. Such a lack of data, especially in areas prone to extreme heat events leaves gaps in our understanding and estimation of weather-related impacts on health outcomes.

Tropical and subtropical climates are characterized by high levels of humidity and heat. Unfortunately, these climates are particularly affected by these data limitations. Missing data in tropical and subtropical regions hinder our ability to comprehend the combined effects of heat and humidity on health outcomes. Therefore, epidemiological studies often do not include locations where humidity plays a significant role due to these data gaps.

Another challenge is the temporal resolution of weather data. Many studies rely on daily mean measurements, which may overlook fluctuations in heat and humidity throughout the day. It is known that heat and humidity fluctuate throughout the day. Using daily mean measurements may underestimate the effects of heat and humidity. Furthermore, correlations between heat and humidity can vary depending on the time of day, potentially impacting the findings of epidemiological studies.

Furthermore, typically weather stations are situated outside urban areas, often at airports. This may not capture urban heat effects. Additionally, as most measurements are taken outdoors, indoor and outdoor variations in heat and humidity are not accounted for.

Another limitation is that typically only relative humidity (RH) is measured at weather stations. This may not provide the most accurate representation of humidity. Deriving mass-based variables like specific or absolute humidity from RH may introduce uncertainties and bias in the estimates as mentioned earlier.

In summary, the lack of weather data can result in an underestimation of the impact of humidity on health outcomes. Researchers should consider these data limitations, especially when generalizing results to different countries and weather conditions. ([Baldwin et al. \(2023\)](#))

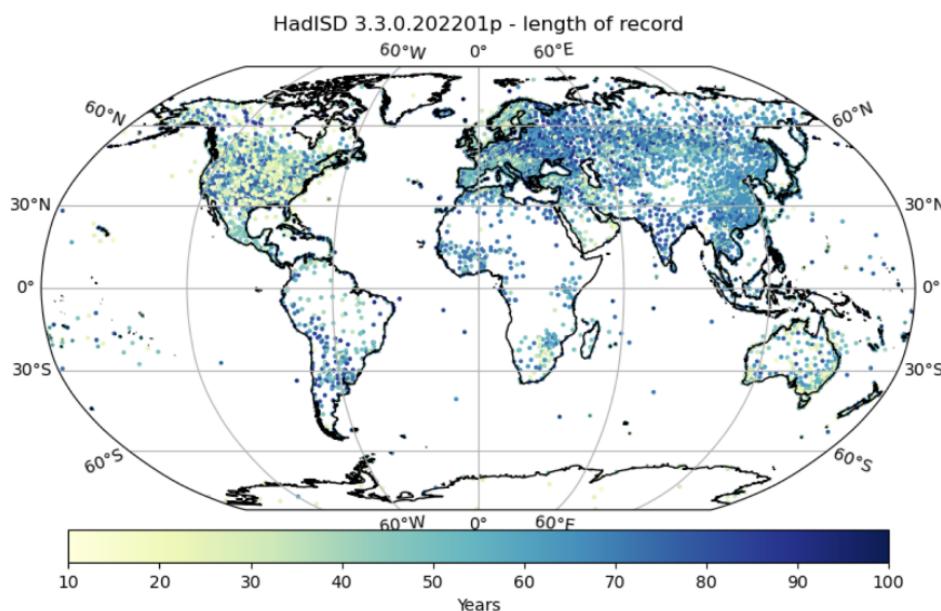


Figure 3: Station coverage and length of record in HadISD ([Raymond \(2024\)](#))

11.4 Examples

Armstrong et al. conducted a study to assess the impact of heat and humidity on daily mortality during summer months. They used a time series regression model with a distributed lag nonlinear model (DNLM) for

heat, using data from 445 cities across 24 countries. Numerous different models were fitted, including different humidity levels and terms and further one model with an interaction term between heat and humidity was modeled. Model comparison using the Akaike information criterion (AIC) revealed that the best fit was achieved when humidity was included as a linear term, i.e., as confounder. One very surprisingly result was that a 23% increase in relative humidity at the 99th percentile was associated with a 1.1% decrease in mortality. This was a non-significant finding contrary to the expectation that humidity leads to an increase in mortality. Sensitivity analyses involving specific humidity and dewpoint yielded similar results. Considering humidity as a potential effect modifier, by including an interaction term, was not favored based on AIC criteria ([Armstrong \(2019\)](#)).

Baldwin et al. criticize solely rely on AIC to determine whether an interaction term should be included in the model. Consequently, as no interaction term was used in the final model, the influence remains uncertain and raises questions about the value of its exclusion. The final model used by Armstrong et al. included relative humidity as a linear term, thereby treating humidity as a confounder. As highlighted earlier, employing humidity as a confounder can introduce several challenges.

Armstrong et al. recognized some limitations of their study. They discussed the data gaps from tropical and less developed countries, the exclusive consideration of mortality as an endpoint, and the possibility of missing the association between humidity and mortality under specific conditions ([Armstrong \(2019\)](#)).

In another epidemiological study from Schwartz et al. the effect of heat and humidity on hospital admissions for heart disease and myocardial infarction among individuals over 64 years from 12 US cities was investigated. They fitted a Poisson model for each city, incorporating regression splines to control for season and barometric pressure, and adjusting for the day of the week. Despite their broad approach, they found no evidence of an effect of humidity on hospital admissions. Possible explanations for this results are the use of relative humidity and the resulting bias. Further, considering humidity as a confounder may also explain that no effect of humidity was found. Additionally, focusing solely on individuals older than 64, who are more vulnerable to heat and sweating impairments, may have biased the study's findings. ([Schwartz \(2004\)](#))

Barreca conducted a study that found an effect of heat and humidity on mortality rates in the US. He used data from the Global Summary of the Day (GSOD) files, which provide detailed weather station data. The study analyzed data from 1973 to 2002, calculating specific humidity using a standard meteorological formula based on mean dew point and mean station pressure.

To estimate the effects of heat and humidity, Barreca employed an ordinary-least-squares model with the following equation:

$$MORT_{cmy} = \sum_b \beta^b * TEMP_{cmy}^b + \sum_{b'} \alpha^{b'} * HUMID_{cmy}^{b'} + X_{cmy} \gamma + \mu_{my} + \phi_{cm} + \delta_{cm} * TIME + \pi_{cm} * TIME^2 + \epsilon_{kym}$$

In this model:

- MORT represents the monthly mortality rate (per 100,000 inhabitants) in county c, year y and calendar month m.
- TEMP is a set of temperature variables indicating the fraction of days county c experiences mean temperatures within specific 10 °F bins (e.g., 50–60 °F).
- HUMID is a set of humidity variables indicating the fraction of days county c experiences mean humidity levels within specific two-grams-of-water-vapor bins (e.g., 2–4 g/kg).
- X is a vector of controls for precipitation and temperature-humidity interaction terms.
- μ represents unrestricted time effects.
- ϕ represents unrestricted county by calendar month fixed effects.
- $\delta_{cm} * TIME$ and $\pi_{cm} * TIME^2$ represent unrestricted county by calendar month quadratic time trends. Barreca clustered the standard errors by the state of residence to account for the possibility that ϵ is correlated within states, both across counties and over time. Additionally, the equation was weighted by the county population in 2000. He fitted different models: one model included only temperature, and another included only humidity. Both the temperature-mortality and humidity-mortality relationships

followed a similar pattern. In a subsequent model, he included both TEMP and HUMID covariates. An interesting finding from this model was that both cold temperatures and low humidity levels remained significant determinants of mortality. Finally, he fitted a model with an interaction term between heat and humidity. The estimates from this model suggest that increasing humidity levels are more dangerous at high temperatures. Therefore, controlling for humidity is according to Barreca particularly important in the context of predicting distributional effects of climate change. ([Barreca \(2012\)](#))

In a separate review, Budd examined the history and limitations of the wet-bulb globe temperature (WBGT) index, widely used since its development for the United States Army and Marine Corps in 1950. The index incorporates air temperature, mean radiant temperature, absolute humidity, and air movement as basic elements of the thermal environment. Budd identifies several limitations of the index, including its underestimation of stress in environments with restricted evaporation, challenges in interpreting its results, and limitations in the accuracy of sweat measurement due to calibration methods and instrumentation. ([Budd \(2008\)](#))

11.5 Discussion

The discrepancies observed in epidemiological studies regarding the impact of humidity on health outcomes may stem from various factors. It is important to understand how to appropriately use humidity in study designs and recognize the different humidity definitions. Prioritizing mass water-vapor based variables such as specific or absolute humidity over relative humidity is a crucial step. As discussed earlier, humidity can be considered as a confounder, effect modifier, in an interaction term with heat or component of a composite index. Each of these considerations of humidity comes with its own set of limitations.

It is essential to provide a clear research question or hypothesis when evaluating the role of humidity in studies. Researchers must be mindful of these considerations to ensure the robustness and validity of their study results.

Efforts to improve data collection are required, especially to include regions with tropical climates and low-income countries that are often underrepresented in studies. Additionally, relying solely on vulnerable populations may introduce bias, underscoring the importance of striving for a representative sample.

Humidity represents an intriguing factor worth studying further, along with heat, especially in studies investigating health effects. Planning and executing appropriate studies to examine the effects of heat and humidity requires a nuanced understanding of humidity. It is appropriate incorporation into models, and knowledge aware of potential limitations and errors.



12

Risk Projections

Author: Author

Supervisor: Helmut Kuechenhoff

Degree: Bachelor

In many studies, the effects of climate change in the future is discussed. In a tutorial paper (?) , statistical methods for modelling future risks due to climate devlopement are presented. In a two other papers paper the future effects of ozone and heat related are presented (?,?).



13

Climate Crisis and Mental Health

Author: Nina Straub

Supervisor: Helmut Küchenhoff

Degree: Master

13.1 1. Introduction

It is widely recognized that climate change (CC) is impacting both economy and society. However, it is also increasingly acknowledged as critical factor influencing various aspects of human health. In 2009, the World Health Organisation (WHO) declared CC as one of the largest threats to human health of the twenty-first century and recognised the need for a research agenda to understand complex linkages between climate change and health (?). In following years, initiatives such as the Lancet Countdown, a global monitoring system designed to track research and support policy decisions on Health and CC, drew resources and attention to this pressing problem (?). Despite these efforts, some health conditions are still little researched and the impact of climate change on them is poorly understood. Among them are mental health (MH) and human well-being (?). According to the OECD, 20-30% of the working population are affected by a MH condition. Mental disorders are the worldwide leading cause of years lived with disability (?, ?) and a considerable burden not only for individuals but also society and economy. In 2010, the direct and indirect costs were estimated to amount to 2493 billion US dollars worldwide (?) and the lost economic value amounts to 4-8% of GDP in different regions (?). However, health care systems spend a disproportional amount on treatment. In England, MH conditions make out about 26% of the burden of disease, but only 13% of health expenditure are attributed to treatment (?). Globally, 80% of people do not get sufficient treatment at all (?). Given the proportion of MH in global burden of disease, it seems surprising that little research has investigated how CC might affect MH. The aim of this chapter is to outline the challenges in researching the impact of CC on MH and to present research approaches on this topic. Emphasis will be placed on the methodological and statistical difficulties.

13.2 2. Challenges of Quantifying Climate Change Impact on Mental Health

Important to notice is that CC does not directly impact MH, but rather the consequences of a changing climate, like the increase in magnitude, frequency and intensity of extreme climate events mediate the effect of CC on MH (?). The Intergovernmental Panel On Climate Change (IPCC) report identified multiple pathways how CC could impact MH. The direct pathway includes exposure to traumatic events such as floods, hurricanes or other severe weather-related events. The indirect pathway consists of the consequences of prolonged events such as drought or sea level rise that lead to displacement, food shortages or economic hardship. The last pathway, vicarious exposure or overarching/diffuse exposure refers to the long-term emotional distress caused by the awareness of the negative consequences of CC and the associated feelings of

helplessness and despair (?). Although in the IPCC report most of the adverse effects of CC on MH are assessed with “very high confidence”, no concrete projections are given as to how many additional people will suffer from mental disorders. This is peculiar because for other health effects that are assessed with very high confidence, figures for additionally affected individuals are given (?). The discrepancy between confidence and tangible figures is caused by multiple challenges of attributing CC to MH, which can be clustered in external and internal difficulties. The external difficulties of quantification mainly consist of the globally insufficient resources devoted to monitoring and treating MH. Although the Lancet Countdown emphasised the need to integrate MH in future CC health considerations, only five out of 16 countries included MH in their health adaption plans (?). In the most recent Countdown report it was noted that the development of an indicator to cover the relations between CC and MH is still ongoing (?). One major problems hereby is that current prevalence estimates for mental illnesses are unreliable, as mental illness is severely underdiagnosed, especially in third world countries (?). With the lack of a baseline, good projection estimates are difficult. The internal difficulties of quantification evolve around the complex nature of MH itself. Hayes et al. (2018) identified four key aspects which complicate the estimation of CC impact: Firstly, the risk of pathologizing normal stress responses due to a changing climate while adverse mental health outcomes are underdiagnosed. Secondly, multiple causal pathways in which one CC correlate can influence different MH outcomes. Thirdly, effects of CC may only become apparent after a considerable time period, which complicates the investigation of causal effects. Lastly, existing complex interactions between MH and other social determinants of health, and their relationship with CC, is insufficiently understood.

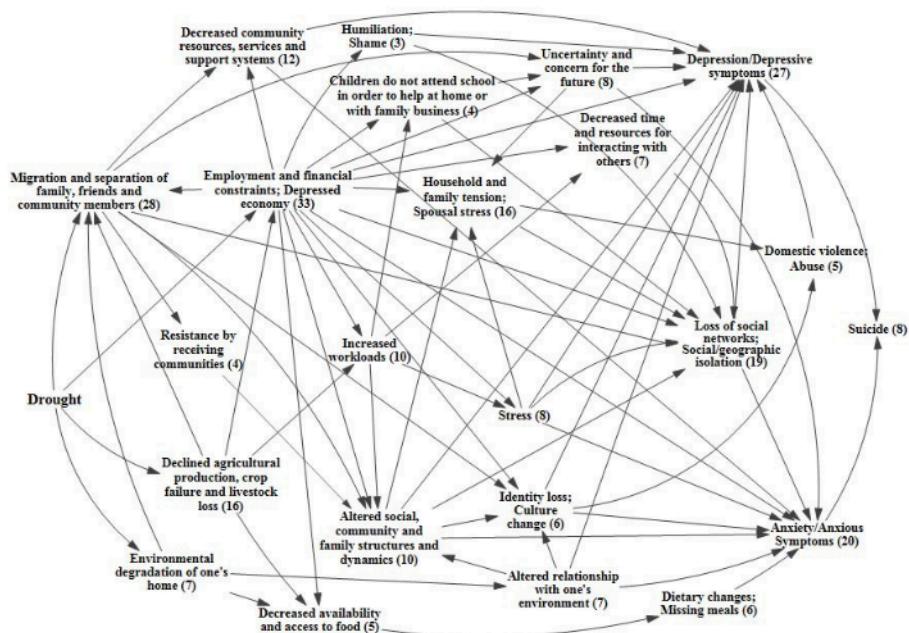


Figure 1. Causal process diagram of the impact of drought on the mental health outcomes depression, anxiety and suicide. Numbers in brackets are number of papers meeting the search criteria for each factor. Figure adapted from Vins et al. (2015).

The four aspects can be illustrated by a causal process diagram, shown in Figure 1, which depicts the multiple pathways by which droughts can influence MH outcomes (?). Most apparent is the economic route: Droughts have a negative impact on crop yield, influencing employment and financial stability in the agricultural sector. Financial hardship has been shown to have a negative impact on MH, leading to depression and, in severe cases, even suicide (?). Another causal route is the one concerned with the decreased availability of food, leading to dietary changes or famine. This can directly influence neurodevelopment of children or anxiety in adults, but also indirectly cause migration, which can lead to social isolation and loss of cultural identity, again having a negative impact on MH outcomes. To estimate the impact of droughts on MH e. g. via the

migration route, it would require not only to estimate the increase and severity of droughts due to CC, but also how much migration this would cause and how many of the affected individuals would consequently develop a mental illness. This is a genuine challenge for accurate climate projections, as estimation of migration movements and the development of mental illness depend on many interrelated factors (?). Some individual linkages between events (e. g. the influence of economic hardship on MH) are well-researched. Future research needs to integrate these pathways and link them to different climate scenarios to better understand the overall causal processes and make accurate projections (?).

13.3 3. Data Sources and Methods Used to Investigate Climate Change Impact

To enable evidence-based policies to protect health from climate change, the WHO has identified five research priorities: (i) Assessing risks with data-based studies and, in a second step, projections and modelling, (ii) identifying the most effective interventions, (iii) guiding health-promoting mitigation and adaption decisions in other sectors, (iv) improving decision-support and (v) estimating costs (?). While this research agenda has advanced in relation to some health outcomes, research on MH and CC is almost exclusively concerned with the first step of the first priority (?). Methods used to describe future MH risks vary, but most quantitative studies can be grouped into a few broad categories relating to the data and measurement instruments, the study design and the statistical analysis applied (?). The two main data sources are either official statistics or (representative) surveys. Official statistics such as mortality data or hospitalisation rates are low in cost for the researcher and have the advantage of being more objective than self-assessment questionnaires or interviews. But there are some considerable downsides. Some health outcomes, e. g. suicide rates, are often underreported, especially in countries where there is high stigmatisation. Furthermore, proxy variables sometimes have to replace variables of interest. For example, when examining the impact of drought on farmer suicide rates, occupation is not captured in the mortality statistics and one is limited to analysing rural males aged 30-60 years, assuming that they are working in agriculture (?). Surveys do not have this downside, as one can collect variables of interest and potential confounders directly, leading to more fine-grained information (?). However, survey data is more prone to self-selection bias and dropouts, particularly in longitudinal designs (?). In addition, measuring instruments can have a low validity and data collection is costly in terms of both time and monetary resources. Which data source is chosen is largely dependent on the research question and study design. The most common study designs are cross-sectional, followed by case-studies. Little studies attempt to model outcomes (?). The specific statistical analysis depends on the data source and the study design. Classic linear regression (in case of a continuous outcome, e. g. general wellbeing) and logit models (in case of binary outcome, e. g. has a mental illness – does not have a mental illness) are most often used, but also more flexible models, allowing for non-linear effects of climatic covariates such as temperature or precipitation, are possible. For a more detailed review on studies about environmental exposure and MH outcomes, see Charlson et al. (2021). It can be summarized that research concerning the impact of CC on MH is less advanced than research on other health outcomes and mostly restricted to assessing risks. A methodological difficulty inherent to research questions regarding MH is firstly the measurement of MH outcomes, relying on psychometric instruments or diagnostic interviews, and secondly the large number of potential unmeasured confounders, especially when working with official statistics. An important motivation for researching the interactions between CC and health is that, based on a good understanding of causality and reliable projections, evidence-based measures can contribute to the adaptation of health care systems and prevent the most adverse outcomes. However, in the most frequently used cross-sectional studies on CC and MH, causal interpretations must be carefully scrutinised, particularly when dealing with complex causal processes. Different designs such as longitudinal or interventionist designs, a shift towards the other WHO research priorities ii)-v) and a greater focus on non-western regions and populations may contribute to a more reliable, causal understanding.

13.4 4. The Impact of Hurricanes on Posttraumatic Stress Disorder (PTSD)

To illustrate how the above theoretical principals are operationalised in practice, an exemplary study and its results will be described in the following. The paper “Trends in Mental Illness and Suicidality after Hurricane Katrina” (?) is concerned with the impact of one of the most destructive natural catastrophes in the history of the United States (?). To investigate the effect on MH, 815 pre-hurricane residents of the most affected areas were interviewed 5-8 months after the event (baseline) and again one year later (follow-up). In these interviews, data on hurricane-related stressors, demographic variables (age, gender, ethnicity, household income, education, etc.) and different MH conditions (anxiety and mood disorders, PTSD, suicidality, severity of mental illness) were collected. As the first screening interview found that individuals with higher hurricane-related stress levels chose to not participate in the study, weights were applied to the baseline hurricane-related stress variable to account for this selection bias. Additionally, the authors differentiated between individuals from the New Orleans Metropolitan area (metro subsample) and individuals from other areas (non-metro subsample). Multiple logistic regression models of the form

$$\pi_i = P(Y_i = \text{mental illness} \mid \eta_i) = h(\eta_i) = \frac{\exp(\eta_i)}{1 + \exp(\eta_i)}$$

with linear predictors

$$\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}$$

were used to investigate the different outcomes of mental illnesses. In the following, only the results of the outcome of PTSD shall be outlined. The overall prevalence of PTSD was high in both subsamples and on both measurement points (14.9% on baseline, 20.9% on follow-up for the total sample), compared to the US-wide prevalence of 6% (?). A two-tailed, within-responded paired t-test showed that prevalence increased significantly over time. Significance was driven by increase of prevalence in the non-metro subsample (11.8 vs 20.0%, $p < 0.001$), while the prevalence of the metro-subsample stagnated over time (25.9 vs 24.1%, $p = 0.37$). Recovery rate in the follow-up was low as 66.4% of baseline cases continued to have PTSD, 16.9% classified for another diagnosis and only 16.7% recovered. In the logistic regression model, only two of the demographic variables, participants age and family income, had a significant influence on PTSD outcome. The authors hypothesised that the significant increase in prevalence might be explained by an increase in hurricane-related stress. However, this hypothesis was not confirmed by the data. Regarding hurricane-related stress as outcome variable, one can find that stress decreased significantly over time in the whole sample (91.7 vs 57.5%, $p = 0.001$), but in the follow-up, stress was significantly higher in the metro subsample than the non-metro subsample (78.3 vs 51.7%, $p < 0.001$). This contrasts with the increase of PTSD in the non-metro subsample. One can conclude that higher levels of residual hurricane-related stress cannot explain the increase of PTSD prevalence over time in the non-metro subsample. In turn, when taking hurricane-related stress of the follow-up into the model as a predictor variable, one can find that it is indeed significant for both subsamples. In this final model, stress exposure was a key predictor which explained substantial variation in PTSD. The results can be seen in Table 1. Remarkable are the odds-ratios of 20.3 for severe and 12.8 for serious stress compared to individuals with no hurricane-related stress, after controlling for socio-demographic variables and diagnosis of PTSD in the baseline interview. The authors argue that, if interpreted causally, the population-attributable risk proportion of PTSD due to hurricane-related stress is 31.9%, meaning that this proportion would be expected to remit if all hurricane-related stress would be resolved. Although the causality assumption may not be fulfilled, this statement is interesting as it could be a methodological approach to estimate additional PTSD diagnoses due to natural disasters.

Note. CI, confidence interval; OR, odds ratio; PTSD, posttraumatic stress disorder. Effects estimated with multiple logistic regression model controlling for sociodemographic variables, baseline value of PTSD and subsample. Reference category were participants with no hurricane-related stress. Significance level was 0.05.

The study has four mayor limitations: First, in addition to the mentioned selection bias after the first

TABLE 13.1: Table 1. Effects of hurricane-related stress on PTSD outcome as shown in Kessler et al. (2008).

Severity	OR	CI
Severe	20.3*	(4.9–84.6)
Serious	12.8*	(3.0–53.7)
Moderate	4.4*	(1.2–16.1)
Mild	3.5	(1.0–12.6)
None	1.0	
χ^2 (\$p\$-value)	21.0	(< .001)

screening, the dropout rate between measurement points was 21.9%. Although a weight was used to account for the selection bias, it is possible that other, non-collected variables might influence participation or dropout. Second, only screening scales instead of diagnostic interviews were used for diagnosis. Hurricane-related stress was measured using self-assessment questionnaires, which is both subjective and retrospective and possibly biased by current emotional functioning. Third, although interpreted causally, hurricane-related stress and PTSD might be influenced by unmeasured common causes, influencing the observed association. Lastly, there is no true pre-hurricane baseline for the sample, and preexisting mental illness was not recorded in the survey. The results of the presented study are partially contradicting similar studies which found a decline in PTSD over time (?). The authors hypothesise that the increase in prevalence in the non-metro subsample may be caused by greater media attention and resources being allocated to the New Orleans metropolitan region, leading to a sense of neglect and slower reconstruction of infrastructure in other areas. It is also remarked that other stressors only indirectly linked to the hurricane, such as economic stress or displacement, might influence MH outcomes. This is an interesting point of discussion, as it shows that the IPCC categorization of climate change exposures impacting MH is not as simple as outlined before. Events that fall into the category of direct exposure, such as hurricanes, might have prolonged effects on mental health outcomes, e. g. via economic hardship. This in turn makes it very difficult to attribute MH outcomes to specific CC events, as effects can be observed directly after the event, but can also have a prolonged onset. The potential causal pathways of short-term and long-term stressors on PTSD are depicted in Figure 2.

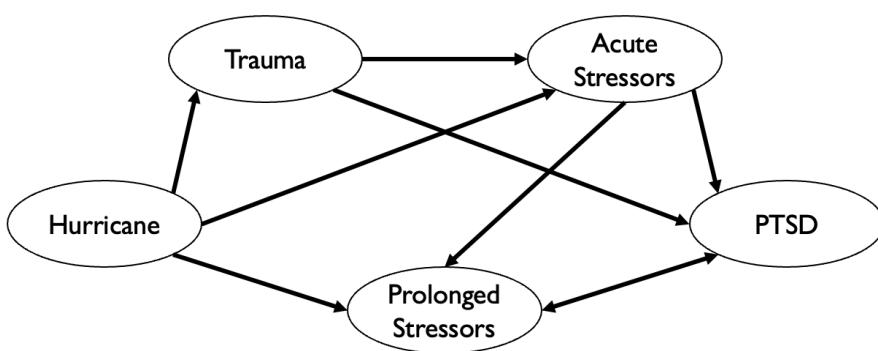


Figure 2. Potential direct and indirect causal pathways through which hurricanes might influence the development of posttraumatic stress disorder (PTSD).

To conclude, the study shows that disaster-related stress does play a significant role in PTSD outcome. Since adverse effects are only weakly related to socio-demographic variables, efforts for sufficient treatment must extend to the entire population. High residual hurricane-related stress after two years and its strong relation to PTSD emphasizes the need for evidence-based measures to address residual stress in multiple ways, ranging from rebuilding infrastructure to providing counselling and financial support. Although the study itself does not link the results to CC, it would be possible to draw a connection between this single, natural disaster and future CC impact on MH. According to the IPCC (2023), hurricanes are projected to

increase in frequency and severity. If, as claimed in the study, hurricane-related stress plays a causal role in the development of PTSD, this increase in hurricanes would have a considerable effect on the number of individuals suffering from PTSD and would need to be taken into account for future health care adaption plans.

13.5 5. Adaption and Mitigation Strategies

Although the IPCC classifies the adverse impact of CC on MH to be of very high confidence, little attention has been directed to mitigation and adaption strategies for mental health and treatment systems (?). In Berry et al. (2018), a possible approach to deal with the high complexity of the matter is outlined. They criticise that today's policies and epidemiological research is focused mainly on individual outcomes, proximate causes and measurements on short timescales. To address the impact of CC on MH more holistically, they propose "systems thinking", which they define as a set of skills used to examine interacting factors that produce outcomes, predict their behaviour and design interventions. More precisely, in research they demand a mixture of CC projections and experimental designs to understand causal influences. In policy, behaviour change should be accomplished through social policy and collective actions instead of aiming at individuals. In direct actions and interventions, they demand a shift from the single-hazard perspective of e. g. extreme weather events to a strategic long-term planning accounting for an increase in those extreme events (?). An example for system thinking would be the protection of infrastructure to benefit MH. This sounds far-fetched, but considering that the cost of damage to critical infrastructure is estimated to increase tenfold (?), and that these costs divert resources from public health, put pressure on the functioning of society and take up individual resources – all factors that have a negative impact on MH – infrastructure protection could be a promotional strategy to protect MH from the impact of CC. The causal diagram of this process is depicted in Figure 3.

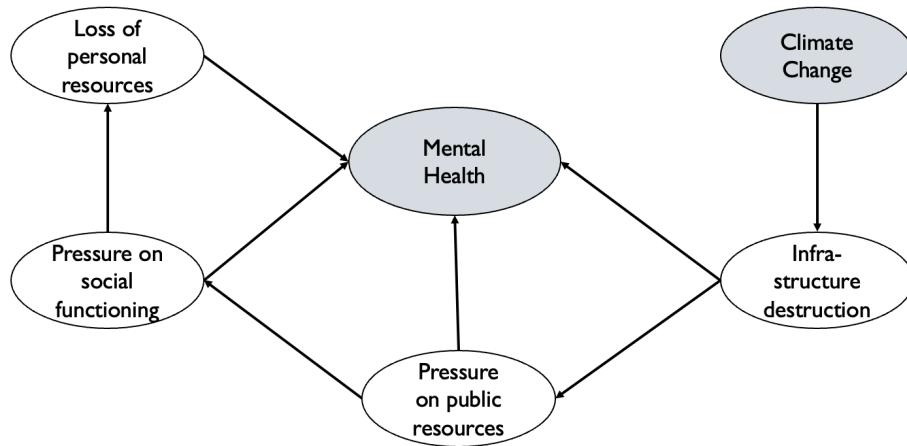


Figure 3. Top-level process diagramm depicting linkages between climate change, infratructure and mental health, as shown in Berry et al. (2018).

The proposal by Berry et al. (2018) is a broad framework that may be useful to approach most of the problems discussed in this chapter. However, the individual components are yet to be tested in future research and policy designs. On the more practical side, Hayes et al. (2018) and Newnham et al. (2020) proposed several practical measures that could protect MH from direct and indirect CC impact. For acute response to disaster, first-aid responders and nurses need to be trained in basic psychological first aid. It has also proven beneficial for subsequent MH outcomes to prepare people living in vulnerable areas (e.g. near rivers that will flood

more frequently in the future) for the possibility that the event might occur (?). Intermediate adaption could include capacity building for therapy and counselling, education and programmes to reduce stigma as well as community-building and preparation training for extreme weather-related events (?). Long-term prevention and mitigation require governments to include MH in their strategic health planning and the development of tools to predict risks, costs and needed resources. Ultimately, a swift transition towards a carbon-neutral economy would be the most cost-effective and efficient way of protecting MH from the impacts of CC.

13.6 6. Conclusion

Climate change will have considerable effects on public health, including MH and wellbeing. As the latter has received little attention in research and politics, scientifically sound methods to quantify the impact of CC on MH are still being developed and research has yet not advanced from assessing the risks to estimating the costs or designing adaption strategies. Considering the importance of MH and wellbeing for individual life as well as society and economic production, more research and policy action is needed to understand this complex topic and prevent or at least attenuate the adverse impact of a changing climate on mental health.



14

References



15

Open issue

Author: Author

Supervisor: Helmut Kuechenhoff

Degree: Bachelor Master

Further aspects of climate change

Economic, political aspects, communication

? ?



16

Acknowledgements

The most important contributions are from the students themselves. The success of such projects highly depends on the students. And this book is a success, so thanks a lot to all the authors! The other important role is the supervisor. Thanks to all the supervisors who participated! Special thanks to Helmut Küchenhoff¹ who enabled us to conduct the seminar in such an experimental way, supported us and gave valuable feedback for the seminar structure. Thanks a lot as well to the entire Department of Statistics² and the LMU Munich³ for the infrastructure.

The authors of this work take full responsibilities for its content.

¹<https://www.stablab.stat.uni-muenchen.de/personen/leitung/kuechenhoff1/index.html>

²<https://www.statistik.uni-muenchen.de/>

³<http://www.en.uni-muenchen.de/index.html>



Bibliography

- Armstrong, B. (2019). The role of humidity in associations of high temperature with mortality: A multi-country, multicity study. *Environmental Health Perspectives*, 127(9):097007.
- Baldwin, J. W., Benmarhnia, T., Ebi, K. L., Jay, O., Lutsko, N. J., and Vanos, J. K. (2023). Humidity's role in heat-related health outcomes: a heated debate. *Environmental Health Perspectives*, 131(5):055001.
- Barreca, A. I. (2012). Climate change, humidity, and mortality in the united states. *Journal of Environmental Economics and Management*, 63(1):19–34.
- Bours, M. J. L. (2021). Tutorial: A nontechnical explanation of the counterfactual definition of effect modification and interaction. *J Clin Epidemiol*, 134:113–124.
- Budd, G. M. (2008). Wet-bulb globe temperature (wbgt)—its history and its limitations. *Journal of Science and Medicine in Sport*, 11(1):20–32.
- Buzan, J. R. and Huber, M. (2020). Moist heat stress on a hotter earth. *Annual Review of Earth and Planetary Sciences*, 48(1):623–655.
- D., W. (2013). The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. *American Journal of Epidemiology*, 177(4):292–298.
- Davis, R. (2016). Humidity: A review and primer on atmospheric moisture and human health. *Environmental Research*, 144:106–116.
- K.L.Ebi (2021). Hot weather and heat extremes: health risks. *The Lancet*, 398(10301):698–708.
- Raymond, C. (2024). *The Climate Data Guide: HadISD: Sub-daily, non-interpolated weather station data*. R Foundation for Statistical Computing.
- Schwartz, J. (2004). Hospital admissions for heart disease the effects of temperature and humidity. *Epidemiology*, 15(6).
- VanderWeele, T. J. (2009). On the distinction between interaction and effect modification. *Epidemiology*, 20(6):863–871.

