Studying extreme heatwaves using novel theoretical approaches

Valeria Mascolo

THESE

en vue de l'obtention du grade de Docteur, délivré par

l'ECOLE NORMALE SUPERIEURE DE LYON

Ecole Doctorale N°52

Physique et astrophysique (PHAST)

Discipline: **Physique**

Soutenue publiquement le 19/09/2024, par :

Valeria Mascolo

Studying extreme heatwaves using novel theoretical approaches

---

Devant le jury composé de :

OPITZ, Thomas	Directeur de recherche, INRAE-BioSP, Rapporteur
RIVIERE, Gwendal	Directeur de recherche, Ecole Normale Supérieure, Rapporteur
FISCHER, Aurelie	Maître de conférences-HDR, Université Paris Cité, Examinatrice
JEZEQUEL, Aglae	Personnalité scientifique, Ecole Normale Supérieure, Examinatrice
VENAILLE, Antoine	Chargé de recherche-HDR, ENS de Lyon, Examinateur
BOUCHET, Freddy	Directeur de recherche, ENS de Lyon, Directeur de thèse

STUDYING EXTREME HEATWAVES USING NOVEL THEORETICAL APPROACHES

VALERIA MASCOLO

Supervisor: Freddy Bouchet

Laboratoire de Physique
Ecole Normale Superieure de Lyon

To our amazing planet Earth.

To that little girl who would have never believed to get this far.

Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.

- Marie Skłodowska-Curie

Au milieu de l'hiver, j'apprenais enfin qu'il y avait en moi un été invincible.

- Albert Camus

Prendete la vita con leggerezza. Che leggerezza non è superficialità, ma planare sulle cose dall'alto, non avere macigni sul cuore.

- Italo Calvino

ABSTRACT

Studying extreme weather events, such as heatwaves, is a fascinating growing field and poses several methodological and computational challenges.

Understanding the physical mechanisms which drive heatwaves and being able to forecast them is of pivotal importance in the context of mitigation and adaptation to the current climate crisis. A new theoretical framework which addresses both tasks is developed in this thesis and applied to mid-latitudes heatwaves. This framework, called the Gaussian Approximation, relies on simple but meaningful assumptions on the statistics of weather fields relevant for heatwaves. It proves to capture well the salient features of atmospheric conditions predominant during those events and to have competitive or even better predictive skills than other more complex weather prediction models.

Another major challenge when dealing with extreme heatwaves, and extreme events in general, is that they suffer from an intrinsic sampling problem, due to their rareness. Historical records are too short and direct simulations are simply too expensive for having a good estimation of these events. Sampling algorithms, such as the rare events algorithms, tackle this task resulting in both improving the quality and the efficiency of extreme events simulation. In this thesis, this class of algorithm has been applied for the study of double jet stream state, an interesting rare phenomenon which is linked to heatwave occurrence at mid-latitudes. The study consists in coupling a climate model (CESM) with a properly chosen score function, which describes the salient features of the phenomenon, in order to sample extreme and unprecedented events.

RÉSUMÉ

L'étude des phénomènes météorologiques extrêmes, tels que les vagues de chaleur, est un domaine fascinant en pleine expansion et pose plusieurs défis méthodologiques et en termes de calcul.

Comprendre les mécanismes physiques à l'origine des vagues de chaleur et être capable de les prévoir est d'une importance capitale dans le contexte de l'atténuation et de l'adaptation à la crise climatique actuelle. Un nouveau cadre théorique qui aborde ces deux tâches est développé dans cette thèse et appliqué aux vagues de chaleur des latitudes moyennes. Ce cadre, appelé Approximation Gaussienne, repose sur des hypothèses simples mais significatives concernant les statistiques des champs météorologiques pertinents pour les vagues de chaleur. Il s'avère qu'il capture bien les caractéristiques principales des conditions atmosphériques prédominantes pendant ces événements et qu'il a des capacités prédictives compétitives, voire meilleures, que d'autres modèles de prévision météorologique plus complexes.

Un autre défi majeur dans le traitement des vagues de chaleur extrême, et des événements extrêmes en général, est qu'ils souffrent d'un problème intrinsèque d'échantillonnage, en raison de leur rareté. Les données historiques sont trop courtes et les simulations directes sont tout simplement trop coûteuses pour bien caractériser ces événements. Les algorithmes d'événements rares s'attaquent à cette tâche, ce qui permet d'améliorer à la fois la qualité et l'efficacité de la simulation des événements extrêmes. Dans cette thèse, cette classe d'algorithmes a été appliquée à l'étude de l'état du double jet stream, un phénomène rare et intéressant lié à l'apparition de vagues de chaleur aux latitudes moyennes. L'étude consiste à coupler un modèle climatique (CESM) avec une fonction de score bien choisie, qui décrit les caractéristiques saillantes du phénomène, afin d'échantillonner des événements extrêmes et sans précédent.

Etenim, iudices, cum omnibus virtutibus me adfectum esse cupio, tum nihil est quod malim quam me et esse gratum et videri.
Haec enim est una virtus non solum maxima sed etiam mater virtutum omnium reliquarum.

In truth, O judges, while I wish to be adorned with every virtue, yet there is nothing which I can esteem more highly than being and appearing grateful.
For this one virtue is not only the greatest, but is also the parent of all the other virtues.

— Marcus Tullius Cicero

ACKNOWLEDGMENTS

PERSONAL STATEMENT

Gratitude is a tricky emotion for the person I am.

Personally, I think of gratitude as an attitude towards life. To me, it ultimately deals with the ability to recognize and accept that beauty, whatever its definition, can arrive and stay in your life. If someone asked me to summarize this PhD journey in one word, "**gratitude**" would be the chosen one.

These past three years have been much more than an incredibly enriching journey of my scientific background. They have been the years that shaped a new version of myself, changing me profoundly, to the point that I could barely recognize the person who started this journey (nothing dramatic here).

Studying climate extremes and being increasingly exposed to the current climate crisis have given me an alternative lens through which to view the reality of our world. Reading about the current and projected effects of climate change has made me particularly sensitive to and touched by a huge variety of subjects, ranging from social justice, geopolitics, economy, agriculture, and has led me to rethink daily personal choices, from my diet regime and ways of travelling to my career plans.

Ultimately, this journey has given me the opportunity to consciously choose a place in this society, rebuilding my personal guiding values with which I now proudly feel aligned. In recognizing all of this, I could not avoid acknowledging the privileged person I am, having had the immense chance of accessing a high level of education and the possibility of freely expressing my political, cultural, and personal visions as a human being - among the many privileges I have recognized in myself. "Grateful" is the only word that fully summarizes how I feel.

Writing this feeling down comes as an *urgent* need, where the urgency is comparable to the burning need to share happiness to feel it real. It also serves as a free space to stimulate self-reflection for the reader who kindly read up to here.

As soon as I started writing the acknowledgments, I realised that it would have been a hard task. Do not get me wrong, the difficult part was choosing a language everyone involved could understand. Spoiler: it was impossible. Thus, ultimately, I did not choose a single language but three, and I will switch among them at some point.

I warmly thank all the members of my PhD committee, and particularly the rapporteurs for their comments and suggestions on this manuscript. It has been both a pleasure and an honour to discuss with you all during my PhD defence. Your insightful questions and comments spanned a wide range of topics, from the core scientific aspects of my research to intellectually stimulating philosophical inquiries: it has been enriching.

I would like to thank my supervisor Freddy for accepting me in his team. Without that, this manuscript and the person I am now would have probably not existed, thus I am grateful for the opportunity you gave me. Your passion, dedication and attitude for research and science in general is of great inspiration as well as your humanity and patience.

A special thanks goes to my collaborators, Corentin, Francesco, and Nili. Our discussions and your constant support have been fundamental to many of the results I present here, as well as to my growth both as a researcher and as a person.

Another great thanks goes to my group members:

- Clément: working with you has been a pleasure, as your enthusiasm and your attention to detail have motivated me at different stages of this journey. I am grateful for the patience and the availability you had every time I asked you for clarifications on our common work and for informal discussions about PhD life and doubts. I really wish you all the best for your future!

- Alessandro: co-authoring a paper with you and writing our manuscripts simultaneously have given me the opportunity of knowing a brilliant scientist and a profound person. We have sailed together in a wind-physics school, tried great restaurants in different European cities, shared the ups and the downs of our doctoral journey in a truly Italian style, building memories I cannot forget, which are hopefully the first of a long series, and creating what I define a true friendship. I am sure that we will hear about you in the future!

- Bastien: summarizing what you have been to me over the last few years is difficult because I can hardly find the words. That shy, truly humble, and gentle guy with a soft voice whom I met three years ago is now my best friend, and I couldn't be more proud to have such a person in my life. Someone who has always calmly listened to me, hugged and supported me in hard times, shared the joy of the best moments, and with whom I know I can share my deepest fears, because I will not be judged but only welcomed. Our list of memories is long, but short compared to the

ones I hope we will continue to create, together with your lovely partner Romane, who is also a true friend to me;

and finally to Amaury and Brice for their incredible sense of humour and the interesting discussions on climate-related topic.

I would like to thank Bastien, Clément, Corentin and Freddy for reading different parts of this manuscript at different stages, giving me the possibility of ameliorating it.

As part of the EDIPI community, I would like to thank Davide Faranda and Gabriele Messori for their constant availability and support. Moreover, I thank from the bottom of my heart the the EDIPI ESRs: the scientific community is lucky to have such brilliant scientists and great people like you on board. Among you, a special thanks for their constant support, hugs, dramas, travel plans, hiking, museums and art tips and complaints, goes to Bianca, Aleksa, Emma and Mireia. You have a special place in my heart.

Ringrazio la mia enorme famiglia, ed in particolare i miei genitori e mia sorella, per aver sempre appoggiato le mie scelte, seppur non capendole fino in fondo, e per aver dimostrato la vostra mancanza non con le parole, ma con i vostri occhi stracolmi di felicità ad ogni mio ritorno, ed il vostro orgoglio nel vedermi costruire passo passo la vita che desidero e la persona che sono.

Ringrazio mia nonna, ovunque lei sia. Avrei voluto tanto stringerle la mano anche per questo traguardo, sentire che è orgogliosa di me e che mi voleva bene, nonostante non riuscisse mai a dirmelo.

Ringrazio i miei amici storici, Claudia, Gennaro e Gioia, per esserci ancora. Non avrei mai creduto che la nostra amicizia potesse resistere a così tanta distanza e sono grata perché ovunque io sia stata, voi ci siete stati sempre e comunque.

Ringrazio Vlad George, per essermi stato vicino durante i momenti più difficili di questa tesi, ed in generale di questi ultimi quattro anni. Sono grata alla vita per i nostri ricordi e per quello che abbiamo condiviso e che siamo diventati insieme.

Ad Alessandro, Bastien, Claudia, Gennaro, Gioia e Vlad George, voglio dedicare queste parole della scrittrice Michela Murgia, che, come sempre, riassume ciò che io arduamente sarei riuscita ad esprimere: 'Le amicizie pluridecennali sono un bene raro che va molto mantenuto, perché sono fonte dell'unica cosa che non si può ripetere: il tempo. Sono quelli che custodiscono il ricordo della ragazza o del ragazzo che eri, che conoscono la fatica che hai fatto per essere la donna o l'uomo che sei, che ricordano l'entusiasmo che avevi e quello che é rimasto, gli errori da cui ti sei salvata/o e quelli da cui ti hanno salvato loro. Non sono solo amici: sono testimoni e complici.'

Merci infiniment au Laboratoire de Physique de l'École normale supérieure de Lyon et tout particulièrement à Fatiha, Erika et Nadine pour m'avoir accompagné avec patience, bienveillance et toujours un sourire au fil des (dizaines) des missions et une série des documents incomprensibles à remplir. Je n'y serais jamais arrivé sans vous!

Enfin et surtout, merci à la France. Au fil des jours j'ai appris à t'aimer du fond de mon coeur, et maintenant que je suis potentiellement à un pas de te quitter, je me sens malheureuse. J'ai du mal à croire que tu m'as autant touchée, et c'est pourtant exactement ce que je ressens. Merci donc, d'être un si beau pays, mais surtout une très belle idée. Il faut que je te remercie également de m'avoir forcé à apprendre ta langue, que je ne maitrise pas encore tout à fait, ou du moins, pas encore au niveau que je voudrais. Affaire à suivre.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AMV Atlantic Multidecadal Variability

CESM Community Earth System Model

CMIP Coupled Model Intercomparison Project

GKLT Giardina-Kurchan-Lecomte-Tailleur

IPCC Intergovernmental Panel on Climate Change

OU Ornstein–Ulhenbeck

PlaSim Planet Simulator

PDF Probability Density Function

SCGF Scaled Cumulant Generating Function

1 INTRODUCTION

A rare event is, by definition, an event that occurs with a low probability. An extreme event is a particular case of a rare event, characterized by its severe intensity and potential high impacts. Rare but impactful events might arise in a plethora of fields, ranging from economics, with extreme fluctuations in global markets, to epidemics, with the spreading of diseases impacting millions of people, to planetary sciences, with the fascinating merging of jets in Jupiter's atmosphere.

In this thesis, I focus on rare and extreme events in the climate system, with a particular emphasis on extreme heatwaves. In the context of the current climate crisis, this class of events has gathered much interest given its obvious societal, environmental, and financial impacts (fig. 1.1). Exposure to prolonged heat causes psychological and physical stress currently classified as heat illness, covering a spectrum of disorders due to increased body temperature (Gauer and Meyers, 2019). More than 70000 people died as a result of the 2003 heatwave (Robine et al., 2008) and socioeconomic factors such as poverty and isolation combined with age and illness augmented the death toll in specific sectors of French society during the same event (Poumadère et al., 2005). During the heatwaves of 2003, 2010, 2015, and 2018 some southern European areas have lost around 1% of their gross domestic product (GDP) and this is expected to increase almost a factor of five by 2060 compared to the historical damages experienced over the period 1981–2010 (García-León et al., 2021). Unfortunately, these are only a few examples among the massive amount of tremendous recorded impacts associated with heatwaves. Record-shattering heatwave events are observed with an abnormal high frequency in several parts of the globe with deleterious effects for the concerned ecosystems. In this class of events, we can certainly include the Western European heatwave of 2003 (García-Herrera et al.,

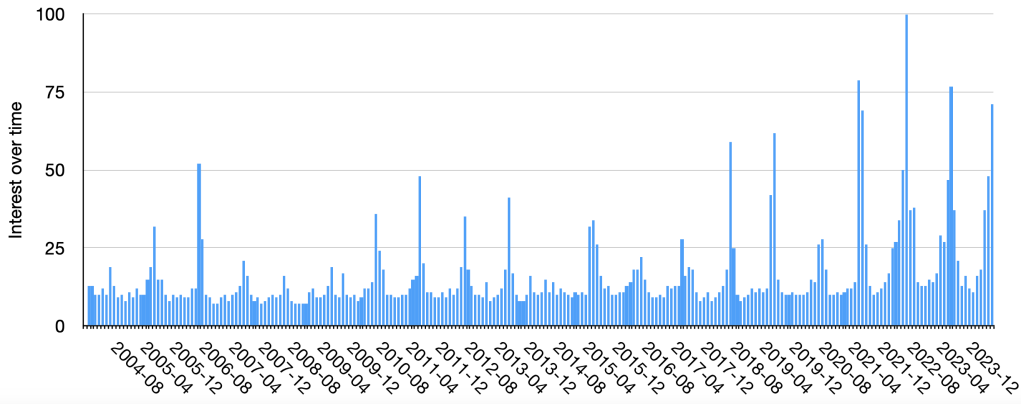


Figure 1.1: Interest in time of the keyword 'heat wave' worldwide. The interest is measured as a proportion of all searches on all topics on Google at that time and location for that keyword. Data taken from Google Trends.

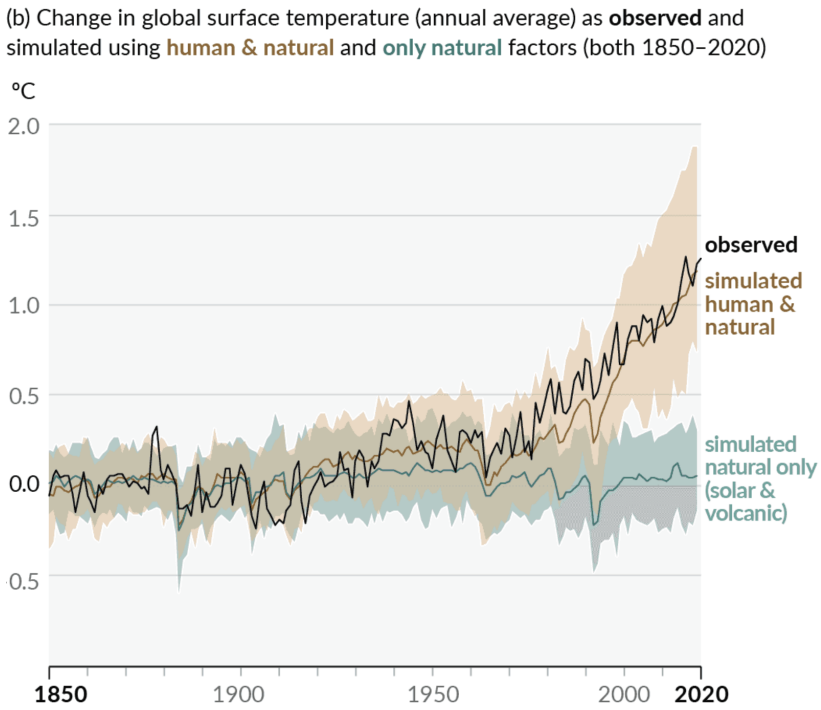


Figure 1.2: Change in global surface temperature over the past 170 years (black line) relative to 1850–1900 and annually averaged, compared to Coupled Model Intercomparison Project (CMIP)6 climate model simulations of the temperature response to both human and natural drivers (brown) and to only natural drivers (solar and volcanic activity, green). Solid colored lines show the multi-model average and colored shades show the *very likely* range of simulations. Figure adapted from (Masson-Delmotte et al., 2023).

2010), the Russian one of 2010 (Barriopedro et al., 2011) and the Western America one of early June in 2021 (White et al., 2023), with a maximal temperature of 49.6 °C reached in Lytton, the highest ever recorded in the area.

Despite being a chaotic system, where random and possibly huge fluctuations arise simply due to the nature of the system itself, the Earth is experiencing an increasing number of climate extremes which could not be explained without considering the influence of human carbon emissions (Seneviratne et al., 2021), fig. 1.2. Given the human-driven warming due to increase concentrations of CO_2 and the current policies, heatwaves (and many other extremes) are projected to increase even further in the future (Seneviratne et al., 2021).

Studying heatwaves turns out to be a major and urgent challenge. A non-exhaustive list of scientific questions associated with their study includes:

- estimate their probability to occur;
- understand the dynamics associated with them;
- forecast them at different time scales;
- understand their thermodynamical and dynamical changes associated with climate change;
- attribute them to climate change.

Despite being interesting and active research fields, in this manuscript I do not concentrate on the latest two points presented above, but I rather focus on the first three, with a particular stress for the first one.

A bottleneck one faces in addressing the above questions is represented by the scarcity of recorded data. In order to characterize and forecast an extreme event, a sufficient amount of realizations of it should be observed to understand its temporal and spatial evolution and correlation. However, due to their rareness, they suffer from a *sampling* problem. Despite the Earth being the most monitored system, the historical records are generally not sufficiently long to have significant samples of rare events in climate simply because the most reliable data only start from the satellite era. As a consequence, climate studies heavily rely on climate model simulations. These huge machines, with different levels of complexity and subsequent realistic representation of the Earth's components, such as atmosphere, land, ocean and ice, simulate the climate by solving fundamental laws of physics constrained with conservation laws and use parametrizations for processes acting at sub-grid scales. One of the drawbacks of those simulations, called *direct simulations*, when applied to the field of extremes study is the enormous time and computation resources needed to gather a statistically significant sample of rare events, such as extreme heatwaves. As an example, in part ii we compared the effect of multi-decadal internal variability of the sea surface temperature of the North Atlantic and of the soil moisture on summer European heatwaves. Using simulation data, we show with the aid of *return time maps*, maps conditioned on the time needed to observe an event with of a certain amplitude, that significant areas are largely limited because of lack of data. However, both factors are known to play a role in favoring and/or maintaining heatwave conditions over Europe, exacerbated by the fact that their effect has the same magnitude as the current warming trend due to climate change.

The question is now: how to sample or generate more extreme events using a lower or limited computational cost than a direct simulation? Posed in these terms, it might seem to be quite an absurd question. Luckily, *de facto*, it is not. In this thesis, we present two methodologies to tackle this task.

The first one is presented in part iii and it is called the Gaussian approximation, because of the hypothesized statistics between the dynamical fields that characterize a heatwave at mid-latitudes and the definition of the heatwave itself. With this methodology, we do not generate more extreme events, but we rather use the information coming from the full dataset and not only from the heatwave subset, which is rather small. This methodology results in being surprisingly skillful in the representation of the average dynamical conditions that led to a heatwave, called *composite maps*. Moreover, when used as a forecasting tool, this approximation is competitive with off-the-shelf neural networks, with a net improvement in the computational time and a better performance on a short reanalysis dataset (a dataset created via a data assimilation procedure between historical records and observational data). Interestingly, the results suggest quasi-stationary Rossby waves and low soil moisture as precursors to extreme heatwaves over France.

The second methodology is a rare event algorithm meant to generate more extreme events based on the theory of large deviations. This theory has a long

history behind in the field of statistical mechanics and, in a nutshell, it deals with the asymptotic estimation of the probability of unlikely events by sums of random variables. In the context of this thesis, I briefly present the fundamental concepts of the theory meant to create a link with the rare event algorithm used in the following, called Giardina-Kurchan-Lecomte-Tailleur (GKLT). Indeed, this algorithm was initially tailored to sample the large deviation function (Giardinà, Kurchan, and Peliti, 2006; Lecomte and Tailleur, 2007), to have ultimately access to the (low) probability of large fluctuations. The algorithm is part of the *cloning algorithm* class (Del Moral and Garnier, 2005; Hidalgo, 2018) and is referred equivalently in the literature as diffusion Monte Carlo (Anderson, 1975). The methodology is based on the modification of the evolution of the system's dynamics to favor large fluctuations to arise more frequently. This is achieved via a selection procedure, which replicates or kills multiple copies or *clones* of the system according to a score function that estimates the tendency of that copy to achieve a large fluctuation. The application of this specific algorithm for climate studies has been recently proposed by (Ragone, Wouters, and Bouchet, 2018) for the sampling of extremely warm summers in Europe. It revealed successful and, since then, it has been applied to a broad range of climate extremes. Applications include precipitation (Wouters, Schiemann, and Shaffrey, 2023), extreme fluctuations of the energy power system (Cozian, 2023), collapsing of the Atlantic Meridional Overturning Circulation (Cini et al., 2024), turbulent flows (Lestang, Bouchet, and Lévêque, 2020), the melting of the Arctic sea ice (Sauer et al., 2024) and finally, Indian extreme warm summers (Le Priol, Monteiro, and Bouchet, 2024). In the context of this thesis, the algorithm is applied to the sampling of extreme double jet summers over the Eurasian continent. This configuration has gathered much attention because of its link to high-impact heatwaves in Europe (Rousi et al., 2022). With the aid of the GKLT algorithm, we arrive to sample events that have a *return time* of $10^4 - 10^5$ years, with only 10^3 simulation years. The return time can be thought of in terms of a probability of observing this event in 1 out of 10^5 years on average. This order of magnitude is absolutely out of range for direct simulations.

The manuscript is organized as follows. Part i is an introductory chapter where I first present briefly the state-of-the-art of our understanding of the dynamics and thermodynamics of heatwaves, with a focus on Eurasia, the geographical zone of interest. I then present some mathematical objects used extensively in the rest of the manuscript to analyze or predict heatwaves and subsequently, I introduce the theory of Large deviations to create a link with the rare events algorithm presented in the latest section of the same chapter.

In part ii I present a study that compares the effects of two slow drivers of heatwaves in Europe, namely the AMV and the soil moisture. In part iii the Gaussian approximation is presented. With this novel theoretical framework, we focus on the analysis and prediction of heatwaves over France. In the same chapter, a comparison between the prediction results with other established methodologies in the field, such as Convolutional Neural Networks, is presented. Finally, in part iv I use the rare events algorithm presented before to study double jet events over Eurasia, an atmospheric configuration that has been linked to heatwaves in some Northern Hemisphere areas.

Part v is a stand-alone chapter where I present a scientific outreach project I coordinate which stands at the intersection between climate and art. In this project, the aim is to show the effects of climate change on art paintings by means of creative repainting following scientific knowledge taken from the Intergovernmental Panel on Climate Change (IPCC) reports.

In chapter 27 I summarize our findings and I give perspective for future works.

Part I

HEATWAVES AND RARE EVENTS

This chapter is an introductory chapter which has two main purposes: firstly, to present the weather phenomenon of interest, namely heatwaves at mid-latitudes and secondly, to introduce the mathematical and physical tools which are used in the following chapters.

In chapter 2 I present a brief overview of the state-of-the-art of the mechanisms connected to heatwaves at mid-latitudes, focusing on the drivers and the projected changes in response to climate change.

In chapter 3 interesting mathematical objects, such as return time, composite maps and committor function, are presented. They are used in the following to analyze and predict heatwaves. Subsequently, the theory of Large deviations is presented in a nutshell.

Finally, in chapter 4 rare events algorithms are introduced. They aim at improving the sampling of rare events both from the efficiency and the quality points of view. Among them, a cloning algorithm, the Giardina-Kurchan-Lecomte-Tailleur GKLT is presented in detail, as it is used in the following chapters. This algorithm is historically connected with Large deviations theory, which is why it was presented before.

2

HEATWAVES AT MID-LATITUDES

This chapter aims at giving a general overview of what are the main known drivers of heatwaves at mid-latitudes. A choice that has been made in order to keep the discussion as simple as possible is not to give a definition of a heatwave in this chapter. Defining this phenomenon could be tricky and several propositions have been made (Perkins, 2015; Barriopedro et al., 2023). In this thesis, I use two of them, which will be introduced in the following. For the moment, I propose to take as definition the one proposed by the IPCC latest report: '[a heatwave is] *a period of abnormally hot weather, often defined with reference to a relative temperature threshold, lasting from two days to months*' (Seneviratne et al., 2021).

Before focusing on heatwaves, in the first part of this chapter I introduce the key dynamical components that shape the climate over Europe (section 2.1), as they affect a wide range of climate events and extremes (not only heatwaves). Once the main actors are presented, I give a brief overview of the physical basis which drives heatwaves at the mid-latitudes, and in particular in Europe and Eurasia (section 2.2), given that I am interested mainly in those geographical areas in the following.

2.1 THE UNIQUE CLIMATE OF EUROPE

Besides the tropics, European weather and climate are one of the hardest to predict at time scales that are larger than a few days (Woollings, 2010). This difficulty is generally related to the large-scale circulation of the atmosphere and to the Atlantic Ocean, which impacts the climate over several timescales. For instance, the south-west north-est tilt of the jet stream and blocking dynamics are widely known for being underrepresented in current climate models (Woollings, 2010; Woollings et al., 2018), making the prediction of European climate particularly challenging and, at the same time, an appealing research field. In this section, we present some of the key processes, from a dynamical point of view, that shape the European climate. The organisation and the key ideas of this section are deliberately taken from (Woollings, 2010).

2.1.1 *The Jet streams*

The jet streams are fast winds which encircle the globe flowing from west to east. In both Hemispheres, we have two types of jets: the subtropical and the eddy-driven one. Both jets ultimately owe their existence to the meridional contrast in solar heating of the Earth.

The *subtropical jet* results from the transport of momentum/vorticity associated with the upper branch of the Hadley cell. Warm air rises at the equator and moves poleward ascending throughout the whole troposphere. At a cer-

tain latitude, around 30° the cold and denser air sinks, creating a descending branch, and, consequently, warming and moving equatorward again. When air moves poleward, it starts moving eastward, because of the Coriolis effect, creating a wind current at an altitude of around 10 km, in the upper troposphere: the subtropical jet is born.

As its name suggests, the *eddy-driven jet* originates from the transients eddies in the atmosphere which change the vorticity and heat transport, resulting into a net acceleration of westerly winds. This kind of jet extends down though the whole depth of the troposphere. During winter the jets are stronger, given the higher contrast of temperature between the polar region and the mid-latitudes. During summers, both jets are weaker and displaced towards the pole. Generally, both jets are merged into a unique jet. However, there exist some situations where two distinct jets are seen, for instance during winter with the subtropical jet crossing North Africa and the the eddy-driven displaced further north. This last configuration will be investigated in part iv of this manuscript.

2.1.2 *The storm tracks*

European climate is strongly affected by the North Atlantic storm track. A storm track is a term used for indicating the preferred regions for mid-latitudes anticyclones and cyclones anomalies to occur. These synoptic scale systems in those preferred regions owe their existence to the baroclinic instability due to difference in heating of the Earth system. These eddies act to reinforce the slow varying large scale flow, or alternatively: the jet stream and the storm track can be considered self-maintaining. Near the surface the storm tracks are mainly concentrated over the Atlantic basin. However, at high altitude the eddy activity coming from the Pacific Ocean continues over North America, joining the Atlantic basin and dissipating its power towards Europe. The effects on the extreme events linked to the storm track are obviously visible during winter and less during summer (the season we are interested in, in this manuscript).

2.1.3 *Atmospheric Blocking*

Beside the quasi-stationary circulation, the climate at the mid-latitudes is dominated by transient eddies interacting with more or less common circulation patterns. They exhibit preferred spatial and temporal distributions and a change in their frequency, location or amplitude could lead to important impacts on the climate. At middle and high latitudes, an important class of weather patterns is that of atmospheric blocking, or simply blocking. The term *blocking* was coined to refer to a situation where the natural and normal west-to-east flow of cyclone systems is blocked or interrupted. Currently, a comprehensive theory of the dynamical features which constitute the onset, maintenance and decay of blocking is still lacking (Woollings et al., 2018). As a consequence, in the meteorological community there is not a unique definition of what a blocking is. Common features, such as persistence, quasi-stationary and obstruction of the westerlies are shared characteristics.

Several and different weather patterns have been associated to this phenomenon, accompanied by as many definitions to classify it (Barriopedro, García-Herrera, and Trigo, 2010). All of them determine a blocking condition based on a so-called blocking index. There have been two traditional paths to define the field in the blocking index: using the absolute meteorological fields or its departure from regional or zonal climatological mean. In both cases, common variables are the geopotential height at 500 hPa and potential temperature at the dynamical tropopause. A novel index mixing both the absolute and the anomaly fields has recently being proposed. A complete review can be found at (Barriopedro, García-Herrera, and Trigo, 2010). Moreover, added features to the definition include a minimum of spatial extension, quasi-stationary and persistence (typically 4-5 days).

Traditionally a block index might be either a one dimensional or a two dimensional index. In the former case, block highs are defined by requesting several constraints in the meridional gradient around a constant latitude representing the climatological position of the jet stream (Tibaldi and Molteni, 1990; Pelly and Hoskins, 2003). To avoid fixing a reference latitude, others designed a two dimensional index by focusing on a latitudinal band. This turns out to be an interesting methodology for characterizing blocks as a high-low dipole (fig. 2.1 (e)) (Masato, Hoskins, and Woollings, 2013). Other methodologies defined blocking as typical circulation patterns, following statistical approaches (Vautard, 1990).

Some examples of blocking patterns are shown in fig. 2.1. A classic, but still contentious for blocking definition, is the one depicted on panel (a), the Summer ridge. In this configuration, air is rapidly advected from the subtropics northward and it is anomalously anticyclonic with respects to its surroundings. If the Rossby wave has a phase speed of zero, then a stationary condition is reached. The onset of a ridge has been attributed to the interaction of the background flow with the Rossby waves or of Rossby waves of different wavelengths (Woollings et al., 2018). This situation resembles to the fig. 2.1 (b), the Omega block. However, there, there is a clear anticyclonic system shifted poleward, which sits between two cyclones advecting poleward air which is warmer than normal.

Panels (c) to (e) of fig. 2.1 are representative of wave breaking events. In such events the meridional potential vorticity (or geopotential height) gradient is reversed and a dipole structure is generated with an anticyclonic anomaly on the poleward side and a cyclone on the equatorward side (Woollings et al., 2018). This is a key feature of the so-called Rex block, after (Rex, 1950).

Traditionally, climate models have always underestimated the frequency and persistence of blockings. In (Woollings et al., 2018) the authors showed that over the Euro-Atlantic sector CMIP5 models underestimate winter blocking frequency by 30-50% and by around 10-30% during the summer season. The reason has to be traced back to the difficulty of models to fully and faithfully represent the general circulation and in particular of the jet stream variability. As a consequence, forecasting their onset and decay remains a challenge to tackle. The latest class of climate models used in the IPCC reports, CMIP6 models, has been shown to improve the representation of blockings, with regional and seasonal variations (Lupo, 2021; Kautz et al., 2022).

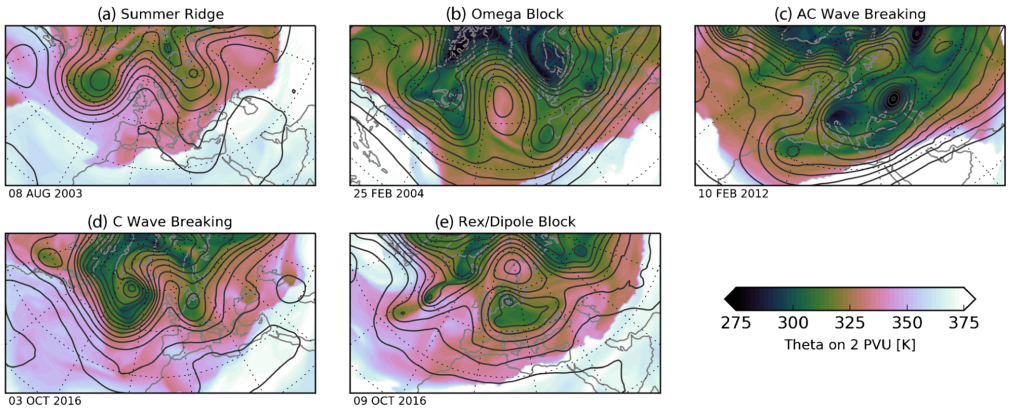


Figure 2.1: Examples North Atlantic blocks. Snapshots of (color shading) potential temperature θ on the dynamical tropopause (PV = 2 PVU) and (contour lines) geopotential height at 500 hPa (contour spacing 60 m) for the dates indicated. Figure taken from (Woollings et al., 2018).

Regardless of the definition, Europe is one of the regions with highest frequency of blocking (Woollings, 2010). In this region, blocking has been linked to heatwaves during summer and cold snaps during the winter season. The impact of this phenomenon on heatwaves is addressed in section 2.2.2.

2.1.4 *Rossby waves*

Rossby waves in the atmosphere are planetary waves associated with the meandering of the extratropical jet stream (Rossby et al., 1939). They result from the conservation of potential vorticity, which is a function of both relative and planetary vorticity, and static stability. Planetary vorticity is due to the rotation and sphericity of the Earth, which increases towards the pole, while relative vorticity comes from the motion of the air relative to the Earth's surface. Because of the meridional gradient of potential vorticity, a parcel of air displacing poleward, in the Northern Hemisphere, to conserve its potential vorticity, has to reduce its relative vorticity by rotating in an anticyclonic way due to the augmentation of its planetary vorticity. This air spinning clockwise drags neighbouring parcels of equatorward. This induces the same mechanism of conservation of potential vorticity but to form cyclones, ultimately leading to a Rossby wave, consecutive regions of anomalous cyclonic and anticyclonic vorticity (White et al., 2022). Large-scale Rossby waves at the mid-latitudes can be decomposed in a. synoptic-scale waves, sometimes referred as free waves, with wavenumber 6 or higher, propagating mainly in the longitudinal direction, b. quasi-stationary waves, with low wavenumber, less than 6, and close to zero phase speed. (White et al., 2022). The large-scale seasonal mean circulation is shaped by stationary Rossby waves, stationarity referring to their persistence over entire seasons. However, typical Rossby waves of interest for extreme phenomena, such as heatwaves, persist for several days to several weeks and have a wavelengths of 3000-6000 km, smaller than stationary waves. At the mid-latitudes it is known that anticyclonic anomalies of Rossby waves are often associated with heatwaves (White et al., 2022), as it is addressed in section 2.2.2.

2.1.5 *The Atlantic Ocean and the low-frequency variability*

The Atlantic Ocean with its mode of variability plays a key role in the European climate at different time scales. The Atlantic meridional overturning circulation (AMOC) is an ocean current responsible for transferring warm masses of water from the Tropics to the North Atlantic. It is involved in keeping relatively warm sea surface temperatures even at high latitudes [1], thus letting the European climate be significantly warmer than other regions at the same latitudes. This warm current travels the globe from south to north, in the Atlantic Ocean, reaching Greenland where it sinks, due to variations of density, and flows back to the south. Variations or slowing of this circulation could potentially lead to a cooling of Europe by some degrees.

From daily to decadal scales, of particular importance is the North Atlantic Oscillation (NAO). The NAO index is the difference in sea level pressure between locations near Iceland and the Azores. The temporal correlation between this index and the sea level pressure at all locations yields the spatial pattern associated with the NAO. Defined in this manner, the NAO was characterized by a dipole sea level pressure field, with anomalies of one sign being centered over Iceland and Greenland, and a band of anomalies of the opposite sign over the subtropical North Atlantic (Feldstein and Franzke, 2017). Negative phases of the NAO are generally associated with cold spells and often with blocking events. The NAO is the dominant teleconnection pattern in the Northern Hemisphere. Teleconnection patterns can be broadly defined as spatial patterns in the atmosphere which last for 1 or 2 weeks (and sometimes even longer) and alter the weather and the climate at global scale. Teleconnection patterns such as the NAO are known for exhibiting a variability on decadal timescales. Of remarkable importance is the shift from strong negative conditions in the 1960s to strong positive conditions in the 1990s which is probably due to changes in the North Atlantic ocean circulation (Woollings, 2010). Unfortunately, there is vast concern about the capability of climate models to reproduce the observed variability (Woollings, 2010).

At interdecadal timescales, the sea surface temperature of the Atlantic Ocean has revealed internally-driven variations. The origin of these variations are not fully understood, representing an active research field (Cassou et al., 2018). Based on the pioneering analyses of the observed data over the North Atlantic region, Bjerknes (1964) hypothesized that the low-frequency AMOC variability plays an active role in the observed basin-scale multidecadal variations in North Atlantic sea surface temperature and associated sea level pressure through related changes in Atlantic meridional heat transport and air-sea coupling. An index has been created to quantify this variability, the Atlantic Multidecadal Variability, AMV (Yeager and Robson, 2017; Zhang et al., 2019). Several studies identified the AMOC as a prominent driver of the AMV (Knight et al., 2005; Yeager and Robson, 2017; Cassou et al., 2018; Zhang et al., 2019). On the atmospheric side, recent studies have underlined how persistent NAO can influence the thermohaline circulation and thus the AMV (Delworth et al., 2016), also at

1 Although it is the South-West-North-East tilt of the jet stream which is mainly responsible for the temperature difference

these longer timescales. In part ii we will investigate the effects of this low frequency variability on summer European heatwaves.

2.2 PHYSICAL DRIVERS OF HEATWAVES

Several factors can lead to favourable conditions for the settling and the development of a heatwave. The thermodynamic ingredients which feature its formation are clear (Horton et al., 2016). A high-pressure system in the upper troposphere is often considered as an essential factor for the occurrence of heatwaves. In response to this high-pressure, air subsides resulting in local adiabatic compression heating. This inhibits convection and cloud formation, resulting in clear sky conditions. Thus, solar radiation reaching the ground surface increases, which in turn leads to upward sensible heating flux from the ground surface. Moreover, warm air advection might favour the increase of temperature (Miralles et al., 2014).

Tracing the causes of heatwave events from a dynamical point of view is all another story. Heatwaves are weather events which is the result of complex interactions among large and small scale processes which affects large areas at several timescales. Given this intricate grid, it is impossible to neatly separate one process to another thus impeding the creation of a clear cause and causality chain. There exist some reviews (Perkins, 2015; Horton et al., 2016; Barriopedro et al., 2023) which draw the contours of the current state of the art on our comprehension of this weather phenomenon. This section is inspired to all of them, but it is by no mean complete. I invite the reader to refer to them for deeper analysis. Following (Barriopedro et al., 2023), I decided to group the main drivers by their scale-dependence, proceeding from planetary-to-large scale factors section 2.2.1, to large-scale-to-regional section 2.2.2 to regional-to-local section 2.2.3. The idea behind this grouping lies in the fact that there exists a relation between spatial and temporal scales of heatwave drivers, as can be seen from fig. 2.2. In this manuscript I define a driver or a factor a thermodynamic or dynamic component of the weather and the climate which influence the formation, the maintaining and the worsening of a heatwave.

In this thesis, I investigate heatwaves at mid-latitudes, with a focus over Europe and Eurasia.

A schematic representation of the main drivers of heatwaves is shown in fig. 2.3. This figure synthesizes the main processes involved in heatwaves formation. Part of them will be discussed in the following, or have been already briefly explained above.

2.2.1 *Planetary-To-Large-Scale Factors*

At decadal and multidecadal time scales there are two main drivers of heatwave events: firstly, the anthropogenic-driven warming caused by increasing concentrations of greenhouse gases in the atmosphere; secondly, long-memory of internal components of the climate system, such as the ocean and the cryosphere with their modes of variability.

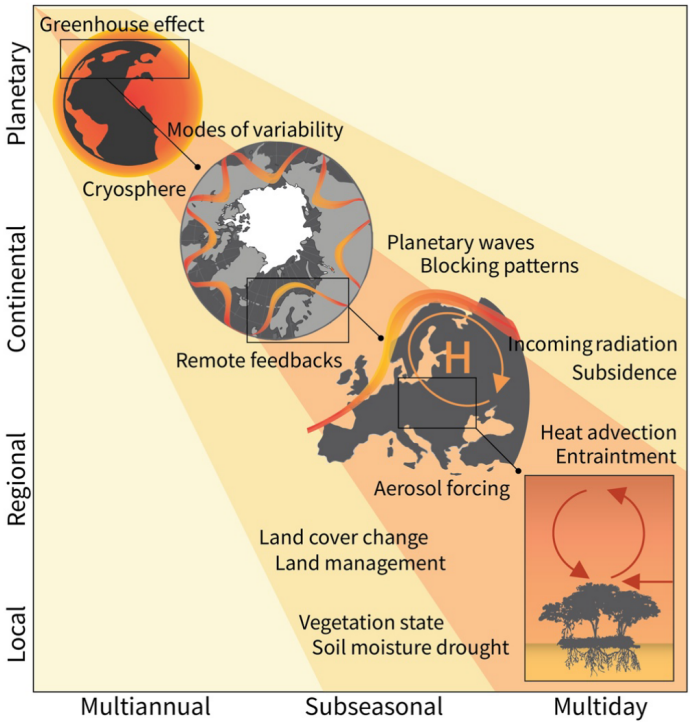


Figure 2.2: Spatial and temporal scales of characteristic heatwave drivers. Diagram identifying the characteristic heatwaves drivers and their relevant scales, from planetary to local spatial scale, and from multiannual to multiday temporal scale. Figure taken from (Barriopedro et al., 2023)

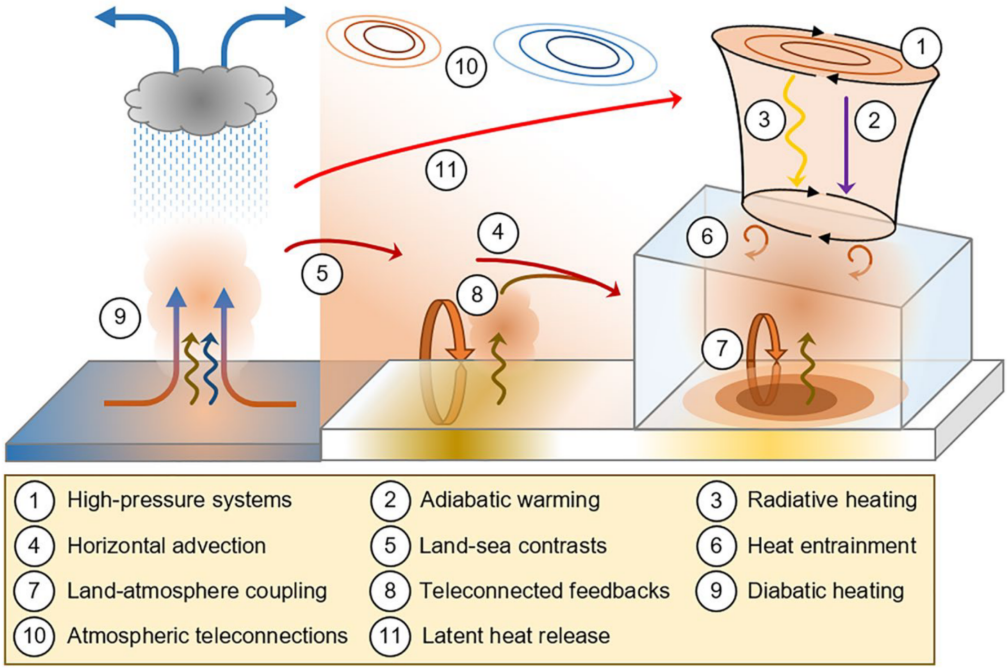


Figure 2.3: Schematics summarizing some of the large-scale-to-regional drivers of heat waves. Figure taken from (Barriopedro et al., 2023)

Nowadays, it is well established that global mean temperature has been rising as a result of enhanced greenhouse gas emissions (Seneviratne et al., 2021). Regional response to this warming is all but homogeneous. Indeed, other factors, such as cloud coverage, water vapour, snow cover, vegetation alter the regional response to global warming leading to different regional responses. Beside the thermodynamic effect due to global warming, a crucial and still open question regards the impacts of that warming on the dynamical response of the circulation. This point is a current research topic and the community does not heave a clear answer.

On the other hand, slowly varying components, such as the ocean and the cryosphere, affect heatwaves over a series of timescales. They act for instance on teleconnections patterns (spatial patterns in the atmosphere which last for 1 or 2 weeks (and sometimes even longer) that link the weather and the climate over large scales (Feldstein and Franzke, 2017)). Decadal to multidecadal internal modes of variability, such as the AMV is well known for playing a role at mid-latitudes on heatwave events (Qasmi et al., 2021; Ruprich-Robert et al., 2017; Ruprich-Robert et al., 2018). This mode is related to the inter-annual variability of the sea surface temperature of the North Atlantic Ocean. In part ii we investigate the effects of this low frequency variability on summer European heatwaves.

At inter-annual scales, there are other factors which are known to modulate the heatwave occurrence in several areas of the globe. Among the more important ones, large-scale tropical warm sea surface temperature patterns, such as El Niño-Southern Oscillation or the tropical Indian Ocean Dipole are known to favour heatwave conditions in several parts of the globe (Barriopedro et al., 2023).

2.2.2 *Large-Scale-To-Regional Factors*

In mid-latitude regions, at synoptic scale, large-scale weather systems are involved in the formation and maintenance of heatwaves. Recent record-shattering events, such as the Russian heatwave of 2010 (Matsueda, 2011) or the Western European one were associated to some weather systems. Particularly relevant for heatwaves formation are blockings, i.e. quasi-stationary high pressure systems. Those slowly moving anticyclones are often referred as 'blocks' as they alter the westerly jet stream flow. Europe is particularly exposed to this kind of phenomenon, as highlighted in section 2.1. Those upper-level tropospheric high pressure systems act to favour the maintenance of high temperatures, ultimately leading to heatwaves. In fact, summer European heatwaves are co-located with anomalous anticyclonic systems extending through the whole depth of the troposphere reaching the ground surface. This kind of pattern strongly correlates with some typical blocking patterns, as explained in section 2.1.3. At the level of the thermodynamics, as we already pointed out, heatwaves are mainly generated by the heat accumulation due to atmospheric advection and diabatic heating through latent and sensible heat fluxes (Kautz et al., 2022; Miralles et al., 2014) and by adiabatic warming by subsidence (Barriopedro et al., 2023). Blocking can be conductive of these processes, thus it can ultimately exacerbate

the severity of the heatwave given the persistence associated to this weather system (Kautz et al., 2022). Both processes, advection and diabatic heating, will be investigated with more details in the next section, as they are part of a land-atmospheric feedback as well. Another consequence of blocking is the reduced level of air quality, as it was measured in Moscow during the heatwave of 2010 (Lupo et al., 2012), as a consequence of the air stagnation.

At the mid-latitudes it is known that anticyclonic anomalies of Rossby waves are often associated with heatwaves (White et al., 2022). The identification of the mechanisms associated with this is an active field of research. Currently, a full theory for the growth and the propagation of Rossby waves at the spatial and temporal scales relevant for weather extremes is lacking (especially in non-linear regimes). Those extremes can rise from Rossby waves being exceptionally stationary/persistent (Röthlisberger et al., 2019; White et al., 2022) and/or exceptionally high in amplitude (Petoukhov et al., 2013; Röthlisberger, Pfahl, and Martius, 2016). One theory which tries to bridge the gap and to establish a causality chain is the Rossby waveguide (White et al., 2022). Within this theory, strong and narrow jets are associated with strong waveguidability which favour less dispersion of Rossby waves energy in the meridional direction, and associated potentially increase of probability of weather extremes. A linear theory on the propagation of Rossby waves has been established from the 80s (Hoskins and Ambrizzi, 1993; Hoskins and Karoly, 1981; Held, 1983). However, it has been criticized because for large-scale waves the eddies are of the same order of magnitude of the background flow, thus violating the limits of applicability of the theory (White et al., 2022; Wirth and Polster, 2021). Other problems are known, such as the definition of the background flow (White et al., 2022; Wirth and Polster, 2021). Quasi-resonance is a mechanism, proposed by (Petoukhov et al., 2013), which creates a causal link between waveguidability and amplified circumglobal Rossby waves through the resonance of Rossby waves of different wavenumbers (free waves and quasi-stationary). There are particular conditions under which this could happen, for instance a double jet: this configuration is analysed in part iv. Despite some studies investigating the relevance of this mechanism in recent heatwave episodes (Coumou et al., 2014; Kornhuber et al., 2019; Petoukhov et al., 2016), this methodology has been criticised as it seems that the causality link is not so strong and dependent on the definition of background flow (Wirth and Polster, 2021). Persistence of non-circumglobal Rossby waves has been associated to Rossby wave packets. They are Rossby waves for which the amplitude has a local maximum and decays to smaller values at large distances (Wirth et al., 2018). They propagate in the zonal direction especially along waveguides. Several connections between those Rossby waves and weather extrema have been documented, especially when the Rossby waves have shown a Rossby wave breaking, which sometimes occur at the decay stage of Rossby wave packets (Wirth et al., 2018). Despite the dynamics at the leading edge of a Rossby wave packet is quite understood, as pointed out in a recent review on the topic (Wirth et al., 2018), more systematic studies need to be done to generalize the identification of the phenomenon, including a unique definition of an Rossby wave packet object, and its causal link to the extremes.

2.2.3 *Regional-To-Local Factors*

At local and regional scales, there are several factors that could modulate the formation and duration of heatwaves. In this section we will present the main ones.

Coupling and feedbacks between land and atmosphere are known to exacerbate heatwave conditions (Seneviratne et al., 2006; Alexander, 2011). In particular, it is the available soil moisture that controls the relative importance of the latent and sensible energy fluxes, which is responsible for the strength of the land-atmosphere coupling and feedback. Indeed, soil moisture is the key factor which affects how the energy absorbed by the surface of the Earth is returned to the atmosphere. Some is returned under the form of infrared radiation and the remaining one, the net radiation is transformed in two energy fluxes: latent heat flux, the loss of heat through evaporation of liquid water, and sensible heat flux which directly heats the surface and the air. Partition of the energy between those two fluxes strongly affects the land and the atmosphere, as explained in the following.

Figure 2.4 summarizes the mechanisms of the land-atmosphere feedback. When the soil is wet, the net radiation arriving on the ground surface is mainly converted into latent heat, through evapotranspiration. Evapotranspiration is the process through which water is transferred to the atmosphere from open water and ice surfaces, bare soil and vegetation. High evapotranspiration cools the ground and enhance the formation of clouds because of the higher concentration of atmospheric water vapour. On the other hand, when the soil is already desiccated, sensible heat is the predominant flux. This creates a warmer and drier air above the ground which inhibits cloud formation, which in turn allows direct heating of the desiccated soil, thus constraining the soil moisture content even further.

This land-atmospheric feedback intensified and lengthened drought conditions and has been proven to be a determinant factor for severe heatwaves in regions where evapotranspiration is limited by the availability of soil moisture, such as some regions in Europe (Fischer et al., 2007b; Fischer et al., 2007a; Vautard et al., 2007). Using either station data or model simulations, the temperatures reached during the heatwave of 2003 could not have been explained without considering the anomalous dry soil in the spring preceding the heatwave event. Moreover, the feedback between land and atmosphere when reinforced with advection of warm air can lead to what in the literature are known as 'mega-heatwaves', with reference to the Western European heat wave of 2003 and the Russian one of 2010 (Miralles et al., 2014; Fischer, 2014). The authors found that the extremely high temperatures reached in those events are explained by the combined multi-day memory of land surface and the atmospheric boundary layer (shown in fig. 2.4). During both events, heat accumulated in the boundary layer causing the underlying soil, which was already in dry conditions, to dry even further. As a result, evapotranspiration decreases favoring high sensible heat flux, thus warming and deepening the boundary layer, creating a feedback loop. Moreover, warm air was anomalously kept warm even during night time, thus exacerbating the effect of the subsequent day warmth. Another interesting point for mid-latitude mega-heatwaves suggested by (Mi-

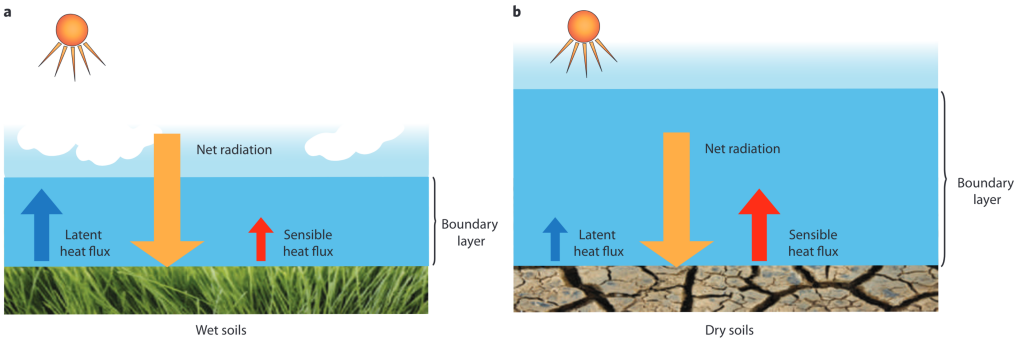


Figure 2.4: Schematic of the net radiation budget at the land surface. Figure taken from (Alexander, 2011).

ralles et al., 2014) is the need of having simultaneously dry soil, heat inflow from warmer region and accumulation of heat during several days to cause temperature exceedance of 40°C. In part ii we investigate the relative contribution of the soil moisture deficit and of the AMV on the amplitude of summer European heatwaves.

Despite the great advances in characterizing the soil moisture impact of heatwaves, our understanding of soil moisture feedback on heatwaves is incomplete. One reason lies in the lack of realistic representations in models of complex vegetation landscapes and dynamics (Barriopedro et al., 2023). In addition to that, more has to be understood on the role of soil condition in the development of large of synoptic and mesoscale systems favorable to heatwave events (Miralles et al., 2019).

Other factors such as land-cover, land-use and anthropogenic aerosols are known to play a role in either the settling or the exacerbation of heatwaves at these scales (Barriopedro et al., 2023).

In this section, I presented some drivers of heatwaves at mid-latitudes acting at different temporal and spatial scales. In the next section, the focus is shifted to the changes in heatwaves frequency and amplitude in response to climate change.

2.3 OBSERVED AND PROJECTED CHANGES IN HEATWAVES

The latest IPCC report confirmed that it is certain that the intensity and frequency of warm days, warm nights and heatwaves have been increasing since 1950 in Europe (particularly over the Mediterranean area) as a consequence of increasing greenhouse gas concentrations (Seneviratne et al., 2021). The projected changes in the magnitude of the extreme temperatures over land are larger than the global mean temperatures and will affect several regions even at a 1.5 °C of global warming. The change in the extreme temperature affects not only the extreme hot ones, but also the extreme colder ones, which are projected to become hotter, leading to more sporadic extreme cold events. Figure 2.5 shows the changes in both the annual maximum and minimal temperatures under different global warming levels. The minimal one has a stronger

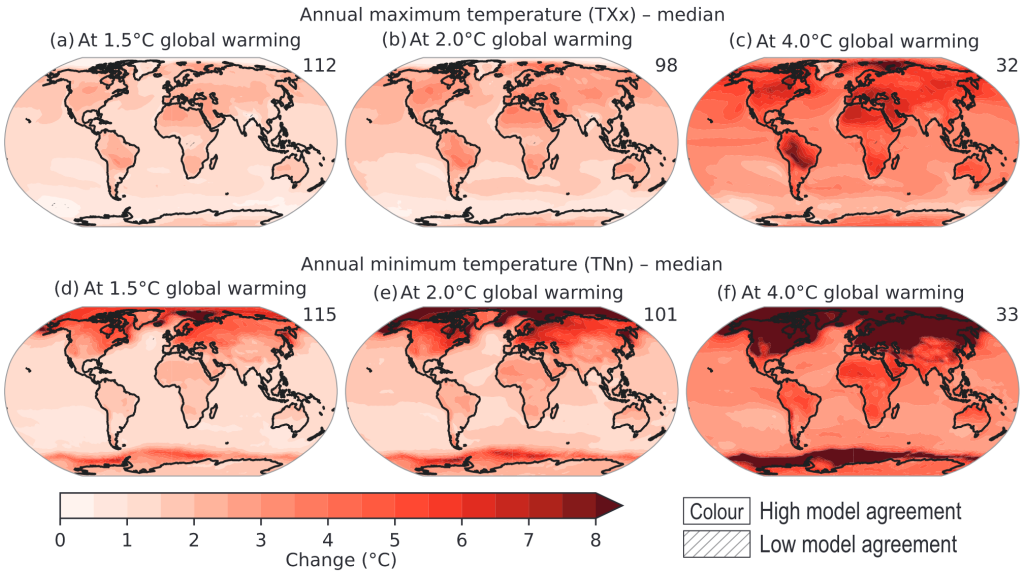


Figure 2.5: Projected changes in (a–c) annual maximum temperature (TXx) and (d–f) annual minimum temperature (TNn) at 1.5°C, 2°C, and 4°C of global warming compared to the 1850–1900 baseline. Results are based on simulations from the Coupled Model Intercomparison Project Phase 6 (CMIP6) multi-model ensemble under the Shared Socio-economic Pathways (SSPs) SSP1-1.9, SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 scenarios. The numbers in the top right indicate the number of simulations included. Figure taken from the AR6 IPCC report, Chapter 11.

equator-to-pole dependency and has a projected higher warming, while for the annual maximum temperature, the warming is spatially more uniform over land areas. An interesting property of projected extreme temperatures is that their intensity does not scale linearly with global warming. For instance, in fig. 2.6, we see that for extreme annual daily maximum temperatures with a return time of 10 and 50 years, the projected intensity is higher than the projected global warming level.

As already mentioned above, the ongoing warming of our planet is all but uniform, with discrepancies among land and ocean territories, as well as among land regions. Understanding the role of soil moisture-temperature feedbacks at regional scale is thus essential for adaptation strategies, especially at the mid-latitudes. Regional changes in the extreme projected temperatures can be decomposed in a mean response to global warming and an additional warming strongly affected by soil moisture-temperature feedback (Vogel et al., 2017; Douville et al., 2016). This response is consistent with the projected soil moisture drying at multidecadal time scale, which has a major impact than the interannual to subseasonal effect of soil moisture variability. In addition, this result points out how much predictions of changes in temperature at regional scale might be ameliorated in climate models with additional inclusion of soil moisture-related processes.

While the thermodynamic effects of climate change on heatwaves are widely known (on a hotter planet the frequency of the extreme temperatures will

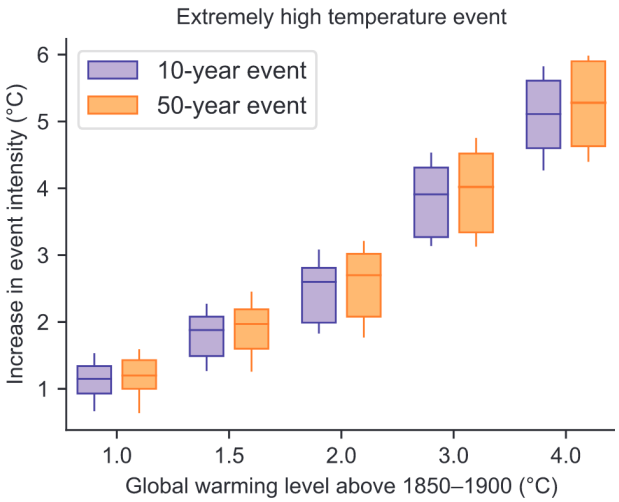


Figure 2.6: Projected changes in the intensity of extreme temperature events under 1°C, 1.5°C, 2°C, 3°C, and 4°C global warming levels relative to the 1850–1900 baseline. Extreme temperature events are defined as the daily maximum temperatures (TXx) that were exceeded on average once during a 10-year period (10-year event, blue) and once during a 50-year period (50-year event, orange) during the 1850–1900 base period. Results are shown for the global land. The figure shows the non linear trend which exists between the projected increase in the intensity of the event (y-axis) and the projected global warming level (x-axis). Figure taken from the AR6 IPCC report, Chapter 11.

change causing more events and more extreme ones to happen), the dynamics associated to this change lacks of a full understanding. For instance, there is no consensus on the future change in blocking (Woollings et al., 2018). One reason for this is the lack of a theory for the explanation of this phenomenon, as introduced in section 2.1.3. Another reason is the continued underestimation of blocking activity by climate models, particularly for the Atlantic/European sector in winter (Woollings et al., 2018; White et al., 2022). Specifically for heatwaves, (Vautard et al., 2023) has shown that CMIP6 simulations do not capture the rapid observed warming of extreme heat over Western Europe. At the mid-latitudes, one relevant phenomenon for potential dynamical changes in response to climate change is the so-called Arctic amplification. This refers to a more rapid increase of temperatures at high latitudes than elsewhere on Earth, causing a less pronounced meridional temperature gradient between the pole and the equator and the jet stream to move equatorward (Coumou et al., 2018). Some people argue that this displacement might cause a change in the weather patterns. In their article (Barnes and Screen, 2015), the authors discuss the impact of the Arctic warming on the jet stream. They state that, despite the fact that model simulations demonstrate that Arctic warming can impact the jet, those simulation responses are of the same magnitude as (or smaller than) the internal variability. Moreover, observations solely, as it is often the case in climate science, cannot provide substantial help because of the short duration of reliable records. An open question is how these responses will effectively change the weather at mid-latitudes.

Different mechanisms have been proposed as possible pathways of the atmospheric response to the Arctic amplification (Barnes and Screen, 2015), summarized by fig. 2.7. The reduced temperature gradient, might cause on one side a reduced baroclinicity and thus a reduced storm activity and on the other side a reduced warming effect, thus temperature extremes, under the effect of climate change (Schneider, Bischoff, and Płotka, 2015). Another effect of this gradient reduction is the hypothesis of a less sinusoidal jet stream, which is expected to slow down because of a reduced thermal wind balance. In their work, (Francis and Vavrus, 2012) hypothesized that this slowing down could cause an amplification of the Rossby waves and more frequent blocking conditions, ultimately leading to more heatwave conditions. Another mechanism leading to the same conclusions has been proposed by (Coumou et al., 2014) in which the reduced meridional temperature gradient might lead to more double jet conditions. This refers to a split of the jet stream in two filaments which would act as barriers to trap the waves in the lower atmosphere. This phenomenon is known in literature as 'quasi-resonance', briefly presented in section 2.2.2. The double jet configuration is the core content of part iv, where we use a rare event algorithm to simulate double jet stream states and analyse the atmospheric response.

In warmer mid-latitudes, the ocean will be warmer as well. This translates into an enhanced frequency and growth of storm activity. Summarising, there is not a clear pathway of the possible effects of climate change on the dynamics of the atmosphere.

This section was a general overview of some of the most important mechanisms at the mid-latitudes, with a focus on Europe, for heatwave formation and maintaining. The overall goal of this part was to present the reader a concise description of both the thermodynamics and dynamics aspects of this fascinating, yet dramatic, weather phenomenon. In the next section I will present a brief history of climate modelling and the models run in the research group I am part of.

2.4 CLIMATE MODELS

In this section I briefly present the history of climate modelling and the current challenges, some of them already mentioned in the previous sections. Numerical models of different complexity are employed to study the climate system at different spatial and temporal scales. These models help in understanding the past, the present and the possible future of our Earth. The highest complex ones are the Earth System Models (ESMs). Climate models embed the mathematical formulations of laws which govern the components of the climate, such as the atmosphere, the ocean, the cryosphere, the land, the biosphere and the carbon cycle. They are built on the fundamentals laws of physics (Navier-Stokes, Clausius-Clapeyron), empirical laws based on observations and finally, when possible, constrained on conservation laws (mass and energy). It is clear that to numerically evaluate the evolution of such a complex system, huge computational power needs to be deployed. Hence, the evolution of the complexity of available models goes hand in hand with the evolution of computing power.

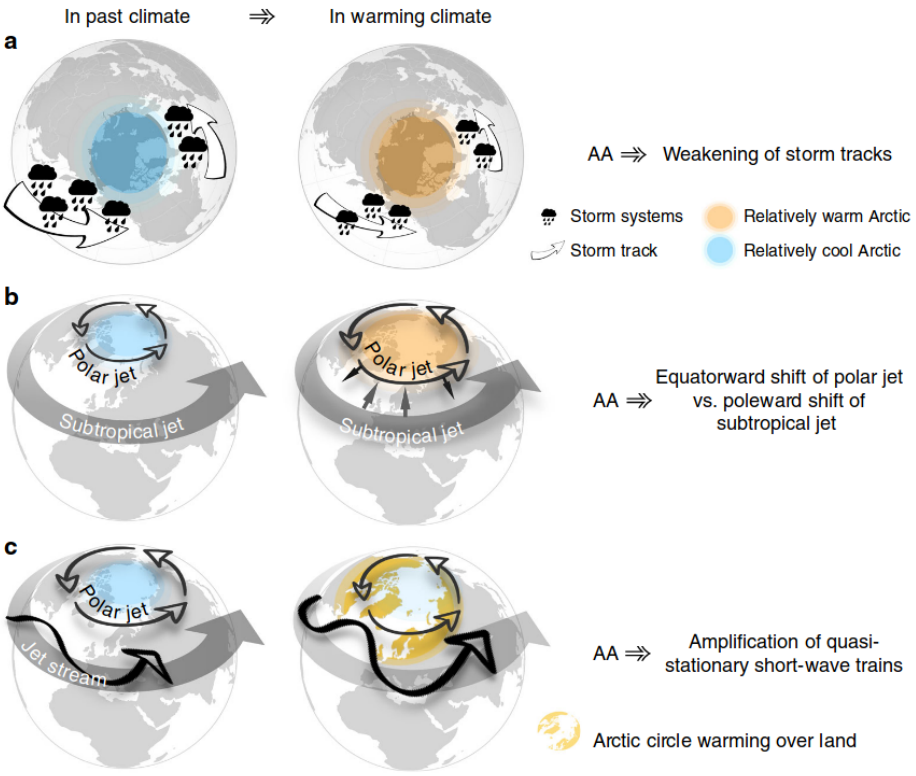


Figure 2.7: Schematic representation of proposed dynamical mechanisms in summer. a) Weakening of storm tracks, b) latitudinal-shift in jet positions, and c) amplification of quasi-stationary waves. Figure taken by (Coumou et al., 2018).

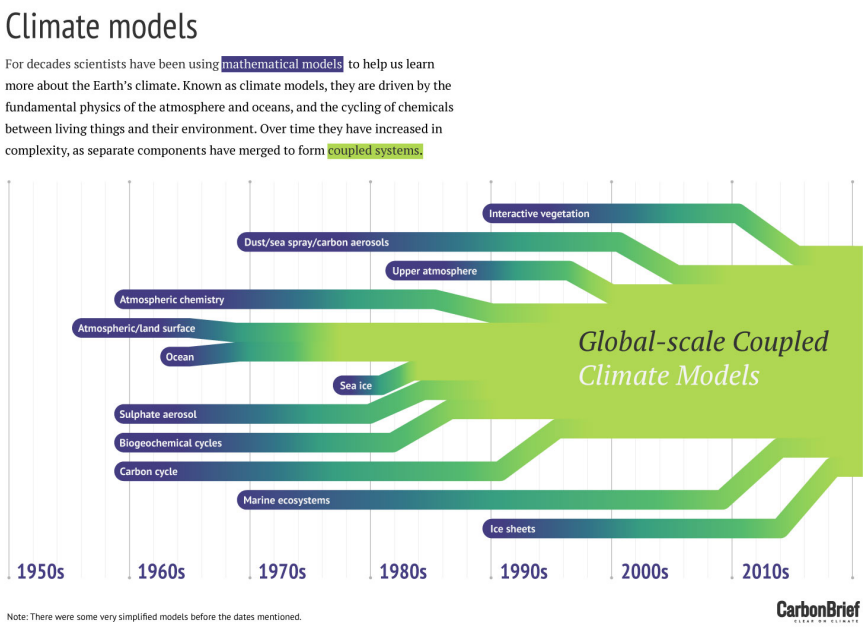


Figure 2.8: Timeline of the history of climate modelling. Figure taken from https://www.carbonbrief.org/timeline-history-climate-modelling/.

In the next section I briefly present the historical path of climate models arriving to today's models and open challenges. I then move to present the climate models run in the research group I am part of and I briefly compare them with the reanalysis dataset ERA5.

2.4.1 *History of climate modelling*

Energy Balance Models are the earliest and simplest attempt of numerical climate models. They actually did not simulate the climate, but rather the energy-balance of the atmosphere. As simple as they are, they only output temperature and consider the Earth as a point. A step along from Energy Balance Models are Radiative Convective Models, which simulate the transfer of energy through the height of the atmosphere. They are generally one dimensional models which calculate the temperature and humidity of different layers of the atmosphere. We have to wait until the creation of General Circulation Models (GCMs), also called Global Climate Models, to have a simulation of the physics of climate, including the flow of air and water masses in the atmosphere and the ocean and the heat transfer. Early GCMs were run in atmosphere-only or ocean-only setup. At a later stage, different models were linked together to provide a comprehensive representation of the climate system. Those are the so-called coupled models. [2]

Going forward in time, more and more aspects of the Earth system were added to GCMs. An infographic of its temporal evolution is presented in fig. 2.8.

2 An interesting overview is given at https://www.carbonbrief.org/timeline-history-climate-modelling/

We need to wait the beginning of the 21st century to have models with surface and deep ocean circulation components, resolved volcanism, aerosols and hydrologic cycle, key for modelling precipitations. Since then, other components such as atmospheric chemistry and interactive vegetation have been included.

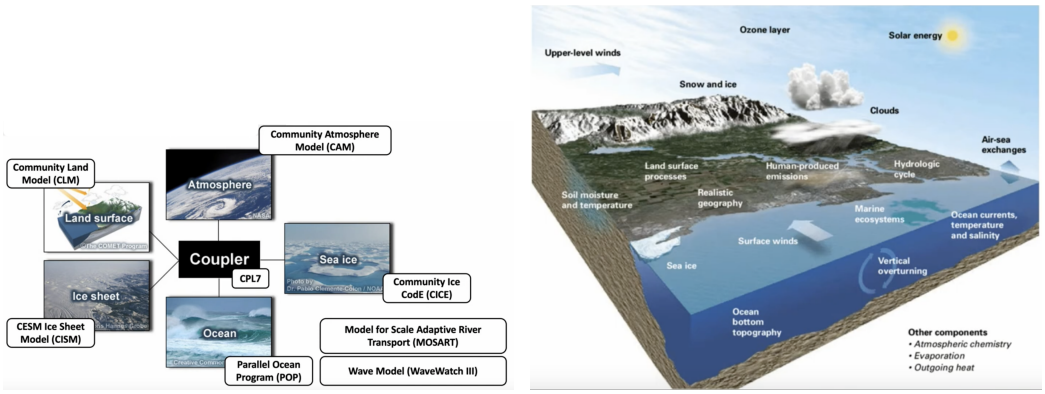
Nowadays, more than 50 operational GCMs are maintained by atmospheric modeling centers around the world, with many of them taking part into the reports of the IPCC. Those reports made use of multi-model ensembles generated through various phases of the World Climate Research Programme CMIP, the latest one being the 6th (CMIP6). In the latest IPCC report Climate Change, 2023, several climate models at global and regional scales, with different complexity, took part into the scientific development which built this report. However, there is still margin for amelioration. In Chapter 1 of the above-mentioned report, the authors highlighted the improvements made from the previous report and the remaining challenges to be addressed, notably for model resolution, parametrizations (for instance for convection, turbulence, clouds physics), land and ocean biosphere and biochemical cycle modelling. Here, we will report the main ones and leave the keen reader to refer to that Chapter for deeper insights.

CMIP6 models have higher horizontal and vertical resolution than its predecessors,to represent finer aspects of the atmospheric circulation and ocean. In the vertical direction, the number of vertical levels in the atmosphere has increased to ameliorate the representation of stratospheric processes. Despite the documented progress of higher resolution, improvements between CMIP5 and CMIP6 are modest at the global scale, highlighting how much the model performance relies on other factors than resolution, such as parametrizations. For the representation of physical and chemical processes, notable effort has been made for radiative transfer, cloud microphysics and aerosols (and effect of aerosols on cloud formation modification). However, it still represents a source of spread among models.

2.4.2 *Models used in this thesis*

The goal of this section is to present the two models run in the research group I am part of and to compare them with the reanalysis dataset ERA5. Between the two climate models I present, namely PlaSim and CESM, I decided to give a more detailed presentation of the latter, as it is the model I have run myself for the study presented in part iv. A longer description of the former is presented in part iii.

PLASIM PlaSim is an intermediate complexity climate model that has a dynamical core that solves the moist primitive equations (Vallis, 2017) and has a resolution of 2.8 degrees both in latitude and longitude with 10 vertical layers. It has fairly simple parametrizations for sub-scale processes, for instance for hydrology over land, with a single-layer bucket model (Manabe, 1969). Sea ice cover and ocean surface temperature are cyclically prescribed for each day of the year. The model reproduces a climate close to that of the 1990s. The simplified parameterizations allow it to run 100 times faster than the models used for CMIP studies, which makes it very suitable to obtain extremely long datasets.



(a) (b)

Figure 2.9: (a) Scheme of the models used in the CESM Panel (b) Solved processes in the model in its fully coupled version. Figure courtesy of COMET program.

Indeed, we dispose of a very long climate simulation of 8000 years with this model. This simulation has been performed by Bastien Cozian, former PhD student in our group, to whom I am really grateful.

COMMUNITY EARTH SYSTEM MODEL The CESM is composed of seven physical components: atmosphere, sea-ice, land, river-runoff, ocean, land-ice and ocean-wave. The coupler is another component which allows the information to flow among the model's components and allows the time evolution of the global model. An overview of the model components can be found in fig. 2.9. The same figure shows the processes resolved in this model in its fully coupled configuration. This model can suit for different combinations, for instance it could be active and solve prognostic equations or the components can be replaced with prescribed data.

In the setup we use, only atmosphere and land are used as active components. More specifically, the atmospheric model is the Community Atmospheric Model version 4 (CAM4), while the Community Land Model version 4 (CLM4) is used for the land. CAM4 (Neale et al., 2013) includes a default dynamical core which solves a finite discretization of primitive equations. It presents a resolution of 0.9 degrees over the latitudes and 1.25 over the longitudes, with 26 vertical layers in a hybrid pressure-sigma coordinate. With these resolutions over latitude and longitude, the model spans roughly over 100km at the mid-latitudes.

The model CLM4 (Lawrence et al., 2011) represents several aspects of the land surface, including surface heterogeneity and several subcomponents to incorporate bio-geophysical and biochemical aspects, hydrology cycle, energy fluxes and vegetation. The hydrology component includes evaporation, surface infiltration, surface and sub-surface runoff, snow and glacier formation and melting, in interaction with the vegetation model. The land model has 15 vertical layers. The version used has a prescribed aerosol concentration, including greenhouse gases, fixed to a level which matches the observed one to reproduce the 2000s climate. Finally, the CO_2 concentration is set at 367 ppmv. In the version used, the other model's components are not active but with prescribed

data. We use historical data to prescribe the ocean and the sea-ice models, for instance. In particular, we use monthly averaged data to obtain daily data via a linear interpolation and then cyclically yearly repeat. In the context of this thesis, I use a long climate simulation of 1000 years of data performed with this model by Francesco Ragone [3,4], to whom I am very grateful. I run the same climate model coupled with a rare events algorithm in the study presented in part iv.

ERA5 ERA5 is a reanalysis dataset (Hersbach et al., 2020). Reanalysis combine past observation records with today's weather models to give a consistent and complete picture of past weather through data assimilation. Observations are unevenly distributed and come with errors. They cannot provide a complete and accurate picture of the state of the Earth system across the globe at a given point in time. Indeed, one could think of the Earth as a puzzle, with observations being the pieces which form that puzzle. Some of them are missing, thus thanks to the laws of physics, one can reconstruct the missing one to create a consistent image. Reanalysis covers not only the Earth surface, but also the atmosphere until the stratosphere. ERA5 dataset is a publicly available of the European Centre for Medium-Range Weather Forecasts service covering the period from January 1940 to present. ERA5 has a resolution of 0.25 degrees in latitude and longitude, with 45 vertical layers.

Beside these different parameters, an important difference is that the first two models are climate models, while ERA5 is a hybrid weather model. Nowadays, the line which separates the two is getting blurred, but traditionally they are created for different purposes. Climate models focus on simulating the energy balance of the Earth's atmosphere over long periods, from decades to centuries. They are designed to predict changes in the climate system by considering the interactions between the atmosphere, oceans, land surface, and ice. The key components include energy balance, dynamical components and parametrizations (to represent processes that occur at scales smaller than the model's resolution). Weather models are designed to forecast atmospheric conditions over short timescales, from hours to weeks. They also include a dynamical core, but with a detailed representation of atmospheric physics to predict the state of the atmosphere with high temporal and spatial resolution. They use parametrizations as well, but to accurately capturing phenomena that affect short-term weather prediction (turbulence, clouds formation, etc..).

In fig. 2.10 I show the zonal wind, i.e. the west-east component of the wind vector at 300 hPa and 2 m air temperature climatology for the months of June, July and August for: ERA5, CESM, PlaSim (from left to right). These maps serve to give the reader an idea of how much idealized are the models used compared with a reanalysis dataset. Looking at the zonal wind, we observe that over the North America and the Atlantic Ocean arriving to Eastern Russia CESM reproduces correctly the latitudinal spread and the amplitude of the zonal wind, while PlaSim tends to overestimate its amplitude and to have a more zonal representation, clearly visible over the Eurasian sector. Over the Pacific Ocean, CESM

3 *Georges Lemaître Centre for Earth and Climate Research (TECLIM), Earth and Life Institute, Université Catholique de Louvain, Louvainla-Neuve, Belgium*

4 *Royal Meteorological Institute of Belgium, Brussels, Belgium*

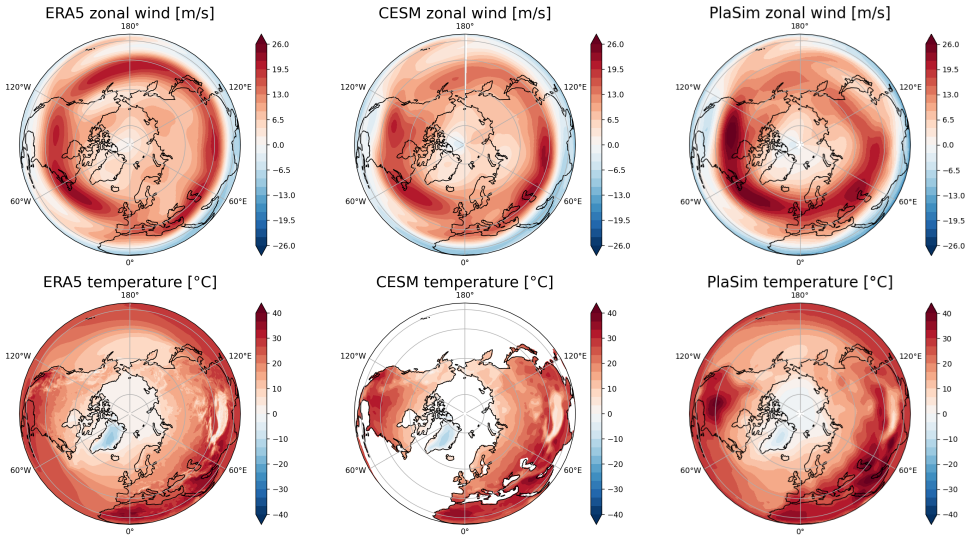


Figure 2.10: Zonal wind, i.e. the west-east component of the wind vector, at 300 hPa and 2 m air temperature for: (from left to right) ERA5, CESM, PlaSim. The average is taken over the months of June, July and August. We recall that the simulation length for CESM is of 1000 years and that it is used in land and atmosphere setup only. PlaSim data has a length of 8000 years and presents prescribed sea ice cover and sea surface temperature.

underestimates the amplitude of the zonal wind, but simulates with a good degree its latitudinal spread, while PlaSim struggles both with the amplitude and the position of the zonal wind. For the 2 m air temperature, there is a large similarity between CESM and ERA5, which is also the case for PlaSim, despite a lower agreement over the central North America, where PlaSim simulates a stronger amplitude of the 2 m air temperature.

Summarizing, this initial chapter has been devoted to present some drivers of heatwaves at mid-latitudes and observed and projected changes in their occurrence and amplitude. Finally, I introduced and compared some of the models which are used in the following chapters.

Let's now turn the page to present some tools which will be used at different stages of this thesis. Finally, we introduce large deviation theory, a well established theory for estimating the probability of rare events.

3

MARKOV PROCESSES, RARE EVENTS AND LARGE DEVIATIONS

When studying climate extremes, we may be interested in two problems. Estimate a posteriori statistics, meaning understanding the average state that has led to them for an analysis purpose of the event or estimating the average time to wait to observe an event with a given amplitude. The first quantity is called *composite map*, while the second one is called *return time* and are two interesting quantities in the context of adaptation and risk limiting policies, and they are recurrent in this thesis. The second question regards the forecasting of extreme events, which deals with the estimation of the probability that the event happens before it actually happens. Namely, I aim at estimating the *committor function*. In this chapter I introduce all the previous mentioned quantities, but before doing so I first present briefly what is a Markov process (Liggett, 2010) and then define *return time* (Lestang et al., 2018) and *composite maps*. Finally, I introduce the *committor function* (Lucente, Herbert, and Bouchet, 2022). In section 3.2 I introduce the *Large deviations theory*, a theory which allows to estimate the probability of a large fluctuation of a given observable. The aim of this section is not to give a complete overview of this theory, as it is not the scope of this thesis. Nevertheless, the rare event algorithm presented in part iv was originally conceived to estimate the large deviation rate function (or actually its Legendre-Fenchel transform), a key quantity to have access to the probability of a rare fluctuation of a given observable. Thus, the ultimate goal of this section is to present the minimal amount of concepts to create a link with the algorithm used in part iv.

3.1 MARKOV PROCESS, COMMITTOR FUNCTION AND RARE EVENTS

Let's define a stochastic process $\{X(t)\}_{t\in\mathcal{T}}$, with $\mathcal{T}$ ordered set, as a collection of random variables defined over a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the sample space, $\mathcal{F}$ is a σ-algebra, $\mathbb{P}$ is a probability measure. The process takes value in the phase space Γ. The index t is representative of time in our case. The process $\{X(t)\}$ is a *Markov process* if it satisfies the following property:

$$\mathbb{P}(X(t+h) \in C | \{X(s)\}_{s\leq t}) = \mathbb{P}(X(t+h)) \in C | X(t)) \tag{3.1}$$

for $h \geq 0$, $C \subset \Gamma$ and $\{X(s)\}_{s\leq t}$ is a collection of the values of the stochastic process up to time t. This means that a Markov process is a process where its future outcome depends only on its nearest past value, i.e. the process does not have memory of all its past history, and — most importantly— such predictions are just as good as the ones that could be made knowing the process's full history. In the following, our study is restricted to stationary and ergodic Markov process, thus we proceed to briefly define them.

STATIONARITY A Markov process is said to be stationary if the joint probabilities of observing a stochastic process are invariant under any time translation τ.

ERGODICITY A Markov process is said to be ergodic if it is stationary and if there is an equivalence between the long-time mean of the process and its ensemble average.

3.1.1 *Return time and composite maps: a posteriori statistics*

The goal of this section is to present two tools which are widely used in this thesis: the return time and the composite maps. Both are common tools for making a posteriori statistics, i.e. they analyze some statistical properties of a rare event once the event has already happened. Given a stochastic process defined as above, a scalar observable which depends on the path $\{O[X(t)]\}$ and a threshold value a which separates between rare and not rare events, it is possible to define the random variable $\tau(a,t) = \min\{\tau \geq t | O[X(\tau)] \geq a\}$. Then the return time is defined as the average time to wait to see an event of magnitude higher than a:

$$r(a) = \mathbb{E}\left[\tau(a,t)\right] \tag{3.2}$$

where $\mathbb{E}$ is the expected value, practically estimated via an empirical average over the data, and a is generally named *return level*. The return time is strictly connected to the probability of observing an event of amplitude a. A well established theory for computing return times is the Extreme Value Theory (EVT) (Coles, 2001). In general, this theory seeks to assess the probability of an unseen event by extrapolating the information from the observed ones. There are two methodologies which are practically used for extreme value analysis: the Block Maxima and the Peaks Over Thresholds. The former one relies on the partition of the observed time series of O in non overlapping blocks, which are then assumed to be statistically identical and independent and long, and on the subsequent evaluation of the maxima within each block. Using the block maxima data one can fit a Generalized Extreme Value distribution which extrapolates beyond the observed data (Ferreira and Haan, 2015; Coles, 2001). The latter approach, instead of blocking, we select all observations that are above a certain threshold a. If a is high enough, then the selected observations can be thought as independent realizations and will be fitted using a Generalized Pareto distribution. However, a limit of this theory is the uncertainty associated with each estimation of the return time or of its probability, which is non-negligible when a is really high (Le Priol, Monteiro, and Bouchet, 2024; Coles, 2001). This limit is common to all theories as there is a lack of data problems. In this thesis, we tackle this by generating data that will effectively and efficiently populate the tails of the observable O in order to sample a statistically significant amount of rare realizations of it. This is done with the aid of a rare events algorithm, which we present in the following. A *return time plot* is a plot in which for each return level a (which measures the rareness of the event) we associate a return time (which measures the probability of observing an event with value a). This

quantity allows to see the evolution of the probability as a function of a and in this thesis is computed in parts ii and iv.

Beside the estimation of the time to wait for an event of amplitude a to occur, i.e. the return time, an interesting question, particularly relevant in the climate community, concerns the dynamics which has let the event to occur (Grotjahn and Faure, 2008; Noyelle, Yiou, and Faranda, 2024; Miloshevich et al., 2023; Teng et al., 2013). *Composite maps* are a statistical tool which estimates the average state of the climate system which has led to an extreme event, such as a heatwave. Formally, if one denotes X as a stochastic process, with $X(t) \in \Gamma = \mathbb{R}^d$, a scalar observable which depends on the path $\{O[X(t)]\}$ and a threshold value a which separates between rare and not rare events, then composite maps are defined as the average state of X lagged of an amount of time τ before the event has happened:

$$C(\tau) := \mathbb{E}[X(t-\tau)|O[X(t)] \geq a] \tag{3.3}$$

where $\mathbb{E}$ is the expected value, practically estimated via an empirical average over the data. The quality of the estimated composite map depends on the number of events we are averaging over. Thus, in principle it suffers from sampling error for the very rare events. However, in part iii, we show that for a properly defined class of extreme events, including heatwaves, the pattern of composite map of very extreme events closely resembles to the one of less extreme ones. They are just rescaled by a factor proportional to the threshold a which defines the events.

The two statistics presented in this section are extremely helpful to characterise events when they already occur. Another interesting question regards the prediction of the event, which is addressed in the next section with the *committor function*.

3.1.2 *Committor function: a priori statistics*

Medium-range forecast is an intriguing subset of forecast which poses several methodological challenges. Indeed, the interesting timescales lie between the deterministic predictability time and the mixing time of the system. In this range, relevant for high-impact events, such as heatwaves, the forecast needs to be probabilistic and the right theoretical tool is the *committor function*, introduced in the field of climate by (Lucente et al., 2019). It is an a priori statistics, meaning that we aim at estimating the probability of an event to happen before it actually happens. In this section, I introduce it in the context of stochastic processes and for the particular case of prediction of heatwaves. Let's consider a stationary process $\{X(t)\}$ and let $\mathcal{D}$ be a subset in the phase space $\mathcal{D} \subset \Gamma$. We define the *first hitting time* $\tau_{\mathcal{D}}(x)$:

$$\tau_{\mathcal{D}}(x) = \inf\{t : X(t) \in \mathcal{D}|X(0) = x\}. \tag{3.4}$$

Let then $\mathcal{A}$ and $\mathcal{B}$ be two disjoint subsets of the phase space Γ. The *committor function* $q(x)$ is the probability that the process started in $X(0) = x$ will visit the set $\mathcal{A}$ before visiting the set $\mathcal{B}$. Mathematically:

$$q(x) = \mathbb{P}(\tau_{\mathcal{B}}(x) > \tau_{\mathcal{A}}(x)). \tag{3.5}$$

The two disjoint subsets $\mathcal{A}$ and $\mathcal{B}$ might be two attractors, or one attractor and one set of extreme events. In this thesis, I deal with one set being typical states of the system and the other one being a set of extreme events, which is visited when rare fluctuations arise. For instance, given a set of predictors, the set of important climate variables for the extreme event under study, we want to know the probability of observing a heatwave in τ days. In that case, the committor is identified as the above-mentioned conditional probability, which reads as:

$$q(x, \tau) = \mathbb{P}(O[X(t)] \geq a \mid X(t - \tau) = x). \tag{3.6}$$

When we are dealing with low dimensional systems, having access to this probability eq. (3.6) can be tricky but not impossible. In the context of climate science, given the high dimensionality of the system of interest, computing this quantity is challenging. A preliminary step is then to map the committor function into a lower dimensional space with a suitable map that retains the salient features which lead to the extreme events of interest. In this thesis, this mapping is estimated both using Deep Convolutional Neural Networks and with a novel methodology that naturally performs the dimensionality reduction and the optimal projection mapping. This framework is called the Gaussian approximation and it is the core content of part iii.

3.1.2.1 *Estimation of the committor function*

This section focuses on the estimation of the committor function via numerical methods. A thorough analysis can be found at (Lucente, 2021). When we consider a diffusion process, the committor function can be analytically obtained by solving the associated Fokker-Planck equation (Lucente, Herbert, and Bouchet, 2022). I will not discuss this approach here, but I rather focus on two that are connected with this manuscript.

DIRECT SIMULATION The straightforward method one could use to estimate the committor function is via a direct simulation approach. This consists of simulating N realizations of a process $\{X(t)\}$, starting with the same initial condition $X(0) = x$ and counting how many have hit the extreme set. In the context presented above, let's assume that we want to estimate the committor at point x for two sets $\mathcal{A}$ and $\mathcal{B}$ using the equation eq. (3.5). If one denotes as $N_{\mathcal{B}}(x)$ the number of realizations that ended up in $\mathcal{B}$ a crude estimation of the committor is simply given by:

$$q(x) = \frac{N_{\mathcal{B}}(x)}{N} \tag{3.7}$$

which converges to the exact value in the limit $N \to \infty$. Despite being technically trivial, this methodology has two main drawbacks. The first one regards the initial conditions that one must generate extremely precisely. Secondly, the computational cost associated with this method could be considerable. Indeed, $N_{\mathcal{B}}(x)$ is not a huge number, simply by definition of rare fluctuations. While both issues could be manageable for rather simple systems, this approach becomes sooner unfeasible for more complex ones, such as climate systems. Rare events algorithms, presented in chapter 4, could overcome this limitation.

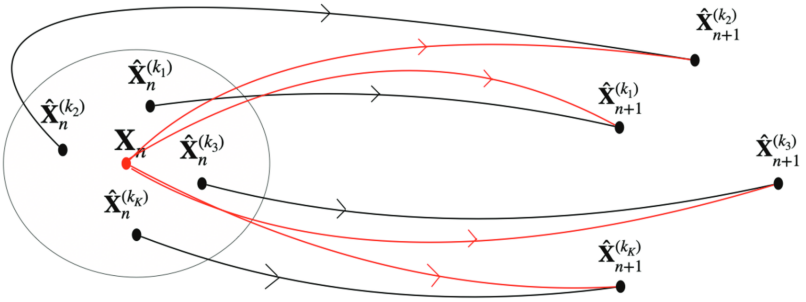


Figure 3.1: Schematic of the analogs method. On the left-hand side of the figure a point X_n is surrounded by its analogs $\{\hat{X}_n^{(k)}\}$ with $K = 4$. On the right-hand side, the observed images of the analogs are shown at one time step forward. The transitions observed in the data are represented by black lines which link the analogue with its corresponding image. Red lines are associated with possible transitions from the state X_n of the analogue Markov chain. Figure taken from (Lucente et al., 2022)

DATA-DRIVEN APPROACH In the context of committor estimation, purely data-driven approaches have been designed. The analogue method was originally proposed by Lorentz (Lorenz, 1969) as a methodology for deterministic prediction. Successive versions of it adapted this method to the context of stochastic prediction, either because the dynamics is itself stochastic or simply because the methodology itself relies on an approximation of a chaotic dynamics (Yiou, 2014; Lucente et al., 2022). A thorough investigation and study of this methodology applied to climate has been done by Dario Lucente, former PhD student of our group (Lucente, 2021) and to which I refer to for a deeper investigation of the concepts presented briefly in this section. Let's consider a dynamical process (it could be either Markovian or not, stochastic or not) $\{X(t)\}$, which takes values in the multidimensional phase space Γ. Let's assume that we observed a realization of this process $\{X_n\}_{1\leq n\leq N_t}$ at regular interval time for a total time $T = N_t\delta t$. We aim at building a Markov chain that is a data-based approximation of the initial process, based on a generalization of the Lorenz analogue method. We search among the available data the K closest neighbours of X_n, i.e. the K analogues of the dynamics, denoted as $\{X_n^{(k)}\}_{1\leq k\leq K}$. We then assume that it exists a transition between the state X_n and the images $\{\hat{X}_n^{(k)}\}_{1\leq k\leq K}$, with a certain probability. This probability could either be uniform and equal to $\frac{1}{K}$ for each analog or for instance be dependent on the distance between X_n and its analogs. Figure 3.1 shows this procedure for $K = 4$ neighbours. On the left side 4 analogs of X_n are identified and on the right side, connected via a black curve, their transition observed in the data.

Choosing K is crucial, and it is ultimately a trade-off in choosing a high value of it (to be sure to properly approximate all the possible transitions from a given state) but not too high (to select analogs that are too distant). Another fundamental choice is the proper distance, which has to be selected on a case basis. This is the basis from which one can then generate an analog Markov Chain. Using the analog Markov chain one can then estimate the committor function via a Monte Carlo direct estimation (presented above), this time having access

to more sampled committor and at a lower computational cost. An alternative and more efficient methodology has been proposed in (Lucente, Herbert, and Bouchet, 2022). A known algorithm that is based on the analog Markov chain is the Stochastic Weather Generator (Yiou, 2014). The method relies on the construction of a catalog of analogs, which are states with similar circulation patterns and other thermodynamic characteristics to generate synthetic time series which are, ultimately, dynamically similar. A recent study compares this methodology with machine learning tools, such as Convolutional Neural Networks, for the prediction of heatwaves over France and Scandinavia using the intermediate complexity model PlaSim (Miloshevich et al., 2024). The result seems to be in favour of the latter, with a better prediction skill, especially when long datasets are employed. In the context of this thesis, Convolutional Neural Networks are employed in part iii for the prediction of heatwaves over France ultimately to compare their performance with a statistical methodology. In recent years, Convolutional Neural Networks, and Deep Learning tools in general, have gathered much attention in many applications of climate science. The classic and historical application of these tools for image recognition can be in principle easily applied to the field of climate studies, where we generally deal with maps of some fields to different purposes. A recent article has summarized the current state-of-the-art of machine learning in weather prediction and climate analyses, with results coming from 500 research articles (Bochenek and Ustrnul, 2022). In the context of this thesis, Convolutional Neural Networks are used as a prediction tool with the task of estimating the committor function, for the application of prediction of extreme heatwaves over France. This is presented in part iii. This tool has already been used for prediction of extreme events, and heatwaves in particular (Miloshevich et al., 2022; Miloshevich et al., 2024; Jacques-Dumas et al., 2022). A review of applications in the field of probabilistic forecast can be found in (Miloshevich et al., 2022). Broadly, this methodology tries to extrapolate meaningful information from an already existing dataset, without relying on the generation of new data. However, it is widely known that it is a data-demanding method, which generally performs in a regime of lack of data, as shown in (Miloshevich et al., 2022). Given that it is carefully explained in the following, I redirect you to part iii for a deeper discussion.

To fix the ideas, in the next section we will illustrate these concepts on a rather simple but interesting Markov process, the Ornstein-Uhlenbeck OU.

3.1.3 *Rare events and committor function for an OU process*

To give an example of the concepts introduced in the previous sections, I present a simple case of study with a well-known process, the OU. This process is a stationary one-dimensional Markov process, which obeys the following stochastic differential equation:

$$dX(t) = \theta(\mu - X(t))dt + \sigma dW(t) \tag{3.8}$$

where $X(t)$ is the process, θ is the friction, μ is the mean, σ is the noise, $W(t)$ is a Wiener process. Originally used to model the velocity of a Brownian particle

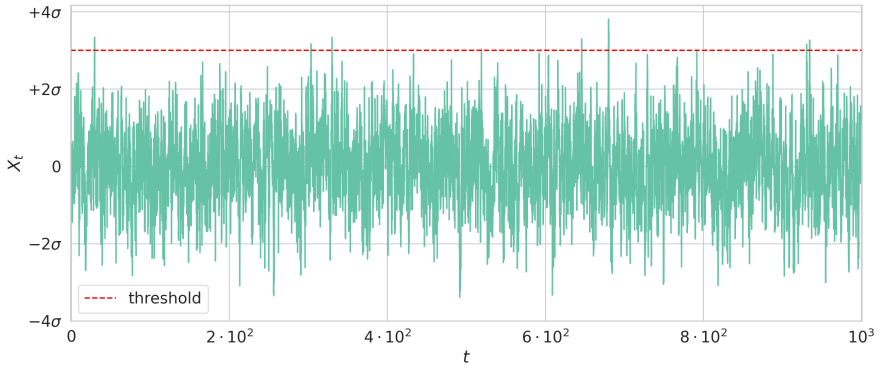


Figure 3.2: Example of realizations of a OU process, see eq. (3.8). The red line is the threshold used for defining a rare fluctuation of the stochastic process (in this case set to 3σ). We chose $\theta = 1$, $\mu = 0$ and $\sigma = 1$.

experiencing friction, this process describes a random walk with a tendency to move back towards the average value and it is widely used in financial mathematics to model interest rates, for instance in the Vasicek model (Vasicek, 1977). In this particular case, we will consider fluctuations arising from the evolution of the process itself. A realization of it is depicted in fig. 3.2. From that figure, we see that the observable fluctuates around its mean μ, which is chosen to be zero in this particular case. However, over timescales that are longer than the correlation time of the process itself, extreme fluctuations arise and they are the ones which overshoot the red dashed line in the figure (they are only 7). Equivalently, one could think of those events as the ones belonging to the queue of the probability density function of the realization of $X(t)$. Given that those events are statistically independent asymptotically because they occur at longer timescales than the correlation time of the process, they can be modeled by a Poisson point process. The good news is that we have access to the probability of those events and to their return time, i.e. the average time to wait to observe a rare fluctuation of this kind. The bad news is the quality of the estimation of either the probability or the return time. In fact, as it can be seen already in fig. 3.2, despite the long integration time, only a few events are sampled. *In se* this is not surprising, as the term *rare* fluctuations already underlines the sampling difficulty. However, this highlights the limitation of direct sampling of extreme events: we need a long integration time of the dynamics to achieve a reasonable amount of statistically independent events. This is absolutely out of reach for application in climate science. Therefore, alternative methods are needed to tackle this task. In this thesis we will explore a statistical algorithm, coupled with the simulation of the dynamics, which enables us to solve the crucial trade-off between a huge numerical cost and a significant sample of rare events: at a reasonable numerical cost, which we will detail, we sample more statistically independent extreme events. The algorithm and its application to transitions from single jet to double jet configuration in the atmosphere, the rare event of interest we studied in this thesis, is discussed in part iv.

In the context of a one-dimensional stochastic process, such as the OU, there always exists an explicit formula (Lucente, Herbert, and Bouchet, 2022) for the

committor function. For the OU with $\mu = 0$ for the sets $\mathcal{A} = (-\infty, a]$ and $\mathcal{B} = [b, +\infty)$, with $a < b$, reads as:

$$q(x) = \frac{\int_a^x \mathrm{e}^{\frac{\theta}{\sigma^2} y^2 dy}}{\int_a^b \mathrm{e}^{\frac{\theta}{\sigma^2} y^2 dy}}. \tag{3.9}$$

3.2 THEORY OF LARGE DEVIATIONS

The theory of large deviations is a theory that focuses on the asymptotic study of rare fluctuations of a given observable described by sums of random variables. After the pioneering works of Cramer (Cramér and Cramér, 1994; Cramér and Touchette, 2022), Donsker and Varadhan (Donsker and Varadhan, 1976), Freidlin and Wentzell (Ventsel' and Freidlin, 1970) and finally of Ellis (Ellis, 1984; Ellis, 1995), nowadays this theory is one of the cornerstones of statistical mechanics and an active field of research for non-equilibrium statistical physics systems. In the following, we present only the concepts which allow linking this theory with the rare events algorithm presented in part iv. We refer to (Dembo and Zeitouni, 2010; Touchette, 2009) for a deeper discussion.

Let $\{X(t)\}_{t\in\mathcal{T}}$ be an ergodic continuous Markov process, with $\mathcal{T} = [0, T]$ and $X(t) \in \Gamma = \mathbb{R}^d$. Let $A_T : \Gamma \to \mathbb{R}$ and $f : \Gamma \to \mathbb{R}$ be a real-valued random variable and a function, respectively, which depend on the path $\{X(t)\}$ via the following relation:

$$A_T := \frac{1}{T}\int_0^T f(X(t))dt. \tag{3.10}$$

For this class of variables, we know that, in the large time limit, the typical value coincides with its ergodic average $\mu := \mathbb{E}(A) = \lim_{T\to+\infty} A_T$. Moreover, by the Central Limit theorem, we know that, in the large time limit, the fluctuations around its typical value are Gaussian and scale as $\frac{1}{\sqrt{T}}$. However, in many applications we might be interested in fluctuations that are beyond the Gaussian ones, for which the scaling factor is of order T, instead of $\sqrt{T}$. To have access to these fluctuations, one could make use of the *Large deviation theory*. In particular, we will focus on the Donsker-Varadhan theory of Large deviations, commonly defined in the literature as large-time large deviations.

Recall that we want to study fluctuations of the long-time averaged observable A_T. We say that A_T follows a *large deviation principle* if the limit:

$$\lim_{T\to+\infty} -\frac{1}{T}\ln \mathbb{P}(A_T \in [a, a + da)) = I(a) \tag{3.11}$$

exists[1]. Then, $I(a)$ is called the *large deviation rate function* and T is the *large deviation rate*. An alternative, yet totally equivalent, formulation of the large deviation principle reads as:

$$\mathbb{P}(A_T \in [a, a + da)) \underset{T\to+\infty}{\asymp} e^{-TI(a)}da \tag{3.12}$$

1 Rigorously, the existence of this function should be proved for both the lim inf and lim sup over open and closed sets. However, this level of detail goes beyond the scope of this thesis and can be found at (Touchette, 2009).

where the equivalence $f_T \underset{T\to+\infty}{\asymp} g_T$ stands for a logarithmic equivalence as T goes to infinity: $\ln(f_T) \underset{T\to+\infty}{\sim} \ln(g_T)$. We are thus implicitly assuming that the prefactor in eq. (3.12) is subdominant in the limit $T \to +\infty$ with respect to the exponential decay of the rate function. Having access to the rate function I translates into having access to the probability of fluctuations of A. Establishing whether for a given variable a large deviation principle holds is the first needed step. For ergodic Markov processes it is already been shown (Chetrite and Touchette, 2015) and we will omit in this section its derivation. When eq. (3.12) holds, the second natural step is the evaluation of the rate function I, which is all but trivial. The Gärtner-Ellis theorem, presented in the next section, provides a recipe for the evaluation of the large deviation rate function I, under some conditions.

3.2.1 *The Gärtner-Ellis theorem*

Instead of working with a probability density function, it is more convenient to use an equivalent function. Let's define the Scaled Cumulant Generating Function (SCGF) as :

$$\lambda(k) := \lim_{T\to+\infty} \frac{1}{T} \ln \mathbb{E}[e^{TkA_T}] \tag{3.13}$$

for $k \in \mathbb{R}$, where

$$\mathbb{E}[e^{TkA_T}] = \int_{\mathbb{R}} e^{TkA_T} \mathbb{P}(A_T \in [a, a + da)). \tag{3.14}$$

The Gärtner-Ellis theorem states that if $\lambda(k)$ exists and is differentiable for all $k \in \mathbb{R}$, then a large deviation principle holds for A_T with rate function $I(a)$ given by the *Legendre-Fenchel* transform:

$$I(a) = \sup_{k\in\mathbb{R}}(ka - \lambda(k)). \tag{3.15}$$

The Legendre-Fenchel transform (eq. (3.15)) is a generalization of the Legendre transform (Touchette, 2005) which holds for functions which are everywhere differentiable and strictly convex. Given the before-mentioned assumptions, not all rate functions can be evaluated using this theorem (Touchette, 2009).

Equation (3.15) can be easily retrieved by noticing that, assuming that a large deviation principle holds for A_T, we could use a *saddle-point approximation* or *Laplace's approximation* for the integral in eq. (3.14)

$$\mathbb{E}[e^{TkA_T}] \underset{T\to+\infty}{\asymp} \int_{\mathbb{R}} e^{T(ka-I(a))} da \underset{T\to+\infty}{\asymp} e^{T \sup_{k\in\mathbb{R}}(ka-I(a))} \tag{3.16}$$

which implies

$$\lambda(k) = \sup_{a\in\mathbb{R}}(ka - I(a)) \underset{\text{Legendre-Fenchel transform}}{\Longleftrightarrow} I(a) = \sup_{k\in\mathbb{R}}(ka - \lambda(k)) \tag{3.17}$$

Equation (3.16) holds if one assumes that the maximum of $ka - I(a)$ exists and is unique. For ergodic Markov processes, this assumption holds. Moreover, the large deviation rate function I has a single minimum and zero, corresponding to the ergodic average of A, μ.

3.2.2 *Cramer's theorem*

Cramer's theorem is an application of the large deviation theory to the case of a sum of independent and identically distributed random variables. In the context of this thesis, it serves as a heuristic justification for the Donsker and Varadhan large deviations, the cornerstone of the rare events algorithm used in part iv.

Let X be a random variable described by a Probability Density Function $\mathbb{P}$, with mean $\mu = \mathbb{E}[X]$ and variance $\sigma^2 = \mathbb{E}[X^2] - \mu^2$. Then, let S_N be the sample mean of N realizations of X: $S_N = \frac{1}{N}\sum_{i=1}^{N} X_i$. Cramer's theorem states that S_N satisfies a large deviation principle:

$$\mathbb{P}(S_N = s) \underset{N\to+\infty}{\asymp} e^{-NI(s)} \tag{3.18}$$

Equation (3.18) actually includes both the Central limit theorem and the Law of large numbers. However, it goes beyond them, given that the fluctuations described by the large deviations principle are beyond Gaussian ones. This can be seen easily by performing a second-order expansion of the right side of eq. (3.18).

To connect with the previous section, coming back to the problem of estimation of the probability of large fluctuations of A_T (eq. (3.12)), assuming that the process $\{X(t)\}$ has a correlation time of τ_c, we might split the integral in the equation for A_T into a sum over N integrals of duration $\Delta T \gg \tau_c$:

$$A_T = \frac{1}{T}\int_0^T f(X(t))dt = \frac{1}{T}\sum_{i=1}^{T/\Delta T}\int_{(i-1)\Delta T}^{i\Delta T} f(X(t))dt \tag{3.19}$$

In this way, we are computing the sum over N identically and independently distributed random variables. Thus, for Cramer's theorem (eq. (3.18)), a large deviation principle holds for A_T:

$$\mathbb{P}\left(\frac{1}{T}\int_0^T f(X(t))dt = a\right) \underset{T\to+\infty}{\asymp} e^{-TI(a)}. \tag{3.20}$$

This result is known in the literature as large-time large deviations or equivalently as Donsker and Varadhan large deviations. In the context of this thesis, it plays a central role in the rare events algorithm, the GKLT which is used in part iv.

3.2.3 *Convergence of large deviations estimators*

This section serves as motivation for the use of rare events algorithms to sample rare and extreme events. As we have pointed out in section 3.2, the theory of Large deviations is a cornerstone when one aims at studying the statistics of rare events. A natural question which could arise is: why do not use this theory to compute empirical estimations of large deviations functions of an observable related to climate, such it could be the averaged temperature over a certain region? As for most of the cases, the evil is in the details. The large deviation scaling applies if the timescale of persistence of the rare events of

interest is large enough such that it belongs to the asymptotic regime. A minimal requirement is that the total time integration of the rare event, T, is much larger than the autocorrelation time τ_c of the time series of the rare event of interest. However, the answer strongly depends on the structure of the autocorrelation function of the time series and on the probability of the process of interest. For these reasons, the convergence of eq. (3.12) and eq. (3.13) needs to be carefully verified case by case. Following closely (Rohwer, Angeletti, and Touchette, 2015), (Ragone and Bouchet, 2020) proved that the large deviations estimators for studying rare fluctuations of the temperature of European summers do not converge. The convergence is limited by two factors. The first one is the so-called *linearization effect*: any estimation of exponential sums, such as eq. (3.14), of unbounded functions, is rapidly dominated, as $k \to +\infty$ by the largest value in the sample. Thus, the tails of the estimated SCGF become *linear* in k. Another limitation comes from the *statistical error* associated with the estimated SCGF. (Rohwer, Angeletti, and Touchette, 2015) showed that the convergence is non-uniform and there is a critical value of k after which the estimated SCGF does not converge. Moreover, in the converged region, we can have access to the statistical error only in half of it, since it is only in that part that the eq. (3.14) converges to a Gaussian distribution. Using a long climate run (Ragone and Bouchet, 2020) showed that, for the specific case of fluctuations of temperature of European summers, the minimal duration of the event to be considered to ensure the convergence of the large deviation estimators is around 3 years. However, due to the seasonal cycle of such observable, it is of interest up to a period of a season, roughly 90 days. In (Ragone and Bouchet, 2020), the authors hypothesized that this slow convergence might be due to the slow decay of the correlation of the soil moisture, which is known to play a role in maintaining extreme temperature events.

In this chapter, I presented the mathematical objects that are used later at different stages of this manuscript. Then, I presented a brief overview of the Large deviations theory, which is presented here because of its connection with the rare events algorithm introduced in the next chapter and used in part iv of this manuscript.

4

SAMPLING RARE EVENTS WITH A RARE EVENTS ALGORITHM

To introduce the difficulty of efficiently and effectively sampling rare events and the subsequent need for alternative approaches to *direct sampling*, let's consider a simple framework. Let X be a random variable defined over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the sample space, $\mathcal{F}$ is a σ-algebra, $\mathbb{P}$ is a probability measure. We denote with f the associated probability density. Our target is to estimate the probability $\gamma_{\mathcal{A}}$ of being in a region $\mathcal{A}$ of low probability in the probability distribution, i.e. $\gamma_{\mathcal{A}} = \mathbb{P}(X \in \mathcal{A})$, under the condition $\gamma_{\mathcal{A}} \ll 1$:

$$\gamma_{\mathcal{A}} = \mathbb{E}_f[\mathbb{1}_{\mathcal{A}}] = \int f(X)\mathbb{1}_{\mathcal{A}}(X)dX \tag{4.1}$$

with $\mathbb{1}_{\mathcal{A}}$ being the indicatrix function for the set $\mathcal{A}$. We aim at estimating this probability. The straightforward approach is to draw many independent realizations of X, $\{X_n\}_{1 \leq n \leq N}$ sampled from f, and count how many have hit the set $\mathcal{A}$. We will obtain a direct estimation of eq. (4.1):

$$\hat{\gamma}_{\mathcal{A}} = \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{\mathcal{A}}(X_n) \tag{4.2}$$

which has an associated relative error:

$$r_{err}(\hat{\gamma}_{\mathcal{A}}) = \frac{\sigma(\hat{\gamma}_{\mathcal{A}})}{\hat{\gamma}_{\mathcal{A}}} \approx \frac{1}{\sqrt{N\hat{\gamma}_{\mathcal{A}}}} \tag{4.3}$$

where $\sigma(\hat{\gamma}_{\mathcal{A}})$ is the standard deviation of $\hat{p}_{\mathcal{A}}$ and the inequality holds because $\hat{\gamma}_{\mathcal{A}}$ is small. A formal proof can be found in (Bucklew, 2010; Rubino and Tuffin, 2009). If $\hat{\gamma}_{\mathcal{A}}$ is of the order of 10^{-2}, for an expected relative error of 1%, N needs to be the order of 10^6. This highlights the tremendous computational cost associated with an accurate sampling of rare events and calls for alternative approaches to *direct sampling*.

Importance splitting and *importance sampling* are two families of approaches for reducing the variance associated with the computation of statistical estimators (Cérou, Guyader, and Rousset, 2019). Despite sharing the same goal, to generate more rare events, they achieve it differently. The basic idea behind the Importance splitting is to draw samples from the original probability density f in a sequential way, such that the samples that move away from the target region $\mathcal{A}$ are discarded and those who get closer to $\mathcal{A}$ are split or branched. A very popular algorithm that is based on this principle is the Adaptive Multilevel Splitting Algorithm (Cérou and Guyader, 2007; Cérou, Guyader, and Rousset, 2019; Rolland, Bouchet, and Simonnet, 2016; Bouchet, Rolland, and Simonnet, 2019; Simonnet, Rolland, and Bouchet, 2021). However, in this thesis we focus on an algorithm which is rather based on importance sampling:

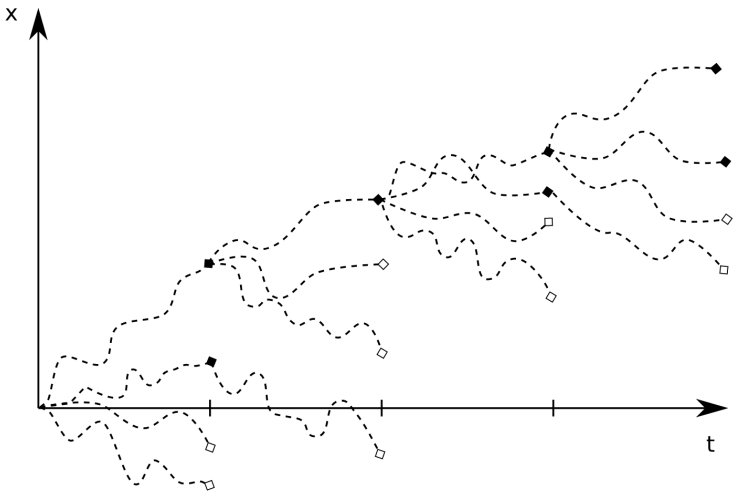


Figure 4.1: Schematic representation of a cloning algorithm with $N = 4$ particles. Figure taken from (Wouters and Bouchet, 2016).

instead of drawing samples from the original probability distribution, we sample them from an appropriately chosen different one (which is related to the original one), where the rare event set is a typical one. In the next section, we formalize this mathematically. Ultimately, both families belong to the class of *cloning algorithms*, or equivalently known in the literature as diffusion Monte Carlo (Anderson, 1975) or populations dynamics algorithms (Bouchet, Rolland, and Simonnet, 2019; Giardinà, Kurchan, and Peliti, 2006; Lecomte and Tailleur, 2007). Within this class, we are interested in the subclass of interacting particles systems (Del Moral and Garnier, 2005; Garnier and Moral, 2006). Their methodology relies on the modification of the evolution of the system dynamics in a way that the rare events responsible for large fluctuations become not rare any more. This involves the evolution of multiples copies or *clones* of the system *in parallel*, which are either replicated or killed according to a *probabilistic weight* which measures how likely each clone could exhibit a large fluctuation at the end of the simulation time. A schematic representation of this class of algorithms with $N = 4$ particles is given in fig. 4.1.

Within this class of algorithms, we focus on a particular one, called GKLT. This method was originally presented by (Giardinà, Kurchan, and Peliti, 2006) in the context of large deviations for computing rate function of an observable in discrete-time Markov chains and then rearranged in continuous-time (Lecomte and Tailleur, 2007). Despite bearing huge similarities, there is a technical detail which distinguishes the classical interacting particles algorithms of (Garnier and Moral, 2006; Del Moral and Garnier, 2005) and the GKLT: the function chosen to perform the importance sampling is for the former an instantaneous value, while an integrated value for the latter. It will become clear in the next section.

In order to make the reader familiar with the algorithm, I start with *importance sampling*. I then present the Giardina-Kurchan-Lecomte-Tailleur GKLT algorithm in section 4.2 and I explain how to retrieve statistical estimators (sec-

tion 4.2.2) and its connection to large deviation theory (section 4.2.3). An implementation of it on an OU process is given in section 4.3.

4.1 IMPORTANCE SAMPLING

In the introduction, I posed the problem of estimating the probability $\gamma_{\mathcal{A}}$ of being in a region $\mathcal{A}$ of low probability in the probability distribution f of a random variable X. An accurate direct estimation of it is costly, as shown from eq. (4.3) and alternative methodologies are needed. Following this line of thought, one could imagine drawing realisations of X from a different probability distribution $\tilde{f}$, doing what it is generally called *a change of measure*. In particular, assuming that $\tilde{f} > 0$ whenever $\mathbb{1}_{\mathcal{A}}(X)f(X) > 0$, we could write:

$$\gamma_{\mathcal{A}} = \int \mathbb{1}_{\mathcal{A}}(X)f(X)dX = \int \mathbb{1}_{\mathcal{A}}(X)\frac{f(X)}{\tilde{f}(X)}\tilde{f}(X)dX = \mathbb{E}_{\tilde{f}}[\mathbb{1}_{\mathcal{A}}(X)L(X)] \quad (4.4)$$

where $L(X) = \frac{f(X)}{\tilde{f}(X)}$ on the set $\{X : \mathbb{1}_{\mathcal{A}}(X)f(X) > 0\}$ and zero otherwise. L is called the *likelihood ratio*. Finally, what we aim is to find $\tilde{f}$ such that $\tilde{f} \gg f$ over the set $\mathcal{A}$ in order to have more realizations of X falling into $\mathcal{A}$. One can show that, the empirical average calculated over N independent realizations of X taken with respect to the density $\tilde{f}$, i.e. $\tilde{\gamma}_{\mathcal{A}} = \frac{1}{N}\sum_{n=1}^{N} L(\tilde{X}_n)\mathbb{1}_{\mathcal{A}}(\tilde{X}_n)$, is an unbiased estimator for $\hat{\gamma}_{\mathcal{A}}$. We could then compare the variances for both estimators:

$$\begin{aligned}\sigma^2(\tilde{\gamma}_{\mathcal{A}}) &= \frac{1}{N}\sigma^2_{\tilde{f}}(\mathbb{1}_{\mathcal{A}}L(X)) = \frac{1}{N}\left[\mathbb{E}_{\tilde{f}}[\mathbb{1}_{\mathcal{A}}L^2(X)] - \gamma^2_{\mathcal{A}}\right] \\ &\ll \frac{1}{N}\left[\mathbb{E}_{\tilde{f}}[\mathbb{1}_{\mathcal{A}}L(X)] - \gamma^2_{\mathcal{A}}\right] = \frac{1}{N}\left[f(X)\mathbb{1}_{\mathcal{A}}(X)dX - \gamma^2_{\mathcal{A}}\right] = \sigma^2(\hat{\gamma}_{\mathcal{A}}) \quad (4.5)\end{aligned}$$

where in the second line we used that $\tilde{f} \gg f$, otherwise the change of measure would not have sense. An important point is: the first equality in eq. (4.5) shows that if we choose $\tilde{f}(X) = \frac{f(X)\mathbb{1}_{\mathcal{A}}(X)}{\gamma_{\mathcal{A}}}$, then the likelihood ratio is $L(X) = \frac{\gamma_{\mathcal{A}}}{f(X)}$, when $f(X) > 0$, and thus:

$$\sigma^2(\tilde{\gamma}_{\mathcal{A}}) = \frac{1}{N}\sigma^2_{\tilde{f}}(\mathbb{1}_{\mathcal{A}}L(X)) = \frac{1}{N}\sigma^2(\gamma_{\mathcal{A}}) = 0 \quad (4.6)$$

This means that there is an optimal change of measure leading to a zero-variance estimator. This is true because in the new density function $\tilde{f}$ all the mass is concentrated on the rare event set, which thus becomes common. An illustration of importance sampling is shown in fig. 4.2. The bad and sad news is that to make this change of measure, we need our target variable $\gamma_{\mathcal{A}}$. However, what we have to retain is that it exists a possible (sub)optimal change of measure. Possibly many of them, waiting out there.

The class of exponential change of measure is the one which turns out to be actually useful for the rest of the discussion. Indeed, it is on that basis of measure that the rare events algorithm I am going to present works, as discussed in the next sections.

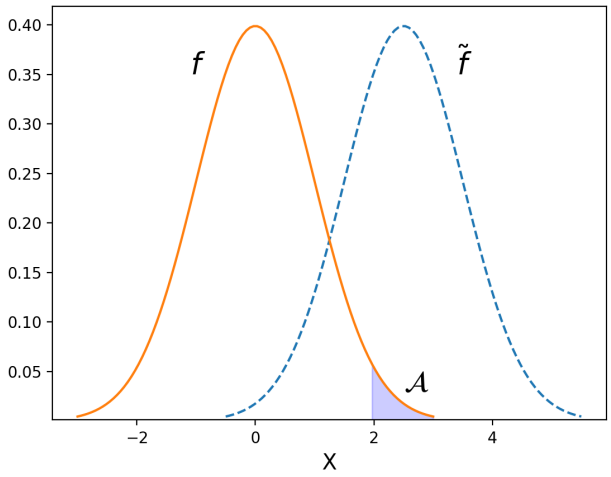


Figure 4.2: Illustration of importance sampling. Under the distribution f, sampling events in $\mathcal{A}$ is hard due to its rareness. However, a better estimation of it can be achieved when $\mathcal{A}$ is drawn from $\tilde{f}$, given that $\mathcal{A}$ becomes common under $\tilde{f}$.

4.1.1 *Importance sampling for dynamical systems*

We are interested in applying the importance sampling technique to the sampling of rare events in climate. Since climate models are non-equilibrium dynamical systems, the change of measure introduced in the section above must be done in the space of the trajectories. In general, a trajectory $\{X(t)\}_{0\leq t\leq T}$ is a realization of the dynamics of the system over a finite integration time, that we will denote as $[0, T]$. The trajectories generated by the model are distributed under an unknown Probability Density Function (PDF) $\mathbb{P}_0\left(\{X(t)\}_{0\leq t\leq T} = \{x(t)\}_{0\leq t\leq T}\right)$. This is the formal way to indicate that the probability of the model variable $\{X(t)\}_{0\leq t\leq T}$ is close to $\{x(t)\}_{0\leq t\leq T}$. Let's consider a generic observable which depends on the path $\{o[X(t)]\}_{0\leq t\leq T}$ and its integrated-time value over the duration T:

$$O_T[\{X(t)\}_{0\leq t\leq T}] = \int_0^T o[X(t)]dt \tag{4.7}$$

We know that sampling O_T is hard, due to the rareness of this event. Thus, in analogy with the previous section, we could introduce an appropriate *likelihood ratio* that gives higher weights to higher values of O_T. This is the core of the GKLT algorithm. For this algorithm, the likelihood ratio consists of an exponential change of measure (Ragone and Bouchet, 2020). In practice:

$$\mathbb{P}_k\left(\{X(t)\}_{0\leq t\leq T} = \{x(t)\}_{0\leq t\leq T}\right) = L_k[\{X(t)\}_{0\leq t\leq T}]\mathbb{P}_0\left(\{X(t)\}_{0\leq t\leq T} = \{x(t)\}_{0\leq t\leq T}\right) = \frac{\exp(k\int_0^T o[X(t)]dt)}{\mathbb{E}_0\left[\exp(k\int_0^T o[X(t)]dt)\right]}\mathbb{P}_0\left(\{X(t)\}_{0\leq t\leq T} = \{x(t)\}_{0\leq t\leq T}\right) \tag{4.8}$$

where $\mathbb{E}_0[.]$ denotes an ensemble average with respect to the stationary path measure $\mathbb{P}_0$. The likelihood ratio L_k depends on the parameter k which controls

the weights given to the high values of O_T. More details on how to arrive to this ratio are given in the following sections.

4.2 THE GIARDINA-KURCHAN-LECOMTE-TAILLEUR ALGORITHM

In the introduction to this chapter, I presented the GKLT algorithm as a rare events algorithm that aims at better sampling, both from the computational cost and from the quality of the sampling, rare events (Giardinà, Kurchan, and Peliti, 2006; Lecomte and Tailleur, 2007; Giardina et al., 2011). This approach has subsequently been applied to a wide variety of contexts (Nemoto et al., 2016; Hidalgo, 2018; Lecomte and Tailleur, 2007). The application of this cloning algorithm to the field of climate science is quite recent and fascinating. The first application in this sense has been made by (Ragone, Wouters, and Bouchet, 2018) to study seasonal heatwaves over Europe using a climate model in perpetual summer conditions. With this methodology, the authors could sample extreme European warm summers with a return time of $10^6 - 10^7$ years with an equivalent computational cost of 10^3 years, which is absolutely unfeasible to sample with a direct simulation approach. Moreover, thanks to the enriched sampling of the rare events, statistically significant areas are wider. Successive works using more realistic climate models, just as the one used for the study presented later in this manuscript, confirmed those results for heatwaves over France and Scandinavia (Ragone and Bouchet, 2021). Following their works, other extreme events have been analyzed using this algorithm, from precipitation (Wouters, Schiemann, and Shaffrey, 2023) to extreme fluctuations of the energy power system (Cozian, 2023), to collapsing of the Atlantic Meridional Overturning Circulation (Cini et al., 2024), to turbulent flows (Lestang, Bouchet, and Lévêque, 2020), to the melting of the Arctic sea ice (Sauer et al., 2024) and finally, to Indian extreme warm summers (Le Priol, Monteiro, and Bouchet, 2024).

In the next section (section 4.2.1) I present the details of the algorithm, how to retrieve the statistics of some quantities of interest, and finally how it is related to the large deviation theory. I refer to (Ragone and Bouchet, 2020) for a deeper description of the algorithm applied to climate science and to (Pérez-Espigares and Hurtado, 2019) for a review on cloning algorithms.

4.2.1 *Description of the algorithm*

The GKLT algorithm is a cloning algorithm which aims at sampling large fluctuations of a time-averaged observable. To achieve this, it relies on the evolution of a set of multiple trajectories run in parallel which are either cloned or killed according to a probabilistic weight which ultimately measures the likelihood of each trajectory to exhibit a rare fluctuation of the observable of interest at the end of the simulation. To create a link with the importance sampling, the cloning and killing procedure acts as an effective likelihood ratio $L(X)$ which allows to sample of the targeted observable from a biased measure $\mathbb{P}_k$ where its large fluctuations are common. In the next, I formalize this mathematically. In this thesis, I run 10 GKLT experiments in order to collect more realizations of

the target observable. However, to keep the discussion easier and more general, I describe in the following the methodology for only one experiment with a target variable O with N trajectories which evolve for a total simulation time of T.

Let's assume that we want to study the fluctuations of a generic time-integrated observable, which depends on the path $\{o[X(t)]\}_{0\leq t\leq T}$:

$$O_T[\{X(t)\}_{0\leq t\leq T}] = \int_0^T o[X(t)]dt. \tag{4.9}$$

where T is the simulation time. The algorithm simulates the evolution of an ensemble of N interacting trajectories of the dynamical system $\{x_n(t)\}_{n=1..N}$ in parallel, from time $t = 0$ up to time $t = T$.

The algorithm can be divided into 3 steps: initialization, integration and, resampling of the trajectories.

INITIALIZATION. The ensemble of N independent trajectories are initialized at time $t_0 = 0$. The trajectories start from N independent initial conditions to sample the model's invariant measure. Each of them carries a weight equal to 1 at this stage. Without a rare event algorithm, this weight is constant and equal for all the trajectories during the simulation.

The next two steps are repeated cyclically until the simulation has reached the chosen integration time T.

INTEGRATION. Simulate the model's dynamics forward in time, until $t_i = i\tau$, $i = 1, 2 \ldots, T/\tau$ where τ is the *resampling time*. Then, for each trajectory $n = 1, \ldots, N$ compute the score variable $O_T[X_n(t_i)]$. Compute the weights with the new integrated value just obtained $w_i^n = \exp\left(k \int_{(i-1)\tau}^{i\tau} o[x_n(t)]dt\right)$, where k is the *biasing parameter* and its role will be elucidated in the following. Finally normalize the weights such that they sum to N:

$$W_i^n = \frac{w_i^n}{Z_i} \quad \text{where} \quad Z_i = \frac{1}{N}\sum_{n=1}^{N} w_i^n. \tag{4.10}$$

where Z_i is a normalizing factor. Note that it is equivalent to the partition function in statistical mechanics.

RESAMPLING. Once the weights are computed, we can move to the cloning and killing procedure. In particular, we select a candidate number of clones for each trajectory $M_i^n = \lfloor W_i^n + U_i^n \rfloor$, where U_i^n is drawn independently from a uniform distribution $\mathcal{U}(0,1)$ and $\lfloor \bullet \rfloor$ is the floor of $\bullet$. To avoid computational issues, we add another step and resize the ensemble members such that they stay constant to N: compute $\Delta N_i = \sum_{n=1}^{N} M_i^n - N$. If $\Delta N_i > 0$, kill randomly and without repetition ΔN_i clones, while if $\Delta N_i < 0$ clone randomly and with repetition $|\Delta N_i|$ clones. For our application, because a climate model is a deterministic dynamical system, after the first resampling step, we add a small perturbation to the cloned trajectories to let them separate enough in the subsequent integration step. Because of the resampling procedure, where a trajectory

can be resampled originating many copies, several trajectories share the same ancestors at the end of the simulation. This aspect will be discussed in the following.

Once the trajectories have reached $t = T$, we are able to back-reconstruct each of the surviving trajectories by attaching their parent at each resampling time t_i. This results in an ensemble of real solutions of the climate model, without any change in the dynamics, but only in the probability of observing them.

A pseudocode of the algorithm is given below. In the next section, we show how to retrieve the statistics of any observable of interest from the GKLT experiment and its relation with large deviation theory.

Algorithm 1 GKLT algorithm

Initialize N trajectories $\{X_n(0)\}$, with $1 \leq n \leq N$ from N independent initial conditions.

for i from 1 to T/τ **do**

for n from 1 to N **do**

Add a small random perturbation to each trajectory n

Compute the dynamics for the n-th trajectory from $t_{i-1} = (i-1)\tau$ to $t_i = i\tau$.

Assign to trajectory n the unnormalized weight $w_i^n = \exp\left(k \int_{(i-1)\tau}^{i\tau} o[x_n(t)]dt\right)$.

end for

Compute the normalization constant and normalize the weights:

$$W_i^n = \frac{w_i^n}{Z_i} \quad \text{where} \quad Z_i = \frac{1}{N}\sum_{n=1}^{N} \exp\left(k \int_{(i-1)\tau}^{i\tau} o[x_n(t)]dt\right) \tag{4.11}$$

for n from 1 to N **do**

Compute the number of copies produced by each trajectory n as

$$M_i^n = \lfloor W_i^n + U^n \rfloor \tag{4.12}$$

where $\lfloor \bullet \rfloor$ is the integer part of $\bullet$ and U^n is a random number drawn uniformly in $[0;1[$.

end for

Compute $\Delta N_i = \sum_{n=1}^{N} M_i^n - N$. If $\Delta N_i > 0$, kill randomly ΔN_i trajectories (without repetition). If $\Delta N_i < 0$, clone randomly $|\Delta N_i|$ trajectories (with repetition).

end for

4.2.2 *Statistics of the Giardina-Kurchan-Lecomte-Tailleur algorithm*

We now show that the path measure $\mathbb{P}_k$ in the rare events algorithm is tilted with respect to $\mathbb{P}_0$ to favour large values of $O_T[\{X(t)\}_{0\leq t\leq T}] = \int_0^T o[X(t)]dt$. Let's assume to stop the simulations after two resampling steps and to con-

sider a trajectory that has not been killed. We could compute the probability of observing it by back-propagating:

$$\mathbb{P}_k(\{x(t)\}_{0\le t\le 2\tau}) = W_1^n W_2^n \mathbb{P}_0(\{x(t)\}_{0\le t\le 2\tau}) =$$

$$= \frac{\exp\left(k\int_0^\tau o[x(t)]dt\right)}{Z_1}\frac{\exp\left(k\int_\tau^{2\tau} o[x(t)]dt\right)}{Z_2}\mathbb{P}_0(\{x(t)\}_{0\le t\le 2\tau}) \quad (4.13)$$

where

$$Z_i = \frac{1}{N}\sum_{n=1}^{N}\exp\left(k\int_{(i-1)\tau}^{i\tau} o[x_n(t)]dt\right) \quad \text{with } i = 1,2. \quad (4.14)$$

Proceeding similarly until the last integration time, one obtains the equation already presented in eq. (4.8):

$$\mathbb{P}_k\left(\{X(t)\}_{0\le t\le T} = \{x(t)\}_{0\le t\le T}\right)$$
$$\underset{N\to+\infty}{\sim} \frac{\exp(k\int_0^T o[X(t)]dt)}{\mathbb{E}_0\left[\exp(k\int_0^T o[X(t)]dt)\right]}\mathbb{P}_0\left(\{X(t)\}_{0\le t\le T} = \{x(t)\}_{0\le t\le T}\right) \quad (4.15)$$

which shows that the tilting of the probability favours large values of the integrated observable we targeted. In the previous formula we have used the mean-field approximation of the partition function Z, which is valid for a large number of initial trajectories:

$$Z(k,T) = Z_1 Z_2 \dots Z_N \underset{N\to+\infty}{\sim} \mathbb{E}_0\left[\exp\left(k\int_0^T o[X(t)]dt\right)\right]. \quad (4.16)$$

By inverting equation eq. (4.15) we can obtain an unbiased estimator for a generic observable $A[\{X(t)\}_{0\le t\le T}]$:

$$\mathbb{E}_0[A[\{X(t)\}_{0\le t\le T}]] \underset{N\to+\infty}{\sim} \frac{1}{N}\sum_{n=1}^{N} p_n A[\{x_n(t)\}_{0\le t\le T}] \quad (4.17)$$

where

$$p_n = \frac{1}{N}\exp\left(-k\int_0^T o[x_n(t)]dt\right)\prod_{i=1}^{T/\tau} Z_i \quad (4.18)$$

with x_n being the n-th reconstructed trajectory. Choosing $\mathbb{1}(\frac{1}{T}\int_0^T o[X(t)]dt > a)$, we could have an estimation of $\mathbb{P}_0\left(\frac{1}{T}\int_0^T o[X(t)]dt > a\right) = \mathbb{E}_0\left[\mathbb{1}\left(\frac{1}{T}\int_0^T o[X(t)]dt > a\right)\right]$ using eq. (4.17). Computing this probability (and other conditioned statistics) is the final scope of using this algorithm.

Given that the trajectories are not independent after the first resampling step, the central limit does not apply. However, it has been shown (Del Moral, 2004) that for a class of algorithms, including the GKLT, the errors for estimating variables linked to the score function used to run the algorithm are of order $1/\sqrt{N}$.

4.2.3 *Computation of large deviation rate functions*

Originally, this algorithm was designed for computing the large deviation function $I(a)$, introduced in section 3.2, or equivalently its Legendre-Transform the Scaled Cumulant Generating Function SCGF $\lambda(k)$. Looking back to the equation for $\mathbb{P}_k$ eq. (4.15), we can obtain an estimation of $\lambda(k)$, in the limit of both large number of trajectories N and integration time T:

$$Z(k,T) = Z_1 Z_2 \dots Z_N \underset{N\to+\infty}{\sim} \mathbb{E}_0\left[\exp\left(k\int_0^T o[X(t)]dt\right)\right] \underset{T\to+\infty}{\sim} \exp\left(T\lambda(k)\right) \tag{4.19}$$

where we have assumed that a large deviation principle holds section 3.2. Inverting the previous formula, we can compute the SCGF:

$$\lambda(k) = \lim_{T\to+\infty} \frac{1}{T} \ln Z(k,T). \tag{4.20}$$

Assuming that the Gartner-Ellis theorem's assumptions are met (see section 3.2), we can Legendre-Fenchel transform $\lambda(k)$ to obtain the large deviation rate function $I(a)$ and finally have access to the probability of observing these rare fluctuations of the time and spatial average O_T/T.

As a final remark, from eq. (4.15) it is possible to show that for a chosen value of k the typical value of the large time average of the observable is

$$\begin{aligned}&\mathbb{E}_k\left[\lim_{T\to+\infty}\frac{1}{T}\int_0^T o[X(t)]dt\right] \underset{N\to+\infty}{\sim} \\ &\lim_{T\to+\infty}\mathbb{E}_0\left[\left(\frac{1}{T}\int_0^T o[X(t)]dt\right)\frac{\exp(k\int_0^T o[X(t)]dt)}{\mathbb{E}_0\left[\exp(k\int_0^T o[X(t)]dt)\right]}\right] = \lambda'(k).\end{aligned} \tag{4.21}$$

Because in the large deviations regime the typical value and the most probable value coincide, the time average of the time-integrated observable fluctuates in the cloned ensemble around a value that depends on the derivative of the SCGF in a neighbourhood of k, which are shown to be the right range of fluctuations for computing large deviation functions (Rohwer, Angeletti, and Touchette, 2015).

In applications that are less computationally expensive than running a climate system, tuning the values of k would result in a point-wise reconstruction of the SCGF, to obtain the large deviation rate function $I(a)$. However, this is unfeasible for a climate model due to the tremendous resources and time needed. Consequently, some prior analysis needs to be performed in order to choose the correct order of magnitude of the parameters for running the algorithm and this analysis is obviously case-dependent. The choice of the parameters used in the algorithm is presented in section 4.4, and it is necessarily guided by the application of this algorithm to the sampling of double jet events, presented in part iv. In the next section I present an illustration of this algorithm to a one-dimensional stochastic process, the Ornstein–Uhlenbeck OU.

4.3 ILLUSTRATION ON THE ORNSTEIN–UHLENBECK PROCESS

We illustrate the performance of the algorithm on a simple process, the Ornstein–Uhlenbeck OU, which was already introduced in section 3.1. Let us recall the equation for the OU process, in which a particle evolves following the stochastic equation:

$$dX(t) = \theta(\mu - X(t))dt + \sigma dW(t) \tag{4.22}$$

where $dW(t)$ is a Weiner process. In the following, we set $\mu = 0$, $\theta = \sigma = 1$. We are interested in large time large fluctuations of the stochastic process itself, namely of :

$$X_T = \frac{1}{T}\int_t^{t+T} X(t)dt \tag{4.23}$$

where T is the total computation time. Solving the associated Fokker-Planck equation, it can be shown that X_T behaves as a Gaussian variable, with mean $\mu = 0$ and standard deviation σ_T. We perform a rare event algorithm run with $N = 5 \times 10^2$ trajectories, resampling time $\tau = 1$, biasing parameter $k = 0.5$, and total integration time of $T = 40\tau$. The score function is the stochastic process itself, i.e. $o[X(t)] = X(t)$. In fig. 4.3 we show the results. In fig. 4.3a we observe that the algorithm has performed the importance sampling of X_T as expected, and events which are rare in the control (or direct) simulation, such as $2\sigma_T$, are common in the rare events algorithm one. Moreover, the algorithm enables to have access to events which were unseen in the control distribution. Another way of observing that is by computing the return time curve (fig. 4.3b), which gives an estimate of the time to wait to see an event of amplitude $X_T = a$. With a computational time which is 100 times lower, we have access to the events with the same return time. In the same figure, with the same computational cost (obtained by simulating $N = 5 \times 10^4$ trajectories) we achieve a return time which is orders of magnitude higher than with the control or direct simulation. The illustration of the GKLT on this process has been carefully investigated by (Wouters and Bouchet, 2016).

For this stochastic process, we have access to an analytical formula for the Scaled Cumulant Generating Function SCGF. In fact, in the limit of large time, $\sigma^2 \underset{T\to\infty}{\sim} \frac{2}{T}$ and

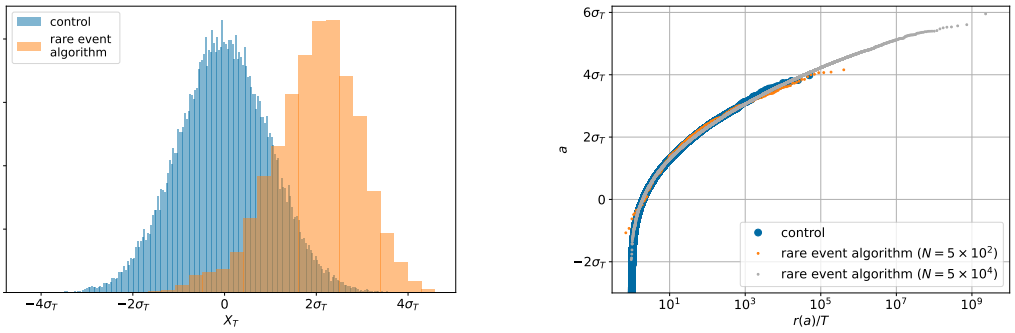
$$\mathbb{P}(X_T = a) \underset{T\to\infty}{\asymp} \exp\left(-\frac{Ta^2}{4}\right). \tag{4.24}$$

It is straightforward to recognize in the exponential the rate function introduced in eq. (3.15):

$$I(a) = \frac{a^2}{4} \tag{4.25}$$

form which, following the Gartner-Ellis theorem section 3.2, one can retrieve the SCGF:

$$\lambda(k) = \sup_a[ka - I(a)] = k^2. \tag{4.26}$$



(a) (b)

Figure 4.3: (a) Histogram and (b) return time curve of the control simulation (blue) and of a rare event algorithm run (orange) with a lower computational cost than the control, while in gray there is a rare event algorithm simulation with the same computational cost.

Given the not-so-huge computational cost associated with the simulation of this process, one can then reconstruct the SCGF by performing several runs of the GKLT with different values of k. This analysis has been deeply investigated by (Lestang, 2018), who showed that for a similar computational cost to direct simulation, the GKLT algorithm provides a far better estimate of the rate function.

4.4 PARAMETERS ESTIMATION

The GKLT algorithm presented above relies on a set of parameters which need to be carefully chosen for its correct implementation. For the peculiar case of study presented in this thesis, some of them are chosen accordingly to achieve a trade-off among available resources, computation time and correctness of the event's representation. Moreover, for some of them, the choice has been guided from the setup of CESM, briefly presented in section 2.4.

4.4.1 *Choice of the observable, resampling time τ and integration time T*

A fundamental choice for running a GKLT algorithm is the one of the observable. In (Chraibi et al., 2018) the authors showed that the optimal potential score function, the one which minimizes the asymptotic variance of the estimator of this cloning algorithm, is strictly connected to the committor function (see section 3.1). In this particular application of this algorithm, we do not aim at doing prediction, but rather at coupling the rare events algorithm with a climate model, CESM, to enhance the sampling of the upper tail of a time-averaged observable. Previous studies have shown that a score function that poorly samples new trajectories leads to the failure of the algorithm (Lestang, 2018).

The resampling time τ should be of the order of the Lyapunov time (Wouters and Bouchet, 2016). If we indicate as τ_c the correlation time of the system, then if $\tau \ll \tau_c$, cloned trajectories likely will not be separate enough to be a potential candidate for a large fluctuation of the observable of interest. Furthermore,

this procedure would lead to a loss of diversity of the trajectories at successive resampling times. On the other hand, τ should not be too large, i.e. $\tau \gg \tau_c$, because this would lead to the loss of the memory of the trajectories before the selection procedure, thus leading to a poor sample of the large long-time average fluctuations. Tests on simpler models (Lestang et al., 2018) have shown that the exact value of τ is not relevant as far as it is of the order of the autocorrelation time.

Finally, for the integration time T: in order to let the potential candidate trajectories yield to the extreme event we need to have $T \gg \tau_c$, but not too large in order not to eventually kill in late resampling steps the trajectories which showed a large fluctuation before and went back to typical values.

4.4.2 *The number of trajectories N*

The results presented in section 4.2.1 are valid asymptotically when the number of trajectories, i.e. the copies of the climate system, tends to infinity. This has an important implication in the expected errors of when computing averages of dynamical observables, as they scale as $1/\sqrt{N}$.

Ultimately, the choice of the number of trajectories has been mainly guided by the available resources. The current setup of the climate model CESM needs 128 cores to run for a single trajectory. Thus, it would be simply not manageable to use $128 \times N$ cores when N is too large. Moreover, the cluster where the climate model is run has a limit for the job duration of 24 hours, which restricts further our choice of N. Finally, we decided to simulate $N = 100$ trajectories. In order to counterbalance this limitation, we run several independent experiments to improve the statistics and estimate the variance of the quantity of interest from one experiment to another.

For a given selection strength, controlled by the parameter k, the degeneration of the trajectories might be important. This means that a trade-off between the level of extremity and the simulated clones needs to be found.

4.4.3 *The biasing parameter k*

In the large deviation regime, the relation eq. (4.21) might be used to find the best value for the biasing parameter k. However, this equation holds in the limit of $T \to \infty$ and we do not have an analytical formula for SCGF, which is in principle the reason why we are performing rare events simulations. The better estimate we could achieve is to assume that the SCGF is Gaussian distributed and derive a rough approximation of k from it. Let us consider the time-averaged observable we wish to sample in the rare events algorithm: $O_T[\{X(t)\}_{0\leq t\leq T}] = \frac{1}{T}\int_0^T o[X(t)]dt$ to be distributed as a Gaussian distribution $O_T \sim \mathcal{N}(\mu, \Sigma^2)$, where μ is the ergodic mean and $\Sigma^2 = 2\tau_c\sigma^2$ is the variance, where σ^2 and τ_c are respectively the variance and the autocorrelation time of the observable $o[\{X(t)\}$. Recall that the equation for the SCGF is given by eq. (3.13),

$\lambda(k,T) = \frac{1}{T}\log\mathbb{E}[\exp(kTO_T)]$. Then we can evaluate the derivative of SCGF:

$$\partial_k\lambda(k,T) = \frac{\mathbb{E}[O_T\exp(kTO_T)]}{\mathbb{E}[\exp(kTO_T)]} = \mathbb{E}_k[O_T] \tag{4.27}$$

which strongly resembles the exponential tilting used in the rare events algorithm, might be approximated by:

$$\partial_k\lambda(k,T) = \mu + kT\Sigma^2 \tag{4.28}$$

To obtain a value for k one has to combine both previous formulae, with a desired value of the time-integrated observable $\mathbb{E}_k[O_T] = a$. Finally:

$$k = \frac{a-\mu}{\Sigma^2} \tag{4.29}$$

The higher the value of k the more stringent the selection is, the more degeneration of the trajectories will be observed, causing in the worst case the phenomenon of ensemble collapse or extinction, i.e. all the trajectories share the same ancestor. This further shows how the finite size of the ensemble limits the rareness of the events one can sample.

4.4.4 *Perturbation of trajectories*

Given that our climate model is deterministic, in order to have an effective separation of the trajectories after each resampling, we need to introduce a small random perturbation after each cloning step $i\tau$. As discussed in (Ragone, Wouters, and Bouchet, 2018), each spherical harmonics of the potential temperature is multiplied by a factor $(1+\epsilon\,\mathrm{r})$, where r is a randomly uniform number between $[-1,1]$ and $\epsilon_{pert} = 10^{-4}$. Thus, each spherical harmonic receives a perturbation proportional to its amplitude. Moreover, the procedure is applied at each vertical level, using the same value of r at each level for a given spherical harmonic. This perturbation is expected to alter the dynamics of about the same order of magnitude, thus resulting in a perturbation that is much lower than the sampling error.

4.4.5 *Initial conditions*

The initial conditions from which we start the algorithm must be independent, in order to avoid unintended biases. This means that, if the simulation starts on June 1st, and we chose $N = 100$ trajectories, we need 100 independent June 1st, the starting date of the algorithm, one for each trajectory. Given that the model we use CESM has only atmosphere and land as active components, we collected these initial conditions from a long simulation that was already available, selecting consecutive June 1st. This procedure could have not been used if the ocean was active in the model, as the relaxation timescales of this component, similar to the ice one, are much longer than one year.

4.4.6 *Computational cost*

An interesting point of using the rare events algorithm is its lower computational cost with respect to an equally long climate simulation. In order to compare the two, a single rare event algorithm run depends on the number of trajectories N and the integration time T. Indeed, we can approximate its cost to $N \times T$. Furthermore, we must distinguish if the initial conditions are already available or not. Assuming they are available and that we simulated three months, say the summer months of June-July and August, for $N = 100$ trajectories, then the cost of a single rare event algorithm run is of 25 years (for each trajectory we only simulated 3 months and not the full year). If we then performed 10 runs, the overall cost for having an equivalent of 1000 summers is of 250 years, which is 4 times lower compared with the 1000 years of direct climate simulation.

In this initial chapter, I presented: the drivers of heatwaves at mid-latitudes, how they are expected to change in response to climate change and some of the models used in the following; some mathematical objects intensively used in this manuscript, and the theory of large deviations; finally, the problem of sampling rare events and the rare event algorithm used in part iv.

The next chapter consists of a study that compares the relative contribution of two slow drivers of heatwaves in Europe. Despite showing both drivers a significant contribution to the heatwave amplitude and duration for typical heatwaves, the results for heatwaves with a return time of a decade are limited because of a lack of data. These conclusions are made more precise in the next chapter.

Part II

COMPARING THE INFLUENCE OF ATLANTIC MULTIDECADAL VARIABILITY AND SPRING SOIL MOISTURE ON EUROPEAN SUMMER HEATWAVES

In this chapter we study and compare the influence of two slow drivers, namely the Atlantic Multidecadal Variability AMV and the spring soil moisture on the duration and intensity of European heatwaves. As we discussed in section 2.2.3, land-atmosphere feedbacks act on the onset and persistence of heatwaves especially in some mid-latitude areas, such as in the Mediterranean basin. Moreover, at interdecadal timescales the sea surface temperature variability of the Atlantic Ocean, strictly connected to the AMV, has been shown to play a role on some climate extremes in different regions on the world, including heatwaves in Europe. I presented this driver in section 2.1. However, which one between the two is a driver of greater importance for European heatwaves ? Using results from global climate model experiments, we show that for heatwaves with return times of a few years spring soil moisture in the Southern Europe has more impact than the AMV on summer European heatwaves. For heatwaves with return times of a decade, this still holds true. However, significant areas are largely limited by the available data. This highlights the lack of data problem faced when dealing with rare and extreme events.

For what concerns the tools used in this chapter, we employ *return time* and *composite maps* for the estimation of the average time to wait to observe an event with a certain amplitude and for the analysis of the dynamical fields after the heatwave event happened, respectively. Both tools were introduced in section 3.1.1.

The material presented in this chapter is an article which has been submitted to the journal Oxford Open Climate Change. I coauthor this study with Clément Le Priol[1] and Fabio D'Andrea[1]. A preprint is available at `https://arxiv.org/abs/2405.10821`. Because of that some concepts might be repeated.

1 *CNRS, LMD/IPSL, ENS, Université PSL, École Polytechnique, Institut Polytechnique de Paris, Sorbonne Université, Paris France*

Abstract. In this work, we study and compare the influence of the Atlantic Multidecadal Variability (AMV) and spring soil moisture in Southern Europe on the duration and intensity of European summer heatwaves. We study heatwaves with return times of a few years and also propose a new methodological approach, *return time maps*, that allows us to study rare heatwaves with return times of 10 years.

We use the outputs from three climate models, namely IPSL-CM6A-LR, EC-Earth3, and CNRM-CM6-1, in which North Atlantic sea surface temperatures are restored towards the observed AMV anomalies. The three models give consistent results, with the exception of EC-Earth simulating a much greater influence of soil moisture.

Spring soil moisture in Southern Europe is a slow driver of greater importance than the AMV for European summer heatwaves, both in terms of the extension of the region of influence and in terms of amplitude. While the influence of the AMV concentrates over the very south of Europe, around the Mediterranean Basin, spring soil moisture influence extends over large parts of Europe. As might be expected, a positive AMV phase or low soil moisture generally induces hotter and longer heatwaves. However, the models suggest that the AMV also induces fewer heatwave days and cooler heatwaves around Poland.

For more extreme events, the influences of the AMV and soil moisture increase, according to regional patterns that seem to be the same as for typical heatwaves. However, confirming this statement would require datasets with more extreme events.

LAY SUMMARY Beyond the daily fluctuations of the weather, the duration and intensity of heatwaves can be modulated by slow drivers. In this work, we study and compare the influence of two slow drivers on the duration and intensity of summer heatwaves in Europe. The first driver is a slow mode of variation of the North Atlantic Ocean sea surface temperature called the Atlantic Multidecadal Variability (AMV). The second one is the quantity of water available in the soil in spring in Southern Europe.

We study typical heatwaves that occur almost every year, but we also introduce a new method to study rare heatwaves that occur only every 10 or 50 years, on average.

Using results from global climate model experiments, we find that a positive AMV phase or low soil moisture generally induces hotter and longer heatwaves, as could be expected. Our main result is that soil moisture is a slow driver of greater importance than the AMV. Indeed, its influence extends over a larger part of Europe and has more amplitude.

5

INTRODUCTION

In a changing climate, extreme hot events are becoming more frequent and intense (Hansen, Sato, and Ruedy, 2012; Dunn et al., 2020; Seneviratne et al., 2021). The impacts of these events are detrimental on many levels, causing damage to our society, the environment, and other living beings (Pörtner et al., 2023). Europe, and especially the Mediterranean basin, is identified as a hot spot for extreme hot events, with magnitudes changing according to future climate scenarios (see (Masson-Delmotte et al., 2023) and references therein). In particular, heat extremes in Western Europe have warmed at a faster pace than elsewhere in the mid-latitudes (Rousi et al., 2022; Vautard et al., 2023).

Besides global warming, different physical drivers influence the formation of heatwaves, acting on different timescales (Perkins, 2015; Horton et al., 2016). The fastest driver is the atmospheric circulation, which can cause heatwaves through persistent high-pressure synoptic systems. The associated time scale corresponds to a few days, which is the typical duration of a heatwave. Slower drivers modulate the occurrence and frequency of heatwaves, acting on seasonal, yearly, and multidecadal time scales (Perkins, 2015).

Regarding summer European heatwaves, an important seasonal driver is soil moisture. A soil moisture deficit in the Mediterranean basin at the beginning of summer has been shown to act as a precondition for some extreme events such as droughts (Vautard et al., 2007; Zampieri et al., 2009) and heatwaves (Fischer et al., 2007a; Fischer et al., 2007b; Alexander, 2011; Materia et al., 2022) over continental Europe. The mechanism is as follows: dry and warm air masses form over the dry soils of the Mediterranean and induce diminished cloudiness. These air masses are advected northward by southerly wind episodes, increasing temperature and evaporative demand over Europe, which in turn leads to drier soils. These drier soils amplify the warming through higher sensible heat emissions and favored upper-air anticyclonic circulation (Zampieri et al., 2009).

Sea surface temperature (SST) anomalies are another slow driver of heatwaves, acting on different timescales. At the seasonal timescale, SST anomalies in the North Atlantic can favor heatwaves through their influence on large-scale atmospheric modes like the NAO (Feudale and Shukla, 2011; Duchez et al., 2016; Wulff et al., 2017; Beobide-Arsuaga et al., 2023). At the multidecadal timescale, SST anomalies are modulated by an internally-driven low-frequency mode of variability known as the Atlantic Multidecadal Variability (AMV). The AMV has been shown to influence the duration of heatwaves in Europe (Sutton and Dong, 2012; Qasmi, Cassou, and Boé, 2017; Qasmi et al., 2021) and to play a role in the occurrence of other extremes, such as droughts and precipitations, in other parts of the globe (Sutton and Hodson, 2005; Ruprich-Robert et al., 2017; Ruprich-Robert et al., 2018).

The Atlantic Meridional Overturning Circulation (AMOC) has been identified as a prominent driver of the AMV (Knight et al., 2005; Yeager and Robson,

2017; Cassou et al., 2018). On the atmospheric side, recent studies have underlined how persistent North Atlantic Oscillation (NAO) can influence the thermohaline circulation and thus the AMV (Delworth et al., 2016), also at these longer timescales.

While previous work has examined the influence of each driver separately, estimating the relative influence of each driver remains an open question. In this paper, we chose to compare the influence of two of these main drivers, namely spring soil moisture and the AMV, on the duration and intensity of European summer heatwaves.

A second question concerns the rarity of the events studied. Heatwaves studied in previous work on the AMV (Ruprich-Robert et al., 2018; Qasmi et al., 2021) actually occur every one or two years. Therefore, they are not such rare events. However, the most harmful events are the largest and rarest ones (Robine et al., 2008). This calls for focusing on rarer, more extreme events. In this paper, we propose a new tool, *return time maps*, to study the influence of the AMV and spring soil moisture on rare European summer heatwaves with return times of a few decades.

We use two definitions of heatwaves, which are complementary. The first one measures the number of heatwave days per year, based on the exceedance of two temperature thresholds (Lau and Nath, 2012; Qasmi et al., 2021). In the second definition, we fix the duration of the events studied and measure their intensity, characterized by the mean temperature anomaly during the event.

The paper is organized as follows. In chapter 6, we present the data used for this study and the two definitions of heatwaves that we use. In chapter 7, we compare the effects of the AMV and spring soil moisture on heatwaves with return times of a few years. In chapter 8, we introduce a methodology to study rarer events with return times of a few decades and assess which driver, between the AMV and spring soil moisture, has the strongest influence on these rare heatwaves. We summarize our findings in chapter 9 and discuss them in a broader perspective in chapter 10.

6 DATA AND METHODS

In this section, we present the data used for this study and the methodology to compare the influence of the AMV and spring soil moisture on heatwaves independently. We then introduce the two different heatwave definitions that we consider in this study.

6.1 DATA

6.1.1 *DCPP-AMV experiments*

Given the relevance of the effects of the AMV on the climate and the potential predictability associated with it, there is ongoing work to deepen the understanding of its dynamical drivers and to improve its representation in models. One source of uncertainty regarding the impact of the AMV is the lack of a full understanding of the phenomenon itself. Additionally, model biases in representing crucial quantities for the AMV, such as the AMOC and teleconnection patterns, contribute to increasing this uncertainty. In this line of thought, it was shown in (Qasmi, Cassou, and Boé, 2017) that CMIP5 models underestimate the ocean-atmosphere coupling at low frequency. The importance of dedicated modeling protocols to study decadal variability, such as the AMV, both at a global and regional scale using a coordinated multi-model approach has also been emphasized (Cassou et al., 2018). To this end, the Decadal Climate Prediction Project (DCPP), part of the CMIP6 project, was established (Boer et al., 2016). Within the DCPP, ensembles of simulations have been conducted to understand the predictability, variability, and impacts of decadal modes of climate variability such as the AMV. We use the outputs of the DCPP-C.1 experiments designed to enhance the understanding of the impact of the AMV on the global climate. In these experiments, the sea surface temperature (SST) of the North Atlantic is restored towards an anomalous SST pattern representative of the observed AMV, shown in fig. 6.1. The detailed procedure of the experiments can be found in the technical note of (Boer et al., 2016).

The outputs of four coupled models that took part in the experiments were available on the Earth System Grid Federation. For one of them, HadGEM3-GC31-MM, the soil moisture outputs were not available, and therefore, we could not use it in our study. We used the remaining three models: IPSL-CM6A-LR (Boucher et al., 2020), EC-Earth3 (Döscher et al., 2021), and CNRM-CM6-1 (Voldoire et al., 2019) (hereafter simply referred to as IPSL, EC-Earth, and CNRM). Further details about the models can be found in section 11.1 and in the corresponding references. For each model, AMV+ and AMV- ensembles consisting of many 10-year members have been computed. In the AMV+ ensemble, the SST is restored towards a positive anomaly pattern, while in the AMV- ensemble, it is restored towards its opposite. For IPSL and CNRM, we also have

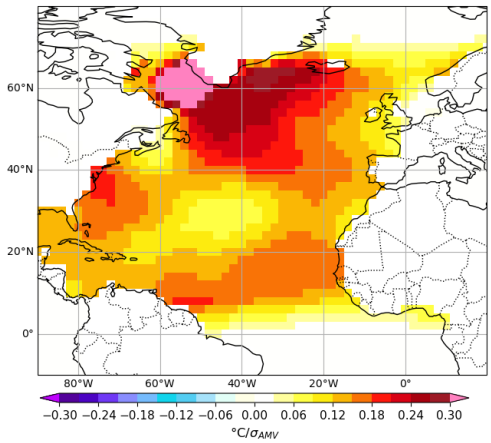


Figure 6.1: SST anomaly pattern used for restoring the SST in the DCPP-AMV experiments. Pattern courtesy C. Cassou.

Dataset \ Model	IPSL	EC-Earth	CNRM
AMV+/-	500	320	390/400
CTRL	250	0	400
Dry/Wet	500	250	450

Table 6.1: *Number of years in all ensembles for each model*. The AMV+/- and CTRL ensembles consist of simulation members of 10 years. The Dry and Wet ensembles are built by sorting the years in all simulation members according to their average May Soil Moisture value over Southern Europe (see sections 6.2 and 11.1 for details). For CNRM, some surface air temperature and soil moisture outputs are missing.

a control run (CTRL) where the SST is nudged towards the climatology. The radiative forcing is set to its 1850 value. Table 6.1 summarizes the number of years available in each ensemble for each model.

6.1.2 *Observation-derived datasets for soil moisture*

We use two observational datasets of land variables, ERA5-Land and GLEAM v3.8a, as benchmarks against which to compare the persistence of soil moisture anomalies in the three models.

ERA5-Land is a reanalysis dataset of land variables that describes the evolution of the water and energy cycles over land in a consistent manner (Muñoz-Sabater et al., 2021). It goes back to 1950 and is produced through global high-resolution numerical integrations of the ECMWF land surface model driven by the downscaled meteorological forcing from the ERA5 climate reanalysis. We use monthly averages of the volumetric soil water in the top 3 (resp. 4) soil layers of the model, corresponding to depth 0-100 (resp. 0-289) cm.

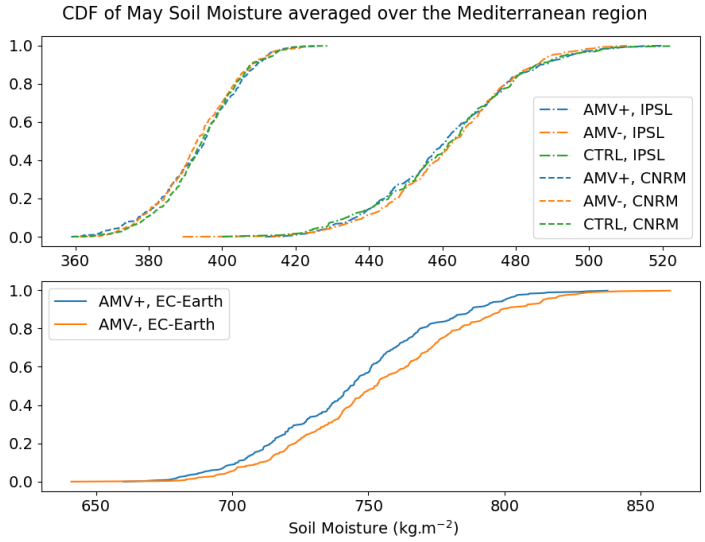


Figure 6.2: Empirical cumulative density function of May soil moisture averaged over the Mediterranean region (35°N-46°N, 10°W-30°E) in each model (IPSL, CNRM, EC-Earth) and each experiment (AMV+, AMV-, CTRL).

The Global Land Evaporation Amsterdam Model (GLEAM) is a set of algorithms dedicated to the estimation of terrestrial evaporation and root-zone soil moisture from satellite data (Miralles et al., 2011; Martens et al., 2017). The surface soil moisture is assimilated from satellite microwave remote sensing data. The model features a multi-layer soil model driven by satellite observations of precipitation with fast and slow draining of water from the surface layer towards the deepest layers. The soil moisture estimates are validated against 2325 soil moisture sensors across a broad range of ecosystems. We use GLEAM v3.8a monthly average datasets of the surface (0-10 cm depth) and root zone (10-100 cm) soil moisture. We perform a weighted average of the two datasets to obtain the soil moisture between 0 and 100 cm. GLEAM datasets run from 1980 to the present.

6.2 INFLUENCE OF THE AMV ON SPRING SOIL MOISTURE AND CREATION OF THE DRY AND WET ENSEMBLES

To assess the influence of spring soil moisture on summer heatwaves, we use the outputs of the three experiments (AMV+/- and CTRL) for each model to build two new ensembles, called Dry and Wet, corresponding to low and high spring soil moisture, respectively. We use the monthly total soil moisture content (mrso), which is the only relevant variable available for all three models. First, we compute the average May soil moisture SM_{av} for each year and each experiment over a domain covering Southern Europe. Following (Vautard et al., 2007), we choose the extent of the domain to be the rectangular box 35°N-46°N, 10°W-30°E. Then, we verify the influence of the AMV on SM_{av} in the models by plotting the cumulative density function (CDF) of SM_{av} for each experiment in fig. 6.2. The influence of a positive versus negative phase of the AMV on SM_{av} is negligible in IPSL and CNRM (-0.8 and $+0.6$ kg.m^{-2} respectively, to be com-

pared to standard deviations of 19.8 and 11.8 kg.m^{-2}) and relatively small in EC-Earth (-10.5 kg.m^{-2} for a standard deviation of 34.4 kg.m^{-2}).

In order to rigorously compare the influence of spring soil moisture with the one of the AMV, we must avoid an imbalance of years from the AMV+ and AMV- experiments in the Dry and Wet ensembles. Indeed, if the Dry ensemble were to contain 200 years from the AMV+ experiment versus 100 from the AMV- experiment, and vice versa for the Wet ensemble, we would observe a partial influence of the AMV when comparing the Dry and Wet ensembles. To ensure that there is no indirect influence of the AMV, we impose the constraint that an equal number of years come from the AMV+ and AMV- experiments in the Dry and Wet ensembles[1]. We then place years with low (resp. high) SM_{av} in the Dry (resp. Wet) ensemble. We refer the reader to section 11.1 for more details about the procedure and fig. 11.1 for an illustration. The mean value of the soil moisture inside the Dry (resp. Wet) ensemble is nearly one standard deviation below (resp. above) the mean value of the soil moisture averaged over the AMV+/- and CTRL experiments all-together. This makes the comparison with the AMV experiments meaningful because the targeted SST pattern for the relaxation corresponds to one standard deviation of the AMV variability: we compare the influence of a one sigma anomaly for both the AMV and spring soil moisture.

6.3 HEATWAVE DEFINITIONS

6.3.1 *Threshold-based definition*

We call the first definition that we use *threshold-based definition* as it relies on two temperature thresholds. It was first introduced in (Lau and Nath, 2012) and has been used in (Qasmi et al., 2021) and (Ruprich-Robert et al., 2018) to study the influence of the AMV on heatwaves in Europe and North America respectively. Using this definition enables us to compare our results with (Qasmi et al., 2021), which used an earlier version of the CNRM and EC-Earth models. According to this definition, a heatwave event is a group of days that satisfies the following three conditions:

(i) T_{max} must exceed T_1 for at least 3 consecutive days,

(ii) T_{max} averaged over the entire event must exceed T_1 and

(iii) T_{max} on each day of the event must exceed T_2.

where T_{max} is the daily maximum 2-meter air temperature and T_1 and T_2 are two temperature thresholds corresponding respectively to the 90th and 75th percentile of the local June-July-August (JJA) T_{max} distribution built from the T_{max} values of all members of the AMV+ and AMV- simulations (for each model). This definition is location-dependent, since the T_{max} distribution varies with latitude and longitude. For each grid point, we count the number of heatwave

1 This constraint also helps address the imbalanced size of the AMV+/- ensembles of the CNRM model. Indeed, 120 years of soil moisture outputs are missing for the AMV- experiment with this model.

days in each year. In this study, we are interested in the response, in terms of heatwave days per year, to an AMV-forcing. In chapter 7, we consider the mean difference between AMV+ and AMV- for each model, and in chapter 8, we condition on the most extreme years, i.e., the ones with the highest number of heatwave days.

6.3.2 *14-day heatwaves*

To quantify the heatwave intensities for several independent durations, heatwave indices based on the combined temporal and spatial averages of the surface or 2-meter air temperature have been adopted in many studies. Notably, seminal studies of the 2003 Western European and 2010 Russian heatwaves considered the averaged temperature over variable long time periods (7 days, 15 days, 1 month, and 3 months) (Schär et al., 2004; Barriopedro et al., 2011; Coumou and Rahmstorf, 2012). Similar definitions have been adopted in a set of recent studies (Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2020; Ragone and Bouchet, 2021; Gálfi, Lucarini, and Wouters, 2019; Galfi and Lucarini, 2021; Gálfi et al., 2021). This viewpoint is expected to be complementary with more classical heatwave definitions (Perkins, 2015) and extremely relevant to events with the most severe impacts. For our second definition, we consider averages of the daily maximal 2-meter air temperature over a period of 14 days. This duration corresponds to long-lasting events that are more impactful than events lasting only a few days, while still being much shorter than the seasonal time scale. Mathematically, the definition reads as follows:

$$\tilde{A}(\mathbf{r}) = \max_{t,t+D\in \mathrm{JJA}} \frac{1}{D} \int_t^{t+D} \left(T_{\max}(\mathbf{r}, t') - \bar{T}_{\max}(\mathbf{r})\right) \mathrm{d}t' \,, \tag{6.1}$$

where $\mathbf{r}$ and t' represent the spatial and time coordinates respectively, and $D = 14$ days is the heatwave duration. We do not perform any spatial averaging because we aim to maintain a local perspective, as with the threshold-based definition. To facilitate the comparison across different locations, we subtract $\bar{T}_{\max}(\mathbf{r})$, the local June-July-August (JJA) average of $T_{\max}(\mathbf{r}, t)$. This allows us to measure a temperature anomaly relative to the local seasonal mean. We compute one value of $\tilde{A}$ for each year and at each grid point.

7

INFLUENCE OF THE ATLANTIC MULTIDECADAL VARIABILITY AMV AND SPRING SOIL MOISTURE ON HEATWAVES WITH RETURN TIMES OF A FEW YEARS

In this section, we compare the effect of the AMV and spring soil moisture on typical heatwaves with return times of one or two years. Following previous studies (Ruprich-Robert et al., 2018; Qasmi et al., 2021), we first use the threshold-based definition introduced in section 6.3.1 to measure the drivers' influence on the mean number of heatwave days per year. We then use the second definition, presented in section 6.3.2, to measure their influence on the 14-day heatwave intensity.

7.1 INFLUENCE ON THE FREQUENCY OF HEATWAVE DAYS

We first study the effect of the AMV and spring soil moisture on heatwaves using the threshold-based definition introduced in section 6.3.1. Figure 7.1 shows the difference in the mean number of heatwave days per year between the Dry and Wet ensembles and the AMV+ and AMV- ensembles for each of the three models. We use a bootstrap to test the significance of the differences at the 95% level (see section 11.1 for details). Both the AMV and spring soil moisture influence significantly the mean number of heatwave days over some areas of Europe but according to different regional patterns.

For the AMV, all models agree on a significant positive influence on southern Europe (Iberian Peninsula, Italy, and Greece) around the Mediterranean basin with positive differences ranging from 1 to 3 heatwave days per year. These numbers must be compared with the mean number of heatwave days per year in the AMV- ensemble, which ranges between 7 and 10 heatwave days per year in most locations (not shown). In EC-Earth, the region of positive influence extends northwards up to approximately 48° North. Differences over the rest of the continent range between -1 and +1 heatwave day per year and do not pass the significance test. We note that all models show a region of small negative influence around Poland (between 0 and -1 heatwave days per year). Although this difference does not pass the test, the consistency between the tree models suggests the presence of a real signal.

Regarding the influence of soil moisture, the models present less consistent results. On one hand, CNRM and IPSL show a positive response over Southern Europe, mostly ranging between 1 and 3 heatwave days per year, with some locations reaching up to 4 heatwave days per year. In these two models, the influence of soil moisture extends further north than the one of the AMV and is slightly larger but of comparable amplitude. On the other hand, EC-Earth simulates a much more extended and much larger positive influence of spring soil moisture deficit. The region of positive difference covers almost all of Europe as well as Northern Africa, with differences ranging from 3 to 6 heatwave days

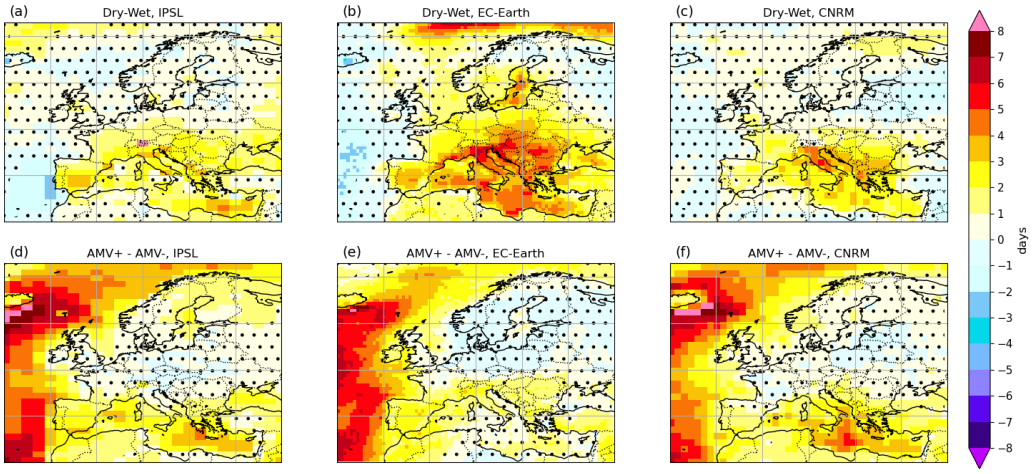


Figure 7.1: *Mean response maps for the frequency of heatwave days.* Differences in the mean number of heatwave days per year between the Dry and the Wet ensembles (top line) and the AMV+ and AMV- ensembles (bottom line) for each of the three models. Stippling denotes area below the 95% significance level according to a bootstrap test.

per year over Spain, Italy, and Southern Central Europe. In this model, spring soil moisture has thus a much stronger influence than the AMV. Once again, regions where the differences are between -1 and +1 heatwave day per year do not pass the significance test.

DISCREPANCY BETWEEN EC-EARTH AND THE TWO OTHER MODELS. EC-Earth presents a response to spring soil moisture deficit that is both more extended and stronger than in the other two models. This stronger response is associated with a strong positive 500 hPa geopotential anomaly and reduced precipitation over Central Europe (not shown). In order to further investigate the cause of this discrepancy, we show in fig. 7.2 the auto-correlation function of the monthly averaged soil moisture averaged over Southern Europe (35°N-46°N, 10°W-30°E).

In IPSL and CNRM, the auto-correlation decays almost linearly from 1 for the May-May correlation to about 0.5 for May-August. On the other hand, in EC-Earth, the correlation levels out for the three summer months and is nearly constant at about 0.9[1]. This larger correlation implies a stronger persistence of soil moisture anomalies in EC-Earth: dry soils in May remain abnormally dry throughout the summer, causing the surface air to heat up more strongly due to a larger sensible heat flux. For comparison, we also plot the auto-correlation of the soil moisture in the observation-derived datasets ERA5-Land and GLEAM

1 The different auto-correlations between models may be due to different land-atmosphere couplings or to intrinsic differences between land models. Among the latter, a notable difference is the difference in soil depth relevant for the total soil moisture content. One may wonder if this depth difference dominates the auto-correlation difference between models. It appears not to be the case, as IPSL and CNRM share similar auto-correlation decays despite having soil depths of 2 and 12 meters respectively.

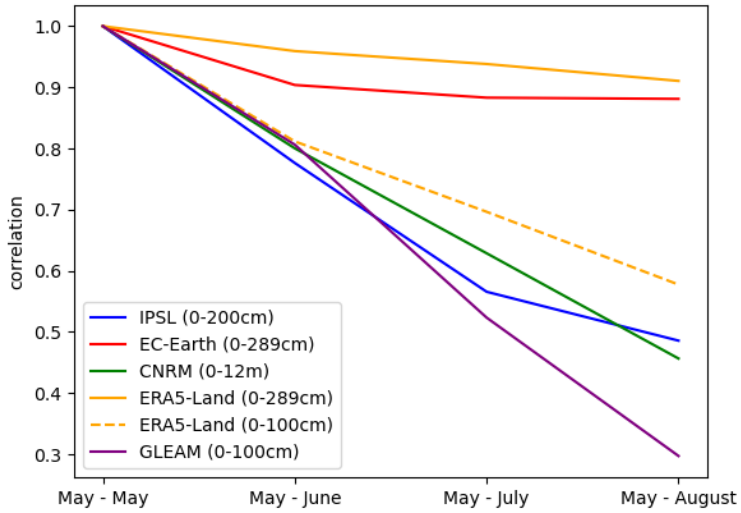


Figure 7.2: Pearson correlation coefficients between the month of May and the months of June, July, and August of the monthly averaged soil moisture averaged over Southern Europe (35°N-46°N, 10°W-30°E).

v3.8a. ERA5-Land shares the same land model, HTESSEL, with EC-Earth and also presents a strongly persistent auto-correlation of soil moisture. This indicates that the large correlation of EC-Earth is not due to an abnormal atmospheric response but rather to the land model itself. In an attempt to estimate if HTESSEL produces too large auto-correlation of soil moisture, we compare the auto-correlation of the 1st meter of soil with GLEAM v3.8a, a product more directly derived from satellite observations than ERA5-Land. The auto-correlation for the months of July and August remains larger in ERA5-Land compared to GLEAM. Thus, it could be that HTESSEL produces too large persistence of soil moisture anomalies. However, it would be hasty to draw any definitive conclusions. While we consider the results of EC-Earth with some caution, we cannot rule it out.

7.2 INFLUENCE ON THE INTENSITY OF 14-DAY HEATWAVES

We turn to the second definition to assess the influence of the AMV and spring soil moisture on the intensity of 14-day heatwaves. Figure 7.3 shows the mean difference in the 14-day averaged temperature $\tilde{A}$ (see section 6.3.2) between the Dry and Wet (top line) and AMV+ and AMV- (bottom line) ensembles. Note that these maps should not be interpreted as the temperature averaged over any given period of the summer. Indeed, the 14 days that contribute to the average are a priori different for each grid point and for each year of the simulations. The patterns of influence are very similar to those observed in fig. 7.1 for heatwave duration.

All three models agree on a positive influence of the AMV that concentrates around the Mediterranean Basin, where differences range from 0.2 to 0.4°C in

most places. EC-Earth simulates higher differences, ranging from 0.5 to 0.8°C, in Northern Spain and Southern France. All models simulate a positive influence over Turkey, and this extends around the Black Sea for IPSL. CNRM and IPSL also simulate positive differences over parts of Scandinavia, but they do not agree on the localization. As for the first definition, all models simulate a region of small negative influence over Poland and neighboring countries that do not pass the significance test but may be a hint of a signal.

The influence of spring soil moisture extends over larger areas of Europe. For CNRM, the region of positive influence extends from the Atlantic Ocean to the Black Sea below 50°N. For IPSL and EC-Earth, it extends over almost all continental Europe. Regions of exception are Scandinavia, the Baltic countries, Poland, and Russia in IPSL, and Northern Scandinavia and Russia in EC-Earth. The influence of soil moisture is also larger than that of the AMV, with differences reaching up to 0.5 or 0.6°C in many locations for IPSL and CNRM. As for heatwave duration, EC-Earth simulates a much larger influence of low soil moisture with positive differences above 0.8°C over all Western and Central Europe and peak differences up to 1.3°C over the Balkans. As for the first definition, we expect the stronger Dry-Wet response of EC-Earth to be due to the larger auto-correlation of soil moisture.

Combining the results of sections 7.1 and 7.2, we conclude that spring soil moisture in Southern Europe is a slow driver of greater importance than the AMV for European summer heatwaves: its region of influence covers larger areas of Europe with a larger amplitude. We note that EC-Earth presents a different result than the other two models, with a much stronger influence of spring soil moisture. Regarding this point, we do not have enough evidence to conclude that one model is more or less biased than the others. We also tried to investigate the mechanisms by which the AMV influence the frequency and intensity of heatwaves over Europe. To this aim, we plotted the mean response maps to the AMV phase for cloud cover, latent heat flux, sensible heat flux, 500hPa geopotential height, precipitation, and soil moisture averaged over the June, July and August period in fig. 11.3 (see section 11.2). There is no significance in the responses for individual models. This lack of significance does not allow us to address the important question of the mechanism by which the AMV influences heatwaves with these datasets. This calls for either longer datasets or the application of stronger SST forcing in the numerical experiments as in (Qasmi et al., 2021), where the authors simulate the influence of one-sigma, two-sigma, and three-sigma AMV anomalies. The experiments with stronger forcing yield significant signals, allowing the authors to conclude that a positive AMV phase is associated with "drier soils and a reduction of cloud cover" around the Mediterranean basin and "an enhancement of the downward radiative fluxes over lands" (Qasmi et al., 2021).

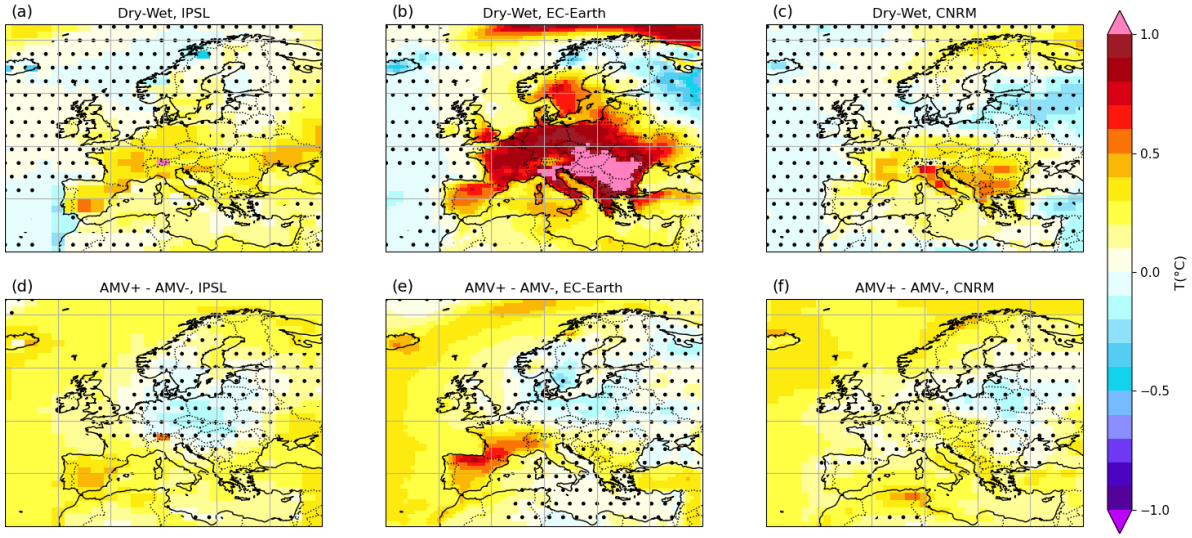


Figure 7.3: *Mean response maps for 14-day heatwaves.* Difference in the mean value of the 14-day averaged temperature $\tilde{A}$ (defined in eq. (6.1)) between (top line) the Dry and the Wet ensembles and (bottom line) the AMV+ and AMV- ensembles for each of the three models. Stippling denotes area below the 95% significance level according to a bootstrap test.

8

INFLUENCE OF THE ATLANTIC MULTIDECADAL VARIABILITY AND SPRING SOIL MOISTURE ON HEATWAVES WITH RETURN TIMES FROM 10 TO 50 YEARS

The heatwaves that we considered in the previous section are not really rare events. Indeed, the second definition gives one value of $\tilde{A}$ for each year, while for the first definition the fraction of years during which at least one heatwave occurs is consistently comprised between 50% and 80% for all models and land grid points (fig. 11.4). This means that such heatwaves occur every one or two years and are thus not so rare events. In this section, we focus on rarer events with return times of a few decades. In section 8.1, we use return time plots to see how the influence of the slow drivers evolves with the return time for a single grid point. However, this provides information only at a local level. To study extreme events at the European level, we introduce return time maps in section 8.2. Based on these maps, we study extreme events with return times ranging from 10 to 50 years.

8.1 STUDY OF RARE EVENTS USING LOCAL RETURN TIME PLOTS

In climate statistics, the probability of an event is often expressed in terms of return time: if an event has a probability $1/Y$ to occur each year, then it has a return time of Y years, which also corresponds to the average duration between two such events. A classical way to visualize the intensity of events of decreasing probability is to build return time curves: events are ranked by decreasing intensity $a_1 > a_2 > ... > a_N$ and the empirical return time associated to the level a_m reads:

$$\hat{r}(a_m) = \frac{N}{m}. \tag{8.1}$$

This is simply the inverse of the empirical probability to have an event of intensity at least as large as a_m which is m/N. By construction, the minimal event's intensity has a return time of one year, the median a return time of two years and the largest event is associated with a return time of N years. In the present study, a will be either the number of heatwave days in a year (following the definition of section 6.3.1) or the 14-day heatwave intensity defined in section 6.3.2. Remember that both quantities are defined at each grid point. We present results for the second definition in this section. The results for the first definition are similar and the corresponding figures are shown in section 11.2.

Figure 8.1 shows the return time curves for the intensity of 14-day heatwaves for a grid point in the Northeast of Spain. We chose this region for illustration purposes, as it is one where we observe a strong response of $\tilde{A}$ to both the AMV and the soil moisture in all models, except for the AMV in CNRM (fig. 7.3). The six upper panels show the empirical return time curves for each ensemble

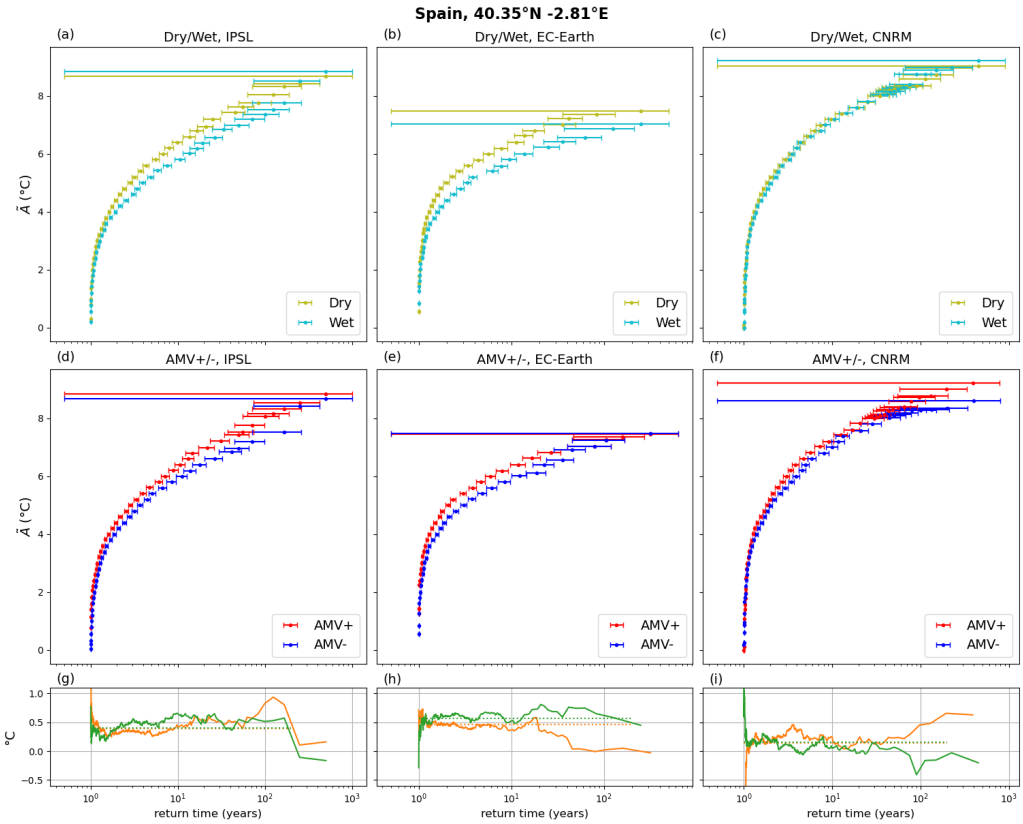


Figure 8.1: *Return time curves of the intensity of 14-day heatwaves*. The two first lines show the temperature anomaly as a function of the return time for the dry and wet ensembles (top row) and the AMV+ and AMV- ensembles (second row) for the three models. Errorbars correspond to one standard deviation of the estimated return time to observe an event of a given amplitude. The third row shows the Dry - Wet (green) and AMV+ - AMV- (orange) differences as a function of the return time. The dotted lines indicate the mean difference values. For each model the chosen grid point covers the coordinate indicated in the figure title which corresponds to a location in the Northeast of Spain, close to the Mediterranean Sea.

(AMV+, AMV-, Dry, Wet) and each model[1]. Several considerations arise from these plots. The temperature anomaly starts at about 0°C for a return time close to 1 year[2], goes up quickly to around 4 or 5°C for a return time of 2 years and then rises at a slower pace, up to anomalies of 7°C to 9°C, depending on the model, for return times of a few centuries. More importantly, these curves allow to visualize how the influence of the drivers on the frequency of events changes as the return level increases. For temperature anomalies from 4°C to 6°C a low soil moisture (resp. a positive AMV phase) double the frequency of occurrence with respect to a high soil moisture (resp. a negative AMV phase) in IPSL and EC-Earth. In CNRM the AMV response is smaller and there is no visible influence of soil moisture.

It is less straightforward to visualize the changes in the influence of the drivers on the intensity of the events. At a first glance, it seems that the influence on the intensity increases with the return time. However, this interpretation is misleading, as it is influenced by the near-vertical alignment of the curves for the lowest return times. Indeed, if two curves are concave and a constant vertical distance away, then the horizontal distance between them grows along the x-abscissa[3]. To properly visualize the evolution of the influence of the drivers as the return time increases, we show the Dry - Wet and AMV+ - AMV- differences as a function of the return time on the third line of fig. 8.1. We see that the influence of both the AMV and soil moisture is roughly constant from a return time of less than two years up to a few decades. There are large fluctuations at very short and large return times which correspond to the two tails of the distribution. The influences of the AMV and spring soil moisture are of the same order within each model. However, they are lower in CNRM than in the two other models. Figure 11.5 shows similar return time plots for the threshold based definition.

While return time plots give a view of the evolution of extreme events intensity when shifting from small to large return times, the information they provide is location dependent. For instance, while EC-Earth produces noticeably lower extreme heatwaves than the other two models at the location examined in fig. 8.1, it turns out to be the opposite in Central Europe (not shown). We present in the next section a method to overcome this limitation and obtain a global picture of the influence of the drivers on extreme events.

8.2 REGIONAL PICTURE OF AMV AND SPRING SOIL MOISTURE INFLUENCE ON HEATWAVES WITH RETURN TIMES OF 10 AND 50 YEARS

In order to synthesize at the continental scale the local results described above, we build on the concept of return time curve to introduce *return time maps*. To put it in a nutshell, these maps allow to visualize the difference in intensity of events with the same return time but belonging to different ensembles. As far

1 The curve for the dry/wet ensembles on one side and for the AMV+/- ensembles on the other side are very similar for each model because they are built from the same dataset.

2 Very few events have an anomaly between -2°C and 0°C. They correspond to extreme cold years. They have been cropped for the sake of plot's readability.

3 As a simple example, one can consider the two curves $y_1(x) = \sqrt{x}$ and $y_2(x) = y_1(x) + \Delta y$. The horizontal distance between them is $\Delta x = 2\Delta y\sqrt{x} - \Delta y^2$ which grows as $\sqrt{x}$.

as we know, we are the first to plot this kind of maps. A similar but different concept is Risk Ratio maps where the changes in probability of events of the same intensity are displayed (Kharin et al., 2018). We first explain how these maps are built before discussing the results.

METHOD For each grid point, we consider the difference between the AMV+ and AMV- or Dry and Wet return time curves averaged over all the events with a return time larger than a threshold RT. We considered threshold values of RT=10 and 50 years, which provide a compromise between studying extreme events and keeping enough events to calculate statistics.

To be more specific, for each dataset $\mathcal{S} \in$ {AMV+, AMV-, Dry, Wet} and for the return times RT=10 or 50 years, we compute at each grid point:

$$a^{\mathcal{S}}_{\mathrm{RT}}(\mathbf{r}) = \frac{1}{K}\sum_{i=1}^{K} a_i(\mathbf{r}) \quad \text{where } K = \frac{N_{\mathcal{S}}}{\mathrm{RT}} \tag{8.2}$$

where $\mathbf{r}$ is the spatial coordinate, $N_{\mathcal{S}}$ is the number of years available in the dataset $\mathcal{S}$ and K is the number of years such that $\hat{r}(a_i) \geq \mathrm{RT}$. We recall that the a_i's are the number of heatwave days in a year or 14-day heatwave intensity, depending on the definition considered. We have assumed that $a_1 > a_2 > a_3 > ... > a_{N_{\mathcal{S}}}$ in each dataset. The interpretation is that $a^{\mathcal{S}}_{10yrs}$ is the empirical mean of 10-year events in the dataset $\mathcal{S}$. We then display for each grid point the differences $a^{\text{AMV+}}_{\mathrm{RT}}(\mathbf{r}) - a^{\text{AMV-}}_{\mathrm{RT}}(\mathbf{r})$ and $a^{\mathrm{Dry}}_{\mathrm{RT}}(\mathbf{r}) - a^{\mathrm{Wet}}_{\mathrm{RT}}(\mathbf{r})$. Note that these differences can be seen as computing the differences between the two curves in panels (a) to (f) of fig. 8.1 (these differences are shown in panels (g) to (i)) and then taking the average of these differences over the points with return time greater than 10 or 50 years only, i.e. on the rightmost part of these curves. We call these maps *return time maps*.

Figure 8.2 shows the 10-year return time maps for the intensity of 14-day heatwaves. Each individual map shows only reduced area of statistical significance compared to the mean response maps. This is because we consider only the most extremes years, which reduces the number of years included in the maps. Regarding the AMV influence, there is almost no region of significance. Larger regions of significance subsist for spring soil moisture: for IPSL, a region of positive influence extends over Western Europe, up to 52°N, while this positive influence covers almost all Europe until about 25°E in EC-Earth. In many areas, the 10-year response is higher than the mean response with an amplitude up to 1°C. Once again, EC-Earth shows a higher response to soil moisture with a peak amplitude of 1.8°C. We note that the peak of this response has shifted northward with respect to the mean response. We note that the regional response patterns are similar to those of the mean response (see fig. 7.3; note that the color scale is different). Therefore, it seems that response patterns for the rarest events are similar to the patterns for less rare events. However, confirming this statement would require datasets with more extreme events.

We show the 50-year return time maps of 14-day heatwaves intensity in fig. 11.6 in section 11.2. These maps are much noisier than the 10-year return time maps. This is because these maps correspond to events that are in the

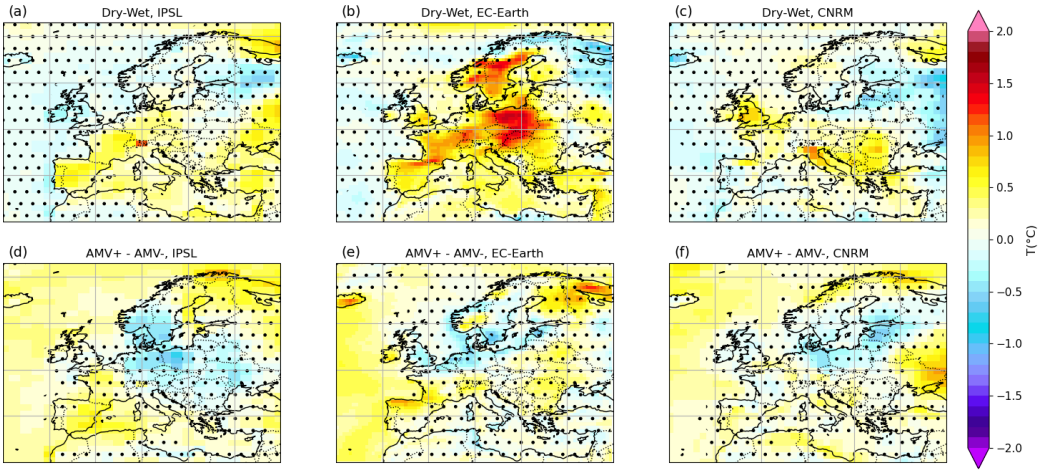


Figure 8.2: *10 years return time maps for 14-day heatwaves.* The maps show the difference of $\tilde{A}$ between (top line) the Dry and Wet ensembles and (bottom line) the AMV+ and AMV- ensembles, conditioned over events with return time greater than 10 years. Stippling denotes area below the 95% significance level according to a bootstrap test.

very tail of the empirical distribution, where the fluctuations of the response are large, as can be seen on panels (g,h,i) of fig. 8.1.

We performed the same analysis for the number of heatwave days per year as measured by the threshold-based definition. The 10 and 50-year return time maps for this definition are shown in figs. 11.7 and 11.8 in section 11.2. The global response patterns are similar to those for the intensity of 14-day heatwaves, leading to the same conclusions.

9 SUMMARY OF RESULTS

In this study, we use numerical experiment outputs from the Decadal Climate Prediction Project to assess which slow driver, between the AMV and spring soil moisture in Southern Europe, has the most influence on European summer heatwaves. We study the influence of a one-sigma deviation from the climatology for both drivers. We use two complementary definitions of heatwaves and study both typical heatwaves with a return time of one or two years and rarer events with a return time of 10 or 50 years.

In order to decouple the effects of the two drivers, we first investigated the influence of the AMV on the distribution of spring soil moisture averaged over Southern Europe. We found it to be negligible in two models (IPSL and CNRM), while there is a small but significant influence in EC-Earth, where the average soil moisture is 10.5 kg.m^{-2} lower in the AMV+ phase compared to the AMV- phase (fig. 6.2). This can be compared to the standard deviation computed over the two ensembles together, which is $\sigma^{\mathrm{SM}}_{\text{EC-Earth}} = 34.4$ kg.m^{-2}. We conclude that the two drivers can be studied independently in IPSL and CNRM. In EC-Earth, there is an indirect influence of the AMV on heatwaves through decreased soil moisture in the positive AMV phase, and we paid attention to removing this influence when building our Dry and Wet ensembles.

Our main conclusion is that spring soil moisture in Southern Europe is a slow driver of greater importance than the AMV for European summer heatwaves, both in terms of extension of the region of influence and in terms of amplitude. While the influence of the AMV concentrates around the Mediterranean Basin, the one of spring soil moisture extends over most of continental Europe, up to 50°N or the Baltic Sea, depending on the model. The amplitude of the response goes up to 4 more heatwave days per year and +0.5°C for the mean intensity of 14-day heatwaves, except for EC-Earth which presents a much larger influence of spring soil moisture than the other two models This can be linked to the higher persistence of this variable along the summer season, as shown in fig. 7.2, and might be due to its land model. However, we could not discriminate with confidence if one model was more biased than the others. Further investigations need to be done in this and other models to gain a more confident understanding of the response of heatwaves to spring soil moisture anomalies.

Consistently with a previous study (Qasmi et al., 2021), a positive AMV phase or low soil moisture generally induce more heatwave days and hotter heatwaves. Regarding the AMV, all models also show a region of negative influence around Poland with responses down to -0.2°C for the mean temperature of 14-day heatwaves (fig. 7.3) and 1 fewer heatwave day per year (fig. 7.1). Although this anomaly does not pass the statistical test, its presence in all three models suggests the existence of a true signal. We speculate that the direct heating effect in phase with the main surface temperature forcing creates a positive geopotential height response at mid troposphere (visible in fig. 11.3). This also

creates a wave-like response in geopotential, which would explain the cold area around Poland. However, we acknowledge that our findings might be biased due to the small number of climate models used and the fact that they show inconsistent responses of physical mechanisms to the AMV (fig. 11.3).

By introducing return time maps, we were able to study extreme heatwaves with return times of 10 years. One main issue in the study of such rare events is the scarcity of data, which causes large areas to be below the statistical significance level in fig. 8.2 and fig. 11.6. This calls not only for new methodologies but also for longer datasets. For events with return times of 10 years, the influences of the AMV and soil moisture increase, according to rather similar regional patterns as for typical events, and remain of similar amplitude. However, the regions where a positive AMV phase or low spring soil moisture induce fewer heatwave days and cooler heatwaves extend. This conclusion is valid for both definitions of heatwaves. Positive AMV phases or spring soil moisture deficit induce a 1°C increase of the temperature. They also increase the number of heatwave days by up to 9 days/year. The return time maps for 50-year events are much noisier due to the fact that corresponding events are in the very tail of the empirical distribution. Once again, the influence of spring soil moisture on these extreme events is greater in EC-Earth than in the other two models.

10 DISCUSSION

Our results show that the AMV can modulate the amplitude of 14-day heatwaves over Southern Europe by 0.5°C to 1°C. This can be put into perspective with the current warming rate of hot extremes in Europe. Vautard et al. (Vautard et al., 2023) showed that the warming rate of hot extremes (TXx) over Europe is comprised between 2°C and 6°C per global warming degree with the fastest warming being over Western Europe. The current global warming rate being about 0.2°C per decade, this corresponds to a warming of 0.4°C to 1.2°C per decades for hot extremes, which is of the same order of magnitude as the modulation by the AMV. This means that the phase of the AMV can either mask or exacerbate the warming trend of hot extremes.

A similar conclusion was already discussed in (Liné et al., 2023) for the seasonally averaged temperature: the internal variability, including the AMV, will modulate the global warming trend at the European scale, either mitigating or exacerbating its effects. Moreover, these authors argued that, in the near term, the internal variability of the climate system will have a stronger influence on the European averaged temperature than the SSP emission scenario that we will actually follow. It is a natural question to ask whether this might be true also for the case of extreme heatwaves, as studied in this paper. However, this is beyond the scope of our article, as it would require a specific study of the impact of the different scenarios on extreme event statistics. We leave this question as an interesting perspective for future works. It has to be mentioned that no distinction is made in this study between seasonal and multidecadal timescale SSTs anomalies which present different patterns. This could also be addressed in future works.

In a warming climate, the Mediterranean basin is projected to become drier (Masson-Delmotte et al., 2023). This suggests that, when considering the effect of climate change, the variability of soil moisture over the zone considered for this study might be reduced, leading to a reduced importance in the modulation of heatwaves. However, the transition zone between dry and wet climates, in which the variability of soil moisture is large, is expected to shift northward with respect to the Mediterranean basin. Thus, it could be relevant to look for the region of largest soil moisture variability in the future and to consider the soil moisture in this region as the relevant slow driver for European heatwaves. Such a prospective study would be a natural follow-up to this work.

As already mentioned above, a limitation of this study arises from the scarcity of data that we face when we want to study really rare events. We are already encountering this data scarcity issue for events with return times of a decade, and the problem would be even more challenging if one wishes to study events with higher return times, for instance, a century. Recently, rare event algorithms have been designed to enhance the sampling of extreme events in numerical simulations at low computation cost. This class of algorithms has been successfully applied to sample heatwaves with return times of a century or even tens

of millennia in some regions of Europe (Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2021) and South Asia (Le Priol, Monteiro, and Bouchet, 2024). Using such rare event simulations could be a promising path to study the influence of slow drivers on heatwaves with return times of a century or more.

The approach presented in this paper is not limited to the study of slow drivers for extreme heatwaves in Europe. It could be useful for studying other rare phenomena with high impacts, and in other world regions.

ACKNOWLEDGEMENTS

The authors thank the modelling groups who took part in the DCPP-C AMV experiment. The model data were accessed through the IPSL mesocenter ESPRI facility which is supported by CNRS, UPMC, Labex L-IPSL, CNES and Ecole Polytechnique. We also thank the Centre Blaise Pascal of the Ecole Normale Superieure de Lyon for the computation resources used to carry on the study and Emmanuel Quemener for his help with the platform. We are grateful to C. Ardilouze for suggesting to plot the soil moisture auto-correlation and bringing GLEAM dataset to our attention. We thank C. Cassou, S. Qasmi and H. Douville for enlightening discussions.

CONFLICT OF INTEREST

The authors declare no conflicts of interests.

STUDY FUNDINGS

V.M. is part of the EDIPI ITN, which has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 956396. C.L.P. is funded by the ANR grant SAMPRACE, Project No. ANR-20-CE01-0008-01.

AUTHOR CONTRIBUTIONS

F.B. and F.D.A. proposed the project and provided guidance for the research. V.M. and C.L.P. conducted the research, performed all analysis and wrote the first draft of the manuscript. All authors contributed to manuscript review and editing.

DATA AVAILABILITY

All data used in this article are publicly available. Model simulation results can be accessed via the ESGF grid, for instance at https://esgf-node.ipsl.upmc.fr/search/cmip6-ipsl/. ERA5-Land monthly dataset, was downloaded from Copernicus Climate Data Store: https://cds.climate.copernicus.eu. The GLEAM evaporation dataset

can be downloaded from https://www.gleam.eu/. The processed data used in this research will be shared on reasonable request to the corresponding author.

11

SUPPLEMENTARY MATERIAL

11.1 METHODS

BRIEF DESCRIPTIONS OF THE MODELS CHARACTERISTICS

- *IPSL* is composed of the LMDZ atmospheric model version 6A-LR, based on a rectangular grid with 144 longitude x 143 latitude equally spaced points, resulting in a resolution of 2.5° × 1.3°. It has 79 vertical levels and extends up to 80 km. The ocean component is the NEMO oceanic model version 3.6, with 71 vertical layers and an horizontal resolution of 1°. The land surface model is the ORCHIDEE version 2.0, with 11 layers for a total of 2 m of soil depth. Further information and details can be found at (Boucher et al., 2020).
- *EC-Earth* is composed of the Integrated Forecast System (IFS) CY36R4 of the European Centre for Medium Range Weather Forecasts (ECMWF) atmospheric model, based on a linearly reduced Gaussian grid equivalent to 512 longitude x 256 latitude points with 91 vertical levels. It includes the land-surface scheme HTESSEL. The ocean and sea-ice model is NEMO-LIM3 version 3.6, with 75 vertical layers. Further information and details can be found at (Döscher et al., 2021).
- *CNRM* is composed of the ARPEGE-Climat atmospheric model, version 6.3 with 91 vertical levels. The ocean component is based on the NEMO version 3.6, while the sea ice component is based on GELATO, version 6. Further information and details can be found at (Voldoire et al., 2019).

CREATION OF THE DRY AND WET ENSEMBLES The Dry and Wet ensembles are made of years coming from the AMV+/- and CTRL ensembles according to SM_{av}. When building the Dry and Wet ensembles we want to make sure that there is no indirect influence of the AMV on the soil moisture through influence of the AMV on SM_{av}. Figure 6.2 shows the distribution of SM_{av} for each ensemble in the three models. We see that in IPSL and CNRM the phase of the AM does not influence the soil moisture, while in EC-Earth, there is a non negligible influence with the average SM_{av} in the AMV- phase being 5.2 $kg.m^{-2}$ than in the AMV+ phase. To make sure that there is no indirect influence of the AMV we choose to enforce to have the exact same numbers of AMV+ and AMV- years in the Dry and Wet ensembles.This choice also allows us to deal with the imbalance of the CNRM ensemble because of missing runs for Soil Moisture data. We have only 380 years of AMV+ and 280 years of AMV-. We will briefly comment on the procedure to create the ensemble. It is sketched in fig. 11.1. Each original ensemble (AMV+, AMV-, CTRL) is sorted according to SM_{av}. The k driest years of the AMV+ and AMV- ensembles are put in the Dry ensemble which is completed with the n driest years from the CTRL run when available.

	IPSL	EC-Earth	CNRM
Dry	443.1 / 442.5	715.2 / 714.3	383.1 / 383.2
Wet	480.2 / 481.0	782.5 / 783.1	404.6 / 405.7

Table 11.1: *Mean value of SM_{av} (in $kg.m^{-2}$) in the Dry and Wet ensembles for each model.* In each case the values are ordered as *realized / target* where *realized* is the computed mean in the Dry or Wet ensemble and *target* corresponds to exactly one standard deviation away from the overall mean value.

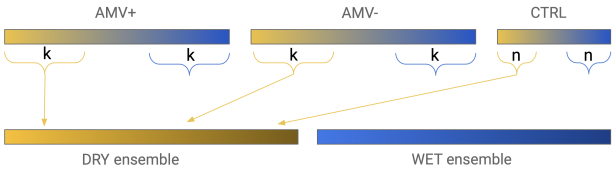


Figure 11.1: *Sketch of the procedure to create the Dry and Wet ensembles.* Data in each experiments are already sorted according to its SM_{av} value.

For this study, the number of years coming from the AMV experiments, namely k, is different from the one coming from the CTRL, namely n. The reason is the different amount of years available for each experiment, which we detail in table 6.1. More details about the values of SM_{av} for each model can be found in table 11.1. We choose the ensemble size such that the mean value of SM_{av} in the Dry (resp. Wet) ensemble is nearly one standard deviation below (resp. above) the mean value SM_{av} averaged over the AMV+/- and CTRL ensembles all-together. This makes the comparison with the AMV forcing sensible because the imposed SST pattern corresponds to one standard deviation of the AMV variability.

COMPUTATION OF THE ERRORBARS ON THE LOCAL RT PLOTS Let A be the random variable of event's amplitudes. Let us fix a return level a. The random variable $X = \mathbf{1}(A \geq a)$ follows a Bernoulli law of parameter $p := \mathbb{P}(A \geq a)$ Let N be the number of years in our dataset and m the number of years where we observe $A \geq a$. Then $\hat{p} = m/N$ is an unbiased estimator of p and we estimate the return time of events larger than a as

$$\hat{r}(a) = \frac{1}{\hat{p}} = \frac{N}{m}. \tag{11.1}$$

We want to compute an estimate of the error on $\hat{r}(a)$. First, note that the empirical variance of $\hat{p}$ reads $\hat{\sigma}_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \hat{p})^2 = \frac{N}{N-1}\hat{p}(1-\hat{p})$. According to the Central Limit Theorem, the variance of $\hat{p}$ reads $\sigma_{\hat{p}}^2 = \frac{p(1-p)}{N}$ which can be estimated by the empirical formula

$$\hat{\sigma}_{\hat{p}}^2 = \frac{\hat{p}(1-\hat{p})}{N}.$$

Let us define $\delta = \hat{p} - p$. Note that $\mathbb{E}[\delta] = 0$ and $\sigma_\delta^2 = \sigma_{\hat{p}}^2$ Let us assume that in all our observations $\delta < p$ (which is likely to be valid if $\sigma(r) \ll r$). Then we can write:

$$\hat{r} = \frac{1}{\hat{p}} = \frac{1}{p} \times \frac{1}{1+\delta/p} = \frac{1}{p}\left(1 - \frac{\delta}{p} + \frac{\delta^2}{p^2} + O\left(\frac{\delta^3}{p^3}\right)\right), \tag{11.2}$$

$$\hat{r}^2 = \frac{1}{n_+^2} = \frac{1}{p^2}\left(1 - \frac{2\delta}{p} + \frac{3\delta^2}{p^2} + O\left(\frac{\delta^3}{p^3}\right)\right). \tag{11.3}$$

Taking the difference of the expectations we get:

$$\sigma_{\hat{r}}^2 = \mathbb{E}[\hat{r}^2] - \mathbb{E}[\hat{r}]^2 = \frac{\mathbb{E}[\delta^2]}{p^4} + O\left(\frac{\mathbb{E}[\delta^3]}{p^3}\right). \tag{11.4}$$

The errorbars plotted on the local return time plots figs. 8.1 and 11.5 corresponds to the standard deviation of $\hat{r}$:

$$\sigma_{\hat{r}} \simeq \frac{\sigma(\hat{p})}{p^2} = \frac{1}{p^2}\sqrt{\frac{p(1-p)}{N}} \simeq \hat{r}\sqrt{\frac{\hat{r}-1}{N}}, \tag{11.5}$$

where we have used that $1/p = r$.

Note that the domain of validity of the this approximation is $\sigma(r) \ll r$ and so these formula is not correct to estimate the error on the most extreme return levels (the ones for which the number of exceedance m is small).

IMPLEMENTATION OF THE BOOTSTRAP TEST We implemented a bootstrap to test the significance of the AMV+ - AMV- and Dry-Wet differences in all maps. The procedure is the same for each model, but was slightly adapted for CNRM to take into account the fact that the AMV+ and AMV- do not have the same size and that the soil moisture of some members is missing.

Concretely, we pooled the results from all experiments (AMV+/- and CTRL) together to obtain a single large dataset for each model. From this large dataset, we draw N=1000 samples of M years, where M is equal to the number of year in each dataset displayed in table 6.1. For each sample we compute the average over the sample, as well as the 10-year and 50-year return time maps. For each of those maps (average, 10-year and 50-year return time maps), we build the empirical distribution of all differences between two distinct samples and compute the quantiles from the distribution. We consider the difference at each grid point to be significant at the 95% significance level whenever it is lower than the quantile 0.025 or higher than the quantile 0.975.

11.2 ADDITIONAL FIGURES

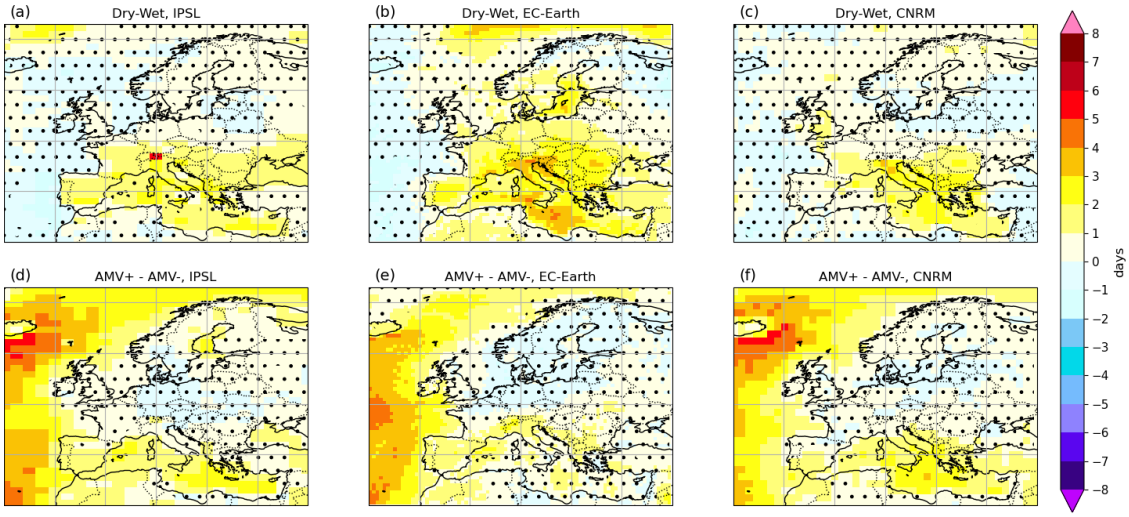


Figure 11.2: *Mean anomaly maps for the threshold based definition*, with thresholds T_1 and T_2 corresponding to the 95th and 80th percentile of the local JJA T_{max} distribution. The maps show the difference of the mean number of heatwaves days per year between (top line) the dry and the wet ensembles and (bottom line) the AMV+ and AMV- ensembles for each of the three models. Hatching denotes area below the 95% significance level according to a bootstrap test.

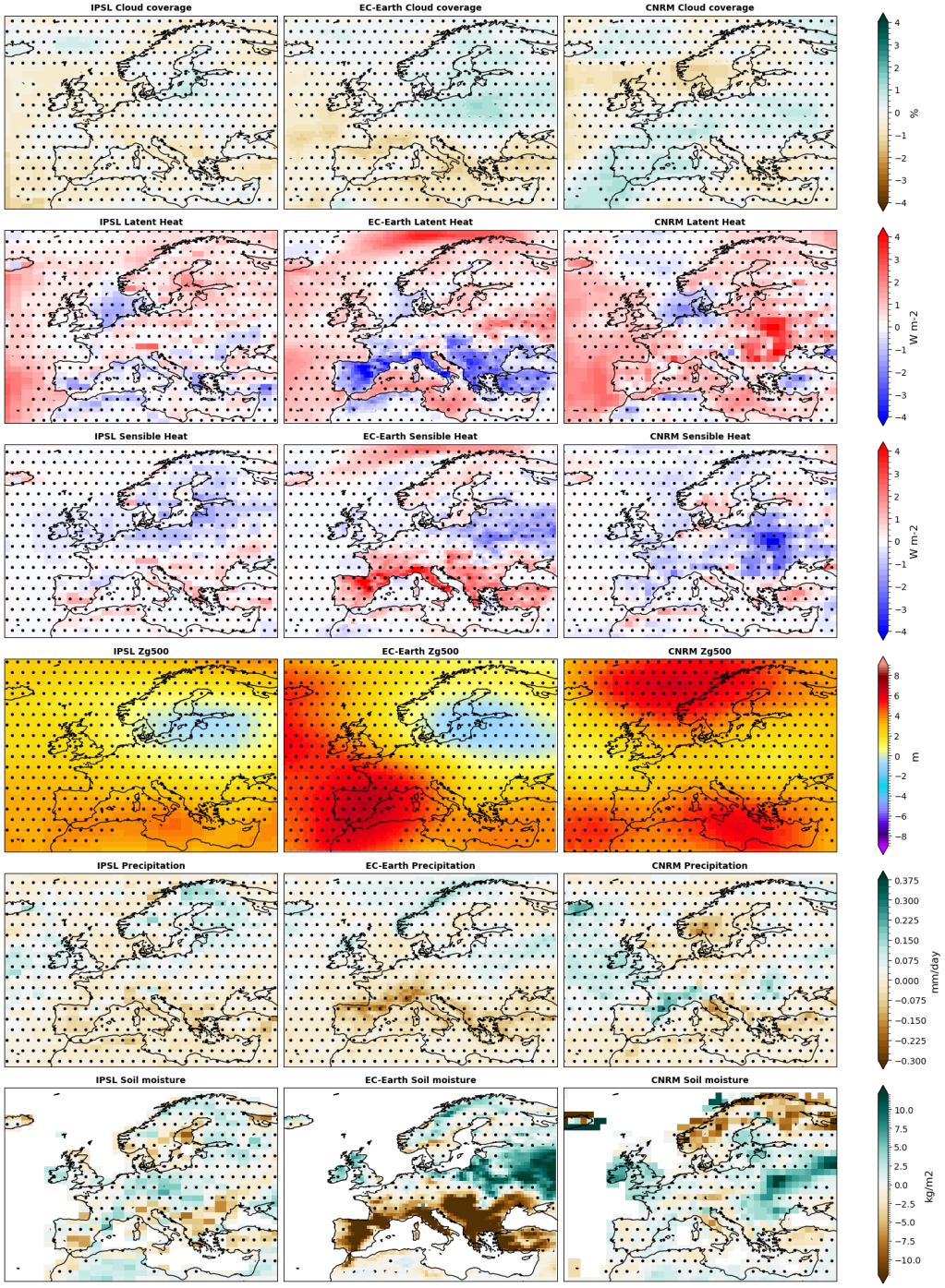


Figure 11.3: *Mean response maps to the AMV phase for different fields.* Each map represents the mean difference between AMV+ and AMV- of (from top to bottom row) percentage of cloud coverage, Latent heat, sensible heat, 500hPa geopotential height, precipitation, soil moisture. The average is taken with respect to the months of June, July and August. Each column represents a model. Stippling denotes area below the 95% significance level according to a bootstrap test.

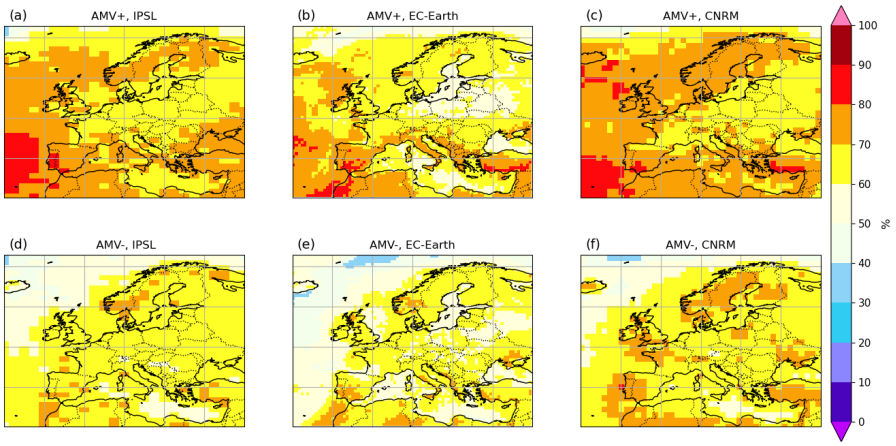


Figure 11.4: *Percentage of years during which at least one threshold-based heat wave occurs.* These maps show the percentage of years during which at least one heat wave (according to the threshold-base definition, see section 6.3.1) occurs. This percentage is above 50% over all land area, in all model and in the two AMV phases. Since these heat waves occur more frequently than every two years, they have a return time of less than two years. The percentage is even larger than 60% over most of Europe and 70% in many locations.

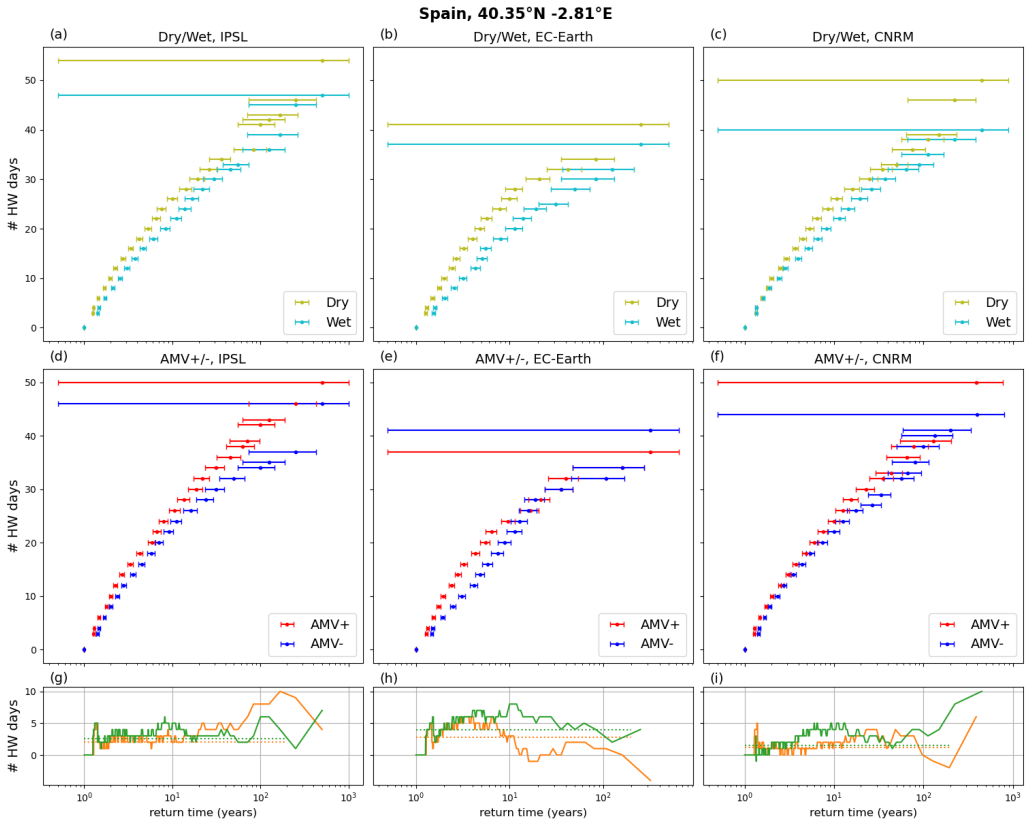


Figure 11.5: *Return time curves for the threshold based definition (section 6.3.1)*. The plots show the yearly number of heatwave days as a function of the return time (in years) for Dry and Wet ensembles (top row) and for AMV+ and AMV- ensembles (bottom row) for the three models. Error bars correspond to one standard deviation of the estimated return time needed to observe an event of a certain amplitude. The third row shows the Dry - Wet (green) and AMV+ - AMV- (orange) differences as a function of the return time. The dotted lines indicate the mean difference values. For each model the chosen grid point covers the coordinate indicated in the figure title which corresponds to a location in the Northeast of Spain, close to the Mediterranean Sea.

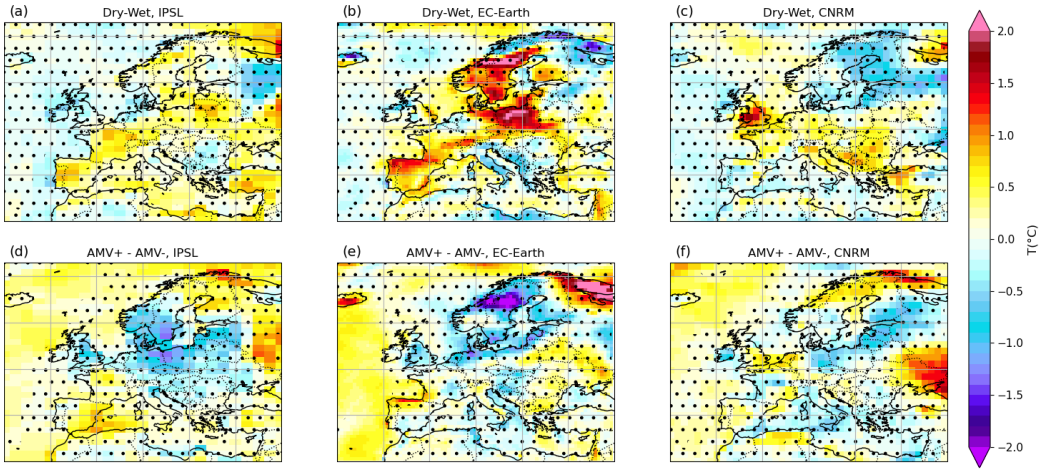


Figure 11.6: *50 years return time maps for 14-day heat waves.* The maps show the difference of $\tilde{A}$ between (top line) the Dry and Wet ensembles and (bottom line) the AMV+ and AMV- ensembles, conditioned over events with return time greater than 50 years. Stippling denotes area below the 95% significance level according to a bootstrap test.

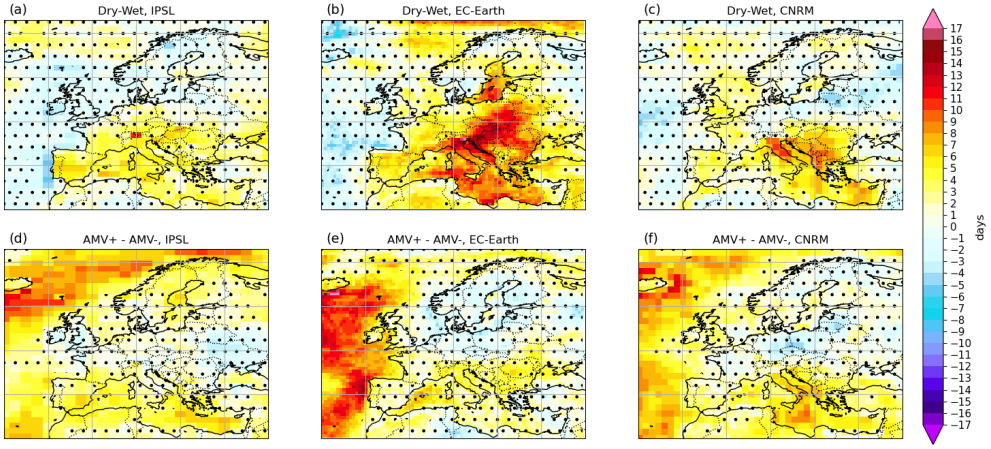


Figure 11.7: *10 years return time maps for the threshold based definition*. The maps show the yearly heatwaves days difference between (top line) the dry and wet ensembles and (bottom line) the AMV+ and AMV- ensembles, conditioned over the return time of 10 years. Stippling denotes area below the 95% significance level according to a bootstrap test.

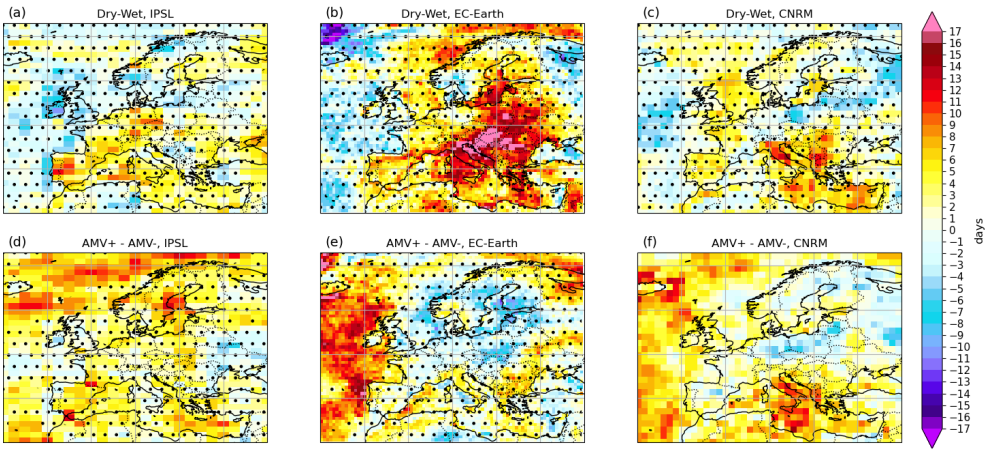


Figure 11.8: *50 years return time maps for the threshold based definition*. The maps show the yearly heatwaves days difference between (top line) the dry and wet ensembles and (bottom line) the AMV+ and AMV- ensembles, conditioned over the return time of 50 years. Stippling denotes area below the 95% significance level according to a bootstrap test.

Part III

GAUSSIAN FRAMEWORK AND OPTIMAL PROJECTION OF WEATHER FIELDS FOR PREDICTION OF EXTREME EVENTS

In the previous chapter (part ii) we analysed and compared two slow drivers of heatwaves acting at sub-seasonal and interdecadal timescales, concluding that soil moisture is a slow driver of greater importance than the Atlantic Multidecadal Variability AMV. For heatwaves with return times of decades, this conclusion is limited by the scarcity of data and calls for alternative methodologies or alternatively for longer datasets. In this chapter we rather focus on the first solution by proposing a new methodology which we call the Gaussian approximation. With this methodology we aim at understanding and at forecasting heatwaves over France. Differently from the previous chapter, we focus on drivers acting solely at sub-seasonal timescales, using both a long climate simulation of 8000 years with the PlaSim climate model and the reanalysis dataset ERA5. They were briefly introduced in section 2.4.

A general conclusion is that this approximation results surprisingly competitive for both tasks, the analysis and the prediction.

For what concerns the tools used in this chapter, we employ *composite maps* for the analysis of the dynamical fields after the heatwave event happened and *committor function* to predict an heatwave with a certain delay time. Both tools were introduced in section 3.1.1 and in section 3.1.2, respectively.

The material presented in this chapter is an article which has been submitted to JAMES (Journal of Advances in Modelling the Earth System). I coauthor this work with Alessandro Lovo[1] and Corentin Herbert[1]. A preprint is available at `https://arxiv.org/abs/2405.20903`. Because of that some concepts and materials might be repeated.

Key points:

- This work presents a new simple framework, called the Gaussian approximation, for a-posteriori and a-priori statistics of extreme events.
- Our method provides an interpretable probabilistic forecast of extreme heatwaves which is competitive with off-the-shelf neural networks.
- The analysis highlights quasi-stationary Rossby waves and low soil moisture as precursors to extreme heatwaves over France.

1 *ENS de Lyon, CNRS, Laboratoire de Physique, F-69342 Lyon, France*

Abstract. Extreme events are the major weather related hazard for humanity. It is then of crucial importance to have a good understanding of their statistics and to be able to forecast them. However, lack of sufficient data makes their study particularly challenging.

In this work we provide a simple framework to study extreme events that tackles the lack of data issue by using the whole dataset available, rather than focusing on the extremes in the dataset. To do so, we make the assumption that the set of predictors and the observable used to define the extreme event follow a jointly Gaussian distribution. This naturally gives the notion of an optimal projection of the predictors for forecasting the event.

We take as a case study extreme heatwaves over France, and we test our method on an 8000-year-long intermediate complexity climate model time series and on the ERA5 reanalysis dataset.

For a-posteriori statistics, we observe and motivate the fact that composite maps of very extreme events look similar to less extreme ones.

For prediction, we show that our method is competitive with off-the-shelf neural networks on the long dataset and outperforms them on reanalysis.

The optimal projection pattern, which makes our forecast intrinsically interpretable, highlights the importance of soil moisture deficit and quasi-stationary Rossby waves as precursors to extreme heatwaves.

PLAIN LANGUAGE SUMMARY

Extreme weather events such as heatwaves are responsible for large financial and human costs and their impact can only be expected to grow in the future. Understanding such events and being able to predict them is therefore of major interest, but suffers from a fundamental problem of lack of data. In this work we present a new framework which addresses this issue by making simple assumptions on the statistics of weather fields relevant for heatwaves. We validate our method using a very long climate simulation. We find that it provides good approximations of atmospheric conditions prevailing during heatwaves, and good prediction capabilities. It even outperforms existing approaches for short datasets, such as those obtained by combining observations and state-of-the-art weather prediction models, which contain much less extreme events than climate simulations but represent more accurately the dynamics of the atmosphere. This approach explains the observed property that more extreme events are simply stronger versions of less extreme ones, and allows to identify the features of atmospheric patterns which are relevant for making predictions. The method is very general and could be applied for many types of extreme events.

12 INTRODUCTION

Extreme weather and climate events, often exacerbated by climate change, have led to major disasters in our recent history (Seneviratne et al., 2012). Heatwaves, in particular, are among the deadliest events. Prolonged exposure to abnormal heat for a certain duration has proven to worsen existing illnesses and to have caused excess deaths during the recent events of the Western European heatwave of 2003 and the Russian heatwave of 2010 (Fouillet et al., 2006; García-Herrera et al., 2010; Barriopedro et al., 2011). Moreover, losses in the agricultural sector with the sequent endangerment of the food production system, together with the endangerment of entire ecosystems, allow to classify heatwaves as events which have critical impacts on the whole society, according to the Intergovernmental Panel on Climate Change (Intergovernmental Panel On Climate Change, 2023).

The intensification and the proliferation of these extreme events in the current climate call for urgent progress in our understanding of the mechanisms that drive them, and for developing prediction tools to anticipate risks. However, the most extreme events are the rarest. For this reason, those two classical tasks of analysis and prediction for extreme event study suffer from large methodological difficulties associated to a lack of both historical and model data (Miloshevich et al., 2022). In this paper we propose a new framework to infer analysis and prediction tools, which is effective with rather short datasets, and efficient for the rare unobserved events up to some approximation we fully characterize. Here, we test thoroughly this framework for extreme heatwaves, but we surmise that it can be applied to a large set of other extreme events.

For the task of understanding which weather conditions led to extreme events, once they have occurred, composite patterns, i.e. maps of averaged dynamical variables conditioned on the outcome of the extreme event, are the most commonly used statistical diagnostic (see for instance (Grotjahn and Faure, 2008; Sillmann and Croci-Maspoli, 2009; Teng et al., 2013; Ratnam et al., 2016; Miloshevich et al., 2023; Noyelle, Yiou, and Faranda, 2024)). As visible in fig. 12.1 for reanalysis data and two other climate models, the composite patterns associated with very extreme events strikingly resemble those for less extreme ones. This fascinating property has not been much commented in the literature before a recent study (Miloshevich et al., 2023) and has never been explained. Whenever this property is relevant, it means that composite maps for rare events can be computed from typical statistics, even if those rare events have not been observed. This is of huge practical interest, and requires understanding. The Gaussian framework we develop in this paper gives a straightforward and enlightening explanation.

For the second task, prediction of future extreme events based on current weather conditions, composite maps are not useful. We clearly demonstrate and explain this in the present paper. The appropriate statistical concept to make predictions is the probability that an extreme event will occur conditioned on

the present state of the climate system, the so-called committor function. However, in order to compute this committor function, one actually has to build a forecasting tool able to estimate this probability. Moreover, the committor function is a function of all the variables which characterize the state of the system, called predictors. For these reasons, it is extremely hard to compute practically and to represent it. Several computations of committor functions have been performed with applications in either geophysical fluid dynamics or in climate sciences (Finkel et al., 2021; Miron et al., 2021; Finkel, Abbot, and Weare, 2020; Lucente et al., 2019; Lucente, Herbert, and Bouchet, 2022; Lucente et al., 2022), using either direct or involved approaches. For climate sciences, methods have been devised using either analogue Markov chains (Lucente, Herbert, and Bouchet, 2022), Galerkin approximations of the Koopman operator (Thiede et al., 2019; Strahan et al., 2021), or neural networks (Lucente et al., 2019; Miloshevich et al., 2022). Neural network seems to be the most efficient and versatile tool. As a matter of fact, there is currently a flourishing literature using neural networks for spatial and temporal predictions of several families of extreme events, such as hurricanes (Racah et al., 2017), tropical cyclones (Giffard-Roisin et al., 2020), droughts (Agana and Homaifar, 2017; Dikshit, Pradhan, and Alamri, 2021), and heatwaves (Chattopadhyay, Nabizadeh, and Hassanzadeh, 2020; Jacques-Dumas et al., 2023; Miloshevich et al., 2022). However, in (Miloshevich et al., 2022) the authors clearly demonstrate that machine learning for rare extreme events is most of the time performed in a regime of lack of data and gives -optimal predictions for typical climate datasets. Moreover, deep learning approaches are, in general, very hard to interpret (Bach et al., 2015; Krishna et al., 2022; Rudin, 2019), and it is extremely difficult to gain some understanding using the forecasting tool.

The main aim of this work is to propose a much simpler alternative method to devise a forecast tool for prediction and to explain the structure of composite maps. This new framework is based on the assumption that the joint probability distribution of the predictors and the extreme event amplitude is Gaussian. Even if this hypothesis is verified only approximately, we show in this paper that the quality of its prediction and its potential for interpretability is extremely high, for extreme heatwaves. We prove that this hypothesis gives a very simple and straightforward explanation of the stability of composite patterns when changing the extreme event amplitude. For the prediction problem, this Gaussian hypothesis leads to a linear regression problem of the heatwave amplitude on the predictor fields. This is in sharp contrast with regression of fields on scalars value, commonly used in climate sciences. In this case, the predictor is a field in very high dimension, and the predicted value is a scalar. The key outcome of this procedure is a regression map, which we call the optimal prediction map for the extreme event. This optimal prediction map is a new concept of this study. It is directly interpretable as it gives, at each geographical location, the importance of the predictor field and its sign to determine the heatwave amplitude. Because of the high dimension of the predictors and because of the not so long dataset length, this regression requires regularization. We analyse thoroughly such optimal prediction maps for extreme heatwaves.

A large part of the work is devoted to the estimate of the accuracy of the results obtained using the Gaussian approximation, compared to the truth. It

turns out that this Gaussian approximation is able to give fully interpretable results which compare very well with the truth. For instance, it computes composite maps up to errors of the order of 20 to 30%, depending on the cases. Moreover, this Gaussian approximation requires much less data, and it can predict composite maps for unobserved events. For prediction, it should often be preferred to neural networks for short datasets. For instance, we prove to have a prediction skill close to convolutional neural networks on very long datasets and to outperform them on short datasets, like the 80-year long ERA5 reanalysis.

This work is organized as follows. In chapter 13 we give the definition of heatwaves used for this study, we present the two datasets used and the set of predictors. In chapter 14 we show with two theoretical examples that composite maps and committor functions are two different probabilistic objects. We then introduce the Gaussian approximation framework, and we derive the formulae for computing composite maps and committor functions. Chapter 15 and chapter 16 are dedicated to a methodological study of the Gaussian framework using the climate model PlaSim. Finally, in chapter 18 we apply our methodology to the reanalysis dataset ERA5. In chapter 19 we summarize our findings and give perspectives for future works.

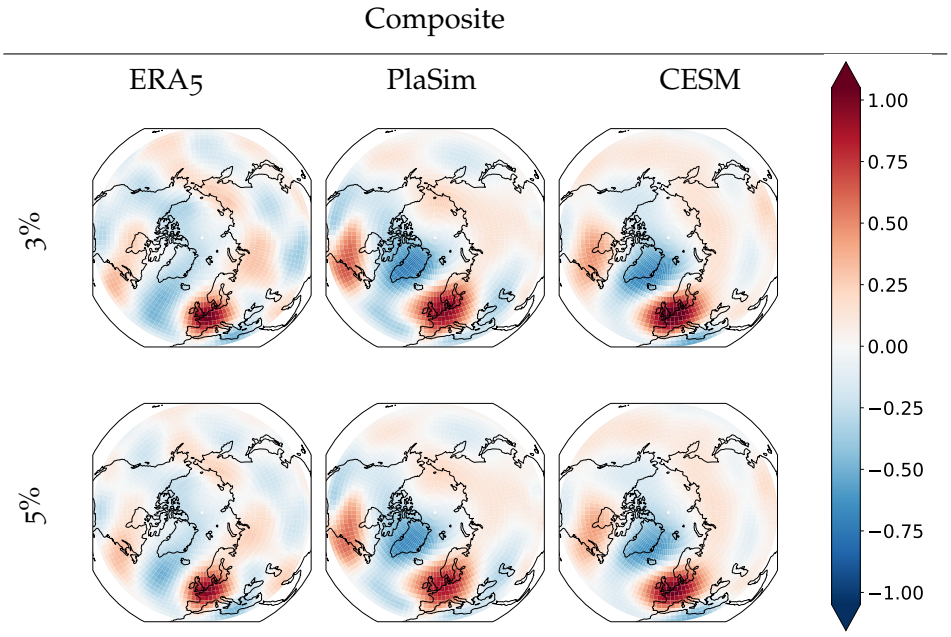


Figure 12.1: First line: maps of 500 hPa geopotential height anomaly for heatwaves over France, defined as situations with the 3% most extreme values of two weeks averaged 2 m temperature anomaly over France (3% composite maps). Second line: the same for a 5% threshold (5% composite maps). The maps are normalized pixel-wise by the climatology standard deviation. Composite maps are estimated respectively on ERA5 (daily data from 1940 to 2022), PlaSim (8000 years of simulation), CESM (1000 years of simulation), datasets. The models reproduce very well ERA5 patterns. Moreover, while the amplitude depends on the threshold defining heatwaves, strikingly the patterns do not. Indeed, we observe in all models and for both thresholds a strong anticyclonic anomaly over Western Europe (which is correctly correlated with the fact that we aim at predicting heatwaves over France). This anticyclonic anomaly is part of a train of a cyclone and an anticyclone which starts over the western part of the United States and continues with a cyclonic anomaly over North Atlantic Ocean for ERA5, while it is northward shifted over Greenland for both PlaSim and CESM.

13 HEATWAVE DEFINITION, DATASETS, AND PREDICTORS

In this section we provide the definition of heatwaves that will be used in the following (section 13.1), we present the datasets (section 13.2), and we identify the weather variables of interest (section 13.3).

13.1 HEATWAVE DEFINITION

In the literature heatwaves have been defined in a plethora of different ways for different analysis purposes (Perkins, 2015). Short and long-lasting heatwaves affect differently our society and environment, but long-lasting ones are the most detrimental (Barriopedro et al., 2011). Despite this, most of the literature on heatwaves focuses on daily events (Seneviratne et al., 2012), as was pointed out in the last assessment report of the Intergovernmental Panel on Climate Change (Intergovernmental Panel On Climate Change, 2023).

Having a definition which measures independently the persistence and the amplitude of heatwaves is thus of primary interest. The simplest way to achieve this is by monitoring the running average of the air temperature field, and this has been applied to the study of heatwaves of different duration (7 days, two weeks, one month) (Barriopedro et al., 2011; Coumou and Rahmstorf, 2012; Schär et al., 2004). In this work, following the recent studies of (Gálfi, Lucarini, and Wouters, 2019; Galfi and Lucarini, 2021; Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2021; Jacques-Dumas et al., 2023; Miloshevich et al., 2022), we use a definition which is based on a time and a spatial average of the 2 m temperature anomaly. We believe that this viewpoint is complementary with the more common definitions (Perkins, 2015) and relevant for our analysis. Such an average-based definition has the advantage of carrying a natural measure of the heatwave amplitude, which can be easily adapted to heatwaves of different duration and intensity or over different regions of the globe. On the contrary, many classical heatwave definitions involve hard thresholds to be reached within specified time frames and are thus less flexible (Perkins, 2015).

Let $\tilde{T}_{2m}$ denote the daily-averaged 2 m air temperature field, which depends on the location $\vec{r}$ and time t. Given that the statistics of $\tilde{T}_{2m}$ are affected by the seasonal cycle, we use temperature anomaly $T_{2m} := \tilde{T}_{2m} - \mathbb{E}_y(\tilde{T}_{2m})$ where $\mathbb{E}_y(\tilde{T}_{2m})$ is the average of $\tilde{T}_{2m}$ over many years for each calendar day, i.e. the climatology. We thus define the heatwave amplitude A as the space and time average of the temperature anomaly:

$$A(t) := \frac{1}{T}\int_t^{t+T}\left(\frac{1}{\mathcal{A}}\int_{\mathcal{A}} T_{2m}(\vec{r},u)d\vec{r}\right)du, \tag{13.1}$$

where T is the duration in days of the heatwave and $\mathcal{A}$ is the spatial region of interest. Both parameters, T and $\mathcal{A}$ can be changed according to the event one wishes to study. In this work, T ranges from one day (short event) to one month (long event), but nothing prevents it from going even to longer, seasonal events.

The region $\mathcal{A}$ typically extends over distances comparable to the synoptic scale, which, in the mid-latitudes, is about 1000 km. This is the order of magnitude of the spatial correlations in tropospheric dynamics, corresponding to the size of cyclones and anticyclones, and of the jet stream meanders. In this study we choose $\mathcal{A}$ to be the equivalent region of France, which is shown for instance in the last column of fig. 15.1. Moreover, as summer heatwaves have higher impacts, we consider only the months of June, July and August.

Following the studies (Jacques-Dumas et al., 2023; Miloshevich et al., 2022), we define an extreme heatwave as an event for which the amplitude A exceeds a threshold a corresponding to rare fluctuations. This threshold can be changed depending on the heatwaves of interest. In this work we will mainly focus on a defined as the 95^{th} quantile of the distribution of A, i.e. we consider as heatwaves the 5% most extreme events in our dataset. For a two-week heatwave, in the PlaSim model (see section 13.2.1), the threshold amounts to $a = 2.76\,\text{K}$. We will also comment briefly on heatwaves that are more or less rare than the 5% most extreme ones.

13.2 DATASETS

In this work we use two datasets. The first is the output of the intermediate complexity climate model called PlaSim, the second is the ERA5 reanalysis data. We use PlaSim to generate an extremely long dataset, over which to train, optimize and test our Gaussian approximation framework (introduced in chapter 14) with little statistical errors. On the other hand, the simplicity of this climate model means that our results may suffer from potentially large biases with respect to the real climate. Hence, after this validation step we also apply our new methods to ERA5 data, which can be expected to suffer from smaller biases and be a more faithful representation of the actual climate.

13.2.1 *PlaSim*

The Planet Simulator PlaSim (Fraedrich et al., 2005b; Fraedrich et al., 2005a) is an intermediate complexity climate model that has a dynamical core that solves the moist primitive equations (Vallis, 2017) in the atmosphere. The model has a T42 horizontal resolution in Fourier space, that in direct space corresponds to a 64×128 grid of 2.8 degrees both in latitude and longitude, with 10 vertical layers and covering the whole globe. The model uses a relatively simplified parametrization of the -grid processes such as radiation, clouds, convection and hydrology over land. For the latter, in particular, PlaSim uses a single-layer bucket model (Manabe, 1969), with soil moisture increased by snow melt and precipitation and depleted by evaporation. Sea ice cover and ocean surface temperature are cyclically prescribed for each day of the year, acting as boundary conditions. By prescribing as well the greenhouse gases concentration and incoming solar radiation, the model is able to run in a steady state that reproduces a climate close to the one of the 1990s.

The fact that PlaSim lacks a dynamic ocean means that, in our study of heatwaves, we cannot investigate the effects of ocean related phenomena such as

El Niño (Hafez, 2017; Zhou and Wu, 2016), or the North Atlantic Oscillation (Hafez, 2017; Li et al., 2020). On the other hand, the representation of the atmosphere of PlaSim is sufficient to properly resolve the large scale dynamics of cyclones, anticyclones and the jet stream, including important teleconnection patterns relevant for heatwaves (Miloshevich et al., 2023). Moreover, the simplified parameterizations used in PlaSim allow it to run 100 times faster than the models used for CMIP studies, which makes it very suitable to obtain extremely long datasets. Here, we use a dataset consisting of 8000 years. It is the same data that was used for previous work on probabilistic forecast of heatwaves using machine learning (Miloshevich et al., 2022). Mode details on the model setup can be found in (Miloshevich et al., 2022).

As we will show, our proposed method for studying heatwaves does not need such a long dataset to achieve good performances. However, we also want to perform comparisons with alternative deep learning methods, and those do require as much data as possible (Miloshevich et al., 2022).

PlaSim resolves the daily cycle and has an output frequency of 3 hours, but we are interested only in daily averages. In particular, we will focus on the anomalies (with respect to the daily, grid point-wise climatology) of 2 m temperature (T_{2m}), 500 hPa geopotential height (Z_{500}) and soil moisture (S).

13.2.2 *ERA5*

In this manuscript we also present an application of our methodology to the ERA5 dataset (Hersbach et al., 2020). We use daily data from the public available dataset of the ECMWF service for summer seasons from 1940 to 2022. ERA5 has a resolution of 0.25 degrees in latitude and longitude. We use this fine resolution to compute the average 2 m temperature anomaly over France and hence the heatwave amplitude A eq. (22.2).

On the other hand, since the dataset is quite small, we reduce the number of predictors (see next section) by using only the 500 hPa geopotential height anomaly field and re-gridding it onto the coarser PlaSim grid.

An important remark is that in our study of heatwaves we assume a stationary climate. We thus need to remove the global warming signal from ERA5 data. This is achieved by means of a parabolic detrending of the averaged temperature over France and of zonal averages of the geopotential height. More technical details on the detrending procedure are given in section 20.1.

13.3 PREDICTORS

To study heatwaves, we focus on a set of climate variables that we call *predictors* and denote it with X. In particular, for a heatwave that starts at time t, we will be interested in the predictors $\tau \geq 0$ days before the event, i.e. $X(t - \tau)$.

For PlaSim, X will be the stack of the anomalies of 2 m temperature (T_{2m}), 500 hPa geopotential height (Z_{500}) and soil moisture (S). The choice of T_{2m} is straightforward given its implication in heatwaves, and the potential of simple persistence and advection of temperature to be useful for prediction. The geopotential height anomaly at the middle of the troposphere (Z_{500}) is a good

representation of the dynamical state of the atmosphere because of its relation with cyclones and anticyclones in the lower troposphere. At that height, the geostrophic approximation applies and thus Z_{500} gives also a good insight into the wind flow. Finally, it has been shown that low soil moisture acts as an important preconditioning factor for the occurrence of extreme summer temperatures in the mid-latitudes, by limiting the evaporative cooling of the surface (Perkins, 2015; Miloshevich et al., 2022; Benson and Dirmeyer, 2021; D'Andrea et al., 2006; Fischer et al., 2007b; Hirschi et al., 2011; Lorenz, Jaeger, and Seneviratne, 2010; Rowntree and Bolton, 1983; Schubert et al., 2014; Shukla and Mintz, 1982; Stefanon, D'Andrea, and Drobinski, 2012; Vargas Zeppetello and Battisti, 2020; Zeppetello, Battisti, and Baker, 2022; Zhou et al., 2019; Vautard et al., 2007).

For the 2 m air temperature and 500 hPa geopotential height fields we will focus on the whole Northern Hemisphere (latitude above 30 degrees North), while soil moisture, instead, is a local variable, and we care only about the values on our region of interest (France). Considering the resolution of the PlaSim model, this will amount to a total of $d = 5644$ scalar predictors.

On the other hand, for ERA5 we use only the 500 hPa geopotential height anomaly field, which yields a total of $d = 2816$ pixels.

For both datasets, as it is commonly done in the machine learning community, we normalize each field value at each grid point independently dividing by its standard deviation. This way, X will be a collection of d (correlated) dimensionless variables with zero mean and unitary standard deviation, which also allows us to easily compare fields with different physical units.

14

OPTIMAL PROJECTION, COMMITTOR FUNCTIONS, COMPOSITE MAPS, AND THE CASE OF GAUSSIAN STATISTICS

As climate scientists, concerned in understanding extreme events, we might ask two classes of questions. The first class is related to prediction or a priori statistics: given the current state of the system (the predictors X), what is the probability to observe an extreme event starting within τ days? The second class of question is related to a posteriori understanding: given that the extreme event actually occurred, what were the probabilities of the system states leading to this event? For instance, composite maps defined as the averaged state given that the event occurred, widely used by climate scientists, are examples of a posteriori statistics. Both a priori and a posteriori statistics are useful and important for the sake of understanding, but only a priori statistics is useful for prediction.

Indeed, the first goal of this section is to stress the difference between a priori and a posteriori statistics. For instance, it is key to understand that in general composite maps do not provide useful information for prediction. At the same time, we define some useful statistical quantities for prediction, namely the committor function (see a definition below). The second goal is to explain the difficulty to compute committor functions, motivating why they are not commonly used. The third and final goal is to devise predictive and simply interpretable statistical models, for instance the regression of the predictors (the state X) on the extreme event observable.

14.1 A POSTERIORI STATISTICS ARE USUALLY NOT USEFUL FOR PREDICTION

In this section we will stress the differences and the links between a posteriori and a priori statistics.

Let's consider two events, F and G, where G happens after F. We will denote with $\mathbb{P}(F|G)$ the *a posteriori* probability of F conditioned on the happening of the future event G. Vice versa, $\mathbb{P}(G|F)$ will be the *a priori* probability of G conditioned on the past event F. In our case the past event will be the predictors being in a particular state $X = x$, while the future event will be the realization of a heatwave $Y = 1$, where Y is the binary random variable

$$Y(t) := \begin{cases} 1 & \text{if } A(t) \geq a \\ 0 & \text{otherwise} \end{cases} \tag{14.1}$$

and a is the threshold which defines an heatwave and will be the quantile of the distribution of A.

14.1.1 *Bayes Formula*

When comparing different conditional probabilities, we can make use of Bayes formula:

$$\mathbb{P}(X = x|Y = 1)\mathbb{P}(Y = 1) = \mathbb{P}(X = x, Y = 1) = \mathbb{P}(Y = 1|X = x)\mathbb{P}(X = x), \tag{14.2}$$

where

- $\mathbb{P}(X = x, Y = 1)$ is the joint probability of being in state x and experiencing a heatwave ($Y = 1$)
- $\mathbb{P}(X = x) =: P_S(x)$ is the *stationary measure* of the predictors, namely the probability of being in state x
- $\mathbb{P}(Y = 1|X = x) =: q(x)$ is the a priori *committor function*: the probability of observing a heatwave, conditioned on being in state x
- $\mathbb{P}(Y = 1) = \int q(x)P_S(x)dx =: p$ is the unconditional (or climatological) probability of having a heatwave, inversely proportional to its *return time*, that tells us how extreme the event is.
- $\mathbb{P}(X = x|Y = 1)$ is the a posteriori probability that the state of the predictors were x given that the heatwave occurred.

Summarising, Bayes formula clearly shows the difference and the relation between a priori and a posteriori statistics. In the next sections we will illustrate a proper tool for the prediction task, namely the committor function, and we will illustrate for what composite maps can be used for, namely a posteriori statistics.

14.1.2 *Definition of Committor Functions*

If one is interested in a prediction task, the proper tool is the committor function $q(x)$, originally introduced in the field of stochastic processes (see section 20.5) for studying transitions between attractors (Bolhuis et al., 2000; Lucente, Herbert, and Bouchet, 2022). In our case we do not have two attractors, but rather a typical state of the climate with no heatwaves ($Y = 0$) and an atypical one ($Y = 1$). In this context the concept of transition gets a bit blurred, and the committor is simply the a priori conditional probability mentioned before. If we expand the notation and introduce back the lead time τ, we can write it as

$$q(x) = \mathbb{P}(A(t) \geq a \mid X(t - \tau) = x). \tag{14.3}$$

where a is the threshold used to define a heatwave. As we will discuss later, committors are extremely hard to compute properly and hence are quite rarely used in the field of climate sciences. However, they are *the* right tool for prediction, and even a very rough estimate of them is better than alternative methods.

14.1.3 *Definition of Composite Maps*

On the other hand, a commonly used tool in the climate community to study a wide range of events, including the extreme ones, is the *composite map* (Grotjahn and Faure, 2008; Sillmann and Croci-Maspoli, 2009; Teng et al., 2013; Ratnam et al., 2016; Miloshevich et al., 2023; Noyelle, Yiou, and Faranda, 2024). It is defined as the average state of the climate τ days before the heatwave happened:

$$C := \mathbb{E}(X(t-\tau)|A(t) \geq a), \tag{14.4}$$

where $\mathbb{E}$ denotes an expectation over event realizations and a is the threshold used to define a heatwave. In practice one would estimate such expectation with an empirical average over all the heatwave events in the dataset, which makes the composite one of the easiest objects to compute and hence motivates its popularity.

It is important to point out that the empirical average will be a good estimate of the true composite provided that the number of heatwave events is enough. This means that, depending on the size of our dataset, a direct estimation of the composite map is useful only for not too rare (extreme) events, because of sampling errors.

Going back to the simpler notation used earlier, we can interpret the composite as the mean of the a posteriori probability distribution

$$C = \mathbb{E}(X|Y=1) := \int x\mathbb{P}(X=x|Y=1)dx, \tag{14.5}$$

and thus, through Bayes theorem, we can relate it to the stationary measure P_S and the committor function q.

$$C = \int x \frac{\mathbb{P}(X=x)\mathbb{P}(Y=1|X=x)}{\mathbb{P}(Y=1)}dx = \frac{\int xP_S(x)q(x)dx}{\int P_S(x)q(x)dx}, \tag{14.6}$$

Equation (14.6) clearly shows that the composite is the mean of a distribution proportional to $P_S(x)q(x)$ and thus not equivalent to $q(x)$. In particular, for rare events, we expect $q(x)$ to be peaked for very atypical values of x, namely in the tail of the stationary measure $P_S(x)$. Thus the composite map may differ significantly from the typical states x associated with a high committor.

14.1.4 *Two Simple Examples which Illustrate that Composites Might be Useless for Prediction*

Now that we have defined the important quantities of interest, we will use some examples to highlight the difference between composites and committor, and in particular how the first may not give us any useful insights on the second.

As a first example, let us assume that our predictor is one dimensional ($X \in \mathbb{R}$), with stationary measure given by a simple normal distribution $P_S(x) \propto \exp\left(-\frac{x^2}{2}\right)$. Similarly, let the committor function be another Gaussian distribution centered in $x^* > 0$ and with standard deviation σ: $q(x) = \mathbb{P}(Y=1|X=x) \propto \exp\left(-\frac{(x-x^*)^2}{2\sigma^2}\right)$. This means that the probability of a heatwave is maximum when we are in state $X = x^*$. We will now compute the composite, and show that it is different from x^*.

From eq. (14.6) we know that the composite is the mean of a distribution proportional to $P_S(x)q(x)$, and with some trivial algebraic manipulations, we find that

$$P_S(x)q(x) \propto \exp\left(-\frac{x^2}{2} - \frac{(x-x^*)^2}{2\sigma^2}\right) \propto \exp\left(-\frac{1}{2}\left(1+\frac{1}{\sigma^2}\right)\left(x - \frac{x^*}{\sigma^2+1}\right)^2\right).$$

Hence, the composite is

$$C = \frac{x^*}{\sigma^2+1},$$

which is strictly smaller than the condition where the heatwave probability is highest. An important consequence is that the probability of having a heatwave when we are in the composite state may be vanishingly small depending on the values of x^* and σ, showing the low predictive power of the composite map:

$$\frac{q(C)}{q(x^*)} = \exp\left(-\frac{1}{2}\left(\frac{x^*}{\sigma+\sigma^{-1}}\right)^2\right).$$

As a second example, let us consider $X = (X_1, X_2) \in \mathbb{R}^2$ with $P_S(x)$ being a distribution that correlates the two components X_1 and X_2, for instance a bi-variate Gaussian with mean $(0,0)$ and covariance matrix $\begin{pmatrix}\sigma_1^2 & \phi \\ \phi & \sigma_2^2\end{pmatrix}$. We will then consider a committor $q(x) = q(x_1)$ that depends only on the first component. Without going into the details (available in section 20.3), it will be clear that the composite map will have a non-zero x_2 component, thanks to the correlation ϕ between x_1 and x_2. However, we know that the committor depends only on x_1, and so the composite will be misleading if we are interested in prediction, as it will draw our attention to variables that do not contain *any* information about the probability of having a heatwave.

In conclusion, the composite map is an average that takes into account both the probability of having a heatwave starting from state x *and* the probability of *being* in state x (eq. (14.6)). This is good to study the statistics of our extreme event, but if we want to know if there is going to be a heatwave tomorrow, we do not care how rare it was to have had today's weather.

14.2 COMMITTOR FUNCTIONS AND OPTIMAL PROJECTION

Now that we have a clear mathematical understanding of committor functions as the proper tool for prediction, we can move to the problem of computing them in practice. In this section we will point out why this is such a complex task as well as provide a way to evaluate how good any approximation of the true committor is. Finally, we will propose the framework of optimal projection of the committor, which will mitigate the problem of high dimensionality as well as make the committor much more interpretable.

14.2.1 *Complexity of Committor Functions*

The committor is a function that maps every point of the phase space x to a number $q(x)$ between 0 and 1 that quantifies the likelihood of having a heatwave. A naive way of estimating the committor would be to initialize many trajectories at the point x and count how many actually lead to a heatwave. This method is called direct numerical simulation, and, if rather inefficient, it is still doable for simple stochastic processes in low dimensional spaces.

In our case, however, $x \in \mathbb{R}^d$, with $d = 5644$ for PlaSim and $d = 2816$ for ERA5 and the dynamics is described by a rather complex climate model. One could argue that we do not need to explore the whole $\mathbb{R}^d$ space, but only the much lower dimensional manifold of *physical* states, which, under ergodic conditions, would be properly sampled by an extremely long trajectory. This argument is absolutely correct, but the task of a thorough and precise sampling of the committor still remains out of reach, even with the help of supercomputers.

Given the importance of committor functions, there is incentive in finding efficient ways to get a reasonable approximation of the committor, potentially also limiting the search to only the physical states that are most likely to yield a heatwave. This makes the task feasible, but far from simple, and attempts have been made using machine learning (Miloshevich et al., 2022), rare event algorithms (Ragone, Wouters, and Bouchet, 2018) or both (Lucente et al., 2022).

In this work, we strive to find an approach which is far simpler than all of the aforementioned, yet still leads to a good enough approximation of the committor.

14.2.2 *Evaluation of Approximations of the Committor Function*

To quantify how good an approximation $\hat{q}$ of the true committor q is, we need a sort of distance between the two. Since committors are probabilities, the natural object to use is the Kullback-Leibler divergence

$$KL(q, \hat{q}) = \int P_S(x) \left(q(x) \log \left(\frac{q(x)}{\hat{q}(x)} \right) + (1 - q(x)) \log \left(\frac{1 - q(x)}{1 - \hat{q}(x)} \right) \right) dx, \quad (14.7)$$

which quantifies the amount of information lost when using $\hat{q}$ instead of q. Expanding the logarithm and removing the terms that depend only on the true committor, we are left with the cross entropy loss.

$$CE(q, \hat{q}) = - \int P_S(x) \left(q(x) \log \hat{q}(x) + (1 - q(x)) \log(1 - \hat{q}(x)) \right) dx. \quad (14.8)$$

Now, since we do not have access to neither the true committor q nor the stationary measure $P_S(x)$, we can replace the first with the heatwave labels Y and the integral over the second with the average over our dataset $\mathcal{D}$. We obtain then the empirical cross entropy loss

$$\mathcal{L} = - \left\langle Y(t) \log \hat{q}(X(t)) + (1 - Y(t)) \log \left(1 - \hat{q}(X(t))\right) \right\rangle_{(X(t), Y(t)) \in \mathcal{D}}, \quad (14.9)$$

which is proven to be the only proper score for a probabilistic forecast (Benedetti, 2010).

$\mathcal{L} = 0$ is the perfect prediction, but $\mathcal{L}$ can be arbitrarily large. To have a reference we can consider the climatological committor, that comes from assuming the only information we have is that we are studying the p-eth most extreme heatwave, for example setting the threshold a to be the 95^{th} quantile of the distribution of A means $p = 0.05$. With only this information, the climatological committor is the constant p, and the associated empirical cross entropy is

$$\begin{aligned}\mathcal{L}_{\text{clim}} &= -\left\langle Y(t)\log p + (1-Y(t))\log(1-p)\right\rangle_{(X(t),Y(t))\in\mathcal{D}} = \\ &= -p\log p - (1-p)\log(1-p).\end{aligned} \tag{14.10}$$

Finally, we can define the *normalized log score* $\mathcal{S}$ as in (Miloshevich et al., 2022), that will quantify the skill of our prediction:

$$\mathcal{S} := 1 - \frac{\mathcal{L}}{\mathcal{L}_{\text{clim}}}. \tag{14.11}$$

A value $\mathcal{S} = 1$ will mean a perfect prediction, namely $\hat{q}(t) = Y(t)\ \forall t$, and $\mathcal{S} < 0$ will mean that our forecast is worse than the climatology.

14.2.3 *Optimal Committor Projection*

Now that we have the tools for evaluating committor approximations, we can tackle the problem of the high dimensionality of $q : \mathbb{R}^d \to [0,1]$. The key idea is to write a surrogate committor $q_\phi = \tilde{q} \circ \phi$, which first applies a projection $\phi : \mathbb{R}^d \to \mathbb{R}^m$ to a space with dimension $m \ll d$, and then represents the committor in this reduced space with function $\tilde{q} : \mathbb{R}^m \to [0,1]$. We want to perform this decomposition in an *optimal* way, which means minimizing the cross entropy defined above, i.e., losing as little information as possible about the original committor.

It is relatively easy to see that, for a given projection function ϕ, the best committor representation is the average of the original committor on the iso-levels of ϕ

$$\tilde{q}^*(f) = \mathbb{E}_{x\in\phi^{-1}(f)} q(x). \tag{14.12}$$

Moreover, the information loss comes from mapping very different values of the original committor onto the same iso-level. Ideally, then, the optimal projection would be the one that has the same iso-levels of q, namely q itself (up to any monotonic rescaling). Of course this is not desirable, as we simply shifted the problem from computing q to computing ϕ. To have something useful, we need to constrain the search space of ϕ, for example to linear maps.

Even with these simplifications, the general problem remains hard to treat in practice. In the next section, we will show the case of Gaussian statistics, which gives an analytic way to compute the optimal linear projection, as well as the reduced committor.

14.3 THE CASE OF A JOINT GAUSSIAN DISTRIBUTION

In this section we present the theory for what we call the Gaussian approximation. We describe the theoretical idea and derive analytically the expressions for the composite map and the committor function.

The Gaussian approximation consists in assuming that the predictor X at time $t-\tau$ and the heatwave amplitude A at time t follow a jointly Gaussian distribution

$$(X(t-\tau), A(t)) \sim \mathcal{N}(0, \Sigma(T,\tau)), \tag{14.13}$$

where X is thought of as a d-dimensional vector, and represents all grid-point values of either a single field or stacked fields. The joint distribution has mean zero because both X and A are anomalies, and it is then solely characterized by the $d+1$ dimensional covariance matrix $\Sigma(T,\tau)$, that depends on the heatwave duration and the lead time.

To simplify the notation, we assume that we work at fixed T and τ, and thus drop the dependencies on them. We can then write Σ as a block matrix of the form $\begin{bmatrix} \Sigma_{XX} & \Sigma_{XA} \\ \Sigma_{AX} & \Sigma_{AA} \end{bmatrix}$, where $\Sigma_{XX} = \mathbb{E}(XX^\top)$ is the $d \times d$ covariance matrix of X, $\Sigma_{XA} = \Sigma_{AX}^\top = \mathbb{E}(XA)$ is the $d \times 1$ correlation map between X and A and $\Sigma_{AA} = \mathbb{E}(A^2)$ is the scalar variance of A.

14.3.1 *Composite Maps Within the Gaussian Approximation*

Under the Gaussian assumption, the composite map can be computed analytically as

$$C_{\mathcal{G}} = \mathbb{E}[X|A \geq a] = \int x \frac{\int_a^{+\infty} \mathbb{P}(x,A)dA}{\int_a^{+\infty} \mathbb{P}(A)dA} dx = \eta\left(\frac{a}{\sqrt{2\Sigma_{AA}}}\right) \frac{\Sigma_{XA}}{\sqrt{\Sigma_{AA}}}, \tag{14.14}$$

with

$$\eta(z) = \sqrt{\frac{2}{\pi}} \frac{e^{-z^2}}{\operatorname{erfc}(z)}, \tag{14.15}$$

where $\operatorname{erfc}(\bullet)$ is the complementary error function and the script $\mathcal{G}$ reminds that the composite is evaluated under the Gaussian assumption. The detailed computation is shown in section 20.4.

From eq. (14.14), we can clearly see that the composite is directly proportional to the correlation map, with the proportionality constant depending only on the threshold a. This has the important implication that the average state of the climate τ days before a heatwave looks like the τ-lagged correlation between the fields and the heatwave amplitude, *regardless* of how extreme the heatwave is. In other words, the composite of a more extreme event has exactly the same pattern as a less extreme one, but amplified according to the function η. The fact that we do observe this effect in the actual data (fig. 12.1) suggests a good validity of our Gaussian approximation. We test it more thoroughly in chapter 15. Moreover, it gives us access to composites of very extreme events, where the direct estimation as the average over the (very small) heatwave set would suffer from huge sampling errors. On the other hand, the correlation map Σ_{XA} is estimated on the whole dataset and thus does not have this issue.

The function η is plotted in fig. 20.3, and has the interesting property that $\eta(z) \sim \sqrt{2}z$ as $z \to \infty$, which means that for very extreme heatwaves the composite map tends to the simple linear regression of X against A.

$$C_{\mathcal{G}} \xrightarrow[a \gg \sqrt{\Sigma_{AA}}]{} a\frac{\Sigma_{XA}}{\Sigma_{AA}} = a\xi, \quad \xi = \arg\min_{\xi} \mathbb{E}\left((X - A\xi)^2\right). \tag{14.16}$$

14.3.2 *Committor Functions Within the Gaussian Approximation*

By definition, the committor is the integral of the a priori distribution of A conditioned on knowing X:

$$q(x) = \mathbb{P}(A \geq a | X = x) = \int_a^{+\infty} \mathbb{P}(A = a | X = x) dA. \tag{14.17}$$

Under the assumption of a joint Gaussian distribution for (X, A), the conditional distribution of A given X is also Gaussian. In particular it has mean $\mu(x)$ that scales linearly with x and constant variance σ^2:

$$\mu(x) = \Sigma_{XX}^{-1}\Sigma_{XA} \cdot x, \qquad \sigma^2 = \Sigma_{AA} - \Sigma_{AX}\Sigma_{XX}^{-1}\Sigma_{XA}. \tag{14.18}$$

For the details of this computation see section 20.4. In fact, $\mu(x) = \tilde{M}^\top x$ is precisely the linear regression of A against X:

$$\begin{aligned} \tilde{M} &:= \Sigma_{XX}^{-1}\Sigma_{XA} = \arg\min_M \left(M^\top \Sigma_{XX} M - 2M^\top \Sigma_{XA}\right) = \\ &= \arg\min_M \mathbb{E}\left((A - M^\top X)^2\right). \end{aligned} \tag{14.19}$$

Then, to obtain the full committor, we just have to compute the Gaussian integral in eq. (14.17), which gives

$$q_{\mathcal{G}}(x) = \frac{1}{2}\mathrm{erfc}\left(\frac{a - \tilde{M}^\top x}{\sqrt{2}\sigma}\right). \tag{14.20}$$

This result can be viewed in light of the framework of optimal committor projection presented in section 14.2.3. In this case, the optimal projection of the high dimensional committor is onto the normalized projection pattern

$$M = \frac{\Sigma_{XX}^{-1}\Sigma_{XA}}{|\Sigma_{XX}^{-1}\Sigma_{XA}|}, \tag{14.21}$$

which condenses all the important information of the high dimensional vector x into the scalar variable $f = M^\top x$. Then the committor in the projected space is simply

$$\tilde{q}(f) = \frac{1}{2}\mathrm{erfc}\left(\alpha + \beta f\right), \tag{14.22}$$

with

$$\alpha = \frac{a}{\sqrt{2}\sigma}, \quad \beta = -\frac{|\tilde{M}|}{\sqrt{2}\sigma}. \tag{14.23}$$

The two operations of linear projection and reduced committor can also be viewed as the architecture of a simple one layer perceptron with the custom activation function $\tilde{q}$. In comparison to other neural network architectures (such as convolutional ones) that may be trained on the same task (Miloshevich et al., 2022), this approach is far simpler, and depends on a much smaller number of parameters.

In addition, we would like to stress that the method is *interpretable* by design: with complex neural networks one may need sophisticated explainable AI techniques to understand *why* they are outputting a particular probability (McGovern et al., 2019; Toms, Barnes, and Ebert-Uphoff, 2020; Delaunay and Christensen, 2022), while in our case the answer is straightforward, namely, it is computing the optimal index f. Furthermore, since the projection pattern M has the same dimension as the predictor X, we can plot it as a map, representing the relative importance of each pixel in our predictor, and providing potential insight in the physical dynamics leading to extreme heatwaves.

Another interesting point to pay attention to is the difference of the two linear regressions for the composite (eq. (14.16)) and for the committor (eq. (14.19)). In the first case, we are doing d independent linear regressions of each pixel in X against the heatwave amplitude A, while for the committor we have a single optimization, regressing A against X. This shows once again the fundamental difference between a posteriori and a priori statistics.

In the following sections, we apply the Gaussian approximation to actual data, see to what extent the assumption of Gaussianity holds and what useful information we are able to extract.

15

VALIDATION OF THE GAUSSIAN APPROXIMATION FOR THE COMPUTATION OF COMPOSITE MAPS FOR EXTREME HEATWAVES

Composite maps are very interesting to understand weather situations that actually led to extreme events (a-posteriori statistics). They are actually defined as the average of weather variables conditioned on the future occurrence of the extreme event.

In section 15.1 we show and compare qualitatively composite maps evaluated empirically and using the Gaussian approximation. In section 15.2 we quantify the error made under the Gaussian approximation, and we distinguish systematic and sampling errors. In section 15.3, using the Gaussian approximation, we give an explanation of the puzzling independence of the empirical composite maps patterns from the threshold a used to define an extreme heatwave. Finally, in section 15.4 we discuss in more detail the effect on the quality of the Gaussian approximation of both the dataset length and the threshold defining extreme events, and conclude that the Gaussian approximation is the best way to estimate composite maps in a regime of lack of data.

In chapter 17, we will use these results to make a physical analysis of extreme events, by varying the heatwave duration T and the lead time τ.

In this section we use the PlaSim dataset with 8000 years of data and predictors $X = (T_{2m}, Z_{500}, S)$ (see section 13.2.1). We show an application of our methodology to the ERA5 dataset in chapter 18.

15.1 COMPARING EMPIRICAL COMPOSITE MAPS WITH COMPOSITE MAPS COMPUTED WITHIN THE GAUSSIAN APPROXIMATION

We now compare the composite maps computed either directly from the data or using the Gaussian approximation, showing that the two are qualitatively very similar, with a relative error of the order of 20%. We consider 14-day heatwaves ($T = 14$), looking at the composites for the first day of the heatwave (lead time $\tau = 0$) and we first focus on the 5% most extreme heatwaves ($a = 2.76\,\mathrm{K}$).

Composite maps C are averages of the predictors X conditioned on the occurrence of a heatwave: $C = \mathbb{E}[X(t)|A(t) \geq a]$ (see chapter 15). We first estimate this conditional expectation as an empirical average $C_{\mathcal{D}} = \frac{1}{N}\sum_{\mu=1}^{N} x_\mu$, where $\{x_\mu\}_{\mu=1}^{N} = \{X(t-\tau)|A(t) \geq a\}$. Figure 15.1 shows the empirical composite maps for the three predictors X (top row). We observe a positive anomaly of both 2 m temperature and 500 hPageopotential height over France and Western Europe, which is expected since we are conditioning over events that are happening over the French region. In the PlaSim grid, France is identified as the 12 pixels shown for the soil moisture field. Soil moisture anomaly displays negative values, as the soil tends to be drier than usual when heatwaves happen. In the rest of the Northern Hemisphere, we see teleconnection patterns in the tem-

perature and geopotential height field, in particular a cyclone over Greenland and an anticyclone over the mid and eastern United States.

All these important features are also visible in the composite map $C_{\mathcal{G}}$ computed with the Gaussian approximation (using eq. (14.14)), represented in fig. 15.1 (middle row), to the point that the only visible discrepancy with the empirical map is slightly darker shades of soil moisture. Indeed, if we take the difference between the two estimates of the composite (fig. 15.1, bottom row), most of the weight is concentrated on the soil moisture field. However, non-trivial patterns are also visible in the temperature and geopotential fields. The latter, in particular, shows a wave zero pattern, with positive values around the polar region and negative ones in the mid-latitudes. The amplitude of the difference, read on the color bar, is on the order of 20% of the amplitude of the composite. To have a more quantitative measure, we compute the ratio $\mathcal{R}$ between the L2 norms of the difference between the two composites and the empirical one:

$$\mathcal{R} = \frac{|C_{\mathcal{D}} - C_{\mathcal{G}}|}{|C_{\mathcal{D}}|}. \tag{15.1}$$

In evaluating the norms, we took into account that we consider grid-cells of different areas. For the parameters considered in this section, the norm ratio is $\mathcal{R} = 0.21$, in agreement with our visual estimate (in chapter 17 we will investigate how this metric varies with the heatwave duration T and the lead time τ).

In the next section, we analyse in more detail the sources of the difference between the two estimates. We will then give an explanation of the striking independence of the pattern from the extreme event threshold a in section 15.3.

15.2 QUANTIFICATION OF THE QUALITY OF THE GAUSSIAN APPROXIMATION FOR COMPOSITE MAPS OF EXTREME HEATWAVES

In the previous section, we showed that the empirical composite map $C_{\mathcal{D}}$ and the Gaussian composite map $C_{\mathcal{G}}$ differ at most by 20% (fig. 15.1, bottom row). A natural interpretation of this difference is that it is an error due to the fact that the Gaussian assumption is not exactly satisfied, and therefore the Gaussian composite map is only an approximation of the true composite map. Indeed, we can investigate the validity of this assumption by visualizing the joint and marginal distributions of the heatwave amplitude A and the predictors at the grid-point level, for regions of low or high error (see section 20.9). For instance, we show in fig. 20.3 that the assumption is poorly satisfied for soil moisture at a grid point over France, where the error is large, while it is a much better assumption for geopotential over Greenland, where the error is small.

However, another source of discrepancy between the two composites is the sampling error affecting $C_{\mathcal{D}}$ due to the limited number of heatwaves in the dataset over which we perform the empirical average. Indeed, if we focus on a single pixel i, and call $\{x_\mu\}_{\mu=1}^{N} = \{X^i(t-\tau)|A(t) \geq a\}$ the set of heatwave events, the central limit theorem tells us that

$$\sqrt{N_{\text{eff}}}\frac{C^i - C^i_{\mathcal{D}}}{\sigma(C^i_{\mathcal{D}})} \xrightarrow[N\to\infty]{} \mathcal{N}(0,1), \tag{15.2}$$

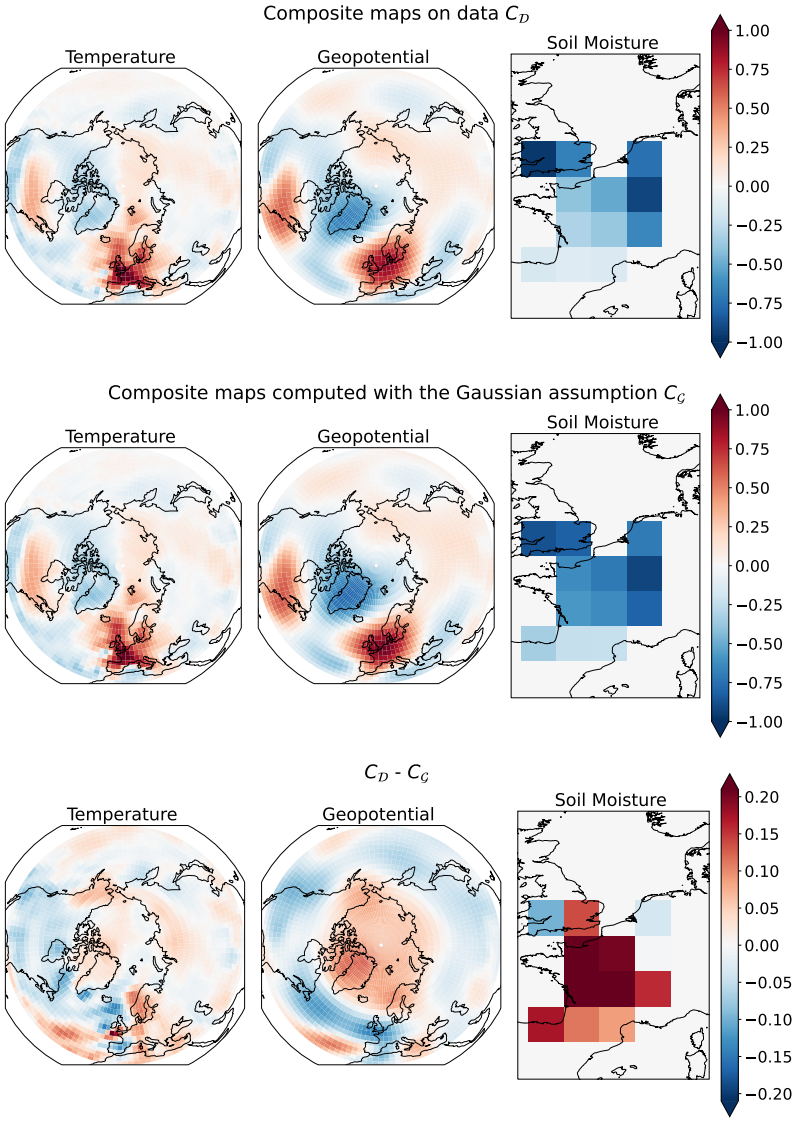


Figure 15.1: Composite maps of normalized 2 mtemperature, 500 hPa geopotential height, and soil moisture anomalies, conditioned on events with the 5% most extremes 14-day temperature over France. Composite maps are computed either directly from PlaSim data (first line), or under the Gaussian approximation (second line). The third line shows the difference between the two. The salient features of both temperature and geopotential are well captured by the Gaussian approximation, with errors of the order of 20% at most.

where C^i is the true composite, $C^i_{\mathcal{D}} = \frac{1}{N}\sum_{\mu=1}^{N} x_\mu$ is the empirical one, $\sigma(C^i_{\mathcal{D}}) = \sqrt{\frac{1}{N}\sum_{\mu=1}^{N}(x_\mu - C^i_{\mathcal{D}})^2}$ is the standard deviation of the heatwave set and N_{eff} is the number of *effectively* independent heatwaves. If all the x_μ were actually independent, we would have $N_{\text{eff}} = N$, but from our definition of heatwave (eq. (22.2)), it is very likely that a series of consecutive days will be all heatwave events, and thus far from independent. In this paper we decide to fix N_{eff} to the number of years with at least one heatwave (equals to 2627 years for 5% most extreme heatwaves of duration $T = 14$ days and lead time $\tau = 0$). The motivation beside this choice can be found in section 20.8.

Equation (15.2) tells us, then, that the distance between the empirical composite and the true one will be of the order of $\frac{\sigma(C^i_{\mathcal{D}})}{\sqrt{N_{\text{eff}}}}$, and thus if the Gaussian composite $C^i_{\mathcal{G}}$ falls much farther than $\frac{\sigma(C^i_{\mathcal{D}})}{\sqrt{N_{\text{eff}}}}$ from the empirical one, we can safely say that is also far from the true composite. In other words, we can define the statistical significance of the error we make as

$$s_i = \frac{\sqrt{N_{\text{eff}}}|C^i_{\mathcal{G}} - C^i_{\mathcal{D}}|}{\sigma(C^i_{\mathcal{D}})}. \tag{15.3}$$

To obtain a global metric for the whole composite map, we can consider the fraction of area $\mathcal{F}$ that have a significance above 2. This allows us to say that, with 95% confidence, a fraction $\mathcal{F}$ of the region of interest has a systematic error, not explainable by the finite size effect of the empirical composite. For the parameters studied here, we obtain the value $\mathcal{F} = 0.37$ (in chapter 17 we will investigate how this metric varies with the heatwave duration T and the lead time τ).

This allows us to conclude that the Gaussian composite suffers from a statistically significant error over roughly half the domain. In spite of this, it gives a reasonable approximation of the empirical composite, within an error of order 20%. However, having 8000 years of data to work with is not common in the climate community, especially when working with observational data or complex model simulations, and we can expect that when data is scarce, the error due to the Gaussian approximation becomes smaller than the sampling error in the empirical composite. In section 15.4 we will address this point on the dataset length and identify a regime where the Gaussian composite gives a better estimation of the true one than the empirical composite.

15.3 COMPOSITE MAPS DO NOT DEPEND MUCH ON THE EXTREME EVENT THRESHOLD

This section aims firstly at giving an explanation for the striking independence of composite maps pattern from the threshold a. Secondly, we show how the norm of the empirical composite maps scales with the threshold a and that this scaling is very close to the one predicted from the Gaussian composite.

In section 14.3.1 we explained that the composite map pattern does not depend on the extreme event threshold a. The independence of the pattern of the empirical composite maps from the threshold a is explained by the Gaussian composite, eq. (14.14). In this equation we see that the threshold intervenes

only in the scaling of the pattern and not on the structure of the pattern itself, which is precisely what we observe in the estimated composite maps. Indeed, in fig. 20.4 we show the difference between the empirical composite and the Gaussian one (evaluated using eq. (14.14)) for the three fields, namely (from the left) 2 m air temperature anomaly, 500 hPa geopotential height anomaly and soil moisture anomaly evaluated for a corresponding to the 1% most extreme temperature 14-day anomaly of A for PlaSim dataset. As predicted by the theory, the observed 500 hPa geopotential height pattern is the same as the one from fig. 15.1. To give a quantitative measure of the error, in fig. 20.5, we evaluate the error using the norm ratio defined in eq. (15.1) for different thresholds a, showing that the error is around the 20%, thus of the same amplitude of the one obtained for a threshold at 5%.

A natural follow up question regards the scaling presented in eq. (14.14). In fig. 15.2 we plot the norm of the empirical composite maps as a function of the threshold a. The gray line corresponds to the total one, the coloured lines are the field-wise norms. The dashed line represents the theoretical scaling η of eq. (14.14). The behavior is very well captured by the 2 m air temperature anomaly, and less well captured by the soil moisture anomaly field. The departure of the empirical scaling from the theoretical one for large values of a might be also due to sampling error.

Due to independence of the composite maps on the parameter a we will omit the sensitivity analysis of this parameter in favor of the other two, which are proven to provide different responses for heatwaves, namely the heatwave duration T and the lead time τ (see chapter 17).

15.4 EFFECT OF DATASET LENGTH ON ESTIMATION OF COMPOSITE MAPS

This section aims at motivating the usage of the Gaussian composite when the estimation of the true composite is highly affected by sampling issues, i.e. when we are in a regime of scarcity of data. For datasets' length of 200 years, the same order of magnitude of ERA5 reanalysis dataset, the Gaussian composite performs much better than the empirical one, for events more extremes than 5%.

Firstly, we use the empirical composite $C_{\mathcal{D}}$ computed on the whole 8000 years dataset as an estimate of the true composite. Then we take a set $\mathcal{P}$ of our data and compute over it the empirical composite $C_{\mathcal{P}}$ and the Gaussian one $C_{\mathcal{G}}^{\mathcal{P}}$.

In fig. 15.3 we see the values of the empirical norm ratio $\mathcal{R}_{\mathcal{P}} = \frac{|C_{\mathcal{P}} - C_{\mathcal{D}}|}{|C_{\mathcal{D}}|}$ (solid lines) and the Gaussian one $\mathcal{R}_{\mathcal{G}} = \frac{|C_{\mathcal{G}}^{\mathcal{P}} - C_{\mathcal{D}}|}{|C_{\mathcal{D}}|}$ (dashed lines), for datasets $\mathcal{P}$ of different lengths. To get confidence intervals, we repeat the experiment 8 times for each dataset length, with 8 independent batches of data.

The Gaussian composites over 1000 years and over 200 years of data show a monotonic increase (in log scale) as function of the heatwave threshold a. The latter shows a plateau for values of p ranging from 50% to 1%, meaning that the error made for typical events is comparable to fairly extreme ones. This is not valid for the composite over 1000 years as there is a constant and more rapid worsening of the Gaussian norm ratio. It is interesting to notice that in the very tail of the distribution of A, thus for small values of p, we achieve

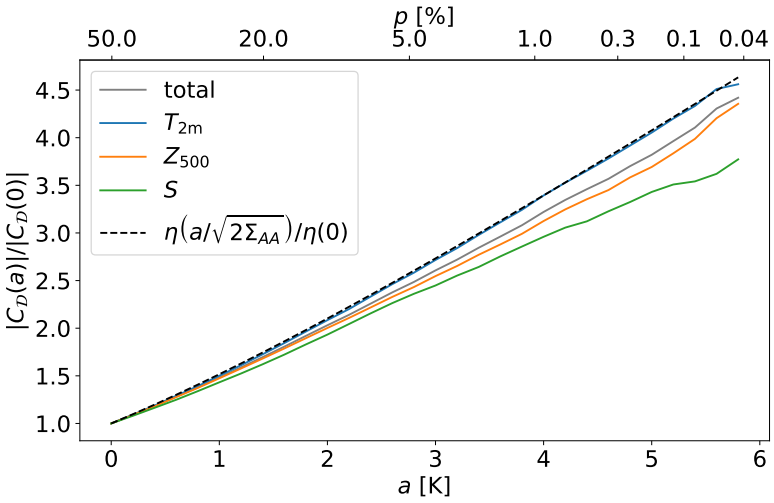


Figure 15.2: Normalized norm of the empirical composite map as function of a, the threshold to define a heatwave. The normalization is the norm of the empirical composite evaluated at $a = 0$. The dashed line represents the theoretical scaling of the composite maps using the Gaussian approximation, see eq. (14.14). The scaling of the empirical composite is not far from the Gaussian one (gray curve) and it is precise for the 2 m air temperature (blue curve). The bottom of x-axis a is the threshold value used to define an heatwave event from the distribution of the temperature anomaly over France, A, for heatwaves of 14 days of duration. On the top of the x-axis, p is the respective percentile value corresponding to a given a.

very similar values of the norm ratio in both datasets. The spread of the norm ratio among the batches is more pronounced for less extreme events than for the most extreme ones. In the case of the Gaussian composite, the main source of error is systematic, as we use the full dataset $\mathcal{P}$ to evaluate the Gaussian composite and not a small set which depends on the threshold (eq. (14.14)).

The empirical composite norm ratio for 200 years of data stays almost constant until $p = 5\%$, after which it starts increasing both in the mean and in the spread of data. For the empirical composite norm ratio over 1000 years we see a less evident constant behavior and a more pronounced minimum of the norm ratio around $p = 5\%$, both in the mean and in the standard deviation. Similar to the 200 years line, there is a worsening of the norm ratio as a increases. It is remarkable that both composites for very small values of p never attain the same value as it happens for the Gaussian ones. Indeed, there is always a constant gap between the two solid lines. As we select fewer and fewer data on the right side of the plot, we see an increase of the spread of the data, mostly due to sampling issues.

Focusing on both composites for 200 years datasets, until $p = 5\%$ both the empirical and the Gaussian have the same values of the norm ratio. For more extreme events, the norm ratio of the empirical one increases drastically, mostly due to the more and more limited data available in the tail, reaching 100% of error at the 0.3% most extreme heatwaves. This is not the case for the Gaussian approximation, whose values of the norm ratios still increase but much more

slowly. Here we can see the power of the Gaussian approximation on smaller datasets.

Indeed, when we are in a regime of scarcity of data, which naturally arises when one wants to study very extreme heatwaves, calculating composite maps using empirical data poses sampling issue. Our methodology overcomes this issue by relying on an estimate of the composite which uses the whole dataset. To confirm this, we see that on longer datasets, such the 1000 years one, where we already have a sufficient amount of data to have a good estimate of the empirical composite, the Gaussian approximation is not a better estimate than computing the composite directly. At least for events up to $a = 4.5\,\mathrm{K}$, after which, due to the sampling issue, the Gaussian estimation performs better than the empirical composite.

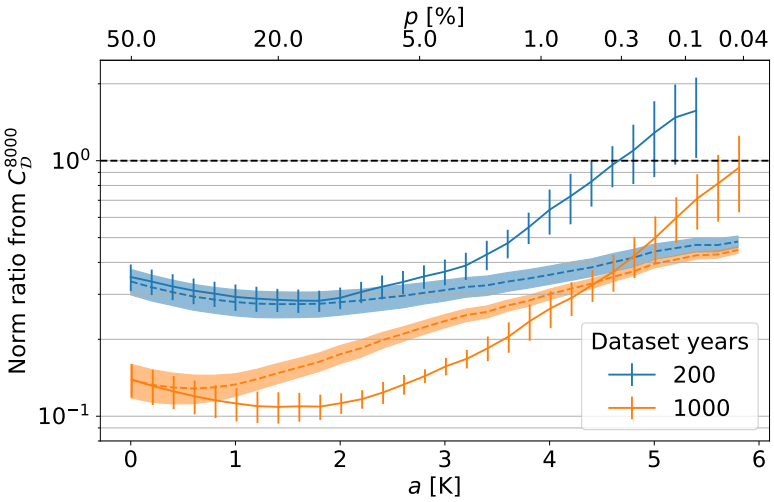


Figure 15.3: Norm of the relative error of the conditional average (composite map) using the Gaussian approximation (dashed lines with shading) and of the composite maps using only a part of the full PlaSim dataset (solid lines with error bars). The error is relative to the empirical conditional average (composite map) evaluated over the full PlaSim dataset of 8000 years. The orange colour indicates a 200 years long dataset. The blue colour a 1000 years one. Shading or error bars indicates the one standard deviation spread obtained from 8 independent batches of either 200 or 1000 years. On the bottom of x-axis a is the threshold value used to define a heatwave event from the distribution of the 2 m temperature anomaly over France, A. On the top of the x-axis, p is the respective percentile value corresponding to a given a. The higher its value, the lower the value of the threshold a, the less extreme are the heatwaves considered. The relative error for dataset of 1000 years is always lower than the one obtained for 200 years simply because of higher amount of available data. The difference is more remarkable, and stays quite stable as a increases, in the relative error obtained with the empirical composite than with the Gaussian approximation. This is not surprisingly because the Gaussian composite uses the information of the full dataset, not just of the set of the heatwave events (see eq. (14.14)). All the curves show an increase in the relative error as a increases due to the lack of data. When we are in this regime, the relative error obtained with the Gaussian composite is lower than the one obtained with the empirical composite. This happens for a p value of around 0.2% for datasets of 1000 years length, and of 5% for datasets of 200 years length. For less extreme events, the Gaussian composite performs worst or similarly than the empirical one.

16

VALIDATION OF THE GAUSSIAN APPROXIMATION FOR COMPUTING COMMITTOR FUNCTIONS ON CLIMATE DATASETS

In chapter 14 we defined committor functions and optimal projection patterns, both generally and within the Gaussian approximation. In this section, we apply the Gaussian approximation of the committor on climate data, the PlaSim dataset described in 13.2.1, and compare its skill with the prediction from a neural network. We then proceed to study the optimal projection pattern, which is given by eq. (14.21). However, we will see in this section that the mathematical expression eq. (14.21), is not directly applicable to high dimensional climate data, where the datasets are usually too short. Indeed, in section 16.2 we show that regularization is necessary to have physically meaningful projection patterns. In sections 16.3 and 16.4 we will show the effect of lack of data on the performance. In the first case lack of data will come from reduced dataset lengths, and in the second from more extreme events.

We illustrate this for the task of predicting heatwaves, but we assume it will generalize well to other prediction problems in climate.

16.1 SKILL OF THE GAUSSIAN APPROXIMATION COMPARED TO PREDICTION WITH NEURAL NETWORKS

We first apply the Gaussian approximation of the committor, defined in eq. (14.20), to the forecast of the 5% most extreme two week heatwaves ($T = 14$), predicted at lead time $\tau = 0$, using the full PlaSim dataset. To have a robust estimate of the performance of our method, we repeat the experiment 10 times in a k-fold cross validation process (see section 20.2). Doing so we get an average validation normalized log score of 0.455 ± 0.010. We can say that the score is much better than the climatology ($S = 0$), but it is very tricky to quantify the maximum achievable score, as $S = 1$ is absolutely unrealistic due to the chaotic nature of the climate system.

However, we can compare to other methods, for instance the prediction using a deep convolutional neural network (Miloshevich et al., 2022). This network takes as input the stack of predictors and produces an estimate of the committor. It is trained on a probabilistic binary classification of the labels Y, i.e. it directly minimizes the loss $\mathcal{L}$ defined in eq. (14.9). More details about the network's architecture can be found in (Miloshevich et al., 2022). Such a network yields a validation score of $S_{CNN} = 0.465 \pm 0.007$.

This is a remarkable result, as the Gaussian approximation is much simpler than a deep neural network, but is able to achieve a result that is only 2% (or less than a standard deviation) worse.

16.2 REGULARIZATION OF THE PROJECTION PATTERN

The simplicity of the Gaussian committor comes with the added benefit of being an interpretable forecast, as we can look at the projection pattern M to obtain some insight into the dynamics leading to a heatwave.

Unfortunately, a direct plot of M looks like the first row of fig. 16.1, from which we cannot extract any meaningful information as no well-defined patterns emerge. This is due to the fact that the covariance matrix Σ_{XX} is very high dimensional ($d^2 \sim 10^7$) and is estimated with a relatively low number of datapoints ($8000 \times 0.9 \times (90 - T + 1) \sim 10^6$). Hence, it will be nearly singular, causing problems when we compute the inverse in eq. (14.21).

A simple solution is the standard Tikhonov regularization, that corresponds to adding an L_2 penalty to the minimization problem:

$$M_\epsilon \propto (\Sigma_{XX} + \epsilon \mathbb{I})^{-1} \Sigma_{XA} = \arg\min_M \left((A - M^\top X)^2 + \epsilon |M|^2 \right), \tag{16.1}$$

where $\mathbb{I}$ is the identity matrix.

However, in our case we can better enforce interpretability of the pattern M by requiring it to be spatially smooth. Namely, we will penalize the squared norm of the spatial gradient, H_2, that we can compute as the weighted sum of the square differences between values of adjacent pixels in the map M. We can then write $H_2(M) = M^\top W M$ (see section 20.6 for the exact formula of matrix W), and hence the regularized pattern will be

$$M_\epsilon \propto (\Sigma_{XX} + \epsilon W)^{-1} \Sigma_{XA} = \arg\min_M \left((A - M^\top X)^2 + \epsilon H_2(M) \right). \tag{16.2}$$

Note that if we tweak the projection pattern M, we should also update the formulas for the coefficients α and β in eq. (14.23). This is relatively straightforward and is discussed in section 20.7.

Varying ϵ yields the different maps shown in fig. 16.1, where indeed we see that the regularization makes the patterns progressively smoother. Unsurprisingly, we note that a higher regularization comes at the price of a lower skill score S (see also table 16.1). It is then up to the user to decide what is a good compromise between performance and interpretability of the pattern. In our case, we argue that the best pattern is the one in the center row of fig. 16.1 ($\epsilon = 1$), as it is smooth enough that we can see some clear structures in the 500 hPa geopotential height field, while a higher regularization does not improve its physical understanding. At this value of the regularization coefficient, the average validation score is 0.418 ± 0.006: three standard deviations or 8% worse than the non regularized case, and five standard deviations or 10% worse than the neural network.

It is important to point out that after proper regularization the skill of the prediction is still much better than climatology, while providing physical insight on the dynamics leading to heatwaves. This latter point is further discussed in chapter 17.

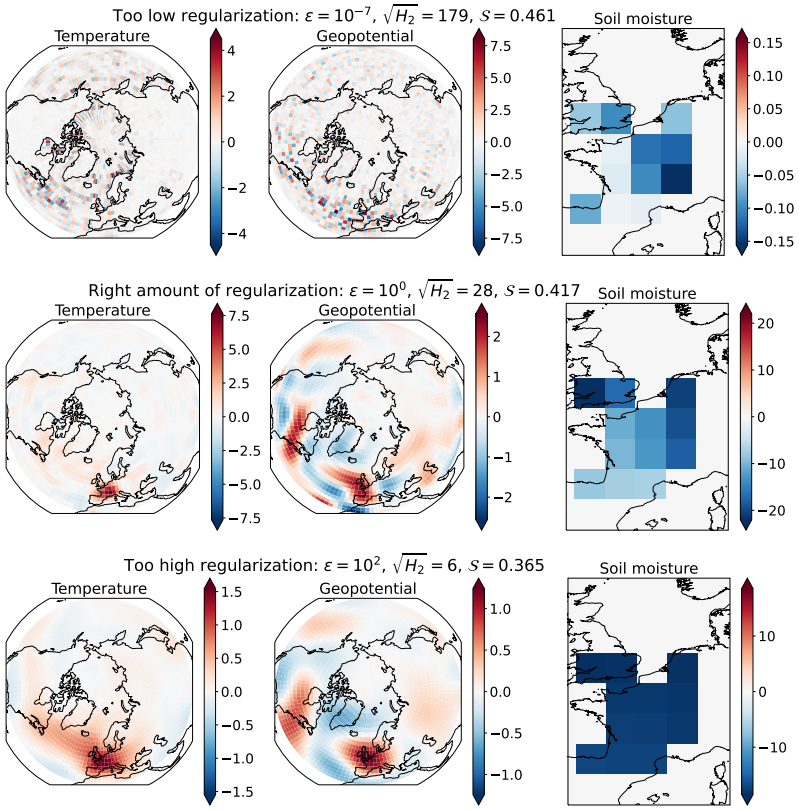


Figure 16.1: Choice of regularization for optimal physical content of the projection map M, using PlaSim data. Each line features the projection map M at different values of the regularization coefficient ϵ. Each map M is represented as its three field components: 2 m air temperature, geopotential height at 500 hPa and soil moisture anomalies (trained on 7200 years of data, for one of the 10 folds). On top of the figures we report the values of ϵ, of the norm of the gradient $\sqrt{H_2}$ and of the normalized log score $\mathcal{S}$. The intermediate value, $\epsilon = 1$, is the best compromise with a very high predictive skill and an excellent readability of the physical fields.

16.3 PERFORMANCE ON SMALLER DATASETS

So far we have applied the Gaussian approximation to an extremely long 8000 year dataset. Such datasets are uncommon in the climate community, especially when dealing with observations or high resolution simulations. To study the effect of the amount of data on the performance of our method, we apply it to gradually smaller and smaller sets of our climate model output.

In the left panel of table 16.1, we can see the behavior of the normalized log score S of the Gaussian committor, as a function of the regularization coefficient and the size of the training set. The first important thing to notice is that the score is not very sensitive to the amount of training data, showing that our method is well suited also for small datasets. By looking at the dependence with respect to ϵ, we see that when we have a lot of data, a stronger regularization means a poorer prediction skill. On the other hand for small datasets the best performance is achieved at a finite value of ϵ. This can be explained by the fact that as we have less and less data to estimate a constant size covariance matrix,

it will become more and more singular, thus requiring a stronger regularization. Also, a smoother pattern is more likely to generalize well when training and validation data are very small.

In any case, we remind that choosing the proper regularization coefficient is not just a matter of score, but also of physical interpretability of the projection pattern, as explained in the previous section. From a qualitative look at projection maps at different values of T, τ and ϵ, $\epsilon = 1$ seemed to be a universally good compromise for the PlaSim dataset. Hence, if not specified differently, in the remainder of this work we will always consider $\epsilon = 1$.

On the right panel of table 16.1 we see the comparison with the skill of the neural network in the form $1 - \mathcal{S}/\mathcal{S}_{CNN}$, which shows that as the dataset gets smaller, the CNN loses its advantage, being outperformed when crossing the 1000 years threshold. An important caveat here is that the many hyperparameters of the CNN where optimized for the biggest dataset (Miloshevich et al., 2022), and then kept constant for the experiments when training on less data. This potentially makes the comparison between the neural network and our method not completely fair. In fact, some experiments (not shown in this work), suggest that by optimizing hyperparameters such as the learning rate and batch size used for training the neural network allow it to prevail even when training only on 450 years of data. The Gaussian approximation, however, is still better when working with 200 years or less, even considering the optimization. So, the qualitative behavior displayed in table 16.1 still holds, and can be ultimately attributed to the higher complexity of the CNN (roughly a million parameters) with respect to the Gaussian approximation (roughly a few thousands of parameters).

Summarizing, our method is well suited to work in a regime of lack of data due to short datasets, where complex neural networks struggle.

16.4 MORE EXTREME HEATWAVES

A question complementary to the one of smaller datasets is the one of more extreme heatwaves, as they both result in very few samples of the event of interest.

First of all, the Gaussian approximation provides a committor that depends on the heatwave threshold a only through the parameter α. This means that, in a similar fashion to the composite maps, the projection pattern M will be the same for all heatwaves independently on how extreme they are. It is thus extremely easy and cheap to get a new committor estimate for a different value of a.

On the other hand, since the neural network we consider is trained on a classification task, as we change a, the labels $Y(t)$ change as well, and hence the whole network needs to be retrained every time. Although transfer learning can reduce the computational cost and avoid retraining from scratch, it still a more complex task than computing a new Gaussian committor. Furthermore, as we focus on more and more extreme heatwaves, the imbalance between the $Y = 0$ and $Y = 1$ classes becomes more and more relevant, eventually hindering the performance of the network (see the gray error-band in fig. 16.2).

Normalized log score

years of training	ϵ = 10^{-2}	10^{-1}	10^{0}	10^{1}	10^{2}
7200	0.43	0.43	0.42	0.40	0.37
3600	0.43	0.42	0.41	0.39	0.37
1800	0.42	0.42	0.41	0.39	0.36
900	0.44	0.43	0.43	0.41	0.38
450	0.43	0.42	0.42	0.40	0.37
180	0.36	0.38	0.39	0.39	0.37

$1 - \mathcal{S}/\mathcal{S}_{CNN}$

ϵ = 10^{-2}	10^{-1}	10^{0}	10^{1}	10^{2}
0.07	0.08	0.10	0.14	0.20
0.03	0.05	0.07	0.11	0.17
0.01	0.02	0.04	0.09	0.15
-0.03	-0.03	-0.01	0.03	0.10
-0.09	-0.08	-0.07	-0.02	0.05
0.01	-0.05	-0.07	-0.07	-0.02

Table 16.1: Left table: normalized log score of the Gaussian approximation (the higher the better), versus training dataset length and the regularization coefficient ϵ. With small datasets, intermediate ϵ values are optimal, while vanishing ones are for large datasets. The apparent peak in performance for 900 years of training is not significant. The Gaussian approximation skill is already nearly optimal for small datasets. Right table: comparison with the skill of the neural network (ϵ affects only the Gaussian approximation). Brown colors mean the CNN performs better, while blue hues mean the Gaussian approximation is better. When the neural network has much data to learn, it can leverage its expressivity potential to outperform the Gaussian approximation. With small datasets, the added complexity of neural networks is detrimental to its score. Both panels are based on PlaSim data.

On the contrary the smaller size of the heatwave class affects the performance of the Gaussian approximation only in its variance, while the mean normalized log score $\mathcal{S}$ has a very weak dependence on the amplitude of the heatwave (blue line in fig. 16.2). This, in turn, suggests that our Gaussian approximation is sufficient to capture well the relationship between the predictors and the heatwave amplitude A even in the most extreme tails of the distribution.

In this section we showed that the Gaussian approximation can be a simple, but powerful, tool for the prediction of extreme heatwaves. Compared to other methods, such as deep neural networks, it does not need as much data to be properly trained. This makes it particularly suited for short datasets, which is typically the case in the climate community. This direction is further expanded in chapter 18, where we apply our method to the ERA5 reanalysis data. Furthermore, and crucially, it is usually very hard to interpret the prediction performed by a deep neural network, while the Gaussian approximation, through the optimal projection pattern, is interpretable *by design*. The study of the projection pattern opens the possibility for insight on the physical processes behind the event under study, and we expand on this in chapter 17.

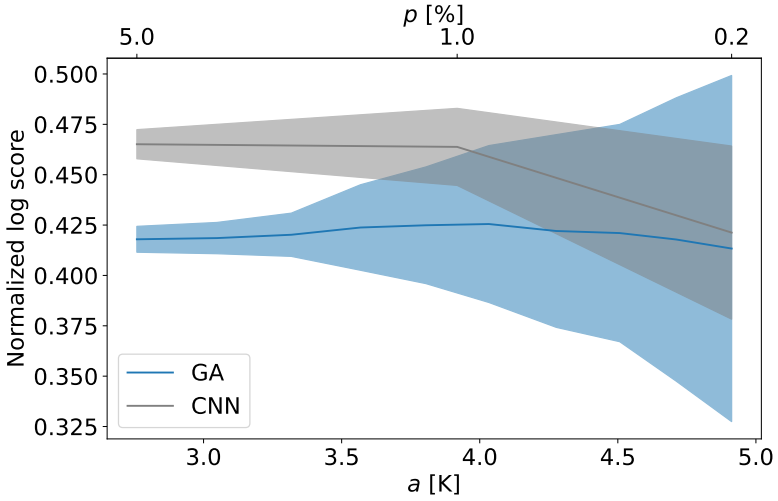


Figure 16.2: Normalized log score of the Gaussian approximation (blue) and CNN (gray) when varying the heatwave threshold a (bottom horizontal axis) or, equivalently, its climatological probability p (top horizontal axis). The solid line is the mean over the 10-fold cross validation process, while the shaded area represents one standard deviation. The experiment is performed with 7,200 years of training with PlaSim data. The regularization coefficient for the Gaussian approximation is kept at the optimal $\epsilon = 1$ value. The CNN is always best but because of the lack of data for rare events, its relative skill decreases with a. The skill of the Gaussian approximation is not much sensitive to the rareness of the event.

17

COMMITTOR FUNCTION AND OPTIMAL PROJECTION FOR EXTREME HEATWAVES

In chapters 15 and 16 we computed composite maps and committor functions for extreme heatwaves. However, in these sections the focus was mainly methodological, with attention to performance and the technical details that influence it. In this section we complement the previous analysis by focusing instead on the physical insight that our method provides on extreme heatwaves. To do so we will compare composite maps and optimal projection patterns at different values of the heatwave duration T and the lead time τ.

17.1 COMPARISON BETWEEN COMPOSITE MAPS AND PROJECTION PATTERNS

In chapter 14 we showed that a-priori and a-posteriori statistics are fundamentally different. Here we proceed to further include some physical reasoning that arises when comparing the two types of statistics. In fig. 17.1 we have the side by side comparison, at different values of the lead time τ, of the Gaussian composite map C_G with the projection pattern M needed for the computation of the committor. As explained in section 14.1.4, the composite map captures the *correlations* between the heatwave amplitude A and the predictors X, while the committor, and thus the projection pattern M, focuses on what is really important for the *prediction*.

A clear example of this is the difference between the 2 m temperature anomaly field in the composite map and in the projection pattern. From fig. 17.1, we can see that the composite shows many teleconnection features, for example over North America, while in the projection map virtually all the weight is over France. This suggests that the relationship between heatwaves and these temperature teleconnections is only of correlation, not causation. Similarly, the 500 hPa geopotential height field anomaly shows a very strong anticyclone over Greenland in the composite maps, which is not present in the projection patterns.

Another remarkable difference between C_G and M is the relative magnitude of the fields. By looking at the colorbars at the bottom of the figure, we see that, in the composite, all the fields have roughly the same order of magnitude, and this makes sense as we work with normalized data and the composite is representative of the typical heatwave event. On the other hand, from the projection patterns we observe that the values of soil moisture are 4 to 10 times higher than the ones of temperature and geopotential, showing that the soil moisture anomaly field is far more important for prediction than one might assume by just looking at the composite.

If we now focus on what happens when we change the lead time τ, we see that in the composites there is essentially just a fading of the structure of the

2 m temperature and 500 hPa geopotential height anomalies apparent at $\tau = 0$, with some minor qualitative changes, such as the connection of the two high pressure systems over the Atlantic at $\tau = 5$. On the other hand, the soil moisture anomaly component remains almost unchanged. This increased prominence of soil moisture as the lead time increases is even more pronounced for the projection pattern M, showing that soil moisture is the key factor for long term heatwave forecast.

Finally, from the evolution of the projection map for the 500 hPa geopotential height field, we see a clear shift of focus from the North-eastern Atlantic at $\tau = 0$ to the United States at $\tau = 5$. At $\tau = 10$ the most prominent feature in the 500 hPa geopotential height projection pattern is a small cyclone over the continental US, something which can barely be seen at all in the composite. These changes in the projection pattern give us insight into the dynamics of atmospheric circulation that leads to heatwaves over France, in particular the dynamics of the jet stream.

17.2 EFFECTS OF CHANGING T AND τ

In this section we analyze more quantitatively how the performance of the Gaussian approximation is affected by the heatwave duration T and the lead time τ, and what physical conclusions we can derive from it. We will first perform this sensitivity analysis on the composite maps (a-posteriori statistics) in section 17.2.1 and then for committor functions (a-priori statistics) in section 17.2.2.

17.2.1 *Composites*

In table 17.1, we see the fraction $\mathcal{F}$ of area for the Gaussian composite that has a significance above 2, as defined in eq. (15.3). The table shows a monotonic trend, with fast and imminent heatwaves having more non-Gaussian features with respect to long and delayed ones. Indeed, for higher values of the heatwave duration T, we expect the statistics of A to be more Gaussian, as we average over a larger number of days. Instead, when we increase the lead time we can think that the chaotic nature of the weather makes the states that led to a heatwave more different from one another. So, both the empirical and the Gaussian composite will tend to 0 as τ increases. Moreover, the higher differences between the states over which we take the empirical average increase the standard deviation. Thus, the significance of each pixel as in eq. (15.3) naturally decreases with τ.

On the other hand if we look at the values for the norm ratio (eq. (15.1)) displayed in table 17.2, we see a rather non-monotonic behavior. In fact, we can gain more understanding if we plot the norm ratio for the three climate variables independently (tables 20.1 to 20.3), which shows that the main contribution to the norm ratio comes from the 500 hPa geopotential height field.

This overall non-monotonic trend can be explained as a competition between the non-linear chaotic dynamics of the weather, that makes the real composite stray more from its Gaussian approximation as τ increases, with the loss of

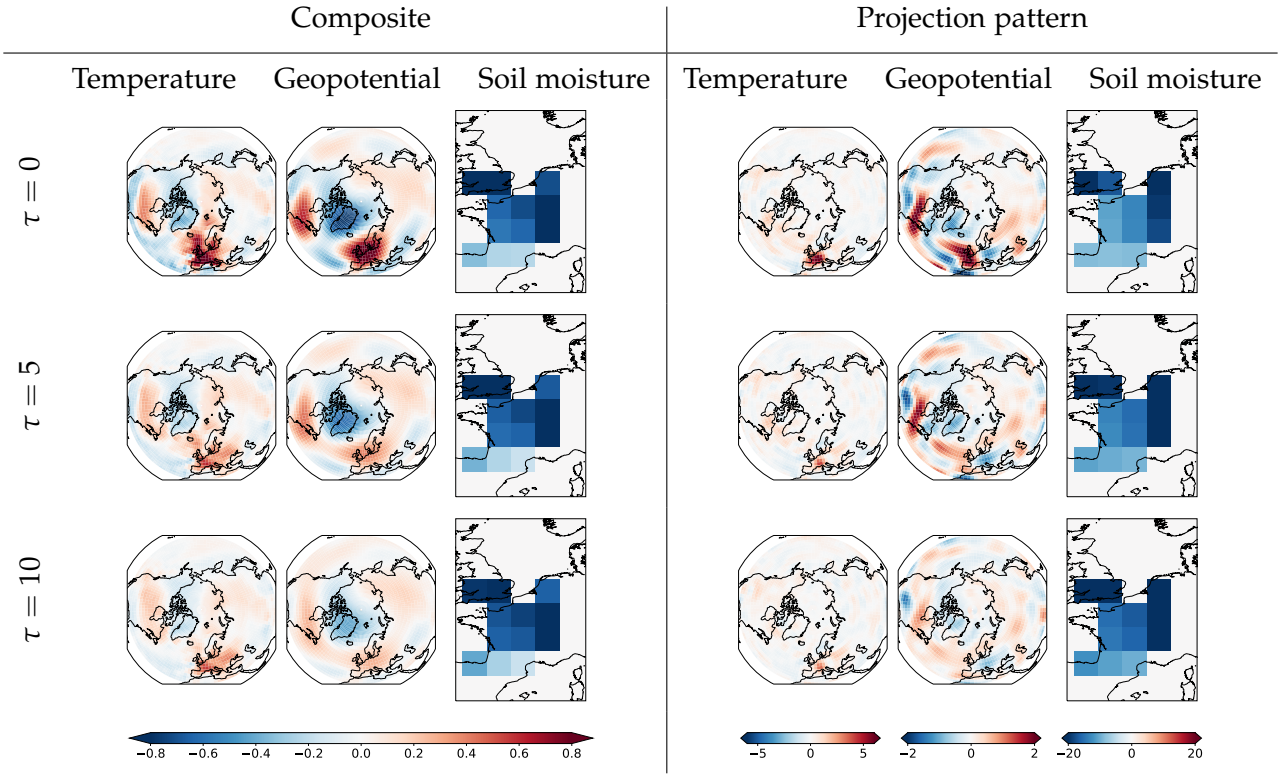


Figure 17.1: Left columns: Gaussian composite maps, temperature, geopotential height at 500hPa and soil moisture, for three different values of the lag time τ. Right column: the optimal projection pattern for prediction, within the Gaussian approximation ($\epsilon = 1$). As expected the two sets of maps are different, characterizing either a-posteriori statistics or best prediction patterns. The composite features hemispheric scale patterns dominated by zonal wave-number zero and zonal wave number three modes. For long lead times, the zonal wave-number zero pattern clearly dominates. The soil moisture composite pattern does not change much with the lag time. The information needed for making an optimal projection, as seen through the projection pattern, is at a finer scale, less global, with a strong meridional structure. Temperature contributes weakly and only through its local values to the projection pattern.

Fraction of area with error above 2σ

T [days] \ τ [days]	0	3	6	9	12	15	18	21	24	27	30
1	0.52	0.50	0.46	0.34	0.27	0.19	0.15	0.08	0.04	0.02	0.02
3	0.52	0.44	0.41	0.28	0.22	0.14	0.11	0.05	0.02	0.01	0.02
7	0.45	0.41	0.34	0.26	0.18	0.15	0.10	0.05	0.03	0.02	0.01
14	0.37	0.30	0.23	0.17	0.11	0.08	0.06	0.05	0.03	0.01	0.01
30	0.15	0.10	0.08	0.07	0.05	0.04	0.03	0.01	0.01	0.01	0.02

Table 17.1: Fraction of significant area in the conditional average (composite map) computed using the Gaussian approximation. The significance in assessed using the fraction of area which is above two standard deviations from the composite map evaluated over the 8000 years PlaSim data (see eq. (15.3)). The threshold used for defining a heatwave is $a = 2.76K$, corresponding to the 5% most extreme values of the distribution of the 2 m temperature anomaly over France, A. The table shows the dependency of the norm of the relative error on T, the heatwave duration, and τ, the lead time. Low value (dark red) means less areas beyond two standard deviations from the empirical composite, thus the Gaussian composite better reproduces the empirical one and the statistical error is a systematic error not due to the size of the dataset. Significant areas are monotonic increasing as T and τ increase.

Norm ratio

T [days] \ τ [days]	0	3	6	9	12	15	18	21	24	27	30
1	0.25	0.28	0.29	0.26	0.27	0.27	0.28	0.26	0.22	0.21	0.21
3	0.24	0.25	0.26	0.24	0.25	0.25	0.26	0.23	0.21	0.19	0.19
7	0.22	0.23	0.24	0.25	0.27	0.29	0.28	0.24	0.23	0.22	0.19
14	0.20	0.23	0.26	0.28	0.28	0.27	0.27	0.26	0.24	0.22	0.22
30	0.21	0.24	0.26	0.28	0.28	0.27	0.27	0.26	0.25	0.25	0.25

Table 17.2: Norm of the relative error of the conditional average (composite map) evaluated using the Gaussian approximation. Relative to the composite value obtained through empirical conditional average over the 8000 years PlaSim dataset. The threshold used for defining a heatwave is $a = 2.76K$, corresponding to the 5% most extreme values of the distribution of the 2 m temperature anomaly over France, A. The higher the value (bright yellow) the worst the Gaussian composite approximates the empirical one. Lower values (dark red) denotes a lower value of the error. The table shows the dependency of the norm of the relative error on T, the heatwave duration, and τ the lead time. There is a non monotonic trend which is due to the different atmospheric fields using in the conditional average. Events which are long lasting and far in time behave more closely to Gaussian distributed events.

Normalized log score

T [days] \ τ [days]	0	5	10	15	20	30
1	0.89	0.27	0.14	0.11	0.09	0.08
7	0.53	0.25	0.18	0.14	0.13	0.12
14	0.42	0.26	0.20	0.18	0.17	0.16
30	0.34	0.26	0.23	0.21	0.21	0.20

$1 - \mathcal{S}/\mathcal{S}_{CNN}$

T [days] \ τ [days]	0	5	10	15	20	30
1	-0.00	0.26	0.22	0.19	0.18	0.15
7	0.11	0.21	0.17	0.14	0.10	0.08
14	0.10	0.17	0.13	0.09	0.06	0.04
30	0.07	0.09	0.05	0.03	0.00	-0.00

Table 17.3: Left table: normalized log score of the Gaussian approximation (the higher the better), versus heatwave duration T and lag time τ. In all cases, we focus on the 5% most extreme heatwaves. As the prediction task gets harder, the skill decreases monotonically with the lag time, faster for shorter heatwaves. Right table: comparison with the skill of the neural network. The CNN is always better, but more so for shorter heatwaves and around $\tau = 5$. This is the regime where the dynamics is more non-linear, and thus the neural network complexity has a better opportunity to make a difference.

memory that averages out the non-linear effects, bringing the empirical composite closer to the Gaussian one. This also would explain why geopotential dominates the norm ratio, as, of the three fields, it is the one with the most non-linear dynamics.

17.2.2 *Committor*

Similarly to what has been done for the composite maps, we can look at how the skill of the prediction is affected by the heatwave duration T and the lead time τ. In the left panel of table 17.3, we can see that the prediction skill decreases monotonically with τ at any level of T. For shorter lead times the skill is best when dealing with shorter heatwaves, while for longer delays, the skill is higher for longer-lasting events. In the limit of $T = 1$ and $\tau = 0$, we are forecasting a one day heatwave that starts today, so we might just look outside the window and see if it is hot. And indeed there is perfect correlation between the temperature anomaly over France and the heatwave amplitude A. However, one day heatwaves are very erratic events, which become very hard to predict for longer lead times. On the other hand, longer lasting events are non trivial to predict for very short delays, but are more influenced by processes with long timescales such as the dynamics of soil moisture, and hence maintain some predictability at higher values of τ (Miloshevich et al., 2022).

On the right panel of table 17.3, we see the skill comparison with the neural network, which is able to capture non-linear and non-Gaussian structures in the data. We can see that our Gaussian committor struggles the most for shorter heatwaves and, more importantly, around $\tau = 5$. We can interpret this region of struggle as the one where the prediction is most *dynamical*, rather than *statistical*. Namely where mere linear correlations are not enough and the complex and non-linear dynamics of the atmosphere plays a significant role.

18

APPLICATION TO THE ERA5 REANALYSIS DATASET

In this article we presented a methodology for estimating composite maps and committor functions using a theoretical framework that we called the Gaussian approximation (see chapter 14) and we tested it over a very long simulation dataset obtained from the climate model PlaSim. The results are really promising.

A key point that we showed in the previous sections is that our Gaussian framework is particularly suited for short datasets. In the case of the composite map (see section 15.4), the empirical average is performed over too few samples to be very accurate. For the committor (see section 16.3), the alternative approach of deep neural networks struggles with the lack of data. It is then natural to try to apply our method the ERA5 reanalysis data (Hersbach et al., 2020), and in this section we show that indeed for this dataset the Gaussian approximation is the best option.

18.1 COMPOSITES

In this section we compute composite maps on the ERA5 dataset, using both the empirical average and the Gaussian approximation. Because the dataset is much shorter than the PlaSim dataset, we do not know the ground truth as precisely as in Chapter 15. We can nevertheless compare the two estimates and see if they qualitatively agree.

Figure 18.1 shows the empirical composite and the composite evaluated within the Gaussian framework for the geopotential height anomaly at 500 hPa. They are both evaluated for $T = 14$, $\tau = 0$ and for heatwaves corresponding to the 5% most extreme value in the distribution of A. The two composites look qualitatively very similar. In both cases we see a clear wave train which starts from the western part of the United States and Canada with an anticyclonic anomaly, continues over the North Atlantic Ocean and finally terminates over Western Europe with another anticyclonic anomaly, stronger than the rest of the wave pattern. This is consistent with the fact that we condition on the temperature anomaly over France. The overall wave structure is well represented by the Gaussian composite, even if it puts a higher weight over the Western Europe anticyclone (fig. 18.1, right panel). The difference between the two composites is larger over Asia and over the Pacific Ocean. Unlike the case of PlaSim data, the difference does not have an annular mode structure but contains a visible wavenumber 6 component. The largest differences between the two composites are on the order of 20%.

As for PlaSim data, we analyzed how the two indices $\mathcal{R}$ (norm ratio, defined in eq. (15.1)) and $\mathcal{F}$ (fraction of area where differences between Gaussian and empirical composites are significant, defined in eq. (15.3)) vary with the parameters T and τ for ERA5 data. Table 18.1 shows the norm ratio as a function of

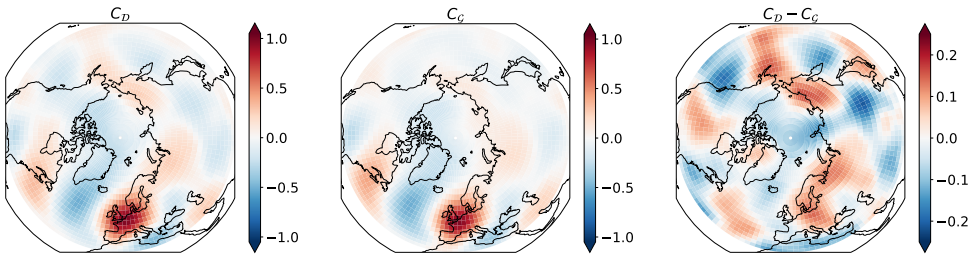


Figure 18.1: Composite maps of normalized 500 hPa geopotential height for 5% most extremes 14-day temperature anomaly over France ($T = 14$). Composite maps are computed either directly from ERA5 data (left map), or under the Gaussian approximation (central map). The right map shows the difference between the first two. The salient features of geopotential are well captured by the Gaussian approximation, with errors of the order of 25% at most.

T and τ for the 5% most extreme heatwaves. We see that, similarly to PlaSim, there is a non-monotonic trend with lowest values for T between 1 and 14 and τ between 0 and 6. Outside this range, the norm ratio has rather high values which are the sign of a great mismatch between the two composites. However, $\mathcal{F}$ (table not shown) assumes values which are almost never above 1% and very often below 0.1%, meaning that we cannot rule out that any discrepancy between the Gaussian and empirical composite is simply a sampling error.

Norm ratio

T [days] \ τ [days]	0	3	6	9	12	15	18	21	24	27	30
1	0.29	0.38	0.47	0.59	0.81	0.89	0.88	0.88	0.78	0.74	0.73
3	0.33	0.40	0.48	0.69	0.90	0.87	0.86	0.81	0.82	0.72	0.81
7	0.35	0.39	0.53	0.73	0.79	0.80	0.83	0.84	0.84	0.76	0.81
14	0.35	0.46	0.61	0.73	0.82	0.86	0.85	0.87	0.83	0.79	0.81
30	0.62	0.81	0.89	0.90	0.84	0.81	0.82	0.82	0.82	0.81	0.83

Table 18.1: Norm of the relative error of the conditional average (composite map) evaluated using the Gaussian approximation. Relative to the composite value obtained through empirical conditional average over the ERA5 dataset. The threshold a used for defining an heatwave corresponds to the 5% most extreme values of the distribution of the temperature anomaly over France, A. The higher the value (bright yellow) the worst the Gaussian composite approximates the empirical one. Lower values (dark red) denotes a lower value of the error. The table shows the dependency of the norm of the relative error on T, the heatwave duration, and τ the lead time.

18.2 COMMITTOR

In this section we deal with the computation of committor functions on the reanalysis dataset. After the necessary technical adaptations to work on this

dataset, we compare the prediction skill of the Gaussian approximation with the one of neural networks, which shows the first is clearly better.

Before discussing any result, we need to define a protocol for the choice of the proper regularization coefficient ϵ, the only hyperparameter of our method. When working with 8000 years of PlaSim data, we had to choose empirically $\epsilon = 1$ to have interpretability in the projection patterns, and this interpretability came at the cost of a lower skill score. On the other hand, on the reanalysis dataset, and more generally when working with small datasets (table 16.1), the value ϵ_{best} of the regularization coefficient that yields the highest skill score also provides an interpretable projection map.

The reanalysis dataset consists of 83 years of data. To have a meaningful cross validation we take the 80 years from 1943 to 2022 and split them in 5 balanced folds (see section 20.2). This way we train on 64 years and validate on 16.

With this choice, for the 5% most extreme two-week heatwaves ($T = 14$) at $\tau = 0$, we obtain a skill score of $\mathcal{S} = 0.16 \pm 0.07$. This number is considerably lower than the skill we have measured for PlaSim (section 16.1). To understand why, we can investigate the impact on the skill score for the PlaSim dataset of the reduced number of predictors (using only the 500 hPa geopotential height field as for reanalysis data) and of the amount of data (training on a set of the same size as the reanalysis data). As can be seen from table 18.2, both the

	Predictor fields	
years of data	T_{2m}, Z, S	Z
8000	0.418 ± 0.006	0.23 ± 0.01
80	0.33 ± 0.07	0.18 ± 0.04

Table 18.2: Skill score on PlaSim, when using different amount of data and different predictor fields.

reduced number of predictor fields and the smaller dataset severely impact the skill score. However, even combining the two effects, the performance remains slightly better than for the reanalysis data. This suggests that the more realistic data of ERA5 have more complexity and variability with respect to PlaSim, and thus it is harder to make a skillful prediction.

Nevertheless, we argue that the result achieved for reanalysis data, albeit humble, is the best we can do. To support this claim, in fig. 18.2 we compare it to the skill of other methods at different values of the lead time τ. In green is the performance of a convolutional neural network with a similar architecture to the one used for PlaSim. It always perform worse than the Gaussian approximation (in blue), and already at $\tau = 3$ days it is consistently below the climatology. On the other hand, the Gaussian approximation manages to extend the predictability margin a few more days. For $\tau \geq 6$ days the latter becomes useless too, and, interestingly, $\epsilon_{\text{best}} \to \infty$, yielding a uniform projection pattern.

In the regime where the prediction is still skillful, the projection patterns look remarkably similar to the composite maps (fig. 18.4), so it is natural to try to project onto the composite itself. This is the orange line in fig. 18.2, which, despite having a smaller error bar than the optimal projection pattern M, on

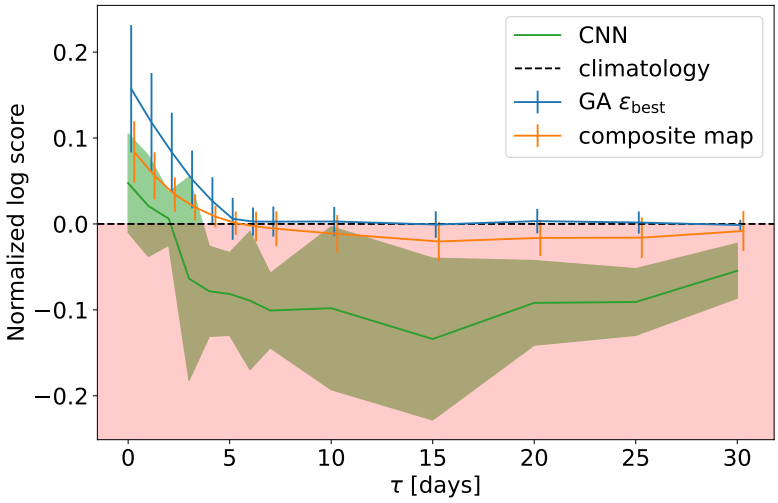


Figure 18.2: Skill score of different prediction techniques for reanalysis data (using geopotential height at 500 hPa anomaly as the only predictor, $T = 14$) changing the lead time. In green the convolutional neural network, in blue the Gaussian approximation, both at their best values for hyperparameters. In orange the Gaussian approximation when using the composite map as projection pattern. Error bars or shaded area indicates the variation among the 5 folds. The red shaded zone below 0 indicates where the prediction is worse than the climatology. The Gaussian approximation is always the best, and gives results better than the climatology only for $\tau \leq 5$.

average yields a worse performance. This once again highlights the fact that composite maps are not the proper tool for prediction.

Now that we showed that the Gaussian approximation is the best option for very small datasets, we can investigate what happens when we vary the heatwave duration. From fig. 18.3 we see that, at any fixed value of τ, the prediction skill decreases with increasing heatwave duration T (solid lines), with shorter heatwaves having a longer predictability horizon. The comparison with 80 years of PlaSim data with only the 500 hPa geopotential height as predictor (dashed lines), shows that predicting heatwaves is harder on the more realistic data. This can be an effect of the lower spatial resolution of the PlaSim model, which yields a more sluggish and less chaotic atmospheric dynamics, and, hence, better predictability. This hypothesis is further reinforced by the fact that, on average, training on PlaSim requires a lower regularization coefficient than the one on reanalysis data (fig. 18.4).

Finally, the dotted lines in fig. 18.3 represent the skill when still training on 80 years of PlaSim data, but with all three predictors. For short lead times and heatwave duration, the increase in skill comes mainly from the direct information of the 2 m temperature field, but the more interesting effect is for longer delays. Here, almost all the predictive power resides in the soil moisture field, and is able to extend the predictability horizon significantly. This effect is enhanced for longer lasting heatwaves. As was already pointed out in (Miloshevich et al., 2022), soil moisture acts as a slow modulator of the chance of a heatwave, that

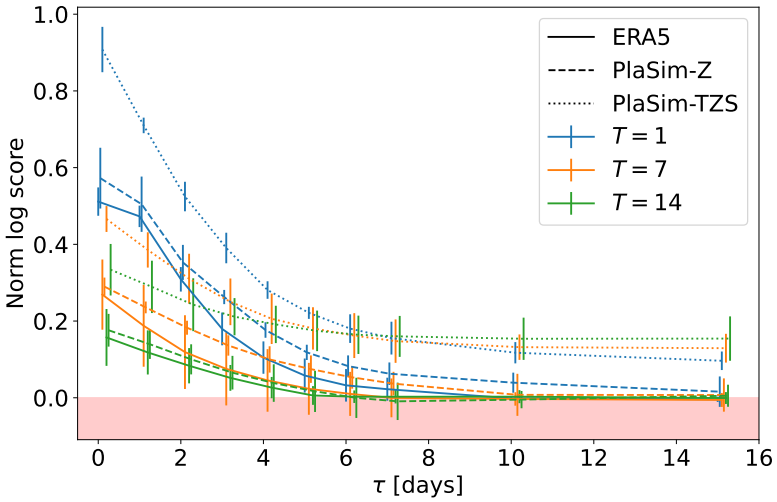


Figure 18.3: Skill score of the Gaussian committor as a function of τ for different values of the heatwave duration T, and three different datasets: ERA5 with only geopotential height at 500 hPa (solid line), 80 years of PlaSim data with 2 m temperature, geopotential height at 500 hPa and soil moisture (dotted lines) and only with geopotential height at 500 hPa (dashed line). PlaSim has a consistently higher predictability than ERA5, and the addition of the slow evolving soil moisture greatly extends the predictability horizon. In the absence of this slow variable, predictability decreases with the heatwave duration.

is still able to give some useful information when the fast predictors, such as the 500 hPa geopotential height, are beyond their de-correlation timescale.

Summarizing, the higher complexity of the ERA5 dataset, its reduced length, and the absence of soil moisture as a slow predictor, all these aspects make so that the forecast skill is much lower than the one for PlaSim. However, the prediction performed by the Gaussian approximation proved to be the best available option, with results that are still remarkable.

18.3 PHYSICAL DISCUSSION

Now that we discussed composite maps and committor functions from the point of view of performance, we proceed to focus more on the physics-oriented analysis of composite maps and optimal projection patterns.

In fig. 18.4, we show the comparison between composite maps and projection patterns computed on reanalysis data and on 80 years of PlaSim data using only the 500 hPa geopotential height as predictor. Interestingly, we observed that both composite maps and projection patterns do not change much with respect to the heatwave duration T (not shown). This is only partially explained by consecutive days with high temperature contributing both to short and long heatwaves. In the future it may be worth investigating this further, but in this work we simply exploit it to discuss the patterns only for a single value of T and still provide a relatively comprehensive picture. In particular, we show

the results for 1-day heatwaves, as they display a clearer evolution of both composites and projection patterns with the lead time τ.

As already mentioned before, one of the main differences between PlaSim and ERA5 is a generally higher signal to noise ratio in PlaSim, that manifests itself in higher norms of the composite maps (left columns) and less smooth projection patterns (right column). At $\tau = 0$, most of the weight of both composite maps and projection patterns is concentrated around France. More precisely, with an anticyclone over France and central Europe to ensure clear skies and a cyclone north of Portugal to advect warm African air northward. This cyclone is more localized in the ERA5 projection map than in composite maps. As the lead time increases, this dipole structure stretches westward into the Atlantic ocean. In the reanalysis dataset there is a clear emergence of a wave-train pattern, and C and M look rather similar. On the other hand, PlaSim's projection patterns stray considerably from the composite maps, and the physics that they hint at is harder to explain. For both datasets, and in both composite maps and projection patterns, a rather strong anticyclonic anomaly is always present over France, getting fainter as τ increases, but remaining always a prominent feature. This suggests that even for very short $T = 1$ heatwaves the most *common* (composites) and the most *likely* (committor) causes of the extreme events are connected with quasi-stationary weather states.

Concerning the reanalysis dataset, the similarity between composite maps and projection patterns may tempt us to use the composite as a prediction tool. However, although both composites and projection maps display the dynamics of a stationary Rossby wave, a careful examination shows a different weight distribution in the projection patterns, for example at $\tau = 6$ the focus is more over North America than in the composite map. The lower prediction skill achieved with the composite map compared to the optimal projection pattern, already discussed above (fig. 18.2), suggests that such differences matter for prediction even if they appear small at first sight. Furthermore, the similarity between the two maps can most likely be attributed to the need for a relatively high regularization coefficient, required to have a prediction that generalizes well when trained on such a short dataset. More technical details are available in section 20.12.

This section has three main conclusions. Firstly, given the size of the ERA5 dataset (and of any other real-data dataset), it is hard to go beyond the Gaussian approximation for both analysis of the averaged weather conditions that led to heatwave events (composite maps) and for a probabilistic forecast of heatwaves, as shown from fig. 18.1 and fig. 18.2. Secondly, the difference between the empirical and the Gaussian composite maps, shown in fig. 18.1 (right panel) has a different wave number with respect to the one observed for PlaSim, for heatwave of the same duration and intensity, (see fig. 15.1 bottom row, central map, which exhibits a wave zero pattern). However, we cannot exclude that this mismatch is due to sampling error. Thirdly, the reduced size of the dataset forces us to strongly regularize the optimal projection patterns, which makes them visually similar to the composite map. However, even if they do not provide any additional qualitative information, they do provide more precise quantitative information, leveraged for prediction skill. Finally, the comparison with the data from PlaSim suggests the importance of predictor fields other than the 500 hPa

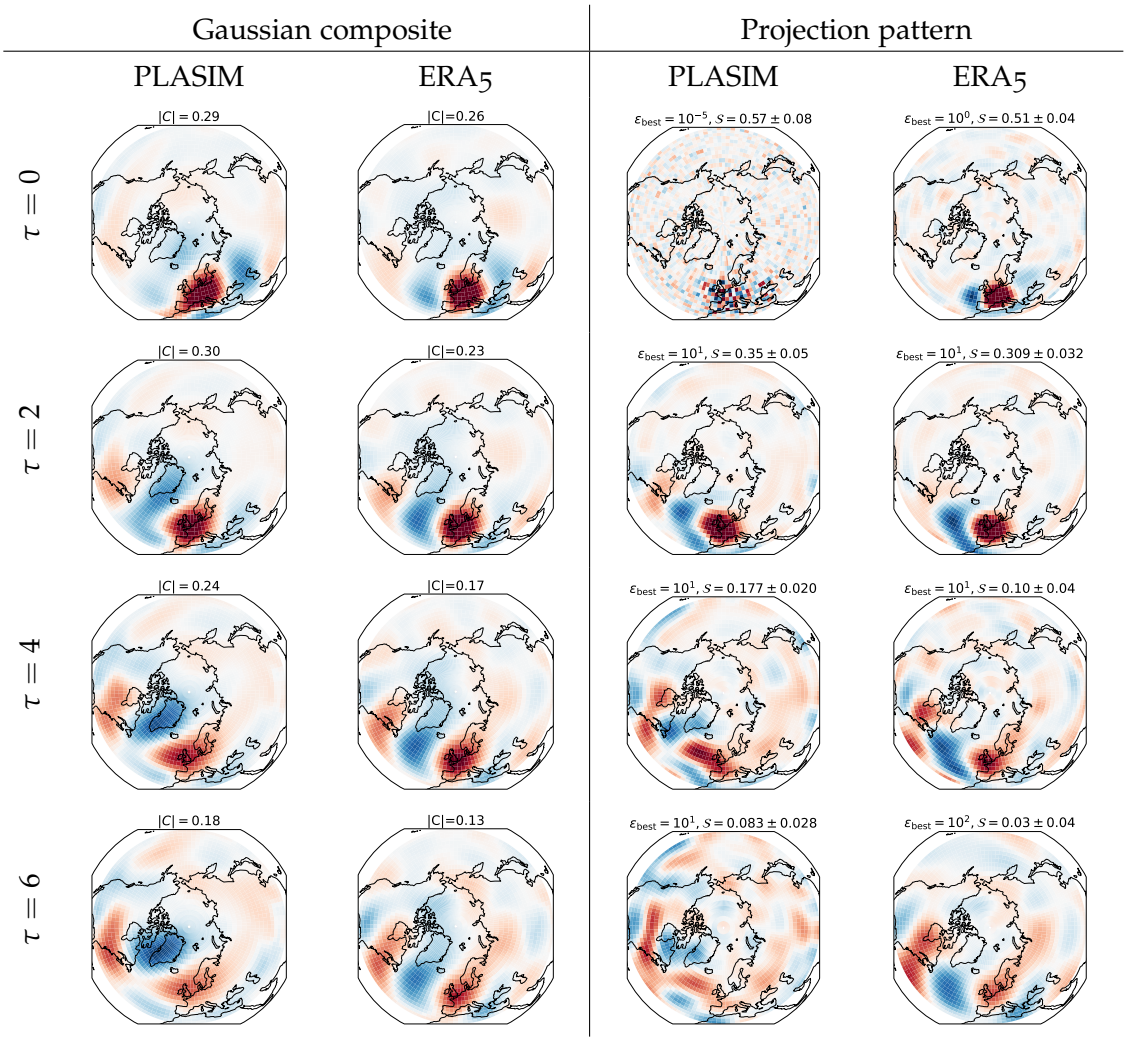


Figure 18.4: Comparison between composite maps and projection patterns of ERA5 and PlaSim (80 years, geopotential height at 500 hPa only), at $T = 1$ and different values of τ. All maps are shown as normalized to unitary L2 norm. The L2 norm of the actual composite maps is reported on top of them, while for the projection pattern we display the regularization coefficient and the skill score. For ERA5 composites and projection patterns *look* qualitatively similar. However, this is a result of the small size of the dataset, which forces us to use high values of the regularization coefficient.

geopotential height, which can significantly improve the prediction skill. For the reanalysis dataset, this opens the possibility to use also ocean variables, like sea surface temperature.

19

CONCLUSIONS AND PERSPECTIVES

In this work, we stressed the important difference between the statistics of climate and weather conditions which led to an extreme event (a-posteriori statistics, for instance composite maps) and the prediction in the future of an extreme event given some knowledge (predictors) of the state of the climate system (a-priori statistics, for instance committor functions). We have highlighted the second as the proper set of tools for any prediction task. At the same time, we provided a simple framework to give easy access to these tools, which is effective even with short datasets of length of the order of several decades to several centuries. In the context of extreme heatwaves over France, we evaluated our method on a very long time series of climate model output data and successfully applied it to a reanalysis dataset.

Concerning a-posteriori statistics, with our Gaussian framework, we were able to provide an explanation of why the composite maps of very extreme heatwaves look qualitatively similar to the ones of less extreme ones. We made this statement quantitative, showing that composite maps are the same up to a rescaling by a non-linear function of the threshold that defines heatwaves. This opens the possibility to estimate composite maps of extremely rare events, even ones that have never been observed in the dataset. For PlaSim data, the computation of composite maps using the Gaussian approximation gives results which are valid up to an error (in L2 norm) of the order of 20 to 30%. We also stress that the deviations from the Gaussian prediction are statistically significant, showing that the statistics is actually not Gaussian and that information beyond the Gaussian approximation can be computed with dataset length of the order of a thousand years or more. On the other hand, on the much shorter reanalysis dataset, errors are larger, but entirely compatible with the imperfect sampling of the empirical composite, and one cannot compute statistically significant deviations from the composite map obtained within the Gaussian approximation.

However, if one is interested in *predicting* heatwaves instead of studying their statistics a-posteriori, composite maps are not the proper tool. The right one is the committor function, and our framework gives probably the easiest non-trivial access to this very complex object. Our method gives very good prediction skill, and is particularly competitive with more complex alternatives, such as neural networks, when working with small datasets, which are very common in the climate community. In fact, for the 80-year long ERA5 dataset, the Gaussian approximation proved to be the method with the highest predictive skill.

As demonstrated in (Miloshevich et al., 2022), too short datasets prevent optimal use of neural networks in many applications in climate sciences. This issue is particularly salient for rare events, for instance extreme events. In this respect, we see the Gaussian framework developed in this paper as a key solution to make the first relevant prediction. It should play an important role

in future studies. For rare events, going beyond the results of the Gaussian approximation may require to have datasets with more rare events. One way is to sample exceptionally rare extreme events using the recently developed rare event simulation techniques, that are able to multiply by several orders of magnitude the number of observed heatwaves with PlaSim model (Ragone, Wouters, and Bouchet, 2018) and with CESM (the NCAR model used for CMIP experiments) (Ragone and Bouchet, 2021). A perspective is to couple these rare event simulations with the Gaussian framework presented in this paper or other machine learning forecast. We have already coupled machine learning simulations to rare event algorithms, for simple academic models (Lucente et al., 2022). Coupling the rare event simulations with machine learning is a very interesting perspective to solve the key fundamental issue of lack of data in the science of climate extremes.

Moreover, beside pure skill, our method provides an optimal index for prediction, which, once properly regularized, makes it easy to interpret our results, giving insight in the dynamics behind our ject of study. This optimal prediction map is one of the key results of this paper. It makes the Gaussian approximation appealing even for applications on long enough datasets so that its skill can be outperformed by neural networks, which are often hard to understand.

From the point of view of understanding the underlying physics, in the case of extreme heatwaves over France, we found that both composite maps and optimal projection maps display a quasi-stationary pattern, that does not depend much on the lead time. In particular, the development of a Rossby wave-train over the Atlantic Ocean plays an important role for the short term prediction. This appears very clearly in the reanalysis data, while PlaSim has a strong competing contribution from a wave number 0 pattern. For longer lead times, instead, the analysis on PlaSim data and the comparison with ERA5 confirmed the key importance of slow drivers, such as soil moisture. The natural next step is then to include these slow drivers in the study on the reanalysis dataset, maybe even using ocean-based variables like sea surface temperature.

As further perspectives, we argue that a deeper analysis at the physical level of optimal projection patterns is needed, turning the qualitative insights presented in this work to more quantitative statements. Moreover, we took as an example extreme heatwaves over France: it would be interesting to apply our method to heatwaves on different geographical locations or to different types of extreme events altogether. Another very interesting direction is, in the cases where the Gaussian approximation is outperformed by neural networks, to interpret where this extra skill comes from. Finally, we suggest that our method can be used as a better baseline than the mere climatology when testing more sophisticated tools for probabilistic prediction.

20 SUPPORTING INFORMATION

20.1 DETRENDING OF ERA5

In this manuscript we present an application of our methodology to the ERA5 dataset (Hersbach et al., 2020). In this section we go over the technical details of handling this dataset.

We start from taking daily averages of the hourly data from the public available dataset of the ECMWF service (https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5) for summer seasons from 1940 to 2022. We decide to use 2 m air temperature for defining the heat wave amplitude and, differently from what has been used for PlaSim, only the 500 hPa geopotential height field as set of predictors X. The reason is that when we will use ERA5 for forecast of heatwaves, it will be impossible to work with the same amount of fields as for PlaSim but with only 83 years of data. In this line of thought, we also performed a regridding of the data over the PlaSim grid to considerably reduce the numbers of features. We then remove the seasonal cycle, so that we can work with anomalies.

In this work we aim at studying the response of climate models in stationary conditions, thus we detrend the ERA5 dataset to remove the climate change signal. For the temperature field, given that we used it just to define our heat wave amplitude A, we performed a spatial averageover the France region, and detrended its seasonal mean with a quadratic fit. The time series of the seasonal mean is shown in fig. 20.1, where the orange line is the trend that we removed for each summer. We tried other sophisticated detrending methodologies, but this one was simultaneously the most simple and effective one.

A similar protocol was also applied for the detrending of the geopotential height field. However, given that we noticed a latitudinal dependence of the trend, we decided to independently detrend via a quadratic fit the seasonal and zonal means of geopotential height. Figure 20.2 shows the contour plot of the trend that we removed as a function of the latitude and year. Indeed, this trend is non-monotonic trend at mid and high latitudes, while it is monotonously increasing at lower latitudes. Given that this non monotonicity is present at the beginning of the dataset, our guess is that it might depend on the quality of the data available before the satellite era.

20.2 BALANCED K-FOLD CROSS VALIDATION

The standard K-fold cross validation process consists in splitting the dataset $\mathcal{D}$ into K disjoint subsets of equal length $\{\mathcal{F}_k\}_{k=1}^{K}$, that can be called *folds*. Then, for each $k = 1, \dots, K$ we define training and validation sets as

$$\mathcal{T}_k = \bigcup_{i \neq k} \mathcal{F}_i, \quad \mathcal{V}_k = \mathcal{F}_k. \tag{20.1}$$

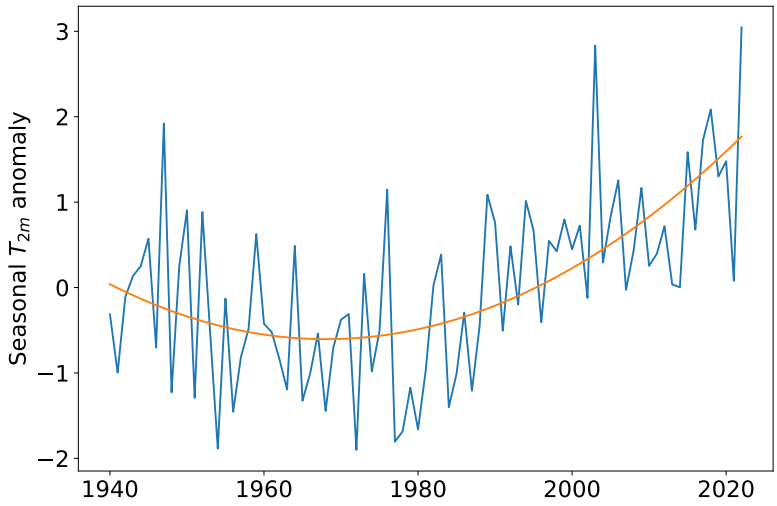


Figure 20.1: Seasonal T_{2m} anomaly averaged over France for the ERA5 dataset. The orange curve is the trend fitted via a second order polynomial.

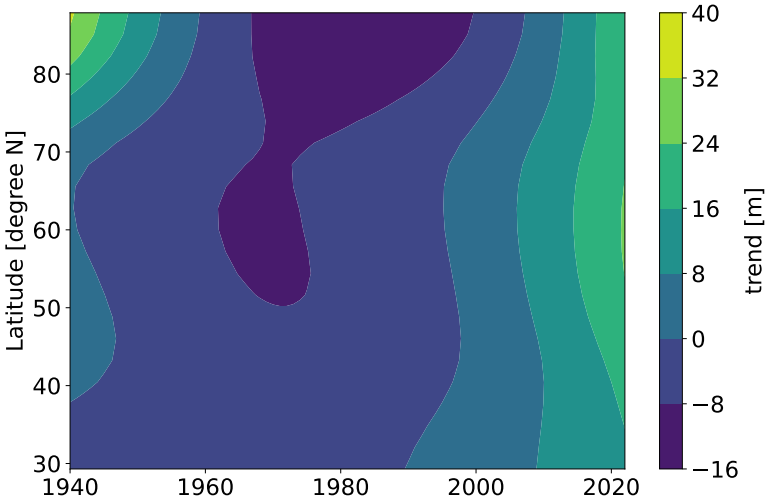


Figure 20.2: Contour plot of the 500 hPa geopotential height trend for ERA5 dataset as function of years and latitude. At latitudes the trend is non-monotonic, while it is monotonically increasing in time at lower latitudes.

To make a *balanced* K-fold cross validation, we ask that the $\mathcal{S}_k$ all contain the same amount of heatwaves. This is essentially equivalent to the classical technique of stratified K-fold cross validation (Hastie, n.d.), but in our case, to avoid contamination between the different folds, we force data belonging to the same summer to end up in only one of the folds.

20.3 DETAILED CALCULATIONS OF THE COMPOSITE MAP IN A 2-DIMENSIONAL GAUSSIAN SYSTEM WITH COMMITTOR DEPENDING ONLY ON ONE VARIABLE

In the second example of section 14.1.4, we use intuition to say that if we have two correlated variables but the committor depends only on one, the composite will still be non zero for the variable upon which the committor does *not* depend. Here we give a formal proof.

According to the assumption of zero mean and Gaussianity for the two variables, we can write the stationary measure as

$$P_S(x_1, x_2) \propto \exp\left(-\frac{1}{2}(x_1, x_2)\begin{pmatrix}\sigma_1^2 & \phi \\ \phi & \sigma_2^2\end{pmatrix}^{-1}\begin{pmatrix}x_1 \\ x_2\end{pmatrix}\right) = \exp\left(-\frac{1}{2}(ax_1^2 + bx_2^2 - 2cx_1x_2)\right), \tag{20.2}$$

where

$$(a, b, c) = \frac{1}{\sigma_1^2\sigma_2^2 - \phi^2}(\sigma_2^2, \sigma_1^2, \phi). \tag{20.3}$$

Then, according to eq. (14.6), the composite value for X_2 is

$$C_2 \propto \int x_2 P_S(x_1, x_2) q(x_1) dx_1 dx_2, \tag{20.4}$$

$$\propto \int dx_1 q(x_1) \int dx_2 x_2 e^{-\frac{1}{2}(ax_1^2 + bx_2^2 - 2cx_1x_2)}, \tag{20.5}$$

$$= \int dx_1 q(x_1) e^{-\frac{1}{2}\left(a - \frac{c^2}{b}\right)x_1^2} \int dx_2 x_2 e^{-\frac{1}{2}b\left(x_2 - \frac{c}{b}x_1\right)^2}, \tag{20.6}$$

$$\propto \int dx_1 q(x_1) e^{-\frac{1}{2}\left(a - \frac{c^2}{b}\right)x_1^2} \frac{c}{b} x_1. \tag{20.7}$$

Now, first we notice that $C_2 \propto c \propto \phi$, so if there is no correlation between x_1 and x_2 we get the expected result that the composite is zero. Otherwise, for a generic committor q, $C_2 \neq 0$. A particular case for which $C_2 = 0$ is when q is an even function. However, this means that the committor must give equal probability to x_1 and $-x_1$, and thus cannot focus on a single tail of the distribution of x_1.

20.4 DETAILED CALCULATION OF THE COMPOSITE MAPS AND OF THE COMMITTOR FUNCTION IN THE GAUSSIAN APPROXIMATION FRAMEWORK

As already presented in the main text, the Gaussian approximation relies on the hypothesis that, at each pixel, the field and the heatwave amplitude follow a jointly Gaussian distribution, namely

$$(X, A) \sim \mathcal{N}(0, \Sigma), \tag{20.8}$$

with Σ being the covariance matrix, $\Sigma = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XA} \\ \Sigma_{AX} & \Sigma_{AA} \end{bmatrix}$. The joint multivariate Gaussian distribution, for given values of X and A, is generally written in the form:

$$\mathbb{P}(x, a) = \frac{1}{Z} \exp\left(-\frac{(x^T \Lambda_{XX} x + 2x^T \Lambda_{XA} a + \Lambda_{AA} a^2)}{2} \right), \tag{20.9}$$

where $Z = \sqrt{(2\pi)^d \det(\Sigma)}$ is the normalization constant d is the dimension of the stack of X and A, $\Lambda = \Sigma^{-1}$, $\Lambda_{XA} = \Lambda_{AX}$. We took advantage of the fact that a is a scalar quantity. Equation (14.14), can be obtained via the following calculation:

$$C = \mathbb{E}[X|A \geq a] = \frac{1}{\mathbb{P}(A \geq a)} \int_a^{+\infty} \left(\int x \mathbb{P}(x, a') \mathrm{d}x \right) \mathrm{d}a' , \tag{20.10}$$

$$= \frac{1}{\mathbb{P}(A \geq a)} \int_a^{+\infty} \mathbb{P}(a') \mathbb{E}[X|A = a'] \mathrm{d}a' , \tag{20.11}$$

$$= \frac{\int_a^{+\infty} \mathbb{P}(a') a' \mathrm{d}a'}{\mathbb{P}(A \geq a)} \frac{\mathbb{E}[XA]}{\Sigma_{AA}}, \tag{20.12}$$

$$= \eta\left(\frac{a}{\sqrt{2\Sigma_{AA}}} \right) \frac{\mathbb{E}[XA]}{\Sigma_{AA}}, \tag{20.13}$$

with

$$\eta(z) = \sqrt{\frac{2}{\pi}} \frac{e^{-z^2}}{\mathrm{erfc}(z)}, \tag{20.14}$$

where $\mathrm{erfc}(\bullet)$ is the complementary error function. Previously, we have used that:

$$\mathbb{E}[X|A = a] = \int x \mathbb{P}(x|a) \mathrm{d}x = \frac{\int x \mathbb{P}(x, a) \mathrm{d}x}{\mathbb{P}(a)} = \frac{a}{\Sigma_{AA}} \mathbb{E}[XA], \tag{20.15}$$

and this completes the proof.

For the committor, the important point is finding the expression for the conditional probability $\mathbb{P}(A = a|X = x)$. This is done by taking a slice at $X = x$ from of $\mathbb{P}(X, A)$ and expressing it as a function of a:

$$\mathbb{P}(A = a|X = x) \propto \exp\left(-\frac{1}{2} \left(2(\Lambda_{XA}^\top x) a + \Lambda_{AA} a^2 \right) \right), \tag{20.16}$$

$$\propto \exp\left(-\frac{1}{2} \Lambda_{AA} \left(a + \Lambda_{AA}^{-1} \Lambda_{XA}^\top x \right)^2 \right), \tag{20.17}$$

which is the expression for a one dimensional Gaussian distribution with variance $\sigma^2 = \Lambda_{AA}^{-1}$ and mean $\mu(x) = -\Lambda_{AA}^{-1}\Lambda_{XA}^{\top}x$. Now from the expressions for inverting a block matrix like Σ, we know that

$$\Lambda_{AA} = \left(\Sigma_{AA} - \Sigma_{AX}\Sigma_{XX}^{-1}\Sigma_{XA}\right)^{-1}, \tag{20.18}$$

and

$$\Lambda_{XA} = -\Sigma_{XX}^{-1}\Sigma_{XA}\Lambda_{AA}. \tag{20.19}$$

Remembering again that Λ_{AA} is a scalar, we immediately get eq. (14.18),

$$\mu(x) = \Sigma_{XX}^{-1}\Sigma_{XA} \cdot x, \qquad \sigma^2 = \Sigma_{AA} - \Sigma_{AX}\Sigma_{XX}^{-1}\Sigma_{XA}. \tag{20.20}$$

After this, getting the full committor is a simple one dimensional Gaussian integral, which is already well explained in the main text.

20.5 COMMITTOR FUNCTION FOR A STOCHASTIC PROCESS

Let's consider a stochastic process $X(t)$ on a phase-space Ω. The *first hitting time* $\tau_{\mathcal{V}}^{'}$ of the set $\mathcal{V} \subset \Omega$, given that the trajectory started at x, is defined as:

$$\tau_{\mathcal{V}}^{'}(x) := \inf\{t : X(t) \in \mathcal{V} \mid X(0) = x\}. \tag{20.21}$$

The committor function q is defined as the probability that the first hitting time of the set $\mathcal{C}$ is smaller than the first hitting time of set $\mathcal{B}$, given the initial conditions x, where $\mathcal{B}, \mathcal{C} \subset \Omega$, $\mathcal{B} \cap \mathcal{C} = \emptyset$:

$$q(x) := \mathbb{P}(\tau_{\mathcal{B}}^{'}(x) > \tau_{\mathcal{C}}^{'}(x)). \tag{20.22}$$

Sets $\mathcal{B}$ and $\mathcal{C}$ can be two attractors of the system or for instance one could correspond to a typical state of the system around which it fluctuates and another one to an atypical state which is visited when rare fluctuations arise. In the context of this paper, we are interested in the second case, where we define the fluctuations of interest based on an observable, namely the heatwave amplitude, defined in eq. (22.2), reaching a given threshold a. It is then natural to rewrite the definition of the committor function as in eq. (14.3).

20.6 SPATIAL GRADIENT REGULARIZATION

To compute the spatial gradient of the projection pattern M, we need to consider that we are working in a spherical geometry, which has two effects. If Λ and Φ are respectively latitude and longitude, the gradient in the local flat geometry x-y (with x pointing eastward and y northward) is

$$\begin{cases} \frac{\partial}{\partial x} & = \frac{\partial \Phi}{\partial x}\frac{\partial}{\partial \Phi} = \frac{1}{\cos\Lambda} \times \frac{\partial}{\partial \Phi} \\ \frac{\partial}{\partial y} & = \frac{\partial \Lambda}{\partial y}\frac{\partial}{\partial \Lambda} = 1 \times \frac{\partial}{\partial \Lambda} \end{cases}. \tag{20.23}$$

The other effect is that the area of a grid cell is

$$d\mathcal{A} = dxdy = \cos\Lambda d\Lambda d\Phi. \tag{20.24}$$

If for simplicity we assume we are dealing with only one climate variable, the total squared spatial gradient of M is

$$H_2(M) = \int \left(\left(\frac{\partial M}{\partial x} \right)^2 + \left(\frac{\partial M}{\partial y} \right)^2 \right) dxdy, \tag{20.25}$$

$$= \int \cos\Lambda \left(\left(\frac{1}{\cos\Lambda} \frac{\partial}{\partial \Phi} \right)^2 + \left(\frac{\partial}{\partial \Lambda} \right)^2 \right) d\Lambda d\Phi, \tag{20.26}$$

$$= \int \left(\frac{1}{\cos\Lambda} \left(\frac{\partial}{\partial \Phi} \right)^2 + \cos\Lambda \left(\frac{\partial}{\partial \Lambda} \right)^2 \right) d\Lambda d\Phi. \tag{20.27}$$

In our case, however, M spans three climate variables, and is sampled on a uniform grid in latitude and longitude. This means we can write the projection pattern as a tensor $M^{\lambda\phi f}$, with indices $\lambda = 1, \ldots, n_\lambda = 22$ for latitude, $\phi = 1, \ldots, n_\phi = 128$ for longitude and $f = 1, \ldots, n_f = 3$ for distinguishing the fields. The discrete version of the gradient is thus

$$\begin{aligned} H_2(M) = \sum_{f=1}^{n_f} \Bigg[& \sum_{\lambda=1}^{n_\lambda - 1} (\cos\Lambda_\lambda) \sum_{\phi=1}^{n_\phi} (M^{(\lambda+1)\phi f} - M^{\lambda\phi f})^2 \\ & + \sum_{\lambda=1}^{n_\lambda} (\cos\Lambda_\lambda) \sum_{\phi=1}^{n_\phi} \left(\frac{M^{\lambda((\phi \mod n_\phi)+1)f} - M^{\lambda\phi f}}{\cos\Lambda_\lambda} \right)^2 \Bigg], \end{aligned} \tag{20.28}$$

where the first row is the meridional gradient and the second row the zonal one, considering also the periodic term. To be precise, we should add the multiplicative term $\Delta\Lambda\Delta\Phi$, but since it is a constant that we can include in the regularization coefficient ϵ, we can ignore it

If we now collapse all the indices of M into a single one $i = i(\lambda, \phi, f)$, it is quite obvious that we can write

$$H_2(M) = M^\top W M = \sum_{ij} W_{ij} M_i M_j. \tag{20.29}$$

To get the expression for W, we can first notice that it is symmetric: $W_{ij} = U_{ij} + U_{ji}$, and, by matching terms, we get

$$\begin{aligned} U_{ij} = & \left(\frac{\cos\Lambda_\lambda + \cos\Lambda_{\lambda-1}}{2} + \frac{1}{\cos\Lambda_\lambda} \right) \delta_{i(\lambda,\phi,f)j(\lambda,\phi,f)} + \\ & - (\cos\Lambda_\lambda) \delta_{i(\lambda+1,\phi f)j(\lambda,\phi,f)} - \frac{1}{\cos\Lambda_\lambda} \delta_{i(\lambda,(\phi \mod n_\phi)+1,f)j(\lambda,\phi,f)}. \end{aligned} \tag{20.30}$$

For simplicity of notation, we assumed a null contribution when one of the indices goes out of range or (in the case of soil moisture) points to a grid cell with no data.

20.7 REGULARIZED GAUSSIAN COMMITTOR

To have the proper coefficients α and β when we deal with a regularized pattern, we can notice that the assumption that X and A follow a jointly Gaussian

distribution implies that, for *any* M, $F = M^\top X$ and A also follow a jointly Gaussian distribution. We can then use the same formulas of eqs. (14.18) and (14.23), but applied to the 2 by 2 covariance matrix between F and A

$$\hat{\Sigma} = \begin{pmatrix} \sigma_F^2 & \mathbb{E}[FA] \\ \mathbb{E}[FA] & \sigma_A^2 \end{pmatrix} =: \hat{\Lambda}^{-1}, \tag{20.31}$$

and simply

$$q_{\mathcal{G}}(x) = \frac{1}{2}\mathrm{erfc}\left(\hat{\alpha} + \hat{\beta} M^\top x\right), \tag{20.32}$$

with

$$\hat{\sigma}^2 = \sigma_A^2 - \left(\frac{\mathbb{E}[FA]}{\sigma_F}\right)^2, \quad \hat{\alpha} = \frac{a}{\sqrt{2}\hat{\sigma}}, \quad \hat{\beta} = \frac{\mathbb{E}[FA]}{\sqrt{2}\hat{\sigma}\sigma_F^2}. \tag{20.33}$$

20.8 EFFECTIVE NUMBER OF INDEPENDENT HEATWAVES

As we said in **??**, estimating the effective number of independent heatwaves can be challenging. The standard way of computing an effective data size for a time-series is the one presented in (Santer et al., 2000), where one uses the lag-1 autocorrelation coefficient r to rescale the total number of data points: $N_{\text{eff}} = N(1-r)/(1+r)$. However, when we consider heatwave events, they are not evenly spaced in time, so the whole approach does not make sense.

We can, tough, easily provide some bounds by observing that surely $N_{\text{y}} \leq N_{\text{eff}} \leq N_{\text{all}}$, where $N_{\text{all}} = N$ is the total number of heatwaves and N_{y} is the number of years that have at least a heatwave. Assuming that heatwaves at least a year apart are independent is definitely reasonable, if rather conservative. In fact, if we indeed compute the lag-1 autocorrelation coefficient for the time-series of $A(t)$, which gives $r = 0.9896$, and then estimate the decorrelation time of A as $\tau_{\text{decorr}} = (1+r)/(1-r)$, we get $\tau_{\text{decorr}} = 191$ days. Namely, it takes half a year to lose memory of the heatwave amplitude, and thus N_y is not only a lower bound for N_{eff}, but likely also very close to it.

If we apply this to our study of 14-day heatwaves we have $N_{\text{y}} = 2627 \lesssim N_{\text{eff}} \leq N_{\text{all}} = 30800$. Considering that we work with 8000 years of data, N_{y} tells us that there is a heatwave at least once every three years, and a year with a heatwave, on average, has $N_{\text{all}}/N_{\text{y}} \approx 12$ days for which $A(t) \geq a$.

20.9 VISUALIZATION OF THE ERROR BETWEEN EMPIRICAL AND GAUSSIAN COMPOSITES ON TWO GRID-POINTS

From fig. 15.1, we see that the biggest error we make when using the Gaussian composite is for the soil moisture variable. To investigate why this is the case, we show in fig. 20.3 (left) the joint and marginal distributions of the heatwave amplitude A (on the y axis) and of one pixel of soil moisture S^i (on the x axis). For comparison, we show the same for a pixel of the 500 hPa geopotential height Z^j in fig. 20.3 (right). While the marginal distributions of A and Z^j are approximately Gaussian (as it is shown from the black curve on the marginal

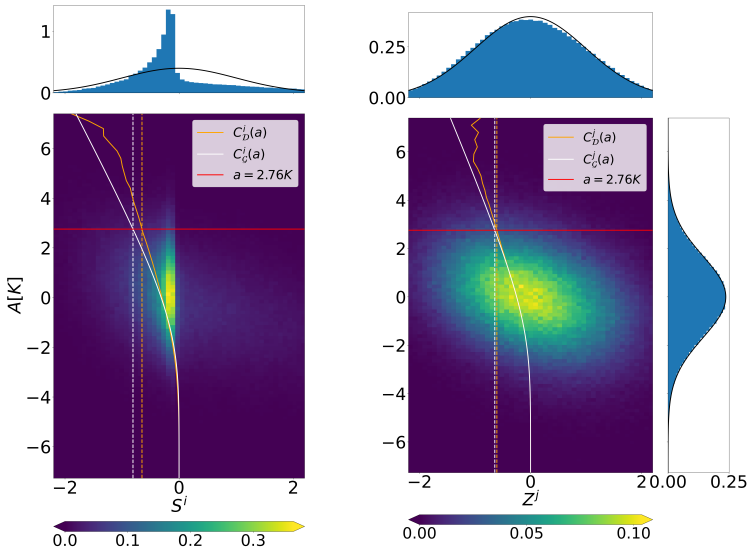


Figure 20.3: Comparison of the quality of the Gaussian and empirical composite map, for two grid-points and at different values of the heatwave threshold a. We show the results for a pixel over France of soil moisture S^i (left) and one over Greenland of 500 hPa geopotential height Z^j (right). For each panel, the main plot is the joint probability density function (PDF) of S^i or Z^j and the heatwave amplitude A, whith marginal distributions displayed on top and to the side. In these plots the black line is a Gaussian fit. In the main plots, the orange line is the empirical composite as a function of a, while the white line is the one estimated through the Gaussian approximation. The red line is the threshold value of $a = 2.76K$ corresponding to the 5% most extreme heatwaves. The dotted vertical white and orange lines indicate the values of the empirical and Gaussian composites at this particular value of a. For Z^j the error is much smaller.

plots of the figure), the one of S^i is clearly not: it is strongly skewed towards negative values of soil moisture anomaly, and exhibits fat tails. This is also reflected in the plot of the joint distribution of A and S^i on the one hand and Z^j on the other hand (heat maps in fig. 20.3), as only the latter has the shape of an ellipsoid.

In both panels of fig. 20.3, the orange curve shows the behavior of the empirical composite $C_{\mathcal{D}}$ as we change the threshold a, while the white curve is the behavior of the Gaussian composite $C_{\mathcal{G}}$. By construction, they both tend to 0 as $a \to -\infty$, since soil moisture S^i and 500 hPa geopotential height Z^j both have zero mean as they are anomalies.

When a is very small, then, the Gaussian composite provides a good, yet useless, approximation of the empirical composite as both are very close to zero. As the threshold increases beyond this trivial region, for soil moisture the two curves start to diverge already at $a \approx 0$, and thus show a significant distance when they reach the 95$^{\text{th}}$ quantile of A, $a = 2.76\,\text{K}$ (red line). On the other hand, the approximation holds quite well in the case of geopotential height, and the two curves separate significantly only when the sampling error kicks in (around $a = 5\,\text{K}$) and throws off the empirical estimate.

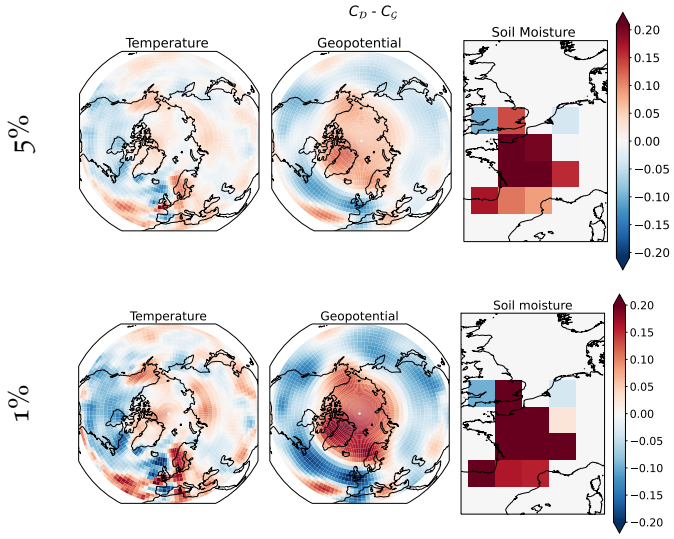


Figure 20.4: Maps of the difference between empirical composite maps and the ones estimated from the Gaussian approximation, for the 5% and 1% most extreme 14-day heatwaves over France. The patterns of this difference do not depend on the threshold a, varying only in intensity. The salient features of both temperature and geopotential height are well captured by the Gaussian approximation, with errors of the order of 30% at most.

20.10 ERROR BETWEEN EMPIRICAL AND GAUSSIAN COMPOSITES AT DIFFERENT HEATWAVE THRESHOLDS

In fig. 20.4 we show the difference between the empirical composite and the Gaussian one (evaluated using eq. (14.14)) for the three fields of 2 meter air temperature, 500 hPa geopotential height and soil moisture evaluated for a corresponding to the 5% and 1% most extreme 14-day heatwaves. The striking result is that the pattern observed changes only in magnitude between extreme and very extreme events.

To give a quantitative measure of the error, in fig. 20.5, we evaluate the error using the norm ratio defined in eq. (15.1), for different threshold level a, showing that the error is around the 20% for the 5% most extreme events. Figure 20.5 gives more details on the behavior of the norm of the error shown in fig. 20.4 for different values of a. On the y-axis we represent the norm ratio introduced in section 15.1 (eq. (15.1)), which measures how distant (in norm) the Gaussian composite is from the empirical composite (normalized by the norm of the empirical composite). We calculated this norm for the three fields independently, and for the whole stack of them (gray line). The norm ratios of the 500 hPa geopotential height and the one of the 2 m air temperature stay pretty close to the norm of the stack, showing values below 0.3 even for events which represents the 1% most extreme ones in the dataset (the x-axis on the top shows the respective percentile of rareness of the a, which is on the bottom x-axis). Soil moisture has a different behavior, showing higher values of the norm ratio for much less extreme events. This is possibly due to a violation of the Gaussian approximation assumption, as we showed for a single pixel in fig. 20.3.

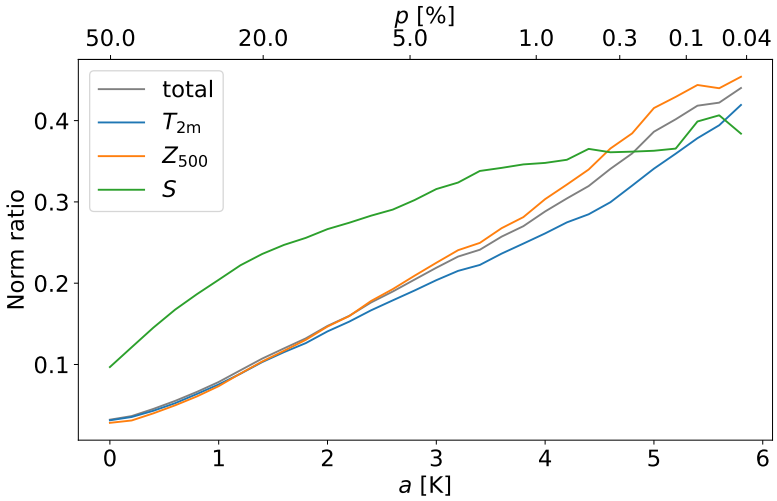


Figure 20.5: Norm ratio (see eq. (15.1) of the main text) of the difference between the empirical composite map and the Gaussian approximated one as a function of the threshold value used to define a heatwave event a. The total norm ratio is in gray, while the colors represent the norm ratio for each of the three fields (namely 2 m temperature anomaly, 500 hPa geopotential height anomaly and soil moisture anomaly). For events which are the 1% most extreme ones of the PlaSim dataset, the relative error is always below the 30%. The bottom x-axis a is the threshold value used to define an heatwave event from the distribution of the 14-day heatwave amplitude A. On the top x-axis, p is the respective percentile value corresponding to a given a.

Temperature

T [days] \ τ [days]	0	3	6	9	12	15	18	21	24	27	30
1	0.26	0.30	0.31	0.22	0.21	0.22	0.22	0.21	0.19	0.19	0.19
3	0.25	0.29	0.26	0.19	0.19	0.20	0.20	0.19	0.18	0.17	0.18
7	0.22	0.24	0.21	0.19	0.19	0.20	0.20	0.19	0.18	0.18	0.16
14	0.19	0.20	0.20	0.19	0.20	0.20	0.21	0.21	0.19	0.18	0.18
30	0.18	0.19	0.20	0.20	0.20	0.20	0.20	0.19	0.19	0.20	0.20

Table 20.1: Values of the norm ratio between Gaussian and empirical composites computed for 2 m temperature anomaly.

Geopotential

T [days] \ τ [days]	0	3	6	9	12	15	18	21	24	27	30
1	0.25	0.26	0.28	0.30	0.31	0.33	0.36	0.34	0.29	0.26	0.27
3	0.24	0.23	0.26	0.28	0.29	0.31	0.34	0.31	0.27	0.25	0.25
7	0.21	0.23	0.27	0.30	0.33	0.38	0.37	0.32	0.30	0.30	0.25
14	0.21	0.24	0.30	0.34	0.34	0.33	0.33	0.34	0.32	0.28	0.28
30	0.22	0.25	0.29	0.32	0.33	0.33	0.34	0.32	0.30	0.31	0.31

Table 20.2: Values of the norm ratio between Gaussian and empirical composites computed for 500 hPa geopotential height anomaly

20.11 FIELD-WISE NORM RATIO OF COMPOSITE MAPS AT DIFFERENT VALUES OF T AND τ

In tables 20.1 to 20.3, we show the norm ratio as defined in eq. (15.1), but computed independently for the three climate variables of the PlaSim dataset. Values related to temperature peak at $T = 1$ and for small delay time. Soil moisture exhibits a clear monotonic trend with respect to T, while not being very sensitive to the the lead time τ. Finally, the geopotential height field shows the more complex structure, with the highest errors happening at intermediate values of both T and τ.

Soil moisture has only 12 pixels, compared to the 2816 of the other two fields, so its contribution to the total norm ratio (table 17.2) is negligible. Outside the small region of low values of T and τ, the norm ratio of the temperature field is almost constant, so the structure visible in table 17.2 is mostly due to the geopotential height field.

Soil moisture

T [days] \ τ [days]	0	3	6	9	12	15	18	21	24	27	30
1	0.10	0.11	0.12	0.12	0.13	0.13	0.13	0.13	0.14	0.14	0.14
3	0.14	0.12	0.12	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.12
7	0.22	0.20	0.19	0.18	0.18	0.17	0.16	0.16	0.15	0.15	0.14
14	0.30	0.28	0.27	0.26	0.25	0.24	0.23	0.22	0.21	0.20	0.19
30	0.40	0.38	0.36	0.35	0.33	0.31	0.30	0.28	0.27	0.25	0.24

Table 20.3: Values of the norm ratio between Gaussian and empirical composites computed for soil moisture anomaly.

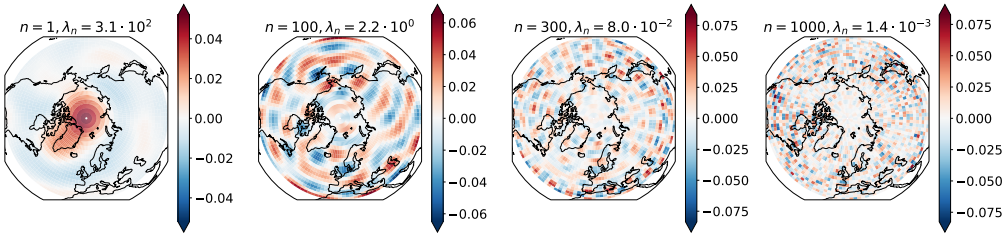


Figure 20.6: Four EOFs e_n. The higher the n the smaller the spatial scales of the characteristic features represented. On top of each plot we report the value of λ_n.

20.12 ASYMPTOTIC BEHAVIOR OF THE REGULARIZED PROJECTION PATTERN

In this section we discuss the effect of the regularization coefficient ϵ on the optimal projection pattern M_ϵ, and in particular why a highly regularized projection pattern may look similar to the composite map.

To do so we move to the basis of Empirical Orthogonal Functions (EOFs) (Hannachi, Jolliffe, and Stephenson, 2007), which diagonalizes the covariance matrix Σ_{XX}. We call its eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq \cdots \geq \lambda_d$, and the corresponding eigenvectors e_n. Here we will use as an example the prediction of $T = 14$ day heatwaves at delay time $\tau = 0$, performed on the ERA5 reanalysis dataset. In this context there are $d = 2816$ degrees of freedom, corresponding to the pixels of the 500 hPa geopotential height anomaly field.

The first important point is that as n increases, the variance λ_n explained by e_n decreases, and so does decrease the typical spatial size of the features that appear in e_n (fig. 20.6). In particular, features with a typical size of the order of the synoptic scale are represented around $n = 100$. Moreover, almost two thirds of the EOFs ($n > 1000$) explain less than 0.05 % of the variance and are extremely noisy.

Second, the Gaussian composite map is proportional to the correlation map Σ_{XA} (eq. (14.14)), and when we write it in the EOF basis,

$$C_{\mathcal{G}} \propto \Sigma_{XA} = \sum_{n=1}^{d} c_n e_n, \tag{20.34}$$

it is dominated by EOFs at low values of n (see black lines in fig. 20.7), and thus it appears spatially smooth.

Third, in the EOF representation the non-regularized ($\epsilon = 0$) projection map is written as

$$M_0 = \sum_{n=1}^{d} M_n^0 e_n \propto \Sigma_{XX}^{-1}\Sigma_{XA} \propto \sum_{n=1}^{d} \frac{c_n}{\lambda_n} e_n, \tag{20.35}$$

and λ_n goes to zero much faster than c_n (dashed dark blue lines in fig. 20.7), resulting in M_0 being dominated by large n, high spatial frequency modes. This is what makes the non regularized pattern utterly non-interpretable.

If we perform L_2 regularization, the aim of the regularization coefficient is to prevent the contribution of these high frequency modes to explode, making them proportional to their values in the composite map:

$$M_\epsilon \propto (\Sigma_{XX} + \epsilon\mathbb{I})^{-1}\Sigma_{XA} \propto \sum_{n=1}^{d} \frac{c_n}{\lambda_n + \epsilon} e_n \approx \sum_{n=1}^{n_\epsilon - 1} M_n^0 e_n + \frac{1}{\epsilon}\sum_{n=n_\epsilon}^{d} c_n e_n, \tag{20.36}$$

where $\lambda_{n_\epsilon} \approx \epsilon$. It is then clear that as ϵ increases $n_\epsilon \to 1$, and the projection map smoothly converges to the composite map (left panel of fig. 20.7).

On the other hand, when we perform H_2 regularization, as we do in this work, we regularize with matrix W, which doesn't share the same eigenvectors of Σ_{XX}. We can in any case write W in the EOF basis as

$$W = \sum_{mn} W_{nm} e_n e_m^\top. \tag{20.37}$$

If we compute the terms W_{mn}, we notice that $W_{nn} \gg \max_{m\neq n} |W_{nm}|$. We can then say that the W is almost diagonal and thus

$$M_\epsilon \propto (\Sigma_{XX} + \epsilon W)^{-1}\Sigma_{XA} \approx \sum_{n=1}^{d} \frac{c_n}{\lambda_n + \epsilon W_{nn}} e_n. \tag{20.38}$$

This lets us apply a similar reasoning to the one explained above for L_2 regularization, where W_{nn} is the norm of the spatial gradient of EOF e_n, which, considering the spatial structure of the EOFs (fig. 20.6), clearly increases with n. For this reason, when we increase ϵ, we remove the high spatial frequencies faster than we would with L_2 regularization (right panel of fig. 20.7). On the other hand, for very high regularization, the approximation of W being diagonal falls apart and the high frequencies are brought back to achieve a spatially uniform pattern, similarly to the Fourier representation of a square wave. So there is no asymptotic convergence to the composite map (brown curve). However, for intermediate values of ϵ (yellow, orange and red curves), the projection pattern is smoothed in a similar way as with L_2 regularization and thus it may look similar to the composite map.

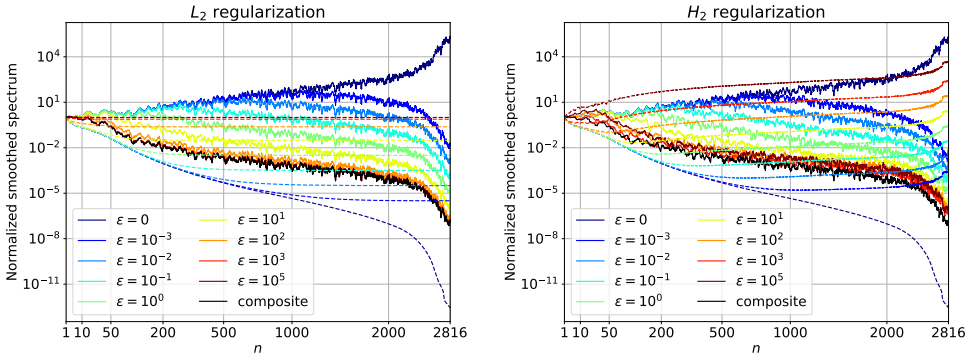


Figure 20.7: EOF spectra $|M_n^\epsilon|$ of the projection pattern at different values of the regularization coefficient ϵ (solid lines) when penalizing the L_2 norm of the pattern (left) or its spatial gradient (right). All spectra are normalized so that the term at $n = 1$ has unitary values. To ease the visualization, the spectra have been smoothed with a running average. On the left panel the dashed lines represent the spectra of the regularized covariance matrix: $\lambda_n + \epsilon$. On the right panel they represent the diagonal part of the gradient regularized covariance matrix in the original EOF basis, i.e. $\lambda_n + \epsilon e_n^\top W e_n$. The black line is the spectrum c_n of the Gaussian composite map. For high n (i.e. EOFs with small spatial scales), the values of λ_n decay faster than those of c_n, which makes the non-regularized pattern extremely noisy. For L_2 regularization, increasing values of ϵ progressively reduce the contribution of EOFs at high n, and the projection pattern converges to the composite. On the other hand, H_2 regularization directly penalizes the spatial gradient of the projection pattern, so the small scales are first suppressed and then brought back to achieve the spatially uniform pattern. These spectra are presented for $T = 14$ and $\tau = 0$ on the ERA5 dataset, where $\epsilon_{\text{best}} = 10$ (light green lines).

Part IV

RARE EVENT ALGORITHM STUDY OF EXTREME DOUBLE JET SUMMERS AND THEIR CONNECTION TO HEATWAVES OVER EURASIA

In the previous chapter (part iii) we presented a novel methodology, the Gaussian approximation, which, despite its simplicity, allows to analyse and predict heatwaves with different duration and delay time. In this chapter we solve the scarcity of data and computational cost associated with the sampling of rare events by generating efficiently more observations of the rare event of interest, namely a summer double jet structure over Eurasia, using a rare event algorithm.

This project has been developed during the secondment periods, for a total of four months, planned as part of my PhD program, in collaboration with Nili Harnik (*Professor at Department of Geosciences, Tel Aviv University, Tel Aviv, Israel*) and Francesco Ragone (*Professor at Royal Meteorological Institute of Belgium, Brussels, Bel- gium* and at *Georges Lemaître Centre for Earth and Climate Research (TECLIM), Earth and Life Institute, Université Catholique de Louvain, Louvainla-Neuve, Belgium*). The project is still an ongoing work and I present here a summary of what I think are the main conclusions of the work gathered up to know. In the months to come, this work will be finalised and will hopefully constitute a third publication achieved during my PhD.

For what concerns the tools used in this chapter, we employ the rare event algorithm, the Giardina-Kurchan-Lecomte-Tailleur GKLT, presented in section 4.2 to perform an importance sampling of rare double jet summers. For the analysis, we use the *return time curves* to estimate the average time to wait for observing events with different amplitude and the *composite maps* for the conditioned analysis of the dynamical fields when there is a double jet. Both tools were introduced in section 3.1.1. For the seek of clarity, some of these concepts will be reintroduced in the following, leading to minor repetitions.

21 INTRODUCTION

Nowadays, heatwaves are observed with an atypical frequency and amplitude in response to the effect of global climate warming (Seneviratne et al., 2021). At the mid-latitudes, several mechanisms have been proven to be involved in the formation, maintaining and worsening of these events (Horton et al., 2016; Barriopedro et al., 2023), including land-atmosphere feedbacks (Miralles et al., 2019; Miralles et al., 2014; Seneviratne et al., 2006), large-scale circulation patterns (Cassou, Terray, and Phillips, 2005; Della-Marta et al., 2007) and sea surface temperature anomalies (Della-Marta et al., 2007; Cassou, Terray, and Phillips, 2005). High-amplitude quasi-stationary Rossby waves are persistent weather regimes known to favour extreme events (Hoskins and Woollings, 2015), including heatwaves during summer (Kornhuber et al., 2019; Teng et al., 2013; Mann et al., 2017). However, a comprehensive fundamental theory for the growth and propagation of Rossby waves at the spatial and temporal scales pertinent to extreme events remains incomplete (White et al., 2022). A proposed mechanism in this sense is quasi-resonance amplification (Petoukhov et al., 2013). Under some conditions, for instance in the presence of a double jet stream, amplified quasi-stationary Rossby waves might result from interactions of Rossby waves of similar wave number (especially between 5 and 7) that are amplified and trapped within a latitudinal band, leading to persistent weather patter. Several studies investigated the relevance of this mechanism for recent heatwaves (Mann et al., 2017; Coumou et al., 2014; Petoukhov et al., 2016; Kornhuber et al., 2019) and the link between double jet and heatwaves in the Eurasian region (Rousi et al., 2022). In this study, we aim at investigating the anomalous summer circulation associated to a persistent double jet occurring over Eurasia and to link its occurrence to heatwaves. However, this persistent configuration of the jet is uncommon.

A difficulty in the prediction and in the understanding of the mechanisms which led to rare events is related to their low frequency of occurrence, or equivalently to their low probability. When dealing with extreme and rare events in the climate field, historical records are too short for observing a significant collection of them and running climate simulations is highly resources demanding. Therefore, alternative methodologies are necessary to overcome the limited amount of available data. In this study, we tackle the issue of computational cost limitations using a rare events algorithm. The primary goal of a rare events algorithm is to efficiently make the rare event of interest less rare, allowing us to observe more occurrences and gather sufficient statistical data at a lower computational cost than direct simulations. The algorithm used here is known in the literature as GKLT and has been historically developed to compute estimators in the context of Large Deviation theory (Giardinà, Kurchan, and Peliti, 2006). Recently, it has been applied in climate science to study warm summers over France and Scandinavia (Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2020; Ragone and Bouchet, 2021), and India (Le Priol, Monteiro, and

Bouchet, 2024), mid-latitudes precipitations (Wouters, Schiemann, and Shaffrey, 2023), weakening and collapse of the Atlantic Meridional Overturning Circulation (Cini et al., 2024), melting of the Arctic sea ice (Sauer et al., 2024) and for energy demand in the power system (Cozian, 2023). Differently from previous applications of this algorithm, in this study we sample directly persistent atmospheric states which are prone to support extreme events. We stress that this was not the case in previous works, where for example the same rare events algorithm was applied for sampling of extreme warm summers and the atmospheric states were retrieved after (Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2021).

In this study, we show that the GKLT efficiently samples double jet summers over the Eurasian continent, using CESM1.2.2 climate model. We compute return times of orders of magnitude larger than what feasible with direct sampling, and statistically significant composite maps of dynamical quantities. Our results show that rare double jet summers are associated with a wave number 3 teleconnection patterns in the Northern Hemisphere and positive temperature anomalies in North Canada, Scandinavia/West Russia and East Russia.

The study is organised as follows. In chapter 22 we present the data used for this study and the two indexes, namely the double jet index which defines the double jet structure and the heatwave index. We then introduce the rare event algorithm and discuss briefly the parameters for this peculiar study. In chapter 23 we first show the performance of the double jet index on a long run of CESM1.2.2 and on the reanalysis dataset ERA5. As a motivation for the study, we show that there is a significant percentage of overlapping days between double jet and heatwaves, and that this percentage augments the more persistent the events considered are. Then, we show that the rare events algorithm effectively and efficiently samples rare and unseen double jet summers, and we analyse associated dynamical quantities. In chapter 24 we summarise our findings and give perspectives.

22 DATA AND METHODS

22.1 DATA

22.1.1 *CESM*

In this study, we use the CESM version 1.2.2 (Hurrell et al., 2013). In the setup we use, only atmosphere and land are used as active components. More specifically, the atmospheric model is the Community Atmospheric Model version 4 (CAM4), while the Community Land Model version 4 (CLM4) is used for the land. The version used has a prescribed aerosol concentration, including greenhouse gases, fixed to a level which matches the observed one to reproduce the 2000s climate, with the CO_2 concentration set at 367 ppmv. Finally, we use historical data to prescribe the ocean and the sea-ice models. The model has a resolution of 0.9 degrees in latitude and 1.25 in longitude, with 26 vertical layers in a hybrid pressure-sigma coordinate.

With this setup we use a long simulation of 1000 years, which has been already used to study heatwaves and energy production in Europe (Ragone and Bouchet, 2021; Cozian, Herbert, and Bouchet, 2024). Without changing the setup, we run novel simulations by coupling this climate model with a rare event algorithm to sample rare double jet events. The procedure is presented in section 22.2.3 and the results in section 23.3 and section 23.4. The model has an output frequency of 3 hours, but we are interested only in daily averages. In particular, we focus on the anomalies (with respect to the daily, grid point-wise climatology) of the zonal wind U, 2 m temperature (T_{2m}) and 500 hPa geopotential height (Z_{500}). For the zonal wind, we are interested in upper level winds, thus we selected 4 vertical levels between 192 and 313 (in an hybrid pressure-sigma coordinate).

22.1.2 *ERA5*

We use the public available reanalysis dataset ERA5 (Hersbach et al., 2020). Specifically, we use daily data for the summer months of June, July and August, starting from 1940 until 2022 for the 2 m air temperature (T_{2m}), 500 hPa geopotential height (Z_{500}), zonal wind U (averaged between 200 hPa and 350 hPa). The dataset has a resolution of 0.25 degree both for latitude and for longitude. As we want to study the response of stationary climate, we need to remove the warming trend present in ERA5. We performed a latitudinal-wise detrending of both 2 m air temperature, 500 hPa geopotential height longitudinally averaged over the Northern Hemisphere (we took into account lands only for the 2 m air temperature). The procedure is detailed in section 25.1.

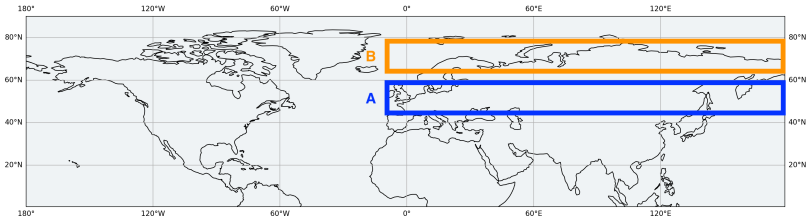


Figure 22.1: Example of the two zones, for B between the latitudinal band 65-78.5°N and A between 46.5-60°N for defining the double jet.

22.2 METHODS

22.2.1 *Double jet index*

Following the work of (Yang and Chang, 2006), we devised the following procedure. We selected and averaged the zonal wind over 4 vertical levels in the upper troposphere (for the levels, the details for CESM and ERA5 are in sections 22.1.1 and 22.1.2, respectively). We calculate the zonal wind daily anomalies (with respect to the daily, grid point-wise climatology) of the months of June, July and August. We define then two zones, according to their latitude: A is a zone comprised between 40°N and 65°N, while B between 60°N and 85°N. We then take a zonal average over the longitudinal sector corresponding to Eurasia, from 10°W to 180°E. Note that given the zonal average, both zones are at this stage only time and latitude dependent. Then, in both zones, we average the daily zonal wind zonal anomaly over a window of $w = 13.5°$ in latitude. We take the difference between pairs of zone B and zone A of $d = 5°$ distance. We define then our double jet index D_{ji} as the daily maximum of this difference:

$$D_{ji}(t) := \max_{\tilde{\lambda}\in(40,52)} \left[\int_{\tilde{\lambda}+w+d}^{\tilde{\lambda}+2w+d} U(t,\lambda)d\lambda - \int_{\tilde{\lambda}}^{\tilde{\lambda}+w} U(t,\lambda)d\lambda \right] \tag{22.1}$$

where $U(t,\lambda)$ is the zonal wind daily anomaly already averaged over the vertical levels and the Eurasian longitudinal sector (from 10°W to 180°E), t denotes the time and λ the latitude. The first integral corresponds to zone B, while the second to zone A. An example of the two zones (before the zonal averaging), for B between the latitudinal band 65-78.5°N and A between 46.5-60°N is shown in fig. 22.1.

The index is robust to changes in many of the features. For instance, we tried to vary the vertical levels over which we evaluated it, the variable used (we computed the same index with the wind speed, instead of the zonal wind component only), we reduced the integrated longitudinal sector up to 70°E and we computed the index with the raw field (without taking the seasonal cycle out). The version presented here is the one that retains most of the double jet structure, and that is at the same time less computationally expensive to be embedded in a rare events algorithm. In fact, given our interest in performing simulations coupled with that algorithm (section 22.2.3), we need a one dimensional index which could successfully represent the double jet splitting. Thus, despite other definitions have been proposed to detect this splitting (Molnos

et al., 2017; Rousi et al., 2022; Gallego et al., 2005; Pena-Ortiz et al., 2013), our choice is a trade-off between a correct representation of the phenomenon and methodological constraints. With this definition, the jet is expected to be split poleward the higher the value of the index, creating a new band of strong westerlies over the desired zone. An application of this index to a long climate simulation performed with the CESM climate model (introduced in section 22.1.1) and to the reanalysis dataset ERA5 (section 22.1.2) is presented in section 23.1.

22.2.2 *Heatwave definition*

In the climate community, heatwaves have been defined in several ways (Perkins, 2015; Barriopedro et al., 2023). In this work, we are interested in events which are persistent in time and space. The natural variable to investigate is the running mean of the 2 m air temperature, which has already been employed for studying events of different duration (from weekly to monthly time windows) (Barriopedro et al., 2011; Coumou and Rahmstorf, 2012; Schär et al., 2004). Following the works of (Gálfi, Lucarini, and Wouters, 2019; Galfi and Lucarini, 2021; Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2021; Jacques-Dumas et al., 2022; Miloshevich et al., 2023), we base the definition of heatwave on spatial and temporal averages of the 2 m temperature anomaly, as will be described later in the section. This definition is suitable for the type of events we wish to study, as it gives a direct measure of the intensity and the duration of the heatwave. Moreover, another interesting point of such definition is that both the temporal duration and spatial extent can be adapted to the type of the event one wishes to study.

Let $\tilde{T}_{2m}$ denote the daily-averaged 2 m air temperature field, which depends on the location $\vec{r}$ and time t. We use the temperature anomaly T_{2m} in order to remove the effect of the seasonal cycle, i.e. $T_{2m} := \tilde{T}_{2m} - \mathbb{E}_y(\tilde{T}_{2m})$ where $\mathbb{E}_y(\tilde{T}_{2m})$ is the climatology. The heatwave amplitude A is then defined as the space and time average of the temperature anomaly:

$$A(t) := \frac{1}{T}\int_t^{t+T}\left(\frac{1}{\mathcal{A}}\int_{\mathcal{A}} T_{2m}(\vec{r},u)d\vec{r}\right)du, \tag{22.2}$$

where T is the duration in days of the heatwave and $\mathcal{A}$ is the spatial region of interest. Both parameters T and $\mathcal{A}$ can be changed according to the event one wishes to study. In this work, T varies from sub-weekly to monthly time scales, enabling to study short and long events. The extension of the spatial region $\mathcal{A}$ should be comparable to the synoptic scales, which is of the order of 1000 km at the mid-latitudes, as these scales correlate to the size of cyclones and anticyclones and jet stream meanders. We decided to focus over 3 regions in the Northern Hemisphere, shown in section 25.6. The choice of these regions is motivated by a teleconnection pattern discussed in section 23.4. For these regions, we know that the most impactful events connected with a double jet occur in summer (Rousi et al., 2022). Thus, in this study we only consider events happening in the months of June, July and August.

22.2.3 *Rare event algorithms: sampling rare events with low computational cost*

The algorithm used in this manuscript has been already employed for resampling different types of events, such as heatwaves in France and Scandinavia (Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2021), in India (Le Priol, Monteiro, and Bouchet, 2024), for energy demand in the electric power system (Cozian, 2023), for collapse of the Atlantic Meridional Overturning Circulation (Cini et al., 2024) and for the Arctic sea ice melting (Sauer et al., 2024). This algorithm is part of the genealogical selection algorithm (Del Moral and Garnier, 2005; Giardinà, Kurchan, and Peliti, 2006; Giardina et al., 2011; Pérez-Espigares and Hurtado, 2019).

The overall goal is to sample more efficiently the upper tail of the distribution of a target observable $O[X(t)]$. To achieve this goal, an ensemble of trajectories is run in parallel and at constant intervals of a priori chosen resampling time τ a weight is assigned to each of them, based on the value of the target observable $O[X(t)]$ reached during the simulation time of duration τ. Based on this weight, trajectories achieving a low value are killed, as this is interpreted as a poor skill of those trajectories in sampling extreme fluctuations of $O[X(t)]$ at the end of the simulation, while trajectories having a high score are cloned, repopulating the ensemble of trajectories. Then, a small perturbation is added to restart the simulation for a time τ and the algorithm is cyclically repeated. The idea behind is that this selection process, when done at the appropriate time scales, will favour the survival of trajectories leading to extreme events characterized by time persistence of large values of $O[X(t)]$, such as the double jet events we are interested in.

In section 4.2 we describe the rare event algorithm in detail, while here we summarize only the main points. We invite the reader to refer to that section for a deeper discussion. Let's consider a realisation of a climate simulation, denoted from now on as a trajectory, as $\{X(t)\}_{t_a \le t \le t_a+T}$, where t_a is the initial starting point of the simulation and T the total duration. We denote as $\mathbb{P}_0\left(\{X(t)\}\right)$ the probability of observing a certain trajectory as a realisation of the dynamics of the climate model and as $\mathbb{P}_k\left(\{X(t)\}\right)$ the probability of observing the same trajectory as a result of the ensemble simulations driven by the rare event algorithm. Following the works of (Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2021), there exists a link between the two probabilities:

$$\mathbb{P}_k\left(\{X(t)\}\right) = \frac{\exp\left(k \int_{t_a}^{t_a+T} O[X(t)]dt\right)}{Z} \mathbb{P}_0\left(\{X(t)\}\right) \tag{22.3}$$

where Z is a normalization constant (the equivalent of the partition function in statistical mechanics), $O[X(t)]$ is the target observable and k the biasing parameter, a parameter which controls the selection strength, i.e. the higher its value, the larger values of the integrated time average of $O[X(t)]$ will be sampled. Equation (22.3) allows having access to the probabilities of the real model statistics (and to related quantities, such as averages) using the rare event algorithm simulated by inverting it. The detailed description with the formulae are in section 4.2.

The target variable for this study is the integrated double jet index averaged over Eurasia, presented in section 22.2.1. Thus, to be consistent with the nota-

tion presented above, $O[X(t)] \equiv D_{ji}(t)$. We perform $M = 10$ ensemble simulations with the rare event algorithm, each of them with $N = 100$ trajectories, a selection strength of $k = 0.01(m/s)^{-1}(day)^{-1}$ for a duration $T = 90$ days, starting from June 1st and ending on August 29th and with a resampling time of $\tau = 5$ days. The value of k has been chosen such that events with a return time of 100 years are common. The resampling time τ should be of the order of the autocorrelation time. An analysis of the autocorrelation function of the integrated double jet index (fig. 25.4) shows that it can be approximated by two exponential functions, suggesting that a good candidate is 5 days. For the integration time T, we wanted to target extreme double jet during the summer months of June, July and August, because of correlation with heatwaves at mid-latitudes. Moreover, in order to let the potential candidate trajectories to yield to the extreme event we need to have $T \gg \tau_c$, where τ_c is the autocorrelation time. However, it should not be too large in order not to eventually kill in late resampling steps the trajectories which shown a large fluctuation before and went back to typical values. Finally, we choose $T = 90$ days. In this way, we perform 18 resampling steps ($T/\tau = 18$). The 10 experiments start from 10 independent sets of 100 independent initial conditions, i.e. we took 1000 independent June 1st, at 1-year interval. This could be done because in the climate model setup there is no active ocean or ice component, thus limiting the long term variabilities to seasonal scales. The computational cost of the experiments with the algorithm is equivalent to simulating 1000 summers in the ensemble control run, but they allow gathering a much richer statistics for the extreme events of interest.

More details on the choice of the parameters of the GKLT are presented in section 4.4.

23 RESULTS

23.1 SAMPLING DOUBLE JET WITH THE DOUBLE JET INDEX

In this section, we use the double jet index defined in section 22.2.1 to sample double jet events using the long climate model CESM run of 1000 years (presented in section 22.1.1) and the reanalysis dataset ERA5 (presented in section 22.1.2). The climatology of the daily zonal wind U for the months of June, July and August, for CESM is presented in fig. 23.2 (a) and (b). Specifically, in (a) the vertical profile of the zonal wind U longitudinally averaged over the Eurasian sector (10°W to 180°E) is shown, while in (b) the climatology map of the zonal wind averaged over the months of June, July and August is presented. The probability density distribution of the daily double jet index is displayed in fig. 23.1, for both CESM (blue) and ERA5 (orange). The index has been deseasonalised and standardized in order to compare both models. The overlapping of both distributions suggests that CESM in this setup is correctly simulating the dynamics of the atmosphere.

As we already mentioned, the splitting of the jet should be evident the higher the value of the index. We define a double jet as an event for which the amplitude D_{ji} exceeds a threshold corresponding to the 95th quantile of the distribution of D_{ji}, i.e. we consider as double jet the 5% most extreme events in our dataset. This threshold is displayed as a vertical black line in fig. 23.1. Conditioning on events with amplitude greater than this value, we observe in the zonal wind in fig. 23.2 (c) and (d) the appearance of a band of strong westerlies which is displaced poleward. These maps are for CESM. Similarly to what has been done before, in fig. 23.3 we repeated this analysis with ERA5. We show the summer climatology of vertical profile of the zonal averaged zonal wind over Eurasia in (a) and conditioned the 5% most extreme double jet values in (c). In the same figure, we show in (b) the climatology map of the zonal wind daily averaged over the months of June, July and August and in (d) the composite map of the zonal wind conditioned over the 5% most extreme value of the daily double jet index. The results between CESM and ERA5 bear great similarity, and we are confident that the proposed index is correctly detecting double jet events over the targeted area. Moreover, this behaviour strongly resembles the one found in (Rousi et al., 2022) using machine learning-based techniques for ERA5.

23.2 DOUBLE JET AND HEATWAVE EVENTS

In this section, we investigate the link between double jet events and heatwaves. A preliminary analysis of the composite maps of T_{2m} and Z_{500} conditioned on the 5% most extreme values of the daily double jet index has shown that positive temperature anomalies and associated anticyclonic anomalies occur prefer-

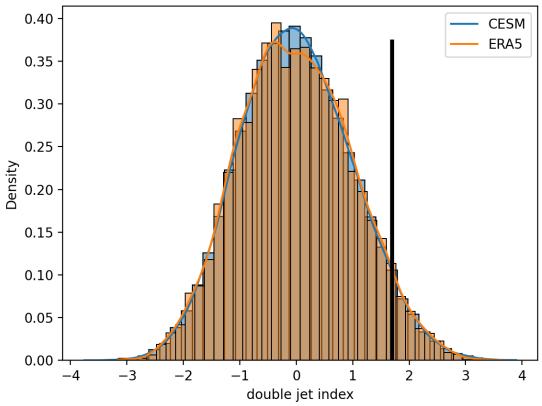


Figure 23.1: Histograms and kernel density estimation of the standardized daily double jet index for CESM (blue) and ERA5 (orange). Both indexes have been deseasonalised and divided by the climatological standard deviation. The vertical black line indicates the threshold corresponding to the 5% most extreme events chosen to define a double jet in figs. 23.2 and 23.3.

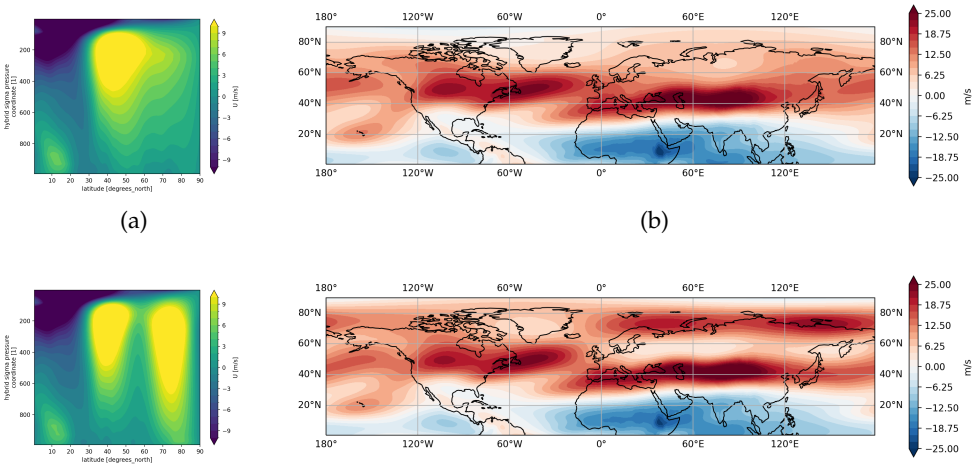


Figure 23.2: (a) Vertical profile of the daily summer zonal wind U for the climate model CESM (1000 years), longitudinally averaged over the Eurasian sector (10°W to 180°E). We considered the months of June, July and August - (b) Zonal wind U climatology for the CESM climate model (1000 years) for the months of June, July and August, (c) - (d) same variable as for (a) - (b) but when selecting days corresponding to the 5% most extreme values of the double jet index. There is a clear sign of a second separate jet appearing at high latitudes.

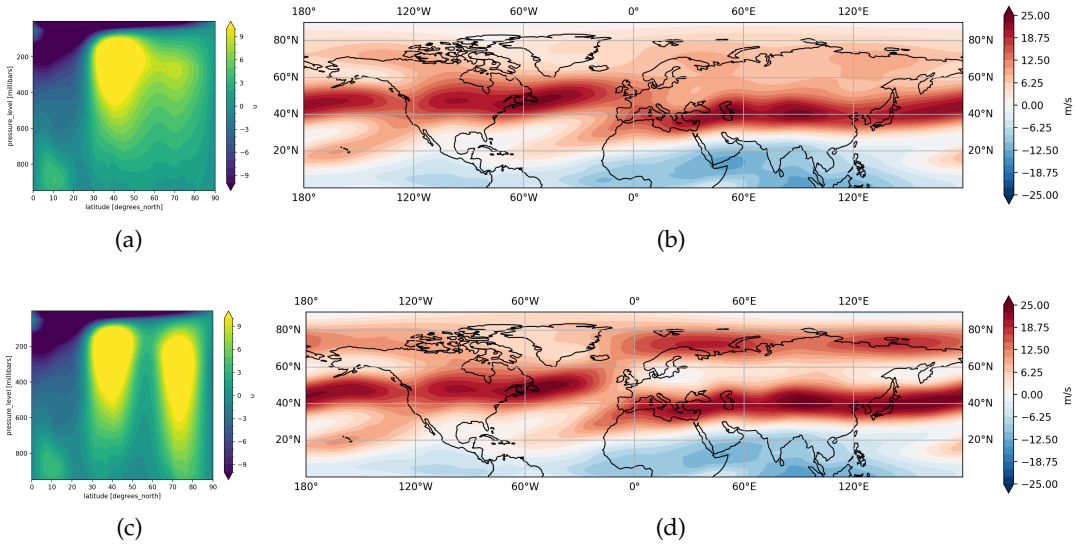


Figure 23.3: (a) Vertical profile of the daily summer zonal wind U for the reanalysis dataset ERA5, longitudinally averaged over the Eurasian sector (10°W to 180°E). We considered the months of June, July and August - (b) Zonal wind U climatology for the reanalysis dataset ERA5 for the months of June, July and August, (c) - (d) same variable as for (a) - (b) but when selecting days corresponding to the 5% most extreme values of the daily double jet index. There is a clear sign of a second separate jet appearing at high latitudes.

ably in three regions: North Canada, Scandinavia and East Russia. The maps are shown in fig. 25.3 for both CESM and ERA5. Thus, we decided to focus on those regions (the exact regions are shown in section 25.6) to assess the link between double jet and heatwaves. According to the definitions presented in sections 22.2.1 and 22.2.2, we computed the double jet and heatwave indexes for different duration, from 1 day to a season (90 days) and we defined a double jet or a heatwave event based on either the 90^{th} or 95^{th} quantile of the distribution of both indexes i.e. we consider as double jet or heatwaves the 10 % or 5% most extreme events in our dataset. We decided to use two thresholds for defining the events for flexibility. The results are shown in fig. 23.4. The top row is for CESM (1000-year control run), the bottom for ERA5. Each column shows the percentage of days in common in (from left to right) respectively North Canada, Scandinavia and East Russia. These plots show a significant number of overlapping days in all the regions, with an increase % of overlapping days for long-lasting events. However, persistent double jet conditions are extremely rare. We propose to overcome this limitation with the aid of a rare events algorithm, which correctly simulate the dynamics of the climate model, but it is tuned in a way to favour dynamical trajectories which most likely lead to rare events. The one used here is a genealogical algorithm, sometimes referred in the literature as Monte Carlo diffusion algorithm (Del Moral, 2004; Del Moral and Garnier, 2005; Giardinà, Kurchan, and Peliti, 2006), presented in section 22.2.3. This particular algorithm works for event which are persistent in space and time, as discussed by (Ragone and Bouchet, 2020), thus it has the potential of

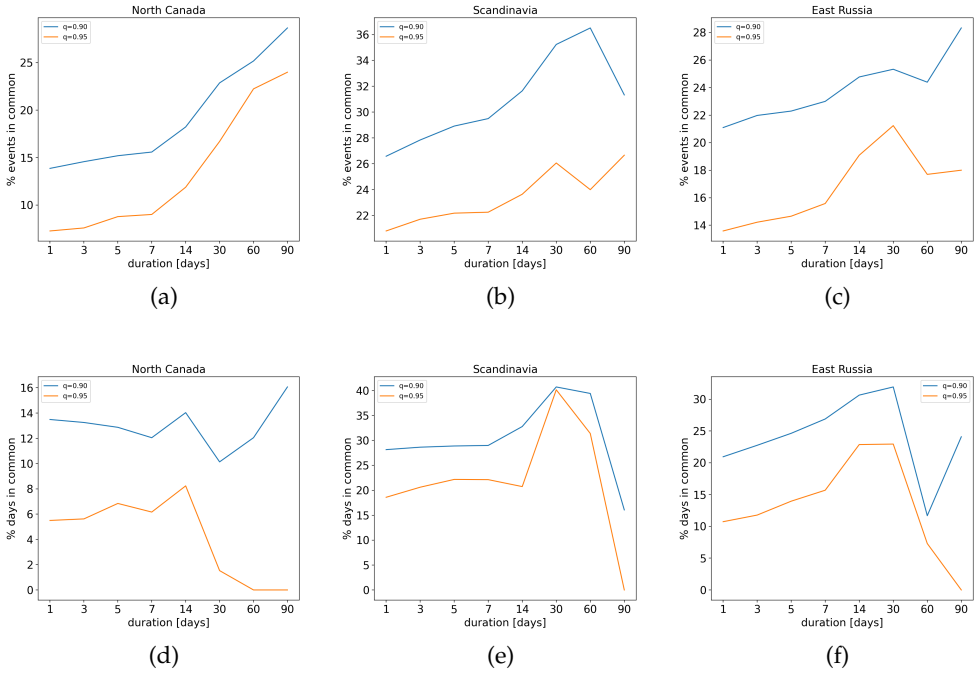


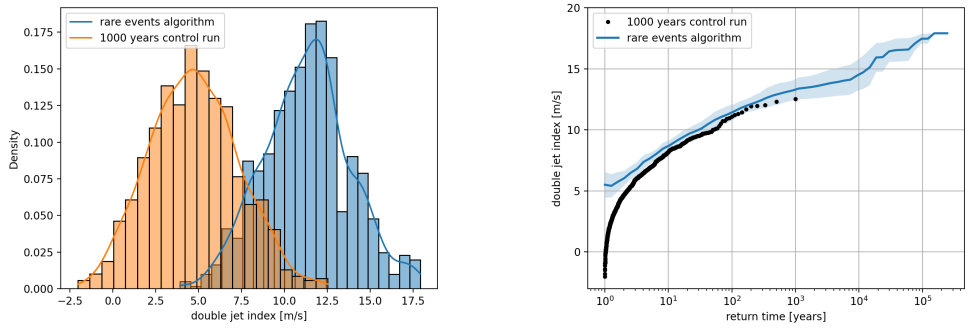
Figure 23.4: Percentage of days in common between double jet (defined in section 22.2.1) and heatwave events (defined in section 22.2.2) as a function of the duration in days, for different regions: (a) North Canada, (b) Scandinavia and (c) East Russia. The spatial areas are shown in section 25.6. Both double jet and heatwaves are defined as events based on either the 90^{th} ($q = 0.90$) or 95^{th} ($q = 0.95$) quantile of the distribution of both indexes, i.e. we consider as double jet or heatwaves the 10 % or 5% most extreme events in our dataset. (d) to (f) same as for (a) to (c) but using the reanalysis dataset ERA5.

capturing the phenomenon we wish to study. For our study we apply it to sample double jet summers (months of June, July and August) over the Eurasia continent. In the next section, we present the results.

23.3 IMPORTANCE SAMPLING OF EXTREME DOUBLE JET SUMMERS

The main goal of the rare event algorithm is to perform an importance sampling of the double jet distribution over the target region, Eurasia. We use the climate model CESM, introduced in section 22.1.1. Figure 23.5a shows the histograms and the kernel density estimation of the seasonal double jet for the control run (orange), the 1000-year long available simulation, and the rare events algorithm (blue). The algorithm effectively performs an importance sampling of the distribution, by sampling more efficiently the upper tail and even events which were not present in the control distribution. The typical value for the double jet index in the resampled distribution is 11.3 m/s, which is extremely rare in the control distribution.

In fig. 23.5b, we plot the return time curve for both the control and the rare events algorithm summers. A detailed description of how these curves are ob-



(a) (b)

Figure 23.5: (a) Histograms and kernel density estimation of the seasonal JJA double jet index in the 1000 years control run of CESM (orange) and with the rare events algorithm (blue) with $k = 0.01(m/s)^{-1}(day)^{-1}$. This histogram shows the power of the rare event algorithm. Indeed, an event in the tail of the orange distribution becomes a typical event in the distribution obtained with the rare event algorithm. We arrive to sample statistically significant events which are not event present in the control run. (b) Return time curves for the control run of CESM (1000 years) in black and the one obtained with 10 runs of the rare event algorithm in blue. The dark blue line represents the ensemble mean, while the shadow blue region corresponds to one standard deviation.

tained can be found in section 25.4. The black curve is evaluated using 1000 years of data and indeed the rarest event has a return time of 10^3 years. The rare events algorithm was tuned such that the typical value in the resampled distribution should correspond to an event with a return time of 100 years in the control run. From the return time curve, it is visible that 11.3 m/s, the typical value of the double jet in the resampled distribution corresponds to a return time of 100 years in the control run. In the same figure, for each of the 10 runs of the rare events algorithm, we compute a return time curve using a similar procedure than for the control simulation. The dark blue shows the mean of the return time curves over 10 runs, while the shading area denotes one standard deviation. Thanks to the rare events algorithm, we sample events which occur one in $10^4 - 10^5$ years, with only 10^3 years of data. In a direct simulation, this is simply impossible.

In the next section, we will investigate some dynamical quantities of interest in both the control simulation of 1000 years of CESM and in the rare events algorithm simulations.

23.4 TELECONNECTION PATTERNS FOR DOUBLE JET SUMMERS

In this section, we analyse the dynamical fields of double jet summers with a return time larger than 100 years (denotes as 100-year double jet summer or events from now on). In fig. 23.6 we show the composite maps for the zonal wind U, the 2 m temperature anomaly (T_{2m}) and 500 hPa geopotential height anomaly (Z_{500}) for values of the double jet index (averaged over the months

of June, July and August) larger than its 100-year double jet summer value, for the control run of 1000 years (left column) and the rare events algorithm (right column). Starting from the top row, we see the appearance of a poleward second filament over Eurasia, which is our target region. Associated with this configuration, we observe the emergence of three zones with positives T_{2m} and a wave number 3 in Z_{500}. There are several interesting conclusions that we want to convey.

Firstly, an important point regards the significance of the presented maps. The first column of fig. 23.6 is obtained via an empirical average of 10 maps solely (given that we have only 1000 years of data available). A Student t-test has been applied and the corresponding t-value maps are shown in fig. 23.7. Details of this statistical test are given in section 25.5. The zonal wind U has the largest amount of significant zones, which is not surprising given that is it strictly connected to the double jet definition. For the other two fields, we can already observe some significant zones in North Canada, East and North Europe and over East Russia, and over the Polar Circle for Z_{500}. Given the higher amount of sampled events in the rare event algorithm simulations, the significant zones were enlarged. This is visible in fig. 23.7 right column, where we plot the t-value maps obtained for the rare events algorithm. The algorithm results are globally significant, despite in some areas, such as in the United States of America. The conclusion in that the algorithm enables to evaluate the existence of significant teleconnections associated with a double jet summer, particularly in North Canada, Europe and East Russia.

The 100-year double jet summer composite maps reveal the presence of a wave 3 pattern for Z_{500} and a correspondent teleconnection pattern in T_{2m} (fig. 23.6). While this pattern was already present for the T_{2m} in the control, with limited significant zones, for the Z_{500} the wave number was not clear in the control simulation, especially over the Atlantic Ocean. This feature becomes clearly visible in the composite maps obtained with the rare events algorithm (right column). This teleconnection pattern is similar to the one shown in (Ragone and Bouchet, 2020), where the authors used the same climate model and rare event algorithm presented here for the sampling of warm summers in both France and Scandinavia. The patterns are significantly similar for Scandinavian heatwaves and a careful comparison with the processes discussed in (Ragone and Bouchet, 2020) might be an interesting future pathway. We also mention that another study from the same authors (Ragone, Wouters, and Bouchet, 2018) which used the same rare events algorithm but a different climate model and area, produced similar results in terms of teleconnection patterns. These results strongly support the idea that teleconnection at subseasonal time scales and the associated large scale dynamics corresponding to a wave number 3 are robust features.

The rare events algorithm allows to sample events which are unseen in the control run. In figure fig. 23.8 we plot the dynamical fields of double jet summers with a return time larger than 1000 years (denotes as 1000-year double jet summer or events from now on). It is interesting to notice that the patterns of the composite maps of 1000-year double jet summer are similar to the ones of the 100-year double jet summer, for all the fields, with an obvious increasing in the amplitude. This is consistent with previous results of the application

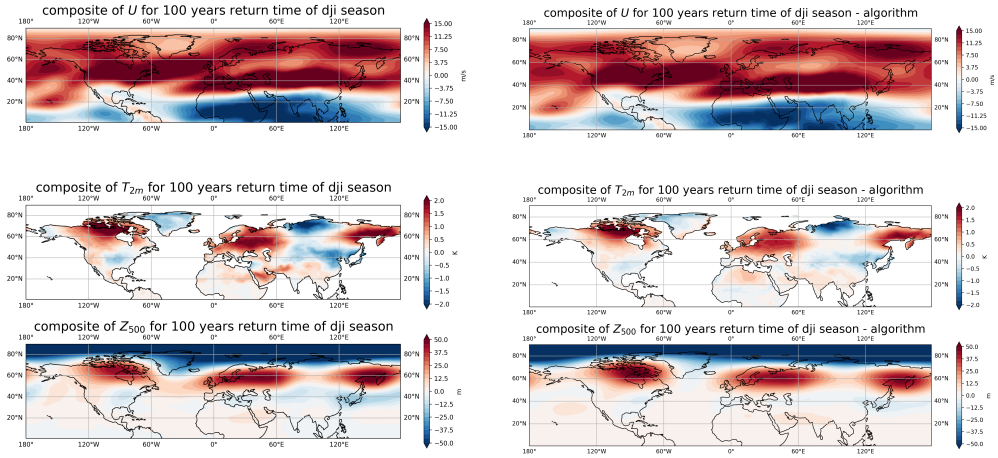


Figure 23.6: Composite maps for U, T_{2m} and Z_{500} for 100 years return time of double jet index averaged over the summer months of June, July and August (dji season in the titles) for (left) CESM control run (average over 10 maps) and (right) the rare events algorithm.

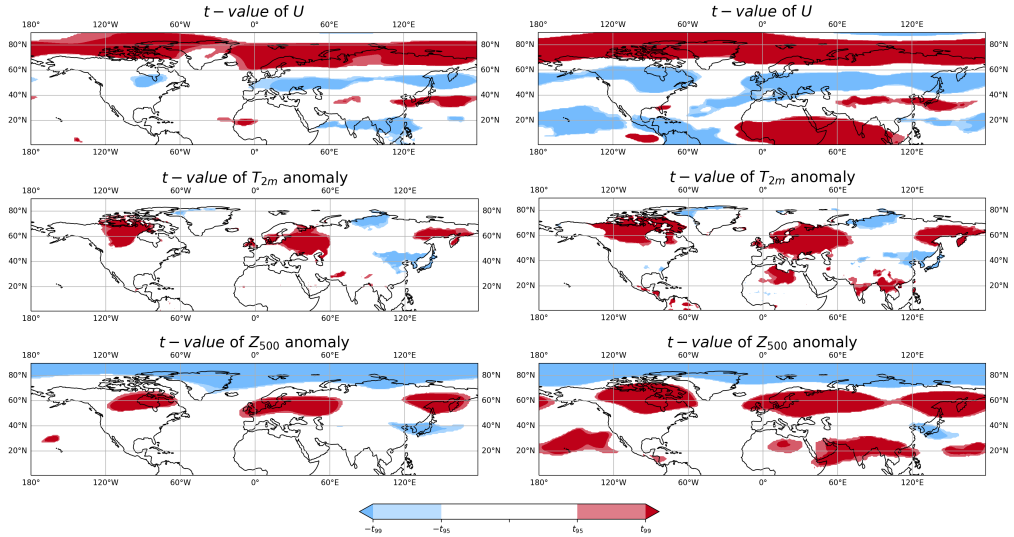


Figure 23.7: Significance maps for fig. 23.6

of this class of algorithms to other climate phenomena (Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2021; Cozian, 2023; Wouters, Schiemann, and Shaffrey, 2023; Cini et al., 2024; Sauer et al., 2024; Lestang, Bouchet, and Lévêque, 2020; Le Priol, Monteiro, and Bouchet, 2024).

A final analysis which is currently missing is to make the link with the extreme events, heatwaves in particular, more robust thanks to the enriched statistics gathered with the rare events simulations.

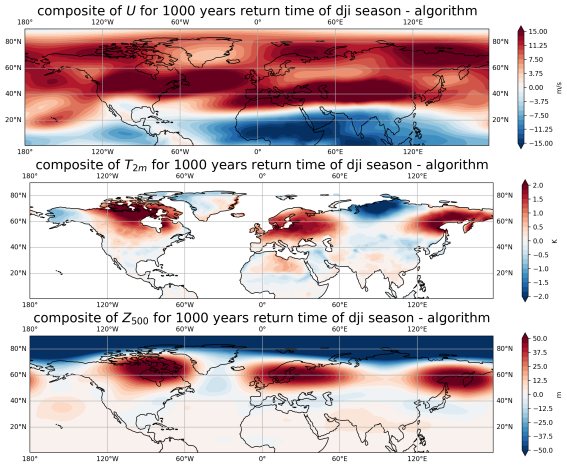


Figure 23.8: Composite maps for U, T_{2m} and Z_{500} for 1000 years return time of double jet index averaged over the summer months of June, July and August (dji season in the titles) for the rare events algorithm.

24

CONCLUSIONS

In this study we apply a rare events algorithm, the GKLT, to simulate double jet stream summers over Eurasia during the months of June, July and August. We decided to focus on the Northern Hemispheric summer season because of the link between double jet events and heatwaves (Rousi et al., 2022).

The GKLT is an algorithm which has been originally designed in the context of Large deviations theory for computing the rate function, which ultimately gives an estimate of the asymptotic probability of observing a rare fluctuation of a given observable. In the context of this study, we couple this algorithm with a climate model to simulate the dynamics of the climate in a condition to favour large fluctuations of a given one dimensional index representative of the double jet. In this study the chosen index is a double jet index for which high values of it translates into a marked splitting of the jet stream over the Eurasian sector. Despite other indexes and methodologies have been used to define a double jet (Rousi et al., 2022; Molnos et al., 2017; Gallego et al., 2005; Pena-Ortiz et al., 2013), in this study we needed a one dimensional object which could retain as much as possible of the double jet structure. This is the reason why we chose to adapt an existing index (Yang and Chang, 2006) in the context of the North Hemisphere and in particular over Eurasia.

With the rare events algorithm GKLT we effectively simulate more extreme double jet seasons, making the extreme observed ones in the control run much more likely. Moreover, we observe events which are unseen in the control simulation. This can be easily seen from the estimated distributions of the double jet index in the control and with the rare events algorithm simulations.

The use of rare events algorithm such as the one employed in this study allows to obtain realizations of the dynamics which led to the extreme event. For this reason, we can empirically compute composite maps of several dynamical quantities related to the double jet summers obtaining wider significant areas than with direct sampling (for which the significant areas are likely limited to the target region). An analysis of composite maps of 2 m temperature anomaly field and of the 500 hPa geopotential height anomaly, conditioned over the occurrence of a 1-in-100 years double jet summer, reveals a significant wave number 3 pattern, with positive temperature and anticyclonic anomalies over North Canada, North Europe/Western Russia and Eastern Russia. Similar teleconnection patterns and with the same wave number were found by (Ragone and Bouchet, 2021) using the same climate model and rare events algorithm as used in this study for simulating warm summer over the Scandinavia peninsula. An interesting future perspective is to compare both studies to assess more quantitatively the role of double jet in Northern European heatwaves. In the literature, double jet have been analysed in correspondence with waves with higher wave numbers (typically between 5 and 7) (Kornhuber et al., 2019; Coumou et al., 2014; Petoukhov et al., 2016). A natural follow-up is then to investigate the appearance of this wave number in the context of double jet.

The rare events algorithm simulations give access to event which are unseen, for instance those occurring with a return time higher than 1000 years. The interesting conclusion is that the patterns of those composite strongly resemble the ones with a return time higher than 100 years. A methodology has been recently developed to motivate this striking feature. In (Mascolo et al., 2024) the authors devised a framework which correctly captures the scaling of the composite maps with the threshold level used to define the extreme events. With simple but meaningful assumptions between the weather fields which characterize the event and the metrics used to define the event, the authors found an analytical expression for this scaling. In that work, the framework was applied to the analysis and forecasting of heatwaves over France. An interesting future perspective is to apply it to the double jet index presented here. The final missing link in the work is to analyse the data obtained with the rare events algorithm for the occurrence of heatwaves. Preliminary analysis of three Northern hemispheric regions shows that there is indeed an overlapping number of days between double jet days and heatwaves, and that this number increases the more the double jet and the heatwaves are persistent. The natural analysis to conclude this study would be interesting to investigate changes in heatwaves with the enriched statistic gathered with the rare events simulations.

25

SUPPLEMENTARY MATERIAL

25.1 DETRENDING OF ERA5 DATASET

In this manuscript, we present an application of the double jet index to the ERA5 reanalysis dataset (Hersbach et al., 2020). The daily data are publicly available at ECMWF website (https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5). Given that we are interested solely in the summer period, we only downloaded the daily values of 2 m air temperature (T_{2m}), 500 hPa geopotential height (Z_{500}), zonal wind U (averaged between 200 hPa and 350 hPa) for the months of June, July and August for the Northern Hemisphere. The first step for using the ERA5 dataset consisted to a detrending process, as we want to study the response of climate in a stationary condition. We only detrend 2 m air temperature (T_{2m}) and 500 hPa geopotential height (Z_{500}). Because we noticed a latitudinal dependency of the trend for both variables, we performed a latitudinal quadratic detrend of the seasonal zonal averages of 2 m air temperature over land only (T_{2m}) and 500 hPa geopotential height (Z_{500}). The contour plots of both trends are shown in fig. 25.1 and fig. 25.2. The latitudinal dependency is stronger for the 500 hPa geopotential height (Z_{500}) and is present at the beginning of the dataset, suggesting a potential bias in the quality of the data before the satellite era.

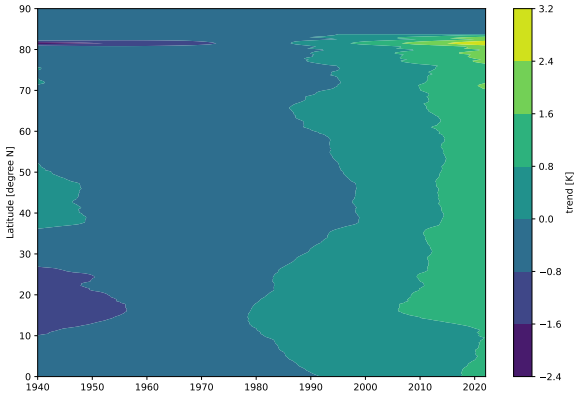


Figure 25.1: Contour plot of the 2 m air temperature trend for ERA5 dataset as function of years and latitude. The oceans are masked.

25.2 TELECONNECTION PATTERNS ASSOCIATED WITH DAILY DOUBLE JET

The composite maps of T_{2m} and Z_{500} conditioned on values of the daily double jet index corresponding to the 5% most extreme values in the distribution (see fig. 23.1) are shown in fig. 25.3. Connected with the double jet we observe the emergence of temperature anomalies and associated anticyclonic anomalies in three preferred zones: North Canada, Scandinavia and East Russia. Those

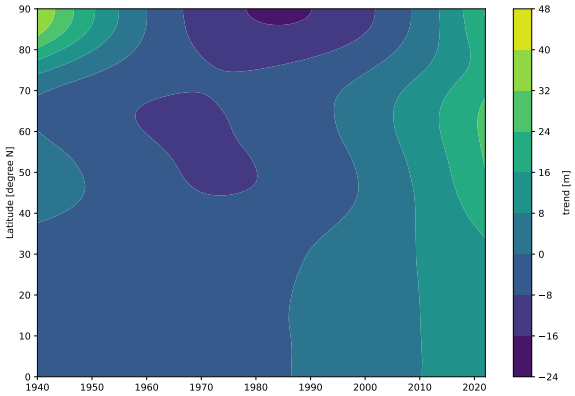


Figure 25.2: Contour plot of the 500 hPa geopotential height trend for ERA5 dataset as function of years and latitude. At high latitudes, the trend is non-monotonic, while it is monotonically increasing in time at lower latitudes.

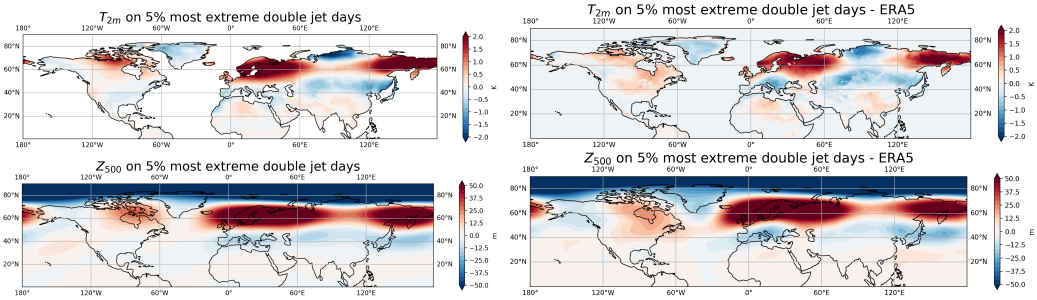


Figure 25.3: Composite maps for T_{2m} and Z_{500} for 5% most extreme double jet days for (left) CESM control run (1000 years) and (right) ERA5.

regions, highlighted in section 25.6 are chosen to assess the link between double jet and heatwaves.

25.3 AUTOCORRELATION FUNCTION

The autocorrelation function of the double jet index is shown in fig. 25.4.

25.4 RETURN TIME CURVES

In this section we detail the computation of the return times presented in fig. 23.5b. We first explain how we compute the return times for the 1000 years long control simulation and then how we obtain them with the rare event algorithm. This section is inspired by the works of (Lestang et al., 2018; Ragone, Wouters, and Bouchet, 2018).

Given a stochastic process $\{X(t)\}$, an observable which depends of the path $\{O[X(t)]\}$ (indicated for simplicity as $\{O(t)\}$ from now on) and a threshold value a which separates between rare and not rare events, we can define the random variable $\tau(a,t) = \min\{\tau \geq t | O(\tau) > a\}$. Then the return time is de-

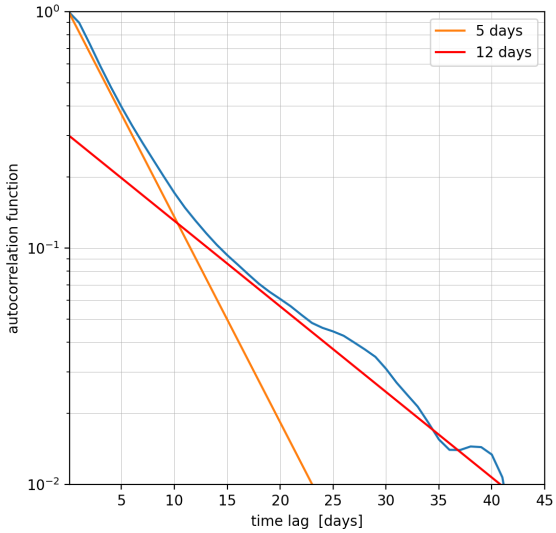


Figure 25.4: Autocorrelation function of the double jet index. The orange and red lines show exponential decays on time scales of 5 and 12 days respectively.

fined as the average time to wait to see an event of magnitude higher than a:

$$r(a) = \mathbb{E}\left[\tau(a,t)\right] \tag{25.1}$$

We can estimate the return time $r(a)$ thanks to the realization of the stochastic process, of simulation length T_d. This means that we have access to a finite time realization of the process and of the path-dependent observable, which we denote as $\{O(t)\}_{0\leq t\leq T_d}$. When we want to study high fluctuations of the stochastic process (or of any quantity which depends on it $O[X(t)]$), namely when a is high, we are interested in time scales which are higher than the typical correlation time τ_c of the process of interest, i.e. $r(a) \gg \tau_c$. Thus, the return time coincides with the time to wait, on average, between two statistically independent events both exceeding the value a. In (Lestang et al., 2018) the authors devised a methodology to correctly sample rare events based on the context presented before. Let's divide the time series of $\{O(t)\}_{0\leq t\leq T_d}$ in M blocks of duration $\Delta T_d \gg \tau_c$, such that $T_d = M\Delta T_d$. For each block, let's define the block maximum:

$$a_m = \max\{O(t)|(m-1)\Delta T_d \leq t \leq m\Delta T_d\} \tag{25.2}$$

and

$$s_m = \begin{cases} 1 & \text{if } a_m \geq a \\ 0 & \text{otherwise} \end{cases} \tag{25.3}$$

for $1 \leq m \leq M$. The variable s_m counts how many rare independent events are observed, i.e. $N(t) = \sum_m s(a)_{\lfloor t/\Delta T_d \rfloor}$, which are well approximate by a Poisson distribution when $r(a) \gg \tau_c$. Then, the probability of that a_m is larger than a

can be estimated as an empirical average of the s_m over the blocks, which gives access to the return time:

$$\hat{r}(a) = -\frac{\Delta T_d}{\ln\left(1 - \frac{1}{M}\sum_{m=1}^{M} s_m(a)\right)} \tag{25.4}$$

In practice, we sort the sequence $\{a_m\}_{1\leq m\leq M}$ in decreasing order $\{\hat{a}_m\}_{1\leq m\leq M}$, such that $\hat{a}_1 \geq \hat{a}_2 \geq \cdots \geq \hat{a}_M$. Using eq. (25.4) we then associate at each threshold $\{\hat{a}_m\}$ its respective return time $r(\{\hat{a}_m\}) = \frac{\Delta T_d}{\ln\left(1-\frac{m}{M}\right)}$. Finally, we can plot the couple $(r(\{\hat{a}_m\}), \hat{a}_m)$ for $1 \leq m \leq M$ as in fig. 23.5b (black curve).

For computing the return time curves for a rare events algorithm, we proceed in a very similar way. The rare event algorithm presented in this manuscript, allow the sampling of rare events from an ensemble of M trajectories, denoted as $\{X_m(t)\}_{0\leq t\leq T}$, with $1 \leq m \leq M$. For each of these trajectories, we will compute $a_m = \max_{0\leq t\leq T} O[X_m(t)]$. Thus, in the particular application of return time estimation for rare events algorithm simulations, each trajectory of the algorithm plays the role of a block in the previous case. However, differently from the previous case, each maxima (trajectory) carries a weight as well. Hence, instead of the sequences $\{a_m\}$, we now have $\{a_m, p_m\}$ for $1 \leq m \leq M$. The generalization of eq. (25.4) in the case of non-equiprobable blocks is:

$$\hat{r}(a) = -\frac{T}{\ln\left(1 - \frac{1}{M}\sum_{m=1}^{M} p_m s_m(a)\right)}. \tag{25.5}$$

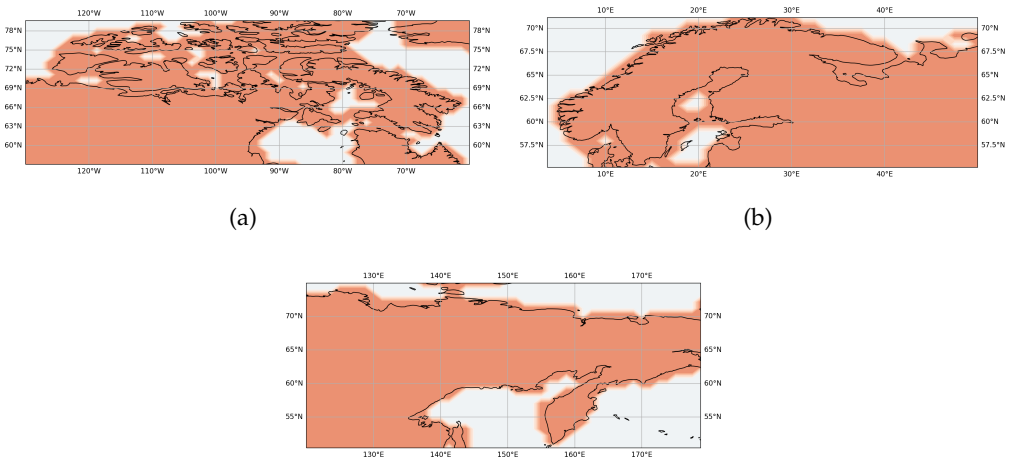
In practice, to plot the return time curve, we sort in decreasing order the sequence $\{\hat{a}_m\}$ to obtain $\{\hat{a}_m, \hat{p}_m\}$ for $1 \leq m \leq M$. We then associate for each couple $\{\hat{a}_m, \hat{p}_m\}$ its respective return time $r(\{\hat{a}_m\}) = \frac{T}{\ln(1-\sum_{l=1}^{m}\hat{p}_l)}$, with $\sum_{l=1}^{m}\hat{p}_l$ being the sum of the weights for events which have an amplitude greater than $\{\hat{a}_m\}$. This is the methodology used for retrieving the blue curve in fig. 23.5b. Note that, to have this curve, we also perform a second average between the 10 run of the rare events algorithm. The shadow corresponds to a standard deviation among the runs.

25.5 SIGNIFICANCE TEST

This section describes the statistical test used for assessing the significance of the composite maps of the zonal wind U, 2 m air temperature (T_{2m}) and 500 hPa geopotential height (Z_{500}). We performed a Student t-test (Student, 1908), to test if the composite map equals the unconditioned mean. We compute the t value:

$$t = \sqrt{N}\frac{\mathbb{E}[X_T|D_{ji,T} > h] - \mu}{S} \tag{25.6}$$

where N are the independent samples, $\mathbb{E}[X_T|D_{ji,T} > h]$ is the empirical average estimator of the composite maps , X_T is the average over a season of each of the fields, $D_{ji,T}$ is the seasonal double jet, h is the threshold, μ is the unconditional mean, S^2 is the sample variance. We compare the t value with the Student t distribution value t_q^N with $N-1$ degrees of freedom at level q. Thus, for a



(a) (b)

(c)

Figure 25.5: Regions in the Northern Hemisphere over which we computed the heatwaves: (a) North Canada, (b) Scandinavia and (c) East Russia.

given value q, then we know that if $|t| \geq t_q^N$ then we reject the null hypothesis with probability q, i.e. the grid-point of the composite map is significant at a level q. In our case $N = 10$. For our study, we want to assess significant areas in the composite maps, thus this test is applied grid point-wise. For the composite maps obtained with the rare events algorithm, seasonal events might still not be independent due to possible common ancestor trajectories. We considered one independent sample for each of the rare event algorithm runs, i.e. 10 samples.

25.6 REGIONS FOR HEATWAVE DEFINITION

The regions are shown in section 25.6.

Part V

EDIPI ART HAZARDS - EARTH - PROJECT

The idea behind the EArtH project was born during a cold and wet Belgian Saturday night while I was watching a documentary of my favourite art current, the Impressionism. The art critic was in the North of France, in Étretat, known for its breathtaking cliffs, in the footsteps of the father of the Impressionists, Claude Monet. The art critic was describing one of its known painted views of the cliffs and the path Monet did to arrive to a remote place to get that viewpoint. Spoiler: not such an easy one.

Suddenly a question just rose in my mind: in a warming world where the oceans are rising, could Monet have ever paint it ? Probably not. Realizing this has made me exceptionally sad, because we are preventing next generations of the purely sense of happiness I felt by watching at this and at many other paintings, and many more emotions. And here comes my idea: can we show climate change effect on art paintings ? Can we take Monet's painting and submerge with water ? Or can we take any other painting and show that other extremes, such has heat, flood, food crops, etc. will prevent all of this beauty to be shown to the world ? From this vague idea, together with the EDIPI team I am part of [1], we built this project I am profoundly proud of. I am and I will ever be sincerely grateful to the EDIPI members who made this true. The project is publicly available at https://edipi-itn.eu/resources/outreach-material/projectearth/.

Here I present the first thematic repainting we posted some months ago. Hopefully the first of a long series. I decided to include this project in this manuscript because it is part of the outreach material published under the EDIPI consortium. Here is the logo of EArtH:

Figure 25.6: EArtH logo designed by me.

1 https://edipi-itn.eu/

26

SNOW

Welcome to the first post of the EArtH team - the EDIPI Art Hazard project, where we bring together **art and climate** to show the effects of climate change on paintings according to scientific evidence taken from the IPCC reports.

Our idea is to modify some paintings using our creativity and scientific facts and imagine them to be painted in 2100 under some emissions scenarios [1].

For every post we will choose some paintings according to the weather event we wish to shade a light on. As the winter season is coming to its end, we thought that it would be logical to start with a fairly wintery and magical one: *snow*!

Figure 26.1: La route de Vetheuil, effet de neige, 1879, Claude Monet.

Winter of 1879 went down in history as one of the coldest in Europe, with Siberian temperatures recorded in Western Europe. It was during these frigid times that Claude Monet, vastly considered as the father of the Impressionism art movement, painted 'The road to Vétheuil, snow effect' . We are in Vétheuil, 70 km northwest from Paris, in the countryside, after a heavy snowfall. The sky is still gray and gives a sense of full winter. To convey that, Monet used a symphony of colors from white to purple, to light brown. *The result is an almost full white landscape, which reflects an immense sense of calm and freezing atmosphere.*

Figure 26.2: Winter landscape in Switzerland near Engadin, 1920, Peder Mørk Mønsted.

Another different sensation is given from the painting 'Winter landscape in Switzerland near Engadin, 1920' by Peder Mørk Mønsted, one of the most renowned Danish painters. *A warm sunlight ray enters the painting from the right side, illuminating a vast part of the Alps landscape, giving us the hope that spring is possibly coming back.*

1 Emission scenarios serve as the basis for working out the possible climate conditions of the future. The scenarios represent possible pathways that society might take regarding the emission and concentration of greenhouse gases, aerosols and chemically active gases as well as emissions from land use/land cover.

The realism which characterizes the vast majority of Mønsted mature works is reflected here in the reproduction of the landscape as if it was a photograph. Indeed, we can still see the influence of Monet works on him, for instance in the choice of the subjects. *Thus, without Monet works, Mønsted could not have been fully the great painter he was and still is.*

Imagine if we, along with Monet and Mønsted, could time travel to 2100 : **could Monet have painted 'The road to Vétheuil, snow effect' under the effect of climate change exactly the same way he did in 1879? Would Peder Mørk Mønsted be able to recreate the same kind of photograph-like painting of a winter mountain landscape if Monet could not have influenced him?**

Unfortunately, the answer is not obvious: it depends on which scenario we, humans, choose to pursue when it comes to our greenhouse gas emissions [2]

Current climate is already different from the one Monet and Mønsted experienced when creating their masterworks. In their times, rapid industrialisation responsible for emitting greenhouse gasses had just begun, and we have greatly changed our climate since then. The IPCC reports represent a comprehensive summary of what the overwhelming majority of the scientific community agrees on topics associated with climate change. In the IPCC's most recent works [3], scientists confirmed that 'global warming has already led to mass loss from ice sheets and glaciers and reductions in snow cover'. In regions such as the European Alps, the snow depth is projected to decrease by around 25% between recent decades and the near future. This corresponds to a continuation of the ongoing decrease in annual snow cover duration (on average 5 days per decade). For small glaciers such as in the European Alps, Pyrenees and Scandinavia, reductions in ice mass of more than 80% by the end of the century are expected under a "worst case scenario" (RCP8.5)[4] compared with 30% under the most optimistic scenario (RCP2.6)[5]. These numbers are best estimates

2 A greenhouse gas (GHG) is a gas that causes the atmosphere to warm by absorbing and emitting radiant energy. Greenhouse gases absorb radiation that is radiated by Earth, preventing this heat from escaping to space. Carbon dioxide (CO_2) is the most well-known greenhouse gas, but there are others including methane, nitrous oxide, as well as water vapor among others. Human-made emissions of greenhouse gases from burning fossil fuels as well as industry, and agriculture are the leading cause of global climate change. Before the Industrial Revolution started in the mid-1700s the atmosphere contained approx. 280 particles of carbon dioxide per 1 million particles (280 ppm CO_2). Human activities led to an increase up to 417 ppm CO2 in 2022.

3 https://www.ipcc.ch/srocc/

4 The Representative Concentration Pathways (RCP) define trajectories representing greenhouse gas, aerosol and chemical active gas concentrations as well as emissions from land use/land cover for particular climate projections (namely, radiative forcing values in the year 2100). One emission scenario thus works on the basis of the global community agreeing to drastically reduce emissions of greenhouse gases. RCP8.5 represents a radiative forcing of 8.5 Watts per square metre by the end of the century. In between these two are other scenarios that entail greater or lesser degrees of technical advancement that would lead to slight reductions in greenhouse gases. The higher the radiative forcing value, the greater the change in climate conditions.

5 The Representative Concentration Pathways (RCP) define trajectories representing greenhouse gas, aerosol and chemical active gas concentrations as well as emissions from land use/land cover for particular climate projections (namely, radiative forcing values in the year 2100). One emission scenario thus works on the basis of the global community agreeing to drastically reduce emissions of greenhouse gases: RCP2.6 scenario assumes an additional radiative of 2.6 Watts per

contained in the findings of the IPCC, and are inherently associated with varying degrees of uncertainty [6]. For more information on the uncertainty with the numbers in this article we refer you to the IPCC summary for policy makers [7].

Imagine that the end of the century has just arrived : we are in 2100 and we have done nothing to curb our emissions. Our imagination combined with our scientific knowledge suggested the following re-paintings.

For Monet's painting we decided to use recycled pieces of clothing to remove some of the snow in order to stress on how human actions have an effect on our surroundings. Using discarded clothes also reminds us of the role of garbage and discarded objects in our lives as consumers:

Figure 26.3: Original and Repainted version of 'La route de Vetheuil, effet de neige', 1879, Claude Monet.

For Mønsted's painting, given the information from the IPCC, we decided to challenge AI and then to post process the image using some graphical tools. We applied that in a corner of Mønsted's painting to give you a feeling of what could be the Alps in the worst case scenario in 2100. Can you imagine the rest of the painting? We challenge you!

square metre by the end of the 21st century. Another scenario represents a situation in which humans continue in the same way as we have to date, emitting ever more greenhouse gases.

6 Projections about possible future changes are always subject to uncertainties. Uncertainties have many types of sources, from data uncertainty to model uncertainty, ambiguously defined concepts or terminology, incomplete understanding of critical processes, or uncertain projections of human behaviour. Scientific literature tries to reduce these uncertainties as best as possible. One example is the usage of simulations by several models to include the impact of model differences.

7 https://www.ipcc.ch/report/ar6/syr/summary-for-policymakers/

Figure 26.4: Original and Repainted version of 'Winter landscape in Switzerland near Engadin', 1920, Peder Mørk Mønsted. Partially done with DallE.

It is important to note that cold and snowy winters like the ones originally painted would still be possible in 2100: they would just be much less common. So, if our time-machine was able to bring Monet and Mønsted back to present times for just one winter, the chances of them seeing the landscapes covered in snow would be much lower than when they have painted them!

However, as we said, the future will depend on the scenario we choose. We would like to dream of a world where another Monet in 2100 could paint another version of 'The road to Vétheuil, snow effect' and that 40 years later, inspired by him/her and other artists, another great future Peder Mørk Mønsted, will impress us with his/her view of 'Winter landscape in Switzerland near Engadin'.

We believe that this world can still be a real possible future one. Join the climate action, do your part!

Dostojevski already said: ' Beauty will save the world'. **Be part of this beauty.**

And do not forget to share and/or comment on this post!

Paint to you soon,
The EArtH team

Part VI

CONCLUSIONS AND PERSPECTIVE

27 CONCLUSIONS AND PERSPECTIVES

27.1 CONCLUSIONS

The aim of this thesis was to study heatwaves with novel theoretical methodologies. Studying this class of extremes is at the same time a fascinating and a challenging field. In the realm of interesting questions associated with the study of heatwaves, this thesis addressed the three following points:

- estimate the probability of heatwaves to occur;
- understand the dynamics associated with them;
- forecast them at intraseasonal time scales.

A bottleneck for answering the above questions is the lack of data, which naturally arises in rare events studies. This was one of the conclusions drawn from part ii. For instance, in that study, we compared the influence of two slow drivers which are known to affect European summer heatwaves: spring soil moisture and the Atlantic Multidecadal Variability AMV. Different studies have proven that both drivers have separately an influence on European heatwaves (Qasmi et al., 2021; Qasmi, Cassou, and Boé, 2017; Fischer et al., 2007b; Fischer et al., 2007a; Alexander, 2011; Materia et al., 2022), but none of them has ever estimated the relative influence of each driver. Using outputs from three different CMIP6 models, we showed that, for two complementary definitions of heatwaves (one measuring the amplitude in terms of degrees and another one the mean number of heatwave days per year) both factors have an influence of the same order of magnitude, but soil moisture is a slow driver of greater importance than the AMV for European summer heatwaves, both in terms of extension of the region of influence and in terms of amplitude. This result is valid both for heatwaves of *return time* of a few years up to heatwaves with a 50-year return time. Studying the effects of both drivers on such rare events was a novelty introduced by our study. Indeed, previous works looked at typical heatwaves, those with a return time of a few years (Qasmi et al., 2021; Qasmi, Cassou, and Boé, 2017). However, we know that the most harmful events are the largest and rarest one (Robine et al., 2008). In our work, we studied the effects of both drivers on much rarer heatwaves with the aid of *return time maps*, maps conditioned on the time needed to observe a heatwave with a certain amplitude. Significance over these maps is already limited when return times of the order of a decade are chosen. However, we proved that slow drivers such as the AMV affects heatwaves in Europe by the same amplitude as the current warming rate of hot extremes in Europe, thus it could either mask or exacerbate this warming trend, with important impacts on society and environment.

In the following two chapters of this manuscript, I presented two different methodologies for tackling the sampling problem connected to extreme and rare events.

In part iii we presented a novel theoretical framework for the analysis and prediction of extreme events. We called this framework the Gaussian approximation, because it relies on the assumption of joint Gaussianity between the dynamical fields which characterize the extreme event under study and the metric used to define the event. With this methodology, we characterized the extreme event of interest using the information coming from the full dataset and not only from the subset of data which define the extreme. We presented an application of this study to heatwaves over France. However, the methodology is fairly general and could be applied to many other extreme events. For the analysis task, we explained theoretically why the composite maps, i.e. maps of averaged dynamical variables conditioned on the outcome of the rare event, associated to very extreme events strikingly resemble those for less extreme ones (fig. 12.1). Indeed, the only visible change regards the amplitude of the maps and not in the structure of the patterns. With the Gaussian approximation, the analytical composite maps are the same up to a rescaling by a non-linear function of the threshold that defines heatwaves and this rescaling predicted by our approximation is the correct one we observe in climate data (fig. 15.2). For the prediction task, with this framework we achieved a prediction skill comparable to that of a Convolutional Neural Network, significantly diminishing the computation time. The methodology naturally computes the committor function, the conditional probability of observing a heatwave (or an extreme event in general) within τ days given the set of dynamical fields (predictors) observed today. This object, despite being the right tool for making prediction, is hard to estimate as it is a function of high dimensional fields. Our framework naturally performs a dimensionality reduction which, once properly regularized, shows interesting and realistic weather patterns. The results suggest quasi-stationary Rossby waves and low soil moisture as precursors to extreme heatwaves over France. Moreover, the Gaussian approximation proved to be the method with the highest predictive skill when applied to the short ERA5 reanalysis dataset, for which the prediction with other methods, such as Convolutional Neural Networks is strongly limited by the available data.

Differently from what has been done in the previous chapter, in part iv we studied the phenomenology of double jet events over Eurasia by means of a rare events algorithm, the Giardina-Kurchan-Lecomte-Tailleur GKLT algorithm (Del Moral and Garnier, 2005; Giardina et al., 2011; Lecomte and Tailleur, 2007): here we overcame the sampling problem by generating more realizations of the event of interest. The double jet refers to a splitting of the jet stream into two filaments, one which follows the natural path of the subtropical jet and the other one which is anomalously displaced poleward. With the aid of a long climate simulation of 1000 years with the CESM climate model, we adapted an already existing definition of the double jet to the target area of study. This was a fundamental step for the study because the chosen rare events algorithm samples large fluctuations of a time-averaged observable, named score function, which in our study was the double jet index defined above. Thus, a proper score function, which shows persistence and good representation of the phenomenon at the right temporal scales, is crucial for a successful implementation of the GKLT algorithm. Moreover, an additional motivation for using this algorithm for this case of study is that the percentage of common days between double

jet events and heatwaves in three different areas of the Northern Hemisphere augment when considering persistent events. Indeed, this algorithm works for event which are persistent in space and time, as discussed by (Ragone and Bouchet, 2020). The obtained results with the algorithm are rather encouraging. Firstly, differently from previous applications of this algorithm, in this study we sampled directly persistent atmospheric states prone to support extreme events. In previous works, for example, the same rare events algorithm was applied for sampling of extreme warm summers and the atmospheric states were retrieved after (Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2021). With the algorithm, we sampled rare and even unseen double jet events, with a return time of $10^4 - 10^5$ using only 10^3 years of simulation. Moreover, an analysis of dynamical fields, conditioned on the occurrence of the double jet, such as the geopotential height at 500 hPa and the 2 m air temperature anomalies showed a significant wave 3 pattern. This teleconnection pattern was found in application of the same rare events algorithm for sampling warm summers in different regions (Ragone, Wouters, and Bouchet, 2018; Ragone and Bouchet, 2020; Ragone and Bouchet, 2021). These results support the idea that teleconnection at subseasonal time scales and the associated large scale dynamics corresponding to a wave number 3 are robust features. The pattern in the 2 m air temperature anomalies, conditioned on the occurrence of the double jet, shows three different zones affected by significant anomalies: North Canada, Scandinavia/West Russia and East Russia. The natural follow-up to finish this study is to investigate the link between double jet and heatwaves further with the rare events simulations obtained.

27.2 PERSPECTIVES

The study on the relative importance of two slow drivers on European heatwaves, the AMV and the spring soil moisture deficit, presented in part ii, might be followed up considering the effects of climate change. Indeed, the data used for that study came from climate models run in stationary conditions. One of the effects of climate change in Europe is the northward shift of climatic zones within the European continent. As a consequence, Central and Eastern Europe becomes a new transitional zone between dry and wet climates (similar to the Mediterranean region in the present climate), becoming susceptible to the effects of land–atmosphere coupling (Seneviratne et al., 2006). In this context, the variability of soil moisture over the zone considered for our study (the Mediterranean basin) might be reduced, leading to a reduced importance in the modulation of heatwaves. This suggests the changing of the considered zone for the analysis in the context of climate change.

A limitation of this study arises from the scarcity of data faced when we aimed at studying really rare events. Specifically, we are already in a lack of data regime for events with return times of a decade, and the analysis would be even more challenging for events with higher return times, for instance a century. The rare events algorithm, the Giardina-Kurchan-Lecomte-Tailleur GKLT, might be used to circumvent this limitation. For instance, coupled climate simulations

imposing different phases of the AMV might be run using as score function the averaged temperature over some parts of Europe.

In part iii we presented a novel framework for the analysis and the prediction of extreme events, focusing on heatwaves over France. A natural follow-up would be to perform the same analysis on the 1000-long run of CESM used in this manuscript. For composite maps, it would be interesting to see how the error pattern (the difference between the estimated composite and the one calculated with the Gaussian approximation) in the 500 hPa geopotential anomaly changes with a model which has a finer atmospheric model than PlaSim. For committor functions, it would be interesting to see how the prediction horizon changes, if it changes, when the delay time and the heatwave duration are changed. The study with the Gaussian approximation is for the moment a methodological study. Thus, for example, an interesting direction would be in the physical understanding of the projection patterns: why do we observe a quasi-stationary pattern, that does not depend much on the lead time?

In part iv we simulated double jet summer seasons over Eurasia with the aid of a rare events algorithm, the GKLT. A follow-up perspective could be to apply the double jet index presented in part iv to the winter season in the Northern Hemisphere. It is widely known that in summer the jet tends to be more zonal than in other periods of the year (Coumou et al., 2018). Applying it to the winter season, when the jet activity is higher, might be an interesting test bench.

A robust feature of the study is the emergence of a wave 3 pattern in response to a double jet. In the literature, this structure of the jet has been associated to higher wave numbers (Rousi et al., 2022; Petoukhov et al., 2016; Coumou et al., 2014; Kornhuber et al., 2019). Further analysis is necessary to understand this discrepancy and how this wave 3 pattern arises.

An interesting future perspective is to link the study of the double jet, presented in part iv, with the methodology presented in the part iii, the Gaussian approximation. The natural starting point is to use the long simulation data of 1000 years with CESM used in this manuscript. Using the Gaussian framework for extreme double jet could also explain why composite maps of extreme events are dynamically indistinguishable from the less extreme ones. This fascinating property has been discussed for heatwave over France in section 15.3. However, the methodology does not rely on the peculiar extreme event and could be applied for double jet events. Another fascinating pathway, within the application of the Gaussian approximation, is the forecast of double jet events. This particular configuration of the jet has been linked to several heatwaves in the Eurasian continent in the latest decades, including the severe ones of 2003 and 2018 (Rousi et al., 2022). In this sense, forecasting this sort of events might be of extreme relevance in the context of risk limiting policies. The Gaussian framework embeds an alternative methodology to more complex weather forecasting models and could be used as a starting point for forecasting.

Another interesting perspective is to study what changes in the double jet state when we consider a non-stationary climate. For instance, it could be interesting to use some simulations in which the global surface temperature is similar to the current values, to simulate the effect of climate change. These simulations have been done by Francesco Ragone using the CESM model in the setup used in this manuscript. I started the post-processing of these simula-

tions, but I could not proceed into the analysis due to time lack. It would be thus interesting to continue on this pathway and to use the double jet index presented in this manuscript to understand the dynamic effects caused by the thermodynamic changes, if any.

An application of GKLT algorithm could be for drought events. Indeed, as for heatwaves and many other extremes, the societal and environmental impacts of these events are higher the longer and the extended they are. Thus, analysing the drought conditions in several parts of the world for entire seasons with the GKLT could be an interesting perspective.

Focusing on the optimization of the rare events algorithm for short events, we could improve the score function used for running it. A preliminary work done at the beginning of my PhD together with Clément Le Priol [1] was to investigate this question in the literature in the context of two classes of algorithms, the Adaptive Multilevel Splitting and the Interacting particles, which embeds the GKLT. Thanks to the work of (Chraibi et al., 2018), we know that there exists an optimal score function for these algorithms. The optimal score function requires the knowledge of the committor function (i.e. the ability to perfectly predict the event) which we do not have and which is in principle the reason why we run these algorithms. In this sense, the Gaussian approximated committor might be an ideal candidate eq. (14.3). This perspective is currently under investigation by Clément Le Priol, using PlaSim climate model, for studying short heatwaves (from 1 to 2 weeks).

Another alternative would be to design a novel algorithm, different from the GKLT, to improve the forecast of extreme events. In our research group, we began a collaborative project to address this task. A novel algorithm has been designed by Tony Lelievre[2] and Julien Reygner[2] and is currently being tested by Alessandro Lovo[3] and Amaury Lancelin [4][5] on toy models. Freddy Bouchet and Corentin Herbert[3] are also involved in direction of the project. My contribution should come at a later stage, in the coupling of this algorithm with a climate model. The goal of the algorithm is to generate new data which are tailored for improving the estimation of the committor. The first part of the work has been centred around the proper theoretical formalism. Practically, we start with an empirical estimate of the committor function given by machine learning (using the same methodology as the one presented for convolutional neural networks in part iii) or by the Gaussian approximation and then we improve its estimate following the theoretical framework developed. Since this model is still under testing, it remains a fully open question to be addressed in the months and probably years to come.

A final long term goal, which is strictly connected to the ones presented above, is to link machine learning and rare event algorithm for prediction pur-

1 *CNRS, LMD/IPSL, ENS, Université PSL, École Polytechnique, Institut Polytechnique de Paris, Sorbonne Université, Paris France*

2 *Centre d'Enseignement et de Recherche en Mathématiques et Calcul Scientifique, École des Ponts ParisTech, 6 et 8, Avenue Blaise Pascal, Cité Descartes—Champs sur Marne, 77455 Marne la Vallée Cedex 2, France*

3 *ENS de Lyon, CNRS, Laboratoire de Physique, F-69342 Lyon, France*

4 *CNRS, LMD/IPSL, ENS, Université PSL, École Polytechnique, Institut Polytechnique de Paris, Sorbonne Université, Paris France*

5 *RTE France*

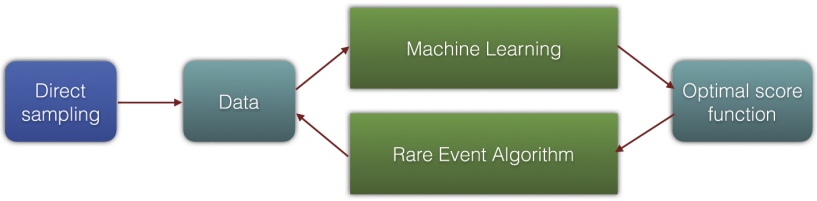


Figure 27.1: Sketch of the coupling process between machine learning and rare event algorithms. Figure taken from (Lucente et al., 2022).

poses. The sketch of the coupling process is shown in fig. 27.1. Starting from the already available data, we could have a primordial estimation of the committor function either via machine learning techniques (such as convolutional neural networks) or via the Gaussian approximation or the new sampling algorithm. Then, with this estimate, we could use rare event algorithms to generate new data with which estimating a new (hopefully better) committor function. This coupling has already been done with a toy model, using Analog Markov Chains to estimate the committor function (Lucente et al., 2022) and seems promising.

BIBLIOGRAPHY

Agana, Norbert A and Abdollah Homaifar (Mar. 2017). "A deep learning based approach for long-term drought prediction." In: *SoutheastCon 2017*. ISSN: 1558-058X, pp. 1–8. DOI: 10.1109/SECON.2017.7925314. URL: https://ieeexplore.ieee.org/document/7925314 (visited on 04/24/2024).

Alexander, Lisa (Jan. 2011). "Extreme heat rooted in dry soils." en. In: *Nature Geoscience* 4.1. Number: 1 Publisher: Nature Publishing Group, pp. 12–13. ISSN: 1752-0908. DOI: 10.1038/ngeo1045. URL: https://www.nature.com/articles/ngeo1045 (visited on 01/30/2024).

Anderson, James B. (Aug. 1975). "A random-walk simulation of the Schrödinger equation: H+3." In: *The Journal of Chemical Physics* 63.4, pp. 1499–1503. ISSN: 0021-9606. DOI: 10.1063/1.431514. URL: https://doi.org/10.1063/1.431514 (visited on 06/27/2024).

Bach, Sebastian, Alexander Binder, Grégoire Montavon, Frederick Klauschen, Klaus-Robert Müller, and Wojciech Samek (July 2015). "On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation." en. In: *PLOS ONE* 10.7. Publisher: Public Library of Science, e0130140. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0130140. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0130140 (visited on 06/17/2024).

Barnes, Elizabeth A and James A Screen (2015). "The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it?" en. In: *WIREs Climate Change* 6.3. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.337, pp. 277–286. ISSN: 1757-7799. DOI: 10.1002/wcc.337. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.337 (visited on 05/03/2024).

Barriopedro, D., R. García-Herrera, C. Ordóñez, D. G. Miralles, and S. Salcedo-Sanz (2023). "Heat Waves: Physical Understanding and Scientific Challenges." en. In: *Reviews of Geophysics* 61.2. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2022RG000780, e2022RG000780. ISSN: 1944-9208. DOI: 10.1029/2022RG000780. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2022RG000780 (visited on 03/25/2024).

Barriopedro, D., R. García-Herrera, and R. M. Trigo (Dec. 2010). "Application of blocking diagnosis methods to General Circulation Models. Part I: a novel detection scheme." en. In: *Climate Dynamics* 35.7, pp. 1373–1391. ISSN: 1432-0894. DOI: 10.1007/s00382-010-0767-5. URL: https://doi.org/10.1007/s00382-010-0767-5 (visited on 05/02/2024).

Barriopedro, David, Erich M. Fischer, Jürg Luterbacher, Ricardo M. Trigo, and Ricardo García-Herrera (Apr. 2011). "The Hot Summer of 2010: Redrawing the Temperature Record Map of Europe." In: *Science* 332.6026. Publisher: American Association for the Advancement of Science, pp. 220–224. DOI: 10.1126/science.1201224. URL: https://www.science.org/doi/10.1126/science.1201224 (visited on 03/06/2024).

Benedetti, Riccardo (Jan. 2010). "Scoring Rules for Forecast Verification." en. In: *Monthly Weather Review* 138.1, pp. 203–211. ISSN: 1520-0493, 0027-0644. DOI:

10.1175/2009MWR2945.1. URL: http://journals.ametsoc.org/doi/10.1175/2009MWR2945.1 (visited on 06/17/2024).

Benson, David O. and Paul A. Dirmeyer (Mar. 2021). "Characterizing the Relationship between Temperature and Soil Moisture Extremes and Their Role in the Exacerbation of Heat Waves over the Contiguous United States." In: *Journal of Climate* 34.6, pp. 2175–2187. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-20-0440.1. URL: https://journals.ametsoc.org/view/journals/clim/34/6/JCLI-D-20-0440.1.xml (visited on 06/17/2024).

Beobide-Arsuaga, Goratz, André Düsterhus, Wolfgang A. Müller, Elizabeth A. Barnes, and Johanna Baehr (2023). "Spring Regional Sea Surface Temperatures as a Precursor of European Summer Heatwaves." en. In: *Geophysical Research Letters* 50.2. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2022GL1007 e2022GL100727. ISSN: 1944-8007. DOI: 10.1029/2022GL100727. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2022GL100727 (visited on 07/11/2024).

Bochenek, Bogdan and Zbigniew Ustrnul (Feb. 2022). "Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives." en. In: *Atmosphere* 13.2. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, p. 180. ISSN: 2073-4433. DOI: 10.3390/atmos13020180. URL: https://www.mdpi.com/2073-4433/13/2/180 (visited on 07/12/2024).

Boer, George J. et al. (Oct. 2016). "The Decadal Climate Prediction Project (DCPP) contribution to CMIP6." en. In: *Geoscientific Model Development* 9.10, pp. 3751–3777. ISSN: 1991-9603. DOI: 10.5194/gmd-9-3751-2016. URL: https://gmd.copernicus.org/articles/9/3751/2016/ (visited on 04/05/2022).

Bolhuis, Peter G, Christoph Dellago, Phillip L Geissler, and David Chandler (Feb. 2000). "Transition path sampling: throwing ropes over mountains in the dark." In: *Journal of Physics: Condensed Matter* 12.8A, A147–A152. ISSN: 0953-8984, 1361-648X. DOI: 10.1088/0953-8984/12/8A/316. URL: https://iopscience.iop.org/article/10.1088/0953-8984/12/8A/316 (visited on 06/17/2024).

Boucher, Olivier et al. (July 2020). "Presentation and Evaluation of the IPSL-CM6A-LR Climate Model." en. In: *Journal of Advances in Modeling Earth Systems* 12.7. ISSN: 1942-2466, 1942-2466. DOI: 10.1029/2019MS002010. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2019MS002010 (visited on 04/05/2022).

Bouchet, Freddy, Joran Rolland, and Eric Simonnet (Feb. 2019). "Rare Event Algorithm Links Transitions in Turbulent Flows with Activated Nucleations." In: *Physical Review Letters* 122.7. Publisher: American Physical Society, p. 074502. DOI: 10.1103/PhysRevLett.122.074502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.074502 (visited on 07/10/2024).

Bucklew, James Antonio (2010). *Introduction to rare event simulation*. eng. softcover version of original hardcover edition 2004. Springer series in statistics. New York, NY: Springer. ISBN: 978-1-4419-1893-2.

Cassou, Christophe, Yochanan Kushnir, Ed Hawkins, Anna Pirani, Fred Kucharski, In-Sik Kang, and Nico Caltabiano (Mar. 2018). "Decadal Climate Variability and Predictability: Challenges and Opportunities." en. In: *Bulletin of the American Meteorological Society* 99.3, pp. 479–490. ISSN: 0003-0007, 1520-0477.

DOI: 10.1175/BAMS-D-16-0286.1. URL: https://journals.ametsoc.org/doi/10.1175/BAMS-D-16-0286.1 (visited on 01/30/2024).

Cassou, Christophe, Laurent Terray, and Adam S. Phillips (Aug. 2005). "Tropical Atlantic Influence on European Heat Waves." EN. In: *Journal of Climate* 18.15. Publisher: American Meteorological Society Section: Journal of Climate, pp. 2805–2811. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI3506.1. URL: https://journals.ametsoc.org/view/journals/clim/18/15/jcli3506.1.xml (visited on 07/17/2024).

Chattopadhyay, Ashesh, Ebrahim Nabizadeh, and Pedram Hassanzadeh (Feb. 2020). "Analog Forecasting of Extreme-Causing Weather Patterns Using Deep Learning." en. In: *Journal of Advances in Modeling Earth Systems* 12.2, e2019MS001958. ISSN: 1942-2466, 1942-2466. DOI: 10.1029/2019MS001958. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019MS001958 (visited on 06/17/2024).

Chetrite, Raphaël and Hugo Touchette (Sept. 2015). "Nonequilibrium Markov Processes Conditioned on Large Deviations." en. In: *Annales Henri Poincaré* 16.9, pp. 2005–2057. ISSN: 1424-0661. DOI: 10.1007/s00023-014-0375-8. URL: https://doi.org/10.1007/s00023-014-0375-8 (visited on 11/04/2023).

Chraibi, H., A. Dutfoy, T. Galtier, and J. Garnier (2018). "Optimal input potential functions in the interacting particle system method." In: Publisher: arXiv Version Number: 2. DOI: 10.48550/ARXIV.1811.10450. URL: https://arxiv.org/abs/1811.10450 (visited on 02/09/2023).

Cini, Matteo, Giuseppe Zappa, Francesco Ragone, and Susanna Corti (Jan. 2024). "Simulating AMOC tipping driven by internal climate variability with a rare event algorithm." en. In: *npj Climate and Atmospheric Science* 7.1. Publisher: Nature Publishing Group, pp. 1–10. ISSN: 2397-3722. DOI: 10.1038/s41612-024-00568-7. URL: https://www.nature.com/articles/s41612-024-00568-7 (visited on 06/24/2024).

Climate Change, Intergovernmental Panel on (2023). *AR6 Synthesis Report: Climate Change 2023*. en. URL: https://www.ipcc.ch/report/ar6/syr/ (visited on 04/24/2024).

Coles, Stuart (2001). *An Introduction to Statistical Modeling of Extreme Values*.

Coumou, D., G. Di Capua, S. Vavrus, L. Wang, and S. Wang (Aug. 2018). "The influence of Arctic amplification on mid-latitude summer circulation." en. In: *Nature Communications* 9.1. Number: 1 Publisher: Nature Publishing Group, p. 2959. ISSN: 2041-1723. DOI: 10.1038/s41467-018-05256-8. URL: https://www.nature.com/articles/s41467-018-05256-8 (visited on 10/31/2023).

Coumou, Dim, Vladimir Petoukhov, Stefan Rahmstorf, Stefan Petri, and Hans Joachim Schellnhuber (Aug. 2014). "Quasi-resonant circulation regimes and hemispheric synchronization of extreme weather in boreal summer." en. In: *Proceedings of the National Academy of Sciences* 111.34, pp. 12331–12336. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1412797111. URL: https://pnas.org/doi/full/10.1073/pnas.1412797111 (visited on 06/22/2023).

Coumou, Dim and Stefan Rahmstorf (July 2012). "A decade of weather extremes." en. In: *Nature Climate Change* 2.7. Publisher: Nature Publishing Group, pp. 491–496. ISSN: 1758-6798. DOI: 10.1038/nclimate1452. URL: https://www.nature.com/articles/nclimate1452 (visited on 03/06/2024).

Cozian, Bastien (2023). *Computing Climate Extreme Events and Extremes of Production of Renewable Energy Using Rare Events Algorithms, PhD thesis*.

Cozian, Bastien, Corentin Herbert, and Freddy Bouchet (Mar. 2024). "Assessing the probability of extremely low wind energy production in Europe at subseasonal to seasonal time scales." en. In: *Environmental Research Letters* 19.4. Publisher: IOP Publishing, p. 044046. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ad35d9. URL: https://dx.doi.org/10.1088/1748-9326/ad35d9 (visited on 05/30/2024).

Cramér, Harald and Harald Cramér (1994). "Sur un nouveau théorème-limite de la théorie des probabilités." In: ed. by Anders Martin-Löf. Book Title: Collected Works II Series Title: Springer Collected Works in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 895–913. ISBN: 978-3-642-39684-7 978-3-642-40607-2. DOI: 10.1007/978-3-642-40607-2_8. URL: http://link.springer.com/10.1007/978-3-642-40607-2_8 (visited on 06/20/2024).

Cramér, Harald and Hugo Touchette (Aug. 2022). *On a new limit theorem in probability theory (Translation of 'Sur un nouveau th\'eor\'eme-limite de la th\'eorie des probabilit\'es')*. en. arXiv:1802.05988 [math]. URL: http://arxiv.org/abs/1802.05988 (visited on 06/20/2024).

Cérou, Frédéric and Arnaud Guyader (Feb. 2007). "Adaptive Multilevel Splitting for Rare Event Analysis." In: *Stochastic Analysis and Applications* 25.2. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07362990601139628, pp. 417–443. ISSN: 0736-2994. DOI: 10.1080/07362990601139628. URL: https://doi.org/10.1080/07362990601139628 (visited on 06/27/2024).

Cérou, Frédéric, Arnaud Guyader, and Mathias Rousset (Apr. 2019). "Adaptive multilevel splitting: Historical perspective and recent results." en. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29.4, p. 043108. ISSN: 1054-1500, 1089-7682. DOI: 10.1063/1.5082247. URL: http://aip.scitation.org/doi/10.1063/1.5082247 (visited on 02/09/2023).

D'Andrea, Fabio, Antonello Provenzale, Robert Vautard, and Nathalie De Noblet-Decoudré (2006). "Hot and cool summers: Multiple equilibria of the continental water cycle." en. In: *Geophysical Research Letters* 33.24. _eprint: https://onlinelibrary.wi ISSN: 1944-8007. DOI: 10.1029/2006GL027972. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2006GL027972 (visited on 06/18/2024).

Del Moral, Pierre (2004). *Feynman-Kac Formulae*. Ed. by J. Gani, C. C. Heyde, and T. G. Kurtz. Probability and its Applications. New York, NY: Springer. ISBN: 978-1-4419-1902-1 978-1-4684-9393-1. DOI: 10.1007/978-1-4684-9393-1. URL: http://link.springer.com/10.1007/978-1-4684-9393-1 (visited on 06/25/2024).

Del Moral, Pierre and Josselin Garnier (Nov. 2005). "Genealogical particle analysis of rare events." In: *The Annals of Applied Probability* 15.4. Publisher: Institute of Mathematical Statistics, pp. 2496–2534. ISSN: 1050-5164, 2168-8737. DOI: 10.1214/105051605000000566. URL: https://projecteuclid.org/journals/annals-of-applied-probability/volume-15/issue-4/Genealogical-particle-analysis-of-rare-events/10.1214/105051605000000566.full (visited on 07/04/2024).

Delaunay, Antoine and Hannah M. Christensen (Aug. 2022). "Interpretable Deep Learning for Probabilistic MJO Prediction." en. In: *Geophysical Re-*

search Letters 49.16, e2022GL098566. ISSN: 0094-8276, 1944-8007. DOI: 10.1029/2022GL098566. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GL098566 (visited on 06/17/2024).

Della-Marta, P. M., J. Luterbacher, H. von Weissenfluh, E. Xoplaki, M. Brunet, and H. Wanner (Aug. 2007). "Summer heat waves over western Europe 1880–2003, their relationship to large-scale forcings and predictability." en. In: *Climate Dynamics* 29.2-3, pp. 251–275. ISSN: 0930-7575, 1432-0894. DOI: 10.1007/s00382-007-0233-1. URL: http://link.springer.com/10.1007/s00382-007-0233-1 (visited on 04/05/2022).

Delworth, Thomas L., Fanrong Zeng, Gabriel A. Vecchi, Xiaosong Yang, Liping Zhang, and Rong Zhang (July 2016). "The North Atlantic Oscillation as a driver of rapid climate change in the Northern Hemisphere." en. In: *Nature Geoscience* 9.7. Number: 7 Publisher: Nature Publishing Group, pp. 509–512. ISSN: 1752-0908. DOI: 10.1038/ngeo2738. URL: https://www.nature.com/articles/ngeo2738 (visited on 02/02/2024).

Dembo, Amir and Ofer Zeitouni (2010). *Large Deviations Techniques and Applications*. Vol. 38. Stochastic Modelling and Applied Probability. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-642-03310-0 978-3-642-03311-7. DOI: 10.1007/978-3-642-03311-7. URL: http://link.springer.com/10.1007/978-3-642-03311-7 (visited on 02/09/2023).

Dikshit, Abhirup, Biswajeet Pradhan, and Abdullah M. Alamri (Feb. 2021). "Long lead time drought forecasting using lagged climate variables and a stacked long short-term memory model." en. In: *Science of The Total Environment* 755, p. 142638. ISSN: 00489697. DOI: 10.1016/j.scitotenv.2020.142638. URL: https://linkinghub.elsevier.com/retrieve/pii/S0048969720361672 (visited on 06/17/2024).

Donsker, M. D. and S. R. S. Varadhan (1976). "Asymptotic evaluation of certain Markov process expectations for large time—III." en. In: *Communications on Pure and Applied Mathematics* 29.4. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.3160, pp. 389–461. ISSN: 1097-0312. DOI: 10.1002/cpa.3160290405. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160290405 (visited on 06/20/2024).

Douville, H., J. Colin, E. Krug, J. Cattiaux, and S. Thao (2016). "Midlatitude daily summer temperatures reshaped by soil moisture under climate change." en. In: *Geophysical Research Letters* 43.2. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015, pp. 812–818. ISSN: 1944-8007. DOI: 10.1002/2015GL066222. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/2015GL066222 (visited on 06/11/2024).

Duchez, Aurélie, Eleanor Frajka-Williams, Simon A. Josey, Dafydd G. Evans, Jeremy P. Grist, Robert Marsh, Gerard D. McCarthy, Bablu Sinha, David I. Berry, and Joël J.-M. Hirschi (July 2016). "Drivers of exceptionally cold North Atlantic Ocean temperatures and their link to the 2015 European heat wave." en. In: *Environmental Research Letters* 11.7. Publisher: IOP Publishing, p. 074004. ISSN: 1748-9326. DOI: 10.1088/1748-9326/11/7/074004. URL: https://dx.doi.org/10.1088/1748-9326/11/7/074004 (visited on 07/01/2024).

Dunn, Robert J. H. et al. (2020). "Development of an Updated Global Land In Situ-Based Data Set of Temperature and Precipitation Extremes: HadEX3."

en. In: *Journal of Geophysical Research: Atmospheres* 125.16. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019JD032263, e2019JD032263. ISSN: 2169-8996. DOI: 10.1029/2019JD032263. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2019JD032263 (visited on 06/11/2024).

Döscher, Ralf et al. (Feb. 2021). *The EC-Earth3 Earth System Model for the Climate ModelIntercomparison Project 6*. preprint. Climate and Earth system modeling. DOI: 10.5194/gmd-2020-446. URL: https://gmd.copernicus.org/preprints/gmd-2020-446/ (visited on 04/05/2022).

Ellis, Richard S. (Feb. 1984). "Large Deviations for a General Class of Random Vectors." In: *The Annals of Probability* 12.1. Publisher: Institute of Mathematical Statistics, pp. 1–12. ISSN: 0091-1798, 2168-894X. DOI: 10.1214/aop/1176993370. URL: https://projecteuclid.org/journals/annals-of-probability/volume-12/issue-1/Large-Deviations-for-a-General-Class-of-Random-Vectors/10.1214/aop/1176993370.full (visited on 06/20/2024).

— (Jan. 1995). "An overview of the theory of large deviations and applications to statistical mechanics." In: *Scandinavian Actuarial Journal* 1995.1. Publisher: Taylor & Francis, pp. 97–142. ISSN: 0346-1238. DOI: 10.1080/03461238.1995.10413952. URL: https://www.tandfonline.com/doi/abs/10.1080/03461238.1995.10413952 (visited on 06/20/2024).

Feldstein, Steven B. and Christian L.E. Franzke (Jan. 2017). "Atmospheric teleconnection patterns." In: *Nonlinear and Stochastic Climate Dynamics*. Cambridge University Press, pp. 54–104. ISBN: 978-1-107-11814-0. DOI: 10.1017/9781316339251.004. URL: http://www.scopus.com/inward/record.url?scp=85040603188&partnerID=8YFLogxK (visited on 05/07/2024).

Ferreira, Ana and Laurens de Haan (Feb. 2015). "On the block maxima method in extreme value theory: PWM estimators." In: *The Annals of Statistics* 43.1. Publisher: Institute of Mathematical Statistics, pp. 276–298. ISSN: 0090-5364, 2168-8966. DOI: 10.1214/14-AOS1280. URL: https://projecteuclid.org/journals/annals-of-statistics/volume-43/issue-1/On-the-block-maxima-method-in-extreme-value-theory/10.1214/14-AOS1280.full (visited on 07/04/2024).

Feudale, Laura and Jagadish Shukla (May 2011). "Influence of sea surface temperature on the European heat wave of 2003 summer. Part II: a modeling study." en. In: *Climate Dynamics* 36.9, pp. 1705–1715. ISSN: 1432-0894. DOI: 10.1007/s00382-010-0789-z. URL: https://doi.org/10.1007/s00382-010-0789-z (visited on 07/11/2024).

Finkel, Justin, Dorian S. Abbot, and Jonathan Weare (July 2020). "Path Properties of Atmospheric Transitions: Illustration with a Low-Order Sudden Stratospheric Warming Model." In: *Journal of the Atmospheric Sciences* 77.7, pp. 2327–2347. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/JAS-D-19-0278.1. URL: https://journals.ametsoc.org/view/journals/atsc/77/7/jasD190278.xml (visited on 03/31/2023).

Finkel, Justin, Robert J. Webber, Edwin P. Gerber, Dorian S. Abbot, and Jonathan Weare (Nov. 2021). "Learning Forecasts of Rare Stratospheric Transitions from Short Simulations." In: *Monthly Weather Review* 149.11, pp. 3647–3669. ISSN: 0027-0644, 1520-0493. DOI: 10.1175/MWR-D-21-0024.1. URL: https:

//journals.ametsoc.org/view/journals/mwre/149/11/MWR-D-21-0024.1.xml (visited on 02/09/2023).

Fischer, E. M., S. I. Seneviratne, D. Lüthi, and C. Schär (2007a). "Contribution of land-atmosphere coupling to recent European summer heat waves." en. In: *Geophysical Research Letters* 34.6. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2006GL029 ISSN: 1944-8007. DOI: 10.1029/2006GL029068. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2006GL029068 (visited on 04/23/2024).

Fischer, E. M., S. I. Seneviratne, P. L. Vidale, D. Lüthi, and C. Schär (Oct. 2007b). "Soil Moisture–Atmosphere Interactions during the 2003 European Summer Heat Wave." en. In: *Journal of Climate* 20.20, pp. 5081–5099. ISSN: 1520-0442, 0894-8755. DOI: 10.1175/JCLI4288.1. URL: http://journals.ametsoc.org/doi/10.1175/JCLI4288.1 (visited on 04/13/2023).

Fischer, Erich M. (May 2014). "Autopsy of two mega-heatwaves." en. In: *Nature Geoscience* 7.5. Publisher: Nature Publishing Group, pp. 332–333. ISSN: 1752-0908. DOI: 10.1038/ngeo2148. URL: https://www.nature.com/articles/ngeo2148 (visited on 04/23/2024).

Fouillet, A., G. Rey, F. Laurent, G. Pavillon, S. Bellec, C. Guihenneuc-Jouyaux, J. Clavel, E. Jougla, and Denis Hémon (Sept. 2006). "Excess mortality related to the August 2003 heat wave in France." en. In: *International Archives of Occupational and Environmental Health* 80.1, pp. 16–24. ISSN: 0340-0131, 1432-1246. DOI: 10.1007/s00420-006-0089-4. URL: http://link.springer.com/10.1007/s00420-006-0089-4 (visited on 06/17/2024).

Fraedrich, Klaus, Heiko Jansen, Edilbert Kirk, Ute Luksch, and Frank Lunkeit (July 2005a). "The Planet Simulator: Towards a user friendly model." en. In: *Meteorologische Zeitschrift*. Publisher: Schweizerbart'sche Verlagsbuchhandlung, pp. 299–304. ISSN: , DOI: 10.1127/0941-2948/2005/0043. URL: https://www.schweizerbart.de/papers/metz/detail/14/54202/The_Planet_Simulator_Towards_a_user_friendly_model?af=crossref (visited on 06/18/2024).

Fraedrich, Klaus, Edilbert Kirk, Ute Luksch, and Frank Lunkeit (Dec. 2005b). "The portable university model of the atmosphere (PUMA): Storm track dynamics and low-frequency variability." en. In: *Meteorologische Zeitschrift*. Publisher: Schweizerbart'sche Verlagsbuchhandlung, pp. 735–745. ISSN: , DOI: 10.1127/0941-2948/2005/0074. URL: https://www.schweizerbart.de/papers/metz/detail/14/54403/The_portable_university_model_of_the_atmosphere_PU?af=crossref (visited on 06/18/2024).

Francis, Jennifer A. and Stephen J. Vavrus (2012). "Evidence linking Arctic amplification to extreme weather in mid-latitudes." en. In: *Geophysical Research Letters* 39.6. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2012GL051000. ISSN: 1944-8007. DOI: 10.1029/2012GL051000. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2012GL051000 (visited on 05/06/2024).

Galfi, Vera Melinda and Valerio Lucarini (July 2021). "Fingerprinting Heatwaves and Cold Spells and Assessing Their Response to Climate Change Using Large Deviation Theory." en. In: *Physical Review Letters* 127.5, p. 058701. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.127.058701. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.058701 (visited on 02/09/2023).

Gallego, David, Pedro Ribera, Ricardo Garcia-Herrera, Emiliano Hernandez, and Luis Gimeno (May 2005). "A new look for the Southern Hemisphere jet stream." en. In: *Climate Dynamics* 24.6, pp. 607–621. ISSN: 0930-7575, 1432-0894. DOI: 10.1007/s00382-005-0006-7. URL: http://link.springer.com/10.1007/s00382-005-0006-7 (visited on 07/18/2023).

García-Herrera, R., J. Díaz, R. M. Trigo, J. Luterbacher, and E. M. Fischer (Mar. 2010). "A Review of the European Summer Heat Wave of 2003." en. In: *Critical Reviews in Environmental Science and Technology* 40.4, pp. 267–306. ISSN: 1064-3389, 1547-6537. DOI: 10.1080/10643380802238137. URL: http://www.tandfonline.com/doi/abs/10.1080/10643380802238137 (visited on 04/13/2023).

García-León, David, Ana Casanueva, Gabriele Standardi, Annkatrin Burgstall, Andreas D. Flouris, and Lars Nybo (Oct. 2021). "Current and projected regional economic impacts of heatwaves in Europe." en. In: *Nature Communications* 12.1. Publisher: Nature Publishing Group, p. 5807. ISSN: 2041-1723. DOI: 10.1038/s41467-021-26050-z. URL: https://www.nature.com/articles/s41467-021-26050-z (visited on 07/03/2024).

Garnier, Josselin and Pierre Del Moral (Nov. 2006). "Simulations of rare events in fiber optics by interacting particle systems." en. In: *Optics Communications* 267.1, pp. 205–214. ISSN: 00304018. DOI: 10.1016/j.optcom.2006.05.066. URL: https://linkinghub.elsevier.com/retrieve/pii/S0030401806005906 (visited on 02/09/2023).

Gauer, Robert and Bryce K. Meyers (Apr. 2019). "Heat-Related Illnesses." eng. In: *American Family Physician* 99.8, pp. 482–489. ISSN: 1532-0650.

Giardina, Cristian, Jorge Kurchan, Vivien Lecomte, and Julien Tailleur (Nov. 2011). "Simulating Rare Events in Dynamical Processes." en. In: *Journal of Statistical Physics* 145.4, pp. 787–811. ISSN: 0022-4715, 1572-9613. DOI: 10.1007/s10955-011-0350-4. URL: http://link.springer.com/10.1007/s10955-011-0350-4 (visited on 02/09/2023).

Giardinà, Cristian, Jorge Kurchan, and Luca Peliti (Mar. 2006). "Direct Evaluation of Large-Deviation Functions." en. In: *Physical Review Letters* 96.12, p. 120603. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.96.120603. URL: https://link.aps.org/doi/10.1103/PhysRevLett.96.120603 (visited on 03/24/2023).

Giffard-Roisin, Sophie, Mo Yang, Guillaume Charpiat, Christina Kumler Bonfanti, Balázs Kégl, and Claire Monteleoni (Jan. 2020). "Tropical Cyclone Track Forecasting Using Fused Deep Learning From Aligned Reanalysis Data." English. In: *Frontiers in Big Data* 3. Publisher: Frontiers. ISSN: 2624-909X. DOI: 10.3389/fdata.2020.00001. URL: https://www.frontiersin.org/articles/10.3389/fdata.2020.00001 (visited on 04/24/2024).

Grotjahn, Richard and Ghislain Faure (June 2008). "Composite Predictor Maps of Extraordinary Weather Events in the Sacramento, California, Region*." en. In: *Weather and Forecasting* 23.3, pp. 313–335. ISSN: 1520-0434, 0882-8156. DOI: 10.1175/2007WAF2006055.1. URL: https://journals.ametsoc.org/doi/10.1175/2007WAF2006055.1 (visited on 06/17/2024).

Gálfi, Vera Melinda, Valerio Lucarini, Francesco Ragone, and Jeroen Wouters (June 2021). "Applications of large deviation theory in geophysical fluid dynamics and climate science." en. In: *La Rivista del Nuovo Cimento* 44.6,

pp. 291–363. ISSN: 1826-9850. DOI: 10.1007/s40766-021-00020-z. URL: https://doi.org/10.1007/s40766-021-00020-z (visited on 03/06/2024).

Gálfi, Vera Melinda, Valerio Lucarini, and Jeroen Wouters (Mar. 2019). "A large deviation theory-based analysis of heat waves and cold spells in a simplified model of the general circulation of the atmosphere." en. In: *Journal of Statistical Mechanics: Theory and Experiment* 2019.3. Publisher: IOP Publishing and SISSA, p. 033404. ISSN: 1742-5468. DOI: 10.1088/1742-5468/ab02e8. URL: https://dx.doi.org/10.1088/1742-5468/ab02e8 (visited on 03/06/2024).

Hafez, Yehia (2017). "On the Relationship between Heat Waves over the Western and Central Europe and NAO, SOI, El-Nino 3.4 in Summer 2015." In: *Journal of Geoscience and Environment Protection* 05.04, pp. 31–45. ISSN: 2327-4336, 2327-4344. DOI: 10.4236/gep.2017.54004. URL: http://www.scirp.org/journal/PaperDownload.aspx?DOI=10.4236/gep.2017.54004 (visited on 06/18/2024).

Hannachi, A., I. T. Jolliffe, and D. B. Stephenson (July 2007). "Empirical orthogonal functions and related techniques in atmospheric science: A review." en. In: *International Journal of Climatology* 27.9, pp. 1119–1152. ISSN: 08998418, 10970088. DOI: 10.1002/joc.1499. URL: https://onlinelibrary.wiley.com/doi/10.1002/joc.1499 (visited on 05/04/2023).

Hansen, James, Makiko Sato, and Reto Ruedy (Sept. 2012). "Perception of climate change." In: *Proceedings of the National Academy of Sciences* 109.37. Publisher: Proceedings of the National Academy of Sciences, E2415–E2423. DOI: 10.1073/pnas.1205276109. URL: https://www.pnas.org/doi/full/10.1073/pnas.1205276109 (visited on 01/26/2024).

Hastie (n.d.). *The Elements of Statistical Learning*.

Held, I. M. (1983). *Stationary and quasi-stationary eddies in the extratropical troposphere: Theory.*

Hersbach, Hans et al. (2020). "The ERA5 global reanalysis." en. In: *Quarterly Journal of the Royal Meteorological Society* 146.730. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1 pp. 1999–2049. ISSN: 1477-870X. DOI: 10.1002/qj.3803. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803 (visited on 06/18/2024).

Hidalgo, Esteban Guevara (June 2018). *Cloning Algorithms: from Large Deviations to Population Dynamics*. arXiv:1806.01943 [cond-mat]. URL: http://arxiv.org/abs/1806.01943 (visited on 06/26/2024).

Hirschi, Martin, Sonia I. Seneviratne, Vesselin Alexandrov, Fredrik Boberg, Constanta Boroneant, Ole B. Christensen, Herbert Formayer, Boris Orlowsky, and Petr Stepanek (Jan. 2011). "Observational evidence for soil-moisture impact on hot extremes in southeastern Europe." en. In: *Nature Geoscience* 4.1. Publisher: Nature Publishing Group, pp. 17–21. ISSN: 1752-0908. DOI: 10.1038/ngeo1032. URL: https://www.nature.com/articles/ngeo1032 (visited on 06/18/2024).

Horton, Radley M., Justin S. Mankin, Corey Lesk, Ethan Coffel, and Colin Raymond (Dec. 2016). "A Review of Recent Advances in Research on Extreme Heat Events." en. In: *Current Climate Change Reports* 2.4, pp. 242–259. ISSN: 2198-6061. DOI: 10.1007/s40641-016-0042-x. URL: http://link.springer.com/10.1007/s40641-016-0042-x (visited on 06/09/2022).

Hoskins, Brian J. and Tercio Ambrizzi (June 1993). "Rossby Wave Propagation on a Realistic Longitudinally Varying Flow." EN. In: *Journal of the Atmospheric Sciences* 50.12. Publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences, pp. 1661–1671. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(1993)050<1661:RWPOAR>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/atsc/50/12/1520-0469_1993_050_1661_rwpoar_2_0_co_2.xml (visited on 07/20/2024).

Hoskins, Brian J. and David J. Karoly (June 1981). "The Steady Linear Response of a Spherical Atmosphere to Thermal and Orographic Forcing." EN. In: *Journal of the Atmospheric Sciences* 38.6. Publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences, pp. 1179–1196. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(1981)038<1179:TSLROA>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/atsc/38/6/1520-0469_1981_038_1179_tslroa_2_0_co_2.xml (visited on 07/20/2024).

Hoskins, Brian and Tim Woollings (Sept. 2015). "Persistent Extratropical Regimes and Climate Extremes." en. In: *Current Climate Change Reports* 1.3, pp. 115–124. ISSN: 2198-6061. DOI: 10.1007/s40641-015-0020-8. URL: https://doi.org/10.1007/s40641-015-0020-8 (visited on 04/02/2024).

Hurrell, James W. et al. (Sept. 2013). "The Community Earth System Model: A Framework for Collaborative Research." EN. In: *Bulletin of the American Meteorological Society* 94.9. Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society, pp. 1339–1360. DOI: 10.1175/BAMS-D-12-00121.1. URL: https://journals.ametsoc.org/view/journals/bams/94/9/bams-d-12-00121.1.xml (visited on 05/31/2024).

Intergovernmental Panel On Climate Change (July 2023). *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. 1st ed. Cambridge University Press. ISBN: 978-1-00-915789-6. DOI: 10.1017/9781009157896. URL: https://www.cambridge.org/core/product/identifier/9781009157896/type/book (visited on 01/30/2024).

Jacques-Dumas, Valérian, Francesco Ragone, Pierre Borgnat, Patrice Abry, and Freddy Bouchet (2022). "Deep Learning-Based Extreme Heatwave Forecast." In: *Frontiers in Climate* 4. ISSN: 2624-9553. URL: https://www.frontiersin.org/articles/10.3389/fclim.2022.789641 (visited on 11/29/2023).

Jacques-Dumas, Valérian, René M. van Westen, Freddy Bouchet, and Henk A. Dijkstra (June 2023). "Data-driven methods to estimate the committor function in conceptual ocean models." English. In: *Nonlinear Processes in Geophysics* 30.2. Publisher: Copernicus GmbH, pp. 195–216. ISSN: 1023-5809. DOI: 10.5194/npg-30-195-2023. URL: https://npg.copernicus.org/articles/30/195/2023/ (visited on 06/18/2024).

Kautz, Lisa-Ann, Olivia Martius, Stephan Pfahl, Joaquim G. Pinto, Alexandre M. Ramos, Pedro M. Sousa, and Tim Woollings (Mar. 2022). "Atmospheric blocking and weather extremes over the Euro-Atlantic sector – a review." English. In: *Weather and Climate Dynamics* 3.1. Publisher: Copernicus GmbH, pp. 305–336. DOI: 10.5194/wcd-3-305-2022. URL: https://wcd.copernicus.org/articles/3/305/2022/ (visited on 04/26/2024).

Kharin, V. V., G. M. Flato, X. Zhang, N. P. Gillett, F. Zwiers, and K. J. Anderson (May 2018). "Risks from Climate Extremes Change Differently from

1.5°C to 2.0°C Depending on Rarity." en. In: *Earth's Future* 6.5, pp. 704–715. ISSN: 23284277. DOI: 10.1002/2018EF000813. URL: https://onlinelibrary.wiley.com/doi/10.1002/2018EF000813 (visited on 04/08/2022).

Knight, Jeff R., Robert J. Allan, Chris K. Folland, Michael Vellinga, and Michael E. Mann (2005). "A signature of persistent natural thermohaline circulation cycles in observed climate." en. In: *Geophysical Research Letters* 32.20. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2005GL024233. ISSN: 1944-8007. DOI: 10.1029/2005GL024233. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2005GL024233 (visited on 01/26/2024).

Kornhuber, Kai, Scott Osprey, Dim Coumou, Stefan Petri, Vladimir Petoukhov, Stefan Rahmstorf, and Lesley Gray (Apr. 2019). "Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern." en. In: *Environmental Research Letters* 14.5. Publisher: IOP Publishing, p. 054002. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ab13bf. URL: https://dx.doi.org/10.1088/1748-9326/ab13bf (visited on 10/31/2023).

Krishna, Satyapriya, Tessa Han, Alex Gu, Javin Pombra, Shahin Jabbari, Steven Wu, and Himabindu Lakkaraju (Feb. 2022). *The Disagreement Problem in Explainable Machine Learning: A Practitioner's Perspective*. arXiv:2202.01602 [cs]. DOI: 10.48550/arXiv.2202.01602. URL: http://arxiv.org/abs/2202.01602 (visited on 06/17/2024).

Lau, Ngar-Cheung and Mary Jo Nath (July 2012). "A Model Study of Heat Waves over North America: Meteorological Aspects and Projections for the Twenty-First Century." en. In: *Journal of Climate* 25.14, pp. 4761–4784. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-11-00575.1. URL: http://journals.ametsoc.org/doi/10.1175/JCLI-D-11-00575.1 (visited on 02/02/2024).

Lawrence, David M et al. (2011). "Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model." en. In: *Journal of Advances in Modeling Earth Systems* 3.1. _eprint: https://onlinelibrary.wiley.com/doi/pd ISSN: 1942-2466. DOI: 10.1029/2011MS00045. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2011MS00045 (visited on 06/20/2024).

Le Priol, Clément, Joy M. Monteiro, and Freddy Bouchet (Apr. 2024). *Using rare event algorithms to understand the statistics and dynamics of extreme heatwave seasons in South Asia*. arXiv:2404.07791 [cond-mat, physics:physics]. DOI: 10.48550/arXiv.2404.07791. URL: http://arxiv.org/abs/2404.07791 (visited on 06/24/2024).

Lecomte, Vivien and Julien Tailleur (Mar. 2007). "A numerical approach to large deviations in continuous time." In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.03, P03004–P03004. ISSN: 1742-5468. DOI: 10.1088/1742-5468/2007/03/P03004. URL: https://iopscience.iop.org/article/10.1088/1742-5468/2007/03/P03004 (visited on 02/09/2023).

Lestang, Thibault (Sept. 2018). "Numerical simulation and rare events algorithms for the study of extreme fluctuations of the drag force acting on an obstacle immersed in a turbulent flow." en. PhD thesis. Université de Lyon. URL: https://theses.hal.science/tel-01974316 (visited on 07/21/2024).

Lestang, Thibault, Freddy Bouchet, and Emmanuel Lévêque (July 2020). "Numerical study of extreme mechanical force exerted by a turbulent flow on a bluff body by direct and rare-event sampling techniques." en. In: *Journal*

of Fluid Mechanics 895, A19. ISSN: 0022-1120, 1469-7645. DOI: 10.1017/jfm.2020.293. URL: https://www.cambridge.org/core/journals/journal-of-fluid-mechanics/article/numerical-study-of-extreme-mechanical-force-exerted-by-a-turbulent-flow-on-a-bluff-body-by-direct-and-rareevent-sampling-techniques/AF7125593B78B17A8CA9958E0985D55B (visited on 06/24/2024).

Lestang, Thibault, Francesco Ragone, Charles-Edouard Bréhier, Corentin Herbert, and Freddy Bouchet (Apr. 2018). "Computing return times or return periods with rare event algorithms." In: *Journal of Statistical Mechanics: Theory and Experiment* 2018.4, p. 043213. ISSN: 1742-5468. DOI: 10.1088/1742-5468/aab856. URL: https://iopscience.iop.org/article/10.1088/1742-5468/aab856 (visited on 05/17/2022).

Li, Muyuan, Yao Yao, Ian Simmonds, Dehai Luo, Linhao Zhong, and Xiaodan Chen (Oct. 2020). "Collaborative impact of the NAO and atmospheric blocking on European heatwaves, with a focus on the hot summer of 2018." en. In: *Environmental Research Letters* 15.11. Publisher: IOP Publishing, p. 114003. ISSN: 1748-9326. DOI: 10.1088/1748-9326/aba6ad. URL: https://dx.doi.org/10.1088/1748-9326/aba6ad (visited on 06/17/2024).

Liggett, Thomas M. (2010). *Continuous Time Markov Processes*.

Liné, Aurélien, Christophe Cassou, Rym Msadek, and Parey Parey (2023). "Assessment of climate change at near-term (2020-2040) over Northern Europe through internal variability storylines." In: *submitted to NPJ*.

Lorenz, Edward N. (July 1969). "Atmospheric Predictability as Revealed by Naturally Occurring Analogues." EN. In: *Journal of the Atmospheric Sciences* 26.4. Publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences, pp. 636–646. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/atsc/26/4/1520-0469_1969_26_636_aparbn_2_0_co_2.xml (visited on 07/12/2024).

Lorenz, Ruth, Eric B. Jaeger, and Sonia I. Seneviratne (2010). "Persistence of heat waves and its link to soil moisture memory." en. In: *Geophysical Research Letters* 37.9. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2010GL042764 ISSN: 1944-8007. DOI: 10.1029/2010GL042764. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2010GL042764 (visited on 06/17/2024).

Lucente, Dario (Nov. 2021). "Predicting probabilities of climate extremes from observations and dynamics." en. PhD thesis. Université de Lyon. URL: https://theses.hal.science/tel-03562535 (visited on 07/19/2024).

Lucente, Dario, Stefan Duffner, Corentin Herbert, Joran Rolland, and Freddy Bouchet (Nov. 2019). *Machine learning of committor functions for predicting high impact climate events*. arXiv:1910.11736 [nlin, physics:physics]. DOI: 10.48550/arXiv.1910.11736. URL: http://arxiv.org/abs/1910.11736 (visited on 06/17/2024).

Lucente, Dario, Corentin Herbert, and Freddy Bouchet (Sept. 2022). "Committor Functions for Climate Phenomena at the Predictability Margin: The Example of El Niño–Southern Oscillation in the Jin and Timmermann Model." EN. In: *Journal of the Atmospheric Sciences* 79.9. Publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences, pp. 2387–2400. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/JAS-D-22-0038.1. URL:

https://journals.ametsoc.org/view/journals/atsc/79/9/JAS-D-22-0038.1.xml (visited on 06/17/2024).

Lucente, Dario, Joran Rolland, Corentin Herbert, and Freddy Bouchet (Aug. 2022). "Coupling rare event algorithms with data-based learned committor functions using the analogue Markov chain." In: *Journal of Statistical Mechanics: Theory and Experiment* 2022.8, p. 083201. ISSN: 1742-5468. DOI: 10.1088/1742-5468/ac7aa7. URL: https://iopscience.iop.org/article/10.1088/1742-5468/ac7aa7 (visited on 02/09/2023).

Lupo, Anthony R. (Nov. 2021). "Atmospheric blocking events: a review." en. In: *Annals of the New York Academy of Sciences* 1504.1, pp. 5–24. ISSN: 0077-8923, 1749-6632. DOI: 10.1111/nyas.14557. URL: https://onlinelibrary.wiley.com/doi/10.1111/nyas.14557 (visited on 07/21/2023).

Lupo, Anthony R., Igor I. Mokhov, Merseid G. Akperov, Alexander V. Chernokulsky, and H. Athar (Oct. 2012). "A Dynamic Analysis of the Role of the Planetary- and Synoptic-Scale in the Summer of 2010 Blocking Episodes over the European Part of Russia." en. In: *Advances in Meteorology* 2012. Publisher: Hindawi, e584257. ISSN: 1687-9309. DOI: 10.1155/2012/584257. URL: https://www.hindawi.com/journals/amete/2012/584257/ (visited on 05/03/2024).

Manabe, Syukuro (Nov. 1969). "CLIMATE AND THE OCEAN CIRCULATION: II. THE ATMOSPHERIC CIRCULATION AND THE EFFECT OF HEAT TRANSFER BY OCEAN CURRENTS." EN. In: *Monthly Weather Review* 97.11. Publisher: American Meteorological Society Section: Monthly Weather Review, pp. 775–805. ISSN: 1520-0493, 0027-0644. DOI: 10.1175/1520-0493(1969)097<0775:CATOC>2.3.CO;2. URL: https://journals.ametsoc.org/view/journals/mwre/97/11/1520-0493_1969_097_0775_catoc_2_3_co_2.xml (visited on 06/18/2024).

Mann, Michael E., Stefan Rahmstorf, Kai Kornhuber, Byron A. Steinman, Sonya K. Miller, and Dim Coumou (Mar. 2017). "Influence of Anthropogenic Climate Change on Planetary Wave Resonance and Extreme Weather Events." en. In: *Scientific Reports* 7.1, p. 45242. ISSN: 2045-2322. DOI: 10.1038/srep45242. URL: https://www.nature.com/articles/srep45242 (visited on 06/26/2023).

Martens, Brecht, Diego G. Miralles, Hans Lievens, Robin van der Schalie, Richard A. M. de Jeu, Diego Fernández-Prieto, Hylke E. Beck, Wouter A. Dorigo, and Niko E. C. Verhoest (May 2017). "GLEAM v3: satellite-based land evaporation and root-zone soil moisture." English. In: *Geoscientific Model Development* 10.5. Publisher: Copernicus GmbH, pp. 1903–1925. ISSN: 1991-959X. DOI: 10.5194/gmd-10-1903-2017. URL: https://gmd.copernicus.org/articles/10/1903/2017/ (visited on 03/28/2024).

Masato, Giacomo, Brian J. Hoskins, and Tim Woollings (July 2013). "Wave-Breaking Characteristics of Northern Hemisphere Winter Blocking: A Two-Dimensional Approach." EN. In: *Journal of Climate* 26.13. Publisher: American Meteorological Society Section: Journal of Climate, pp. 4535–4549. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-12-00240.1. URL: https://journals.ametsoc.org/view/journals/clim/26/13/jcli-d-12-00240.1.xml (visited on 07/09/2024).

Mascolo, Valeria, Alessandro Lovo, Corentin Herbert, and Freddy Bouchet (May 2024). *Gaussian Framework and Optimal Projection of Weather Fields for Pre-*

diction of Extreme Events. arXiv:2405.20903 [physics]. DOI: 10.48550/arXiv.2405.20903. URL: http://arxiv.org/abs/2405.20903 (visited on 06/24/2024).

Masson-Delmotte, Valérie et al. (July 2023). *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. 1st. Cambridge University Press. ISBN: 978-1-00-915789-6. DOI: 10.1017/9781009157896. (Visited on 06/13/2024).

Materia, Stefano, Constantin Ardilouze, Chloé Prodhomme, Markus G. Donat, Marianna Benassi, Francisco J. Doblas-Reyes, Daniele Peano, Louis-Philippe Caron, Paolo Ruggieri, and Silvio Gualdi (Apr. 2022). "Summer temperature response to extreme soil water conditions in the Mediterranean transitional climate regime." en. In: *Climate Dynamics* 58.7, pp. 1943–1963. ISSN: 1432-0894. DOI: 10.1007/s00382-021-05815-8. URL: https://doi.org/10.1007/s00382-021-05815-8 (visited on 03/25/2024).

Matsueda, Mio (2011). "Predictability of Euro-Russian blocking in summer of 2010." en. In: *Geophysical Research Letters* 38.6. _eprint: https://onlinelibrary.wiley.com/doi/p ISSN: 1944-8007. DOI: 10.1029/2010GL046557. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2010GL046557 (visited on 05/03/2024).

McGovern, Amy, Ryan Lagerquist, David John Gagne, G. Eli Jergensen, Kimberly L. Elmore, Cameron R. Homeyer, and Travis Smith (Nov. 2019). "Making the Black Box More Transparent: Understanding the Physical Implications of Machine Learning." EN. In: *Bulletin of the American Meteorological Society* 100.11. Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society, pp. 2175–2199. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-18-0195.1. URL: https://journals.ametsoc.org/view/journals/bams/100/11/bams-d-18-0195.1.xml (visited on 06/17/2024).

Miloshevich, George, Bastien Cozian, Patrice Abry, Pierre Borgnat, and Freddy Bouchet (2022). "Probabilistic forecasts of extreme heatwaves using convolutional neural networks in a regime of lack of data." In: Publisher: [object Object] Version Number: 2. DOI: 10.48550/ARXIV.2208.00971. URL: https://arxiv.org/abs/2208.00971 (visited on 04/04/2024).

Miloshevich, George, Dario Lucente, Pascal Yiou, and Freddy Bouchet (Jan. 2024). "Extreme heat wave sampling and prediction with analog Markov chain and comparisons with deep learning." en. In: *Environmental Data Science* 3, e9. ISSN: 2634-4602. DOI: 10.1017/eds.2024.7. URL: https://www.cambridge.org/core/journals/environmental-data-science/article/extreme-heat-wave-sampling-and-prediction-with-analog-markov-chain-and-comparisons-with-deep-learning/B07FBA3C3B59DFF98D4C7CEF1A51BC1B (visited on 07/12/2024).

Miloshevich, George, Philippine Rouby-Poizat, Francesco Ragone, and Freddy Bouchet (Aug. 2023). "Robust intra-model teleconnection patterns for extreme heatwaves." en. In: *Frontiers in Earth Science* 11. arXiv:2306.11769 [physics], p. 1235579. ISSN: 2296-6463. DOI: 10.3389/feart.2023.1235579. URL: http://arxiv.org/abs/2306.11769 (visited on 05/14/2024).

Miralles, D. G., T. R. H. Holmes, R. a. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman (Feb. 2011). "Global land-surface evaporation estimated from satellite-based observations." English. In: *Hydrology and Earth System*

Sciences 15.2. Publisher: Copernicus GmbH, pp. 453–469. ISSN: 1027-5606. DOI: 10.5194/hess-15-453-2011. URL: https://hess.copernicus.org/articles/15/453/2011/hess-15-453-2011.html (visited on 03/28/2024).

Miralles, Diego G., Pierre Gentine, Sonia I. Seneviratne, and Adriaan J. Teuling (2019). "Land–atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges." en. In: *Annals of the New York Academy of Sciences* 1436.1. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/nyas.13912, pp. 19–35. ISSN: 1749-6632. DOI: 10.1111/nyas.13912. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/nyas.13912 (visited on 04/23/2024).

Miralles, Diego G., Adriaan J. Teuling, Chiel C. van Heerwaarden, and Jordi Vilà-Guerau de Arellano (May 2014). "Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation." en. In: *Nature Geoscience* 7.5. Publisher: Nature Publishing Group, pp. 345–349. ISSN: 1752-0908. DOI: 10.1038/ngeo2141. URL: https://www.nature.com/articles/ngeo2141 (visited on 04/23/2024).

Miron, P., F. J. Beron-Vera, L. Helfmann, and P. Koltai (Mar. 2021). "Transition paths of marine debris and the stability of the garbage patches." In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 31.3, p. 033101. ISSN: 1054-1500. DOI: 10.1063/5.0030535. URL: https://doi.org/10.1063/5.0030535 (visited on 06/17/2024).

Molnos, Sonja, Tarek Mamdouh, Stefan Petri, Thomas Nocke, Tino Weinkauf, and Dim Coumou (Feb. 2017). "A network-based detection scheme for the jet stream core." en. In: *Earth System Dynamics* 8.1, pp. 75–89. ISSN: 2190-4987. DOI: 10.5194/esd-8-75-2017. URL: https://esd.copernicus.org/articles/8/75/2017/ (visited on 11/02/2023).

Muñoz-Sabater, Joaquín et al. (Sept. 2021). "ERA5-Land: a state-of-the-art global reanalysis dataset for land applications." English. In: *Earth System Science Data* 13.9. Publisher: Copernicus GmbH, pp. 4349–4383. ISSN: 1866-3508. DOI: 10.5194/essd-13-4349-2021. URL: https://essd.copernicus.org/articles/13/4349/2021/ (visited on 03/07/2024).

Neale, Richard B., Jadwiga Richter, Sungsu Park, Peter H. Lauritzen, Stephen J. Vavrus, Philip J. Rasch, and Minghua Zhang (July 2013). "The Mean Climate of the Community Atmosphere Model (CAM4) in Forced SST and Fully Coupled Experiments." EN. In: *Journal of Climate* 26.14. Publisher: American Meteorological Society Section: Journal of Climate, pp. 5150–5168. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-12-00236.1. URL: https://journals.ametsoc.org/view/journals/clim/26/14/jcli-d-12-00236.1.xml (visited on 06/20/2024).

Nemoto, Takahiro, Freddy Bouchet, Robert L. Jack, and Vivien Lecomte (June 2016). "Population-dynamics method with a multicanonical feedback control." en. In: *Physical Review E* 93.6, p. 062123. ISSN: 2470-0045, 2470-0053. DOI: 10.1103/PhysRevE.93.062123. URL: https://link.aps.org/doi/10.1103/PhysRevE.93.062123 (visited on 04/07/2023).

Noyelle, Robin, Pascal Yiou, and Davide Faranda (Feb. 2024). "Investigating the typicality of the dynamics leading to extreme temperatures in the IPSL-CM6A-LR model." en. In: *Climate Dynamics* 62.2, pp. 1329–1357. ISSN: 0930-

7575, 1432-0894. DOI: 10.1007/s00382-023-06967-5. URL: https://link.springer.com/10.1007/s00382-023-06967-5 (visited on 06/17/2024).

Pelly, J. L. and B. J. Hoskins (Mar. 2003). "A New Perspective on Blocking." EN. In: *Journal of the Atmospheric Sciences* 60.5. Publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences, pp. 743–755. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(2003)060<0743:ANPOB>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/atsc/60/5/1520-0469_2003_060_0743_anpob_2.0.co_2.xml (visited on 04/29/2024).

Pena-Ortiz, Cristina, David Gallego, Pedro Ribera, Paulina Ordonez, and Maria Del Carmen Alvarez-Castro (2013). "Observed trends in the global jet stream characteristics during the second half of the 20th century." en. In: *Journal of Geophysical Research: Atmospheres* 118.7. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10 pp. 2702–2713. ISSN: 2169-8996. DOI: 10.1002/jgrd.50305. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50305 (visited on 11/03/2023).

Perkins, Sarah E. (Oct. 2015). "A review on the scientific understanding of heatwaves—Their measurement, driving mechanisms, and changes at the global scale." en. In: *Atmospheric Research* 164-165, pp. 242–267. ISSN: 01698095. DOI: 10.1016/j.atmosres.2015.05.014. URL: https://linkinghub.elsevier.com/retrieve/pii/S0169809515001738 (visited on 04/08/2022).

Petoukhov, Vladimir, Stefan Petri, Stefan Rahmstorf, Dim Coumou, Kai Kornhuber, and Hans Joachim Schellnhuber (June 2016). "Role of quasiresonant planetary wave dynamics in recent boreal spring-to-autumn extreme events." In: *Proceedings of the National Academy of Sciences* 113.25. Publisher: Proceedings of the National Academy of Sciences, pp. 6862–6867. DOI: 10.1073/pnas.1606300113. URL: https://www.pnas.org/doi/full/10.1073/pnas.1606300113 (visited on 07/17/2024).

Petoukhov, Vladimir, Stefan Rahmstorf, Stefan Petri, and Hans Joachim Schellnhuber (Apr. 2013). "Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes." en. In: *Proceedings of the National Academy of Sciences* 110.14, pp. 5336–5341. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1222000110. URL: https://pnas.org/doi/full/10.1073/pnas.1222000110 (visited on 06/22/2023).

Pörtner, Hans-Otto et al. (June 2023). *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. 1st. Cambridge University Press. ISBN: 978-1-00-932584-4. DOI: 10.1017/9781009325844. (Visited on 06/13/2024).

Poumadère, Marc, Claire Mays, Sophie Le Mer, and Russell Blong (2005). "The 2003 Heat Wave in France: Dangerous Climate Change Here and Now." en. In: *Risk Analysis* 25.6. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1539-6924.2005.00694.x, pp. 1483–1494. ISSN: 1539-6924. DOI: 10.1111/j.1539-6924.2005.00694.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6924.2005.00694.x (visited on 07/03/2024).

Pérez-Espigares, Carlos and Pablo I. Hurtado (Aug. 2019). "Sampling rare events across dynamical phase transitions." en. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29.8, p. 083106. ISSN: 1054-1500, 1089-7682. DOI: 10.

1063/1.5091669. URL: https://pubs.aip.org/cha/article/29/8/083106/1058598/Sampling-rare-events-across-dynamical-phase (visited on 09/08/2023).

Qasmi, Saïd, Christophe Cassou, and Julien Boé (2017). "Teleconnection Between Atlantic Multidecadal Variability and European Temperature: Diversity and Evaluation of the Coupled Model Intercomparison Project Phase 5 Models." en. In: *Geophysical Research Letters* 44.21. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10. pp. 11,140–11,149. ISSN: 1944-8007. DOI: 10.1002/2017GL074886. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/2017GL074886 (visited on 01/30/2024).

Qasmi, Saïd, Emilia Sanchez-Gomez, Yohan Ruprich-Robert, Julien Boé, and Christophe Cassou (Feb. 2021). "Modulation of the Occurrence of Heatwaves over the Euro-Mediterranean Region by the Intensity of the Atlantic Multidecadal Variability." In: *Journal of Climate* 34.3, pp. 1099–1114. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-19-0982.1. URL: https://journals.ametsoc.org/view/journals/clim/34/3/JCLI-D-19-0982.1.xml (visited on 04/05/2022).

Racah, Evan, Christopher Beckham, Tegan Maharaj, Samira Ebrahimi Kahou, Prabhat, and Christopher Pal (Dec. 2017). "Extreme weather: a large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events." In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., pp. 3405–3416. ISBN: 978-1-5108-6096-4. (Visited on 04/24/2024).

Ragone, F. and F. Bouchet (June 2021). "Rare Event Algorithm Study of Extreme Warm Summers and Heatwaves Over Europe." en. In: *Geophysical Research Letters* 48.12, e2020GL091197. ISSN: 0094-8276, 1944-8007. DOI: 10.1029/2020GL091197. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020GL091197 (visited on 02/19/2024).

Ragone, Francesco and Freddy Bouchet (June 2020). "Computation of Extreme Values of Time Averaged Observables in Climate Models with Large Deviation Techniques." en. In: *Journal of Statistical Physics* 179.5-6, pp. 1637–1665. ISSN: 0022-4715, 1572-9613. DOI: 10.1007/s10955-019-02429-7. URL: http://link.springer.com/10.1007/s10955-019-02429-7 (visited on 02/09/2023).

Ragone, Francesco, Jeroen Wouters, and Freddy Bouchet (Jan. 2018). "Computation of extreme heat waves in climate models using a large deviation algorithm." en. In: *Proceedings of the National Academy of Sciences* 115.1, pp. 24–29. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1712645115. URL: https://pnas.org/doi/full/10.1073/pnas.1712645115 (visited on 02/09/2023).

Ratnam, J. V., Swadhin K. Behera, Satyaban B. Ratna, M. Rajeevan, and Toshio Yamagata (Apr. 2016). "Anatomy of Indian heatwaves." en. In: *Scientific Reports* 6.1, p. 24395. ISSN: 2045-2322. DOI: 10.1038/srep24395. URL: http://www.nature.com/articles/srep24395 (visited on 05/31/2022).

Rex, Daniel F. (1950). "Blocking Action in the Middle Troposphere and its Effect upon Regional Climate." en. In: *Tellus* 2.3. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2153-3490.1950.tb00331.x, pp. 196–211. ISSN: 2153-3490. DOI: 10.1111/j.2153-

3490.1950.tb00331.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2153-3490.1950.tb00331.x (visited on 05/02/2024).

Robine, Jean-Marie, Siu Lan K. Cheung, Sophie Le Roy, Herman Van Oyen, Clare Griffiths, Jean-Pierre Michel, and François Richard Herrmann (2008). "Death toll exceeded 70,000 in Europe during the summer of 2003." In: *Comptes Rendus Biologies* 331.2, pp. 171–178. ISSN: 1631-0691. DOI: https://doi.org/10.1016/j.crvi.2007.12.001. URL: https://www.sciencedirect.com/science/article/pii/S1631069107003770.

Rohwer, Christian M., Florian Angeletti, and Hugo Touchette (Nov. 2015). "Convergence of large-deviation estimators." en. In: *Physical Review E* 92.5, p. 052104. ISSN: 1539-3755, 1550-2376. DOI: 10.1103/PhysRevE.92.052104. URL: https://link.aps.org/doi/10.1103/PhysRevE.92.052104 (visited on 02/28/2023).

Rolland, Joran, Freddy Bouchet, and Eric Simonnet (Jan. 2016). "Computing Transition Rates for the 1-D Stochastic Ginzburg–Landau–Allen–Cahn Equation for Finite-Amplitude Noise with a Rare Event Algorithm." en. In: *Journal of Statistical Physics* 162.2, pp. 277–311. ISSN: 0022-4715, 1572-9613. DOI: 10.1007/s10955-015-1417-4. URL: http://link.springer.com/10.1007/s10955-015-1417-4 (visited on 02/09/2023).

Rossby, C., H. Willett, Messrs Holmboe, J. Namias, L. Page, and R. Allen (Jan. 1939). "Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the permanent centers of action atmosphere and the displacements of the permanent centers of action." In: *Journal of Marine Research* 2.1. URL: https://elischolar.library.yale.edu/journal_of_marine_research/544.

Rousi, Efi, Kai Kornhuber, Goratz Beobide-Arsuaga, Fei Luo, and Dim Coumou (Dec. 2022). "Accelerated western European heatwave trends linked to more-persistent double jets over Eurasia." en. In: *Nature Communications* 13.1, p. 3851. ISSN: 2041-1723. DOI: 10.1038/s41467-022-31432-y. URL: https://www.nature.com/articles/s41467-022-31432-y (visited on 10/11/2022).

Rowntree, P. R. and J. A. Bolton (1983). "Simulation of the atmospheric response to soil moisture anomalies over Europe." en. In: *Quarterly Journal of the Royal Meteorological Society* 109.461. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.4 pp. 501–526. ISSN: 1477-870X. DOI: 10.1002/qj.49710946105. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.49710946105 (visited on 06/17/2024).

Rubino, Gerardo and Bruno Tuffin, eds. (Mar. 2009). *Rare Event Simulation using Monte Carlo Methods*. en. 1st ed. Wiley. ISBN: 978-0-470-77269-0 978-0-470-74540-3. DOI: 10.1002/9780470745403. URL: https://onlinelibrary.wiley.com/doi/book/10.1002/9780470745403 (visited on 07/21/2024).

Rudin, Cynthia (May 2019). "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead." en. In: *Nature Machine Intelligence* 1.5. Publisher: Nature Publishing Group, pp. 206–215. ISSN: 2522-5839. DOI: 10.1038/s42256-019-0048-x. URL: https://www.nature.com/articles/s42256-019-0048-x (visited on 06/17/2024).

Ruprich-Robert, Yohan, Thomas Delworth, Rym Msadek, Frederic Castruccio, Stephen Yeager, and Gokhan Danabasoglu (May 2018). "Impacts of the Atlantic Multidecadal Variability on North American Summer Climate and Heat Waves." en. In: *Journal of Climate* 31.9, pp. 3679–3700. ISSN: 0894-8755,

1520-0442. DOI: 10.1175/JCLI-D-17-0270.1. URL: https://journals.ametsoc.org/doi/10.1175/JCLI-D-17-0270.1 (visited on 04/05/2022).

Ruprich-Robert, Yohan, Rym Msadek, Frederic Castruccio, Stephen Yeager, Tom Delworth, and Gokhan Danabasoglu (Apr. 2017). "Assessing the Climate Impacts of the Observed Atlantic Multidecadal Variability Using the GFDL CM2.1 and NCAR CESM1 Global Coupled Models." en. In: *Journal of Climate* 30.8, pp. 2785–2810. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-16-0127.1. URL: https://journals.ametsoc.org/doi/10.1175/JCLI-D-16-0127.1 (visited on 06/09/2022).

Röthlisberger, Matthias, Linda Frossard, Lance F. Bosart, Daniel Keyser, and Olivia Martius (June 2019). "Recurrent Synoptic-Scale Rossby Wave Patterns and Their Effect on the Persistence of Cold and Hot Spells." EN. In: *Journal of Climate* 32.11. Publisher: American Meteorological Society Section: Journal of Climate, pp. 3207–3226. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-18-0664.1. URL: https://journals.ametsoc.org/view/journals/clim/32/11/jcli-d-18-0664.1.xml (visited on 07/14/2024).

Röthlisberger, Matthias, Stephan Pfahl, and Olivia Martius (2016). "Regional-scale jet waviness modulates the occurrence of midlatitude weather extremes." en. In: *Geophysical Research Letters* 43.20. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1 pp. 10,989–10,997. ISSN: 1944-8007. DOI: 10.1002/2016GL070944. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/2016GL070944 (visited on 07/14/2024).

Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor (2000). "Statistical significance of trends and trend differences in layer-average atmospheric temperature time series." en. In: *Journal of Geophysical Research: Atmospheres* 105.D6. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/1999JD901105, pp. 7337–7356. ISSN: 2156-2202. DOI: 10.1029/1999JD901105. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/1999JD901105 (visited on 06/17/2024).

Sauer, Jerome, Jonathan Demaeyer, Giuseppe Zappa, François Massonnet, and Francesco Ragone (Apr. 2024). "Extremes of summer Arctic sea ice reduction investigated with a rare event algorithm." en. In: *Climate Dynamics*. ISSN: 1432-0894. DOI: 10.1007/s00382-024-07160-y. URL: https://doi.org/10.1007/s00382-024-07160-y (visited on 06/24/2024).

Schneider, Tapio, Tobias Bischoff, and Hanna Płotka (Mar. 2015). "Physics of Changes in Synoptic Midlatitude Temperature Variability." EN. In: *Journal of Climate* 28.6. Publisher: American Meteorological Society Section: Journal of Climate, pp. 2312–2331. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-14-00632.1. URL: https://journals.ametsoc.org/view/journals/clim/28/6/jcli-d-14-00632.1.xml (visited on 05/06/2024).

Schubert, Siegfried D., Hailan Wang, Randal D. Koster, Max J. Suarez, and Pavel Ya. Groisman (May 2014). "Northern Eurasian Heat Waves and Droughts." en. In: *Journal of Climate* 27.9, pp. 3169–3207. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-13-00360.1. URL: http://journals.ametsoc.org/doi/10.1175/JCLI-D-13-00360.1 (visited on 06/09/2022).

Schär, Christoph, Pier Luigi Vidale, Daniel Lüthi, Christoph Frei, Christian Häberli, Mark A. Liniger, and Christof Appenzeller (Jan. 2004). "The role of increasing temperature variability in European summer heatwaves." en.

In: *Nature* 427.6972. Publisher: Nature Publishing Group, pp. 332–336. ISSN: 1476-4687. DOI: 10.1038/nature02300. URL: https://www.nature.com/articles/nature02300 (visited on 03/06/2024).

Seneviratne, S. I. et al. (2021). "Weather and Climate Extreme Events in a Changing Climate." In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, pp. 1513–1766.

Seneviratne, Sonia I., Daniel Lüthi, Michael Litschi, and Christoph Schär (Sept. 2006). "Land–atmosphere coupling and climate change in Europe." en. In: *Nature* 443.7108, pp. 205–209. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature05095. URL: http://www.nature.com/articles/nature05095 (visited on 06/09/2022).

Seneviratne, Sonia et al. (2012). "Changes in climate extremes and their impacts on the natural physical environment." en. In: pp. 109–230. DOI: 10.7916/d8-6nbt-s431. URL: https://doi.org/10.7916/d8-6nbt-s431 (visited on 06/18/2024).

Shukla, J. and Y. Mintz (Mar. 1982). "Influence of Land-Surface Evapotranspiration on the Earth's Climate." In: *Science* 215.4539. Publisher: American Association for the Advancement of Science, pp. 1498–1501. DOI: 10.1126/science.215.4539.1498. URL: https://www.science.org/doi/10.1126/science.215.4539.1498 (visited on 06/18/2024).

Sillmann, Jana and Mischa Croci-Maspoli (2009). "Present and future atmospheric blocking and its impact on European mean and extreme climate." en. In: *Geophysical Research Letters* 36.10. _eprint: https://onlinelibrary.wiley.com/doi/pdf/1 ISSN: 1944-8007. DOI: 10.1029/2009GL038259. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2009GL038259 (visited on 06/18/2024).

Simonnet, Eric, Joran Rolland, and Freddy Bouchet (June 2021). "Multistability and Rare Spontaneous Transitions in Barotropic -Plane Turbulence." EN. In: *Journal of the Atmospheric Sciences* 78.6. Publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences, pp. 1889–1911. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/JAS-D-20-0279.1. URL: https://journals.ametsoc.org/view/journals/atsc/78/6/JAS-D-20-0279.1.xml (visited on 07/10/2024).

Stefanon, Marc, Fabio D'Andrea, and Philippe Drobinski (Mar. 2012). "Heatwave classification over Europe and the Mediterranean region." In: *Environmental Research Letters* 7.1, p. 014023. ISSN: 1748-9326. DOI: 10.1088/1748-9326/7/1/014023. URL: https://iopscience.iop.org/article/10.1088/1748-9326/7/1/014023 (visited on 09/26/2022).

Strahan, John, Adam Antoszewski, Chatipat Lorpaiboon, Bodhi P. Vani, Jonathan Weare, and Aaron R. Dinner (May 2021). "Long-Time-Scale Predictions from Short-Trajectory Data: A Benchmark Analysis of the Trp-Cage Miniprotein." In: *Journal of Chemical Theory and Computation* 17.5. Publisher: American Chemical Society, pp. 2948–2963. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.0c00933. URL: https://doi.org/10.1021/acs.jctc.0c00933 (visited on 06/18/2024).

Student (1908). "The Probable Error of a Mean." In: *Biometrika* 6.1. Publisher: [Oxford University Press, Biometrika Trust], pp. 1–25. ISSN: 0006-3444. DOI:

10.2307/2331554. URL: https://www.jstor.org/stable/2331554 (visited on 06/20/2024).

Sutton, Rowan T. and Buwen Dong (Nov. 2012). "Atlantic Ocean influence on a shift in European climate in the 1990s." en. In: *Nature Geoscience* 5.11, pp. 788–792. ISSN: 1752-0894, 1752-0908. DOI: 10.1038/ngeo1595. URL: https://www.nature.com/articles/ngeo1595 (visited on 01/31/2024).

Sutton, Rowan T. and Daniel L. R. Hodson (July 2005). "Atlantic Ocean Forcing of North American and European Summer Climate." In: *Science* 309.5731. Publisher: American Association for the Advancement of Science, pp. 115–118. DOI: 10.1126/science.1109496. URL: https://www.science.org/doi/10.1126/science.1109496 (visited on 01/26/2024).

Teng, Haiyan, Grant Branstator, Hailan Wang, Gerald A. Meehl, and Warren M. Washington (Dec. 2013). "Probability of US heat waves affected by a subseasonal planetary wave pattern." en. In: *Nature Geoscience* 6.12. Publisher: Nature Publishing Group, pp. 1056–1061. ISSN: 1752-0908. DOI: 10.1038/ngeo1988. URL: https://www.nature.com/articles/ngeo1988 (visited on 06/18/2024).

Thiede, Erik H., Dimitrios Giannakis, Aaron R. Dinner, and Jonathan Weare (June 2019). "Galerkin approximation of dynamical quantities using trajectory data." In: *The Journal of Chemical Physics* 150.24, p. 244111. ISSN: 0021-9606. DOI: 10.1063/1.5063730. URL: https://doi.org/10.1063/1.5063730 (visited on 06/18/2024).

Tibaldi, Stefano and Franco Molteni (1990). "On the operational predictability of blocking." en. In: *Tellus A* 42.3. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1034/j.1600-0870.1990.t01-2-00003.x, pp. 343–365. ISSN: 1600-0870. DOI: 10.1034/j.1600-0870.1990.t01-2-00003.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1034/j.1600-0870.1990.t01-2-00003.x (visited on 05/02/2024).

Toms, Benjamin A., Elizabeth A. Barnes, and Imme Ebert-Uphoff (2020). "Physically Interpretable Neural Networks for the Geosciences: Applications to Earth System Variability." en. In: *Journal of Advances in Modeling Earth Systems* 12.9. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS002002, e2019MS002002. ISSN: 1942-2466. DOI: 10.1029/2019MS002002. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2019MS002002 (visited on 06/18/2024).

Touchette (2005). "Legendre-Fenchel transforms in a nutshell." In.

Touchette, Hugo (July 2009). "The large deviation approach to statistical mechanics." en. In: *Physics Reports* 478.1-3, pp. 1–69. ISSN: 03701573. DOI: 10.1016/j.physrep.2009.05.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S0370157309001410 (visited on 02/09/2023).

Vallis, Geoffrey K. (2017). *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. eng. 2nd ed. Cambridge: Cambridge university press. ISBN: 978-1-107-06550-5.

Vargas Zeppetello, L. R. and D. S. Battisti (2020). "Projected Increases in Monthly Midlatitude Summertime Temperature Variance Over Land Are Driven by Local Thermodynamics." en. In: *Geophysical Research Letters* 47.19. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL090197, e2020GL090197. ISSN: 1944-8007. DOI: 10.1029/2020GL090197. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL090197 (visited on 06/18/2024).

Vasicek, Oldrich (Nov. 1977). "An equilibrium characterization of the term structure." In: *Journal of Financial Economics* 5.2, pp. 177–188. ISSN: 0304-405X. DOI: 10.1016/0304-405X(77)90016-2. URL: https://www.sciencedirect.com/science/article/pii/0304405X77900162 (visited on 06/13/2024).

Vautard, R. (Oct. 1990). "Multiple Weather Regimes over the North Atlantic: Analysis of Precursors and Successors." EN. In: *Monthly Weather Review* 118.10. Publisher: American Meteorological Society Section: Monthly Weather Review, pp. 2056–2081. ISSN: 1520-0493, 0027-0644. DOI: 10.1175/1520-0493(1990)118<2056:MWROTN>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/mwre/118/10/1520-0493_1990_118_2056_mwrotn_2_0_co_2.xml (visited on 06/12/2024).

Vautard, R., P. Yiou, F. D'Andrea, N. de Noblet, N. Viovy, C. Cassou, J. Polcher, P. Ciais, M. Kageyama, and Y. Fan (Apr. 2007). "Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit." en. In: *Geophysical Research Letters* 34.7, p. L07711. ISSN: 0094-8276. DOI: 10.1029/2006GL028001. URL: http://doi.wiley.com/10.1029/2006GL028001 (visited on 06/03/2022).

Vautard, Robert et al. (Oct. 2023). "Heat extremes in Western Europe increasing faster than simulated due to atmospheric circulation trends." en. In: *Nature Communications* 14.1. Publisher: Nature Publishing Group, p. 6803. ISSN: 2041-1723. DOI: 10.1038/s41467-023-42143-3. URL: https://www.nature.com/articles/s41467-023-42143-3 (visited on 03/29/2024).

Ventsel', A. D. and M. I. Freidlin (Feb. 1970). "ON SMALL RANDOM PERTURBATIONS OF DYNAMICAL SYSTEMS." en. In: *Russian Mathematical Surveys* 25.1. Publisher: IOP Publishing, p. 1. ISSN: 0036-0279. DOI: 10.1070/RM1970v025n01ABEH001254. URL: https://iopscience.iop.org/article/10.1070/RM1970v025n01ABEH001254/meta (visited on 06/20/2024).

Vogel, M. M., R. Orth, F. Cheruy, S. Hagemann, R. Lorenz, B. J. J. M. van den Hurk, and S. I. Seneviratne (2017). "Regional amplification of projected changes in extreme temperatures strongly controlled by soil moisture-temperature feedbacks." en. In: *Geophysical Research Letters* 44.3. _eprint: https://onlinelibrary.wiley.com/, pp. 1511–1519. ISSN: 1944-8007. DOI: 10.1002/2016GL071235. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/2016GL071235 (visited on 06/11/2024).

Voldoire, A. et al. (July 2019). "Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1." en. In: *Journal of Advances in Modeling Earth Systems* 11.7, pp. 2177–2213. ISSN: 1942-2466, 1942-2466. DOI: 10.1029/2019MS001683. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001683 (visited on 04/05/2022).

White, Rachel H., Kai Kornhuber, Olivia Martius, and Volkmar Wirth (Mar. 2022). "From Atmospheric Waves to Heatwaves: A Waveguide Perspective for Understanding and Predicting Concurrent, Persistent, and Extreme Extratropical Weather." In: *Bulletin of the American Meteorological Society* 103.3, E923–E935. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-21-0170.1. URL: https://journals.ametsoc.org/view/journals/bams/103/3/BAMS-D-21-0170.1.xml (visited on 10/11/2022).

White, Rachel H. et al. (Feb. 2023). "The unprecedented Pacific Northwest heatwave of June 2021." en. In: *Nature Communications* 14.1. Publisher: Nature

Publishing Group, p. 727. ISSN: 2041-1723. DOI: 10.1038/s41467-023-36289-3. URL: https://www.nature.com/articles/s41467-023-36289-3 (visited on 06/19/2024).

Wirth, Volkmar and Christopher Polster (Oct. 2021). "The Problem of Diagnosing Jet Waveguidability in the Presence of Large-Amplitude Eddies." In: DOI: 10.1175/JAS-D-20-0292.1. URL: https://journals.ametsoc.org/view/journals/atsc/78/10/JAS-D-20-0292.1.xml (visited on 07/13/2024).

Wirth, Volkmar, Michael Riemer, Edmund K. M. Chang, and Olivia Martius (July 2018). "Rossby Wave Packets on the Midlatitude Waveguide—A Review." In: DOI: 10.1175/MWR-D-16-0483.1. URL: https://journals.ametsoc.org/view/journals/mwre/146/7/mwr-d-16-0483.1.xml (visited on 07/13/2024).

Woollings, Tim (Aug. 2010). "Dynamical influences on European climate: an uncertain future." In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368.1924. Publisher: Royal Society, pp. 3733–3756. DOI: 10.1098/rsta.2010.0040. URL: https://royalsocietypublishing.org/doi/10.1098/rsta.2010.0040 (visited on 04/02/2024).

Woollings, Tim, David Barriopedro, John Methven, Seok-Woo Son, Olivia Martius, Ben Harvey, Jana Sillmann, Anthony R. Lupo, and Sonia Seneviratne (Sept. 2018). "Blocking and its Response to Climate Change." en. In: *Current Climate Change Reports* 4.3, pp. 287–300. ISSN: 2198-6061. DOI: 10.1007/s40641-018-0108-z. URL: https://doi.org/10.1007/s40641-018-0108-z (visited on 05/02/2024).

Wouters, J and F Bouchet (Sept. 2016). "Rare event computation in deterministic chaotic systems using genealogical particle analysis." In: *Journal of Physics A: Mathematical and Theoretical* 49.37, p. 374002. ISSN: 1751-8113, 1751-8121. DOI: 10.1088/1751-8113/49/37/374002. URL: https://iopscience.iop.org/article/10.1088/1751-8113/49/37/374002 (visited on 02/09/2023).

Wouters, Jeroen, Reinhard K. H. Schiemann, and Len C. Shaffrey (Apr. 2023). "Rare Event Simulation of Extreme European Winter Rainfall in an Intermediate Complexity Climate Model." en. In: *Journal of Advances in Modeling Earth Systems* 15.4, e2022MS003537. ISSN: 1942-2466, 1942-2466. DOI: 10.1029/2022MS003537. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022MS003537 (visited on 10/05/2023).

Wulff, C. Ole, Richard J. Greatbatch, Daniela I. V. Domeisen, Gereon Gollan, and Felicitas Hansen (2017). "Tropical Forcing of the Summer East Atlantic Pattern." en. In: *Geophysical Research Letters* 44.21. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.100 pp. 11,166–11,173. ISSN: 1944-8007. DOI: 10.1002/2017GL075493. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/2017GL075493 (visited on 07/11/2024).

Yang, Xiaosong and Edmund K. M. Chang (Feb. 2006). "Variability of the Southern Hemisphere Winter Split Flow—A Case of Two-Way Reinforcement between Mean Flow and Eddy Anomalies." en. In: *Journal of the Atmospheric Sciences* 63.2, pp. 634–650. ISSN: 1520-0469, 0022-4928. DOI: 10.1175/JAS3643.1. URL: https://journals.ametsoc.org/doi/10.1175/JAS3643.1 (visited on 06/27/2023).

Yeager, S. G. and J. I. Robson (June 2017). "Recent Progress in Understanding and Predicting Atlantic Decadal Climate Variability." en. In: *Current Climate Change Reports* 3.2, pp. 112–127. ISSN: 2198-6061. DOI: 10.1007/s40641-017-0064-z. URL: https://doi.org/10.1007/s40641-017-0064-z (visited on 02/02/2024).

Yiou, P. (Apr. 2014). "AnaWEGE: a weather generator based on analogues of atmospheric circulation." English. In: *Geoscientific Model Development* 7.2. Publisher: Copernicus GmbH, pp. 531–543. ISSN: 1991-959X. DOI: 10.5194/gmd-7-531-2014. URL: https://gmd.copernicus.org/articles/7/531/2014/ (visited on 07/12/2024).

Zampieri, Matteo, Fabio D'Andrea, Robert Vautard, Philippe Ciais, Nathalie De Noblet-Ducoudré, and Pascal Yiou (Sept. 2009). "Hot European Summers and the Role of Soil Moisture in the Propagation of Mediterranean Drought." en. In: *Journal of Climate* 22.18, pp. 4747–4758. ISSN: 1520-0442, 0894-8755. DOI: 10.1175/2009JCLI2568.1. URL: http://journals.ametsoc.org/doi/10.1175/2009JCLI2568.1 (visited on 01/26/2024).

Zeppetello, Lucas R. Vargas, David S. Battisti, and Marcia B. Baker (Apr. 2022). "The Physics of Heat Waves: What Causes Extremely High Summertime Temperatures?" EN. In: *Journal of Climate* 35.7. Publisher: American Meteorological Society Section: Journal of Climate, pp. 2231–2251. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-21-0236.1. URL: https://journals.ametsoc.org/view/journals/clim/35/7/JCLI-D-21-0236.1.xml (visited on 06/18/2024).

Zhang, Rong, Rowan Sutton, Gokhan Danabasoglu, Young-Oh Kwon, Robert Marsh, Stephen G. Yeager, Daniel E. Amrhein, and Christopher M. Little (2019). "A Review of the Role of the Atlantic Meridional Overturning Circulation in Atlantic Multidecadal Variability and Associated Climate Impacts." en. In: *Reviews of Geophysics* 57.2. _eprint: https://onlinelibrary.wiley.com/doi/pdf/1 pp. 316–375. ISSN: 1944-9208. DOI: 10.1029/2019RG000644. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2019RG000644 (visited on 05/14/2024).

Zhou, Sha, A. Park Williams, Alexis M. Berg, Benjamin I. Cook, Yao Zhang, Stefan Hagemann, Ruth Lorenz, Sonia I. Seneviratne, and Pierre Gentine (Sept. 2019). "Land–atmosphere feedbacks exacerbate concurrent soil drought and atmospheric aridity." In: *Proceedings of the National Academy of Sciences* 116.38. Publisher: Proceedings of the National Academy of Sciences, pp. 18848–18853. DOI: 10.1073/pnas.1904955116. URL: https://www.pnas.org/doi/10.1073/pnas.1904955116 (visited on 06/18/2024).

Zhou, Yefan and Zhiwei Wu (2016). "Possible impacts of mega-El Niño/Southern Oscillation and Atlantic Multidecadal Oscillation on Eurasian heatwave frequency variability." en. In: *Quarterly Journal of the Royal Meteorological Society* 142.697. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2759, pp. 1647–1661. ISSN: 1477-870X. DOI: 10.1002/qj.2759. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.2759 (visited on 06/18/2024).

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/

Final Version as of December 17, 2024 (`classicthesis` version 4.2).