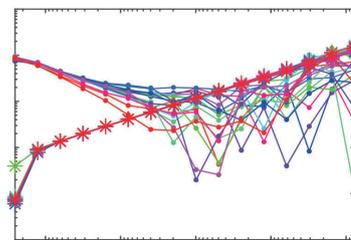
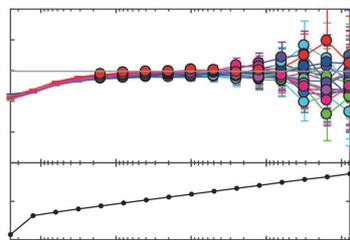
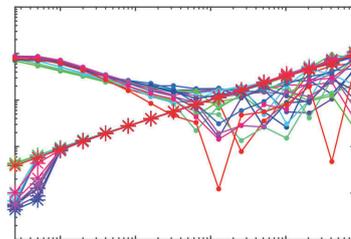
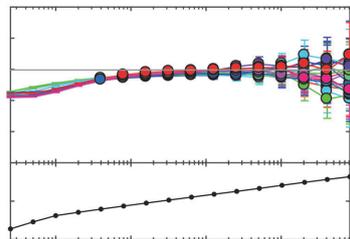




Mathematics of extreme events in atmospheric models



Vera Melinda Galfi

Hamburg 2018

Hinweis

Die Berichte zur Erdsystemforschung werden vom Max-Planck-Institut für Meteorologie in Hamburg in unregelmäßiger Abfolge herausgegeben.

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Die "Berichte zur Erdsystemforschung" führen die vorherigen Reihen "Reports" und "Examensarbeiten" weiter.

Anschrift / Address

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg
Deutschland

Tel./Phone: +49 (0)40 4 11 73 - 0
Fax: +49 (0)40 4 11 73 - 298

name.surname@mpimet.mpg.de
www.mpimet.mpg.de

Notice

The Reports on Earth System Science are published by the Max Planck Institute for Meteorology in Hamburg. They appear in irregular intervals.

They contain scientific and technical contributions, including Ph. D. theses.

The Reports do not necessarily reflect the opinion of the Institute.

The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.

Layout

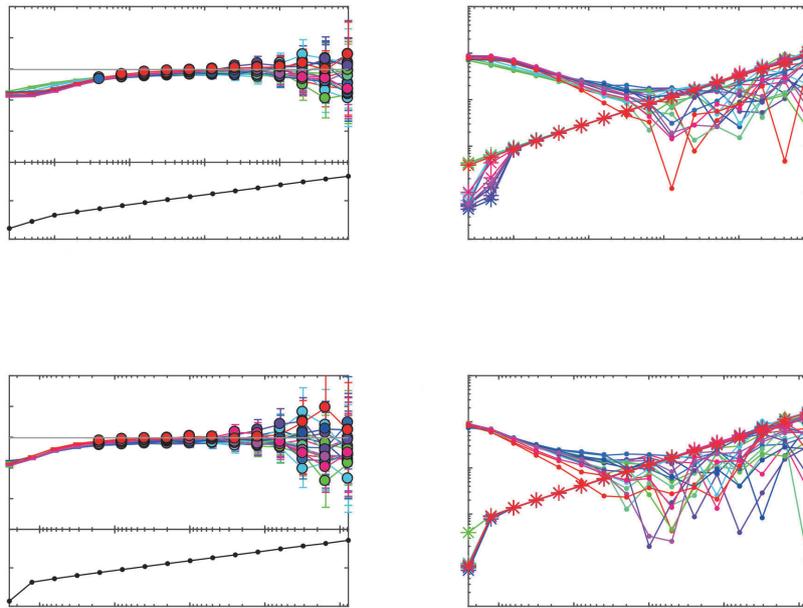
Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Photos below: ©MPI-M
Photos on the back from left to right:
Christian Klepp, Jochem Marotzke,
Christian Klepp, Clotilde Dubois,
Christian Klepp, Katsumasa Tanaka



Mathematics of extreme events in atmospheric models



Dissertation with the aim of achieving a doctoral degree at the
Faculty of Mathematics, Informatics and Natural Sciences
Department of Earth Sciences of Universität Hamburg
submitted by

Vera Melinda Galfi

Hamburg 2018

Vera Melinda Galfi

Max-Planck-Institut für Meteorologie
The International Max Planck Research School on Earth System Modelling
(IMPRS-ESM)
Bundesstrasse 53
20146 Hamburg

Universität Hamburg
Geowissenschaften
Meteorologisches Institut
Bundesstr. 55
20146 Hamburg

Tag der Disputation: 29.10.2018

Folgende Gutachter empfehlen die Annahme der Dissertation:

Prof. Dr. Valerio Lucarini
Prof. Dr. Johanna Baehr

Abstract

The understanding of extreme events is important for several fields of research. From a practical point of view, the prediction of extreme events is necessary in order to allow for adaptation and damage reduction. From a theoretical point of view, extreme fluctuations of the system under study provide valuable information about the system itself, thus revealing processes behind these events.

In this thesis we adopt the theoretical point of view. We study extremes of atmospheric observables (total energy and near-surface temperature) in two numerical models, a quasi-geostrophic (QG) two-layer atmospheric model and a simplified general circulation model of the atmosphere (PUMA - Portable University Model of the Atmosphere). Our general aim is to provide a mathematical background for the study of extreme events in atmospheric flows. To achieve this we use two rigorous mathematical frameworks: Extreme Value Theory (EVT) and Large Deviation Theory (LDT). In both cases, we connect properties of extreme or rare events to general properties of the system (or of the system state) producing these events, by taking perspectives based on dynamical systems theory and statistical mechanics.

In case of energy extremes in the QG model, we search for the universality of extreme events, and find a connection between statistical properties of extremes of different energy observables and general dynamical properties of the system. This is in accordance with theoretical results developed for chaotic dynamical systems. In the case of extremes of temperature averages in the PUMA model, we aim at analysing persistent temperature extreme events (heat waves or cold spells) based on LDT. Here, we also find universal characteristics in form of a connection between temporal and spatial (or spatio-temporal) averages, while the spatial averaging is performed along latitudes. However, we are able to explore universality in both cases only if the asymptotic limit demanded by the theories is valid, i.e. on large temporal and/or spatial scales.

On smaller scales, the effect of correlations prevents universal results, producing pre-asymptotic deviations. Additionally, if the system state is not chaotic enough, asymptotic levels cannot be reached, at least on realistic finite scales. Furthermore, we realise that analysing persistent events based on asymptotic theories is far from being a trivial task, however, it can be done by averaging in two dimensions: first, on intermediate scales in space and, second, obtaining a LD limit in time.

Zusammenfassung

Die wissenschaftliche Untersuchung von Extremwerten ist wichtig für zahlreiche Forschungsgebiete. Aus praktischer Sicht ist die Vorhersage von Extremwerten für Anpassungsmaßnahmen und Schadensreduzierung unerlässlich. Aus theoretischer Sicht liefern extreme Fluktuationen des untersuchten Systems wertvolle Informationen über das System selbst und offenbaren dadurch die Prozesse hinter diesen Fluktuationen.

In dieser Doktorarbeit wählen wir die theoretische Herangehensweise. Wir analysieren Extremwerte von atmosphärischen Variablen (totale Energie und oberflächennahe Temperatur) in zwei numerischen Modellen, ein quasi-geostrophisches (QG) Zwei-Schichten-Modell und ein vereinfachtes, globales Zirkulationsmodell der Atmosphäre (PUMA - Portable University Model of the Atmosphere). Unser Hauptziel ist die Erstellung eines mathematischen Rahmens für die Untersuchung von Extremwerten in atmosphärischen Strömungen. Um dies zu erreichen verwenden wir zwei rigorose, mathematische Theorien: Extremwertstatistik (EWS) und die Theorie der großen Abweichungen (TGA). In beiden Fällen untersuchen wir - aus der Perspektive der Theorie dynamischer Systeme und der statistischen Mechanik - die Verbindung zwischen den Eigenschaften von extremen und seltenen Ereignissen und allgemeinen Systemeigenschaften (oder Systemzustandseigenschaften), die diese Ergebnisse hervorrufen.

Im Rahmen der Untersuchung der Energieextreme im QG-Modell forschen wir nach der Universalität von Extremwerten, und wir finden eine Verbindung zwischen den statistischen Eigenschaften von Extremwerten und generellen, dynamischen Eigenschaften des Systems. Dies entspricht theoretischen Ergebnissen die für chaotische dynamische Systeme hergeleitet wurden. Im Falle der Extremereignisse von gemittelten Temperaturwerten im PUMA-Modell untersuchen wir andauernde Extremereignisse (wie Hitze- oder Kältewellen) basierend auf TGA. Auch in diesem Fall entdecken wir universale Eigenschaften in Form einer Verknüpfung zwischen zeitlich und räumlich (oder in beiden Dimensionen) gemittelten Größen entlang eines Breitengrades. Nichtsdestotrotz können wir in beiden Fällen die Universalität nur dann beobachten, wenn die theoretisch vorgeschriebene, asymptotische Annäherung gilt, d.h. für große Zeit- und/oder Raumskalen.

Allerdings verhindert die Wirkung der Korrelationen universale Ergebnisse auf kleineren Skalen und verursacht präasymptotische Abweichungen. Außerdem kann die

asymptotische Annäherung nicht erreicht werden, zumindest auf realistischen, endlichen Skalen, wenn das System nicht chaotisch genug ist. Darüber hinaus merken wir, dass die Untersuchung der andauernden Extremereignissen alles andere als einfach ist. Wir zeigen dennoch dass es möglich ist, wenn man in zwei Dimensionen mittelt: erstens im Raum über mittleren Skalen, und zweitens in der Zeit, wobei eine Annäherung für große Abweichungen erreicht werden soll.

Contents

Abstract	I
Zusammenfassung	III
1 Introduction	1
1.1 Some properties of chaotic dynamical systems	2
1.2 The atmosphere as a chaotic dynamical system	6
1.3 Extreme events and general system properties	11
1.4 Large deviations and persistent extreme events	14
1.5 This thesis	17
2 Extreme events in chaotic dynamical systems	23
2.1 Extreme Value Theory for dynamical systems	23
2.2 Large Deviation Theory	29
3 Atmospheric models	33
3.1 The quasi-geostrophic two-layer atmospheric model	34
3.2 The simplified global circulation model PUMA	39
4 Extreme events of energy observables in the two-layer quasi-geostrophic atmospheric model	45
4.1 Strong forcing	45
4.2 Weak forcing	56
4.3 Summary and discussion	60
5 Large deviations of temperature in the general circulation model of the atmosphere PUMA	65
5.1 The link between temporal and zonal large deviations	69
5.2 Spatio-temporal large deviations	73
5.3 Return levels of large deviations	78
5.4 Summary and discussion	80
6 Conclusions and outlook	85
References	i

List of acronyms	xi
List of important symbols in Chapter 4	xiii
List of important symbols in Chapter 5	xv

1 Introduction

Extreme values of a certain observable are usually defined as values that lie in the tail of the probability density function (pdf). This has two major consequences: first, extreme values have a low probability of occurrence meaning that they are rare and, second, their magnitude is unusually large or low. Due to these two major properties, *extreme events* affecting different areas of human life are perceived as “unexpected” events. Climate or weather extremes - like storms, tropical cyclones, floods, forest fires - can have a huge impact, producing severe financial, human, biological damages (Easterling et al., 2000; IPCC, 2012; Lobell et al., 2011). However, these negative effects can become even more serious if the extreme event is persistent. During a *persistent extreme event* - like a heat wave, cold spell, or drought - the extraordinary circumstances continue to exist over a prolonged period. This long-lasting stress created by a persistent extreme event can push whole societies, ecosystems, infrastructures to the limit (Robinson, 2001; WHO, 2004; IPCC, 2012).

The present uncertainty in the future projections of climate extremes makes their study even more urgent and crucial. According to IPCC (2013), the changes in temperature extremes over most land areas since the mid-20th century are consistent between numerical models and observations, but models perform worse in simulating changes of other extremes, like precipitation extremes, droughts, as well as tropical cyclone tracks and intensity.

The investigation of extreme events is extremely relevant for a range of disciplines in mathematical, natural, social sciences and engineering. Understanding the large fluctuations of the system of interest is of great importance from a practical point of view - when it comes to assess the risk associated to low probability and high impact events in order to prevent from or prepare for damages - but also from a theoretical point of view when one tries to explore the connection between “unexpected” events and the properties of the system generating them. In this thesis we adopt the theoretical point of view, the system of interest being represented by numerical atmospheric models. We seek answers for questions like the following ones. Which general dynamical properties represent the system producing the extreme events? Or the other way around: What can extreme, i.e. “extraordinary”, events reveal about general dynamical properties of the system under investigation? What are the main characteristics of system states leading

to persistent extreme events?

This introduction is meant to prepare the ground for the relatively complex mathematical and physical theories and concepts used to obtain the results of the thesis. We start with introducing the basic concepts of dynamical systems theory applied in this work, after which we describe general notions of the atmospheric circulation, and argue that the atmosphere can be considered a high dimensional dynamical system. We define the main objectives and the structure of the thesis at the end of this chapter, after discussing the connection between extreme events and general properties of the dynamical system, as well as, between persistent extreme events and large deviations from the mean system state. These two relationships represent at the same time the two main topics around which the thesis is built.

1.1 Some properties of chaotic dynamical systems

Before discussing the arguments according to which the atmosphere can be interpreted as a high dimensional chaotic dynamical system, we introduce the basic concepts of dynamical systems theory necessary to understand the applied methods and theories. This section is mostly based on the work of Ott (1993). We wish to remark that we do not strive for mathematical rigour nor for completeness, our aim being to provide a concise and general overview of the relevant concepts.

The state of a *dynamical system* evolving forward in time t in an N -dimensional Euclidean space, called the *phase space* Γ , from an initial condition \mathbf{x}_0 at $t = 0$ can be described as $\mathbf{x}(t, \mathbf{x}_0) = \Pi^t(\mathbf{x}_0)$, where Π^t is the evolution operator. If time is a continuous variable, the system is called a *flow*, and can be described by a set of N ordinary differential equations:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{F}(\mathbf{x}(t)). \quad (1.1)$$

Whereas if time is a discrete variable, the system is a *map*, and evolves based on the following set of N equations:

$$\mathbf{x}_{n+1} = \mathbf{M}(\mathbf{x}_n), \quad (1.2)$$

with n denoting the discrete time variable $n = 0, 1, 2, \dots$. Dynamical systems can exhibit different types of behaviour, such as stationary, periodic, and *chaotic*. A chaotic dynamical system is sensitive against arbitrarily small perturbations of the initial conditions, exhibiting a significantly different temporal evolution for different initial conditions.

In the N -dimensional phase space Γ , a specific state of the system is represented by a single point, and a state varying in accordance with the evolution equation is described

by a *trajectory*. Hence, the dynamics of the system is connected to geometry in a multidimensional space (Lorenz, 1984). Some dynamical systems are characterised by the presence of attracting invariant sets, i.e. bounded subsets of Γ representing a state of permanent regime to which regions of initial conditions of nonzero phase space volume converge asymptotically as time increases (Ott, 1993). We refer to these invariant sets as *attractors*, since trajectories become “trapped” after reaching them. The existence of an attractor is characteristic for *dissipative systems*, in which the phase space volume contracts as time evolves, i.e. $\nabla \cdot \mathbf{F} < 0$.

The geometry of the attractor is related to the dynamical behaviour of the system: the attractor of a stationary system is a fixed point, the attractor of a periodic system is a closed curve, called limit cycle, whereas the attractor of a chaotic dynamical system is an object with fractal geometry, called *strange attractor*. Fig. 1.1 represents schematically the three attractor types.

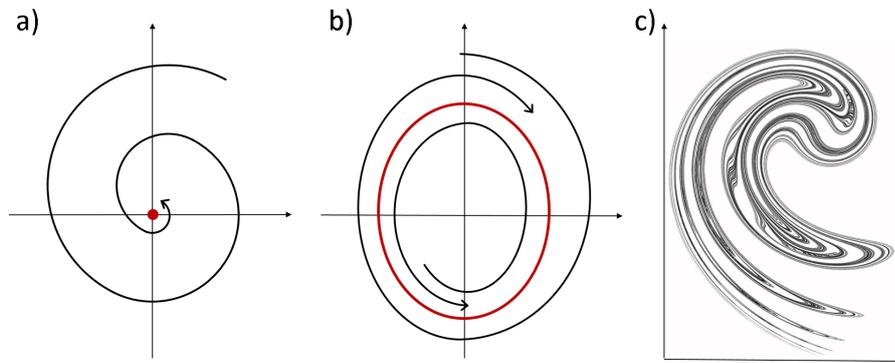


Figure 1.1: Schematic representation of a) a fixed point, b) a limit cycle, and c) a strange attractor, reproduced from Prokofiev (2010). In a) and b) the red point and the red line represent the attractors, and the black curves illustrate possible trajectories.

Another important concept in dynamical systems theory is the measure μ corresponding to a probability density. A probability measure can be defined, even if a density cannot, thus it is a more general concept (Ott, 1993). In the context of invariant sets one deals with *invariant measures* as well. The probability measure μ is invariant if satisfies $\mu[f^{-t}(E)] = \mu(E)$, $t > 0$, where E is a subset of points of \mathbb{R}^N , and $f^{-t}(E)$ is the set obtained by evolving these points backwards in time (Eckmann and Ruelle, 1985).

A dynamical system is *ergodic* with respect to an invariant measure if and only if all invariant sets have either full measure or none (Ott, 1993). This statement basically prohibits that the attractor consists of two or more disjoint sets, i.e. that trajectories with different initial conditions get attracted by different disjoint regions of the attractor. Analogously, an invariant measure μ is ergodic, if it is indecomposable. According to the *ergodic theorem*, long-term integrals in time and integrals in phase space are equal

for almost all initial conditions x_0 with respect to an ergodic measure μ (Eckmann and Ruelle, 1985):

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \varphi[f^t x_0] dt = \int \varphi(x) \mu(dx), \quad (1.3)$$

where φ represents a continuous function, and f^t is the time operator. The above equation considers flows. In the case of maps, the temporal integral can be replaced by a long-term average in time.

We can study properties of dynamical systems by analysing geometrical properties of their attractors or by focusing on characteristics of their dynamics. To describe the dynamics of the system, one often uses the so-called *Lyapunov exponent*, which, in the one dimensional case, describes the average exponential divergence rate of trajectories due to an infinitesimally small perturbation of the initial condition. Now, let us consider an N-dimensional map \mathbf{M} where \mathbf{x}_0 represents the initial condition and \mathbf{x}_n ($n = 0, 1, 2, \dots$) the corresponding orbit. If the initial displacement from \mathbf{x}_0 is in the direction of a tangent vector \mathbf{y}_0 , the Lyapunov exponent can be defined as (Ott, 1993):

$$\lambda(\mathbf{x}_0, \mathbf{u}_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln(|\mathbf{y}_n|/|\mathbf{y}_0|) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln |J_{\mathbf{M}}^n(\mathbf{x}_0) \cdot \mathbf{u}_0|, \quad (1.4)$$

where $|\mathbf{y}_n|/|\mathbf{y}_0|$ represents the factor by which the initial displacement grows, and $\mathbf{u}_0 = \mathbf{y}_0/|\mathbf{y}_0|$ is the initial orientation of the displacement. $J_{\mathbf{M}}^n(\mathbf{x}_0) = J_{\mathbf{M}}(\mathbf{x}_{n-1}) \cdot J_{\mathbf{M}}(\mathbf{x}_{n-2}) \cdot \dots \cdot J_{\mathbf{M}}(\mathbf{x}_0)$, with $J_{\mathbf{M}}(\mathbf{x}) := \left(\frac{\partial M_i}{\partial x_j} \right)_{i=1, \dots, N, j=1, \dots, N}$ denoting the Jacobian matrix of \mathbf{M} at \mathbf{x} .

Oseledets' multiplicative ergodic theorem guarantees the existence of N Lyapunov exponents in an N-dimensional ergodic system, under very general circumstances (Oseledets, 1968). If the system is ergodic, the set of values is the same for almost every x_0 (Ruelle, 1989). Each exponent describes the stretching or contracting properties of the attractor in the respective dimension. The Lyapunov exponents are usually sorted in an increasing order $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_N$, so that for a chaotic system $\lambda_1 > 0$. The value of λ_1 and the number of positive exponents are important indicators for the chaoticity of the dynamical system.

An intuitive way to describe a geometrical object embedded in an N-dimensional space is to specify its dimension. To achieve this, one usually covers the object of interest, in this case the attractor, with a number n of N-dimensional cubes with side length ϵ , and looks at the change in n by reducing ϵ . One usually obtains that the number of cubes n scales as $n \sim \epsilon^{D_0}$ if ϵ is small enough. For $\epsilon \rightarrow 0$, D_0 is called the *box-counting dimension*. However, this is a purely geometrical indicator, and in case of chaotic dynamical systems we are not only interested in the number of cubes needed to cover the attractor, but additionally we would like to know how often and for how much time each cube is visited by trajectories. Therefore, it is reasonable to weight each cube based on a probability measure.

To express the change of information amount contained in the cubes as ϵ is reduced, we define the *information dimension*:

$$D_1 = \lim_{\epsilon \rightarrow 0} \frac{-H_\epsilon}{\ln \epsilon}. \quad (1.5)$$

$H_\epsilon = -\sum_{i=1}^{n(\epsilon)} \mu_i \ln \mu_i$ is the Shannon entropy and measures the amount of information needed to describe the system state with a certain accuracy $1/\epsilon$. μ_i is referred to as the natural probability measure, representing the fraction of time a long orbit spends in a given region of the attractor, in this case inside an N-dimensional cube (Ott, 1993).

The above dimensions are global characteristics of the attractor. However, one might be interested in local attractor properties at a certain point ζ . We define the *local information dimension*, also called point-wise dimension, of a probability measure μ :

$$D_\zeta = \lim_{\epsilon \rightarrow 0} \frac{\ln \mu(B_\epsilon(\zeta))}{\ln \epsilon}. \quad (1.6)$$

where $B_\epsilon(\zeta)$ is a ball with radius ϵ centred at ζ (Ott, 1993).

A further important and more complex concept of dimension is the so-called *Kaplan-York* (or Lyapunov) *dimension*, which expresses dynamical properties of the system. It is obtained based on the already introduced Lyapunov exponents λ :

$$D_{KY} = m + \frac{\sum_{i=1}^n \lambda_i}{|\lambda_{m+1}|} \quad (1.7)$$

where m is such that $\sum_{i=1}^n \lambda_k$ is positive and $\sum_{i=1}^{n+1} \lambda_i$ is negative (Ott, 1993). D_{KY} measures the dimension of volume elements, which neither grow nor shrink as they evolve in phase space, thus estimating the dimension of the invariant set in a dynamical way.

We notice that although the above quantities are all called dimensions, they are different definitions of this notion. Whereas D_0 characterises the attractor from a purely geometrical perspective, D_1 considers additionally the probability measure of recurrences on the attractor. D_ζ is a local version of D_1 , and D_{KY} is defined based on the dynamics. However, one intuitively feels that they should be somehow connected. In some cases, as we see below, some of the dimensions are equal.

Considering that an ergodic measure μ is indecomposable, it also means that μ is the same at almost every point of the attractor (Ott, 1993). Therefore, in ergodic systems, locally defined invariant measures are equal to global ones, thus $D_1 = D_\zeta$. Showing the connection between the differently defined dimensions D_1 and D_{KY} is a little bit more complicated, however, it has been conjectured that $D_1 = D_{KY}$ in "typical systems".

In search for a mathematically more justified framework instead of “typical systems”, one finds the concept of *SRB* (Sinai, Ruelle, Bowen) *measures*. These are measures with smooth density in the stretching or unstable direction of the dynamical system, providing an equivalence between the rate of creation of information and the mean rate of expansion of volume elements (Gallavotti and Cohen, 1995; Eckmann and Ruelle, 1985). In a system where all the information is contained by the unstable manifold already, where, at the same time, the dynamics takes place, one obtains the equality $D_1 = D_{KY}$. For other systems, $D_1 \leq D_{KY}$. Additionally, SRB measures are stable against weak stochastic perturbations (Eckmann and Ruelle, 1985; Ruelle, 1989). As a result of this stability, SRB measures can actually be observed, and thus are also called “physical measures”.

Axiom A systems are a special class of dynamical systems that possess an SRB measure. They are uniformly hyperbolic on their attractor, so that stable and unstable directions are well separated. Due to these convenient characteristics, one of the great merits of Axiom A dynamical systems is that they represent the perfect framework for deriving rigorous and robust statistical mechanical properties for purely deterministic background dynamics. Despite having deterministic dynamics, when looking at their observables, they behave just like generators of stochastic processes.

While Axiom A systems are rather special and indeed not generic, they have great relevance for applications if one takes into account the *chaotic hypothesis*, which indicates that high-dimensional chaotic systems behave at all practical purposes as if they were Axiom A (Gallavotti and Cohen, 1995; Gallavotti, 2014). Thus, Axiom A systems represent good effective models for chaotic systems with many degrees of freedom. Lucarini et al. (2014b) point out that, whenever we perform numerical simulations, we implicitly assume that the system is Axiom A-equivalent.

1.2 The atmosphere as a chaotic dynamical system

The object of study in this thesis is the *Earth’s atmosphere*, which we consider to be a very high dimensional chaotic dynamical system. Besides the hydrosphere, cryosphere, the land surface and the biosphere, the atmosphere is part of an even more complex system, the climate system. The processes in the atmosphere are extremely complex due to the non-trivial interactions with other system components, however, its intrinsic motions are already difficult to describe precisely, as we show in the following.

The source of potential energy necessary for atmospheric motions comes from the Sun: the net radiation at the Equator is higher than the one at the Poles, thus creating zonal

mean available potential energy. According to basic physical laws, the atmosphere tries to compensate this meridional difference in radiative forcing. Thus, hot air rises at the Equator and flows polewards. During this process it is directed to the west by the Coriolis force, cools down and sinks at around latitude 30° , and flows again equatorwards near to the surface. The Coriolis force, which is an apparent force acting on a motion due to a rotating frame of reference, deviates the equatorward flow to the east, creating the so-called trade winds. These are the basic features of the so-called *Hadley cell*, which represents atmospheric motions averaged along latitudes and over a long time period. An equator-to-pole Hadley cell is mathematically possible, and can be encountered for example on slowly rotating planets, like the Venus, on Earth, however, it represents a baroclinically unstable circulation.

Baroclinic instability evolves in rotating stratified fluids with a horizontal temperature gradient, and is responsible for generating the dominant weather systems of the mid-latitudes, the *cyclones* and *anticyclones*, also called *eddies*. Baroclinic instability grows by converting available potential energy maintained by the horizontal temperature gradient, which is associated to the vertical shear of the zonal mean wind through the *thermal wind balance*. Hence, it is responsible for the existence of the strong wind band, the *jet stream*, near to the tropopause (Holton, 2004). The vertically sheared zonal mean wind in the mid-latitude becomes unstable generating baroclinic waves, which transport heat against the horizontal temperature gradient, i.e. polewards. Baroclinic waves convert the available potential energy into kinetic energy, and, thus, maintain the atmospheric circulation against dissipation. On the one hand, baroclinic instability leads to a negative feedback assuring the overall stability of the atmospheric circulation, on the other hand, it is also the main reason for the error growth in the short-term weather forecast for the mid-latitudes (Lucarini et al., 2014a). Figure 1.2 illustrates the main components of the global atmospheric circulation in an idealised form (Stull, 2015-2016). For an extensive treatment of baroclinic instability and the global circulation of the atmosphere the reader is referred to Holton (2004).

From a mathematical point of view, atmospheric motions are described by a set of *non-linear partial differential equations*. The basic atmospheric equations express the temporal and spatial evolution of the three-dimensional wind vector $\mathbf{U} = (u, v, w)$, density ρ , and potential temperature θ^1 . They guarantee for the conservation of momentum, mass, and energy (Ott, 1993):

$$\frac{D\mathbf{U}}{Dt} = -2\boldsymbol{\Omega} \times \mathbf{U} - \frac{1}{\rho} \nabla p + \mathbf{g} + \mathbf{F}_r, \quad (1.8)$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{U}) = 0, \quad (1.9)$$

¹ $\theta = T(p_s/p)^{(R/c_p)}$ is the temperature that a parcel of dry air at pressure p and temperature T would have if it would be brought adiabatically, i.e. without any heat exchange with the surroundings, to standard pressure p_s . R is the gas constant for air and c_p the specific heat capacity at a constant pressure.

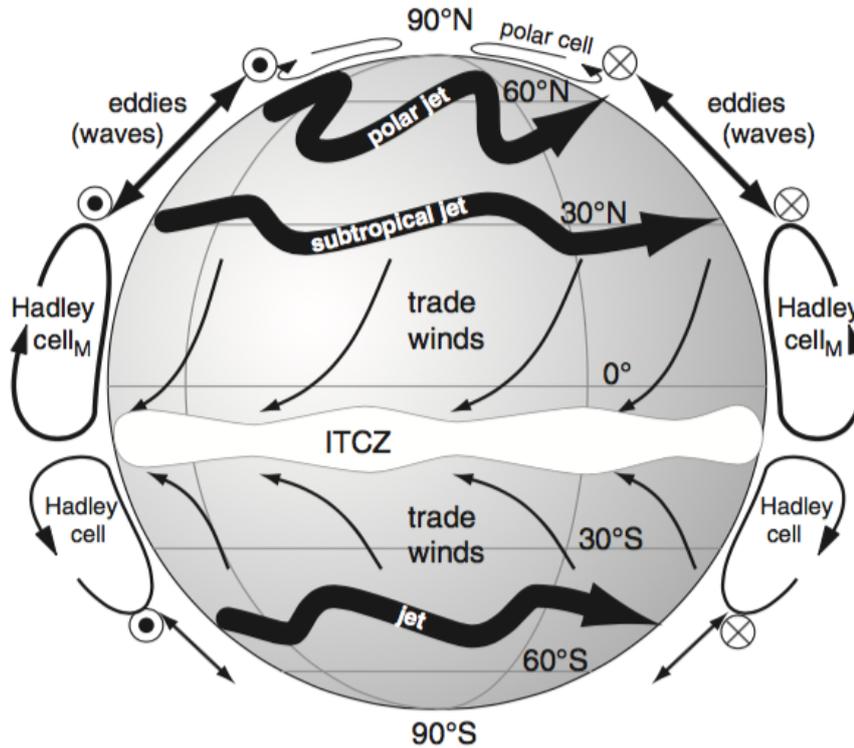


Figure 1.2: Idealised global circulation for Northern Hemisphere winter. The bulls-eye symbol indicates the jet stream coming out of the page toward the reader, and the "X" with a circle around represents the jet-stream flowing into the page. Reproduced from Stull (2015-2016).

$$\frac{D\theta}{Dt} = F_{\theta}, \quad (1.10)$$

with Ω : angular velocity of the Earth, p : pressure, \mathbf{g} : gravitational acceleration, \mathbf{F}_f : frictional force, F_{θ} : diabatic force. $\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla + \mathbf{v} \cdot \nabla + \mathbf{w} \cdot \nabla$ represents the total derivative. The terms on the right hand side of the equation of motion (1.8) express the forces acting in a rotating frame of reference: Coriolis force, pressure-gradient force, gravity, and friction. Eq. (1.9) is called the continuity equation, and states that the net rate of mass inflow in a fixed control volume must equal the rate of mass accumulation within the volume. According to the thermodynamic energy equation (1.10), changes in the potential temperature are caused by diabatic forces. In numerical models, \mathbf{F}_f and F_{θ} represent unresolved processes and are parameterized.

These equations cannot be solved analytically, thus *numerical models* are needed to predict the weather and climate. Such models require the replacement of the original continuous equations by a finite size of *ordinary differential equations*, which can be solved numerically. Lorenz (1963) was the first one to show that a low order truncated model of an atmospheric process (Rayleigh-Bénard convection) based on ordinary differential equations exhibits *irregular or chaotic behaviour* under specific configurations, even if the

forcing is constant. Thus he demonstrated that irregularity is not necessarily externally prescribed to the system, but it is rather an inherent characteristic coming from non-linearity. Several years later, *Lorenz (1984)* presented another simplified model meant to represent some basic features of a zonally symmetric circulation, the Hadley circulation. This model consists of only three equations, representing the temporal evolution of a symmetric globe-encircling westerly flow X and sine and cosine phases of a chain of large-scale eddies, Y and Z respectively:

$$\frac{dX}{dt} = -Y^2 - Z^2 - aX + aF, \quad (1.11)$$

$$\frac{dY}{dt} = XY - bXZ - Y + G, \quad (1.12)$$

$$\frac{dZ}{dt} = bXY + XZ - Z, \quad (1.13)$$

where F and G represent zonally symmetric and asymmetric thermal forcings, the non-linear mixed terms express interactions of the eddies with the zonal mean current, the linear terms stand for mechanical and thermal damping, and a and b are constants. Lorenz shows in his paper that, if one varies G starting with a value near 0 and increasing it gradually, the system goes over from a stable periodic behaviour to an unstable non-periodic behaviour which settles on a strange attractor. He also found multiple equilibria for certain values of G , representing, for example, a Hadley circulation and a blocking situation. Fig. 1.3 illustrates the Lorenz-84 attractor for $G = 1$ and for two different forcings $F = 6$ and $F = 8$. We see that, while in case of $F = 6$ the attractor is a closed curve in the phase space indicating a periodic behaviour, for $F = 8$ the attractor is a complex manifold according to the chaotic behaviour of the system.

Regarding the evolution of the atmosphere, we talk very often about “*weather*” and “*climate*”. The usual way to differentiate between them is to define the weather as the short term evolution of the atmosphere, i.e. starting from a specific point on the attractor and following the trajectory in phase space for a relatively short time, whereas “*climate*” represents the long term evolution of the atmosphere, in which case the starting point does not matter anymore, because the trajectory is supposed to be so long that it populates the whole attractor. What matters, however, are the boundary conditions under which we let the system evolve. In the case of climate, it is assumed that the systems can be described statistically, since the exact temporal evolution is not relevant anymore.

We analyse the atmosphere in this work from the point of view of *dynamical systems theory*. However, we wish to mention that, in the case of the atmosphere, we do not certainly know whether the real system is a complex quasi-periodic or a chaotic dynamical system, since we do not have infinitely long observations about its evolution. What we can say, however, is that on realistic finite time scales it does exhibit an irregular

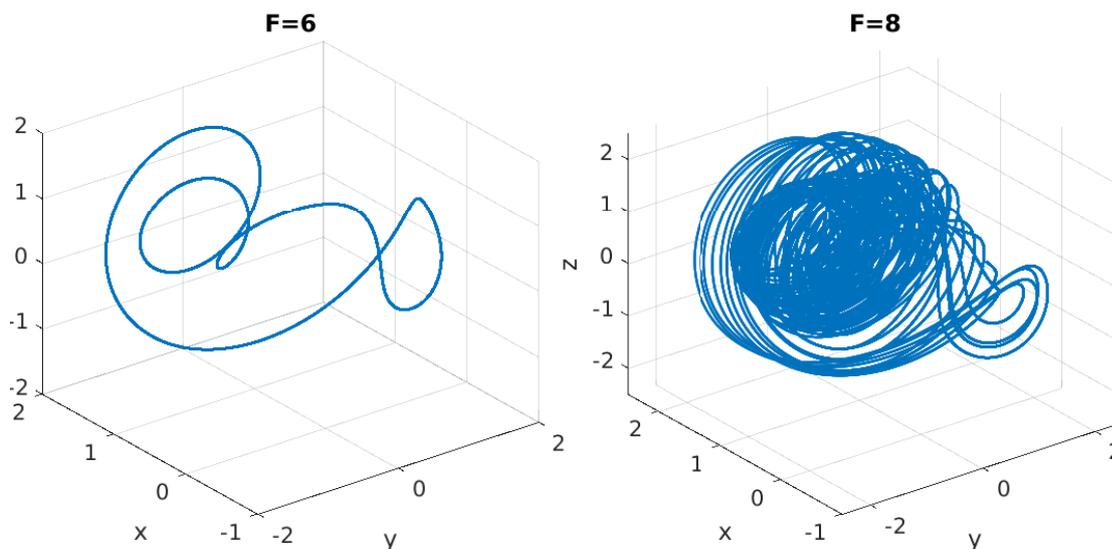


Figure 1.3: Attractors of the Lorenz-84 model for $F = 6$ and $F = 8$, with $G = 1$, $\alpha = 0.25$, and $b = 4$. The integration consists of 146 model time units corresponding to two years, and is performed by using the MATLAB-function `ode45` (MATLAB, 2015).

behaviour.

From a *statistical-mechanical* perspective, the atmosphere can be interpreted as a high dimensional *forced-dissipative* chaotic system in an *out-of-equilibrium state*. It is possible to find a system in a *non-equilibrium steady state* if the forcing acting on it is constant, and, at the same time, the system features some regulatory mechanisms between processes driving it towards equilibrium and processes pushing it away from equilibrium. In the case of the atmosphere, for example, the previous ones are represented by emission of radiation and frictional dissipation, whereas the later ones by absorption of radiation emitted at a higher temperature and meridional temperature as well as density gradients (Kleidon, 2010). Main characteristics of these kind of systems are contraction of the phase space, generation of entropy, and a finite predictability horizon (Lucarini et al., 2014a).

The complexity of the dynamics of the atmosphere is reflected in the fact that its variability is characterised by the peculiar co-existence of waves, turbulence, and particle-like features (e.g. cyclones) (Ghil and Roberston, 2002). Additionally, it is well known that in the climate system there is a non-trivial *relationship between spatial and temporal scales* of variability - with large spatial scales associated, in a nontrivial way, to longer time scales. The existence of such relationships (which we will explore also in this thesis) comes from the fact that one can loosely identify different dynamical regimes, each characterised by specialised *dynamical balances* between the forces acting on the fluid components (Lu-

carini et al., 2014a). Such balances can be rigorously derived via asymptotic analysis applied on the Navier-Stokes equations on a rotating frame of reference (Klein, 2010). A schematic representation of the approximation based on dynamical balances as function of the considered temporal and spatial scales is given by Fig. 1.4 (Klein, 2010).

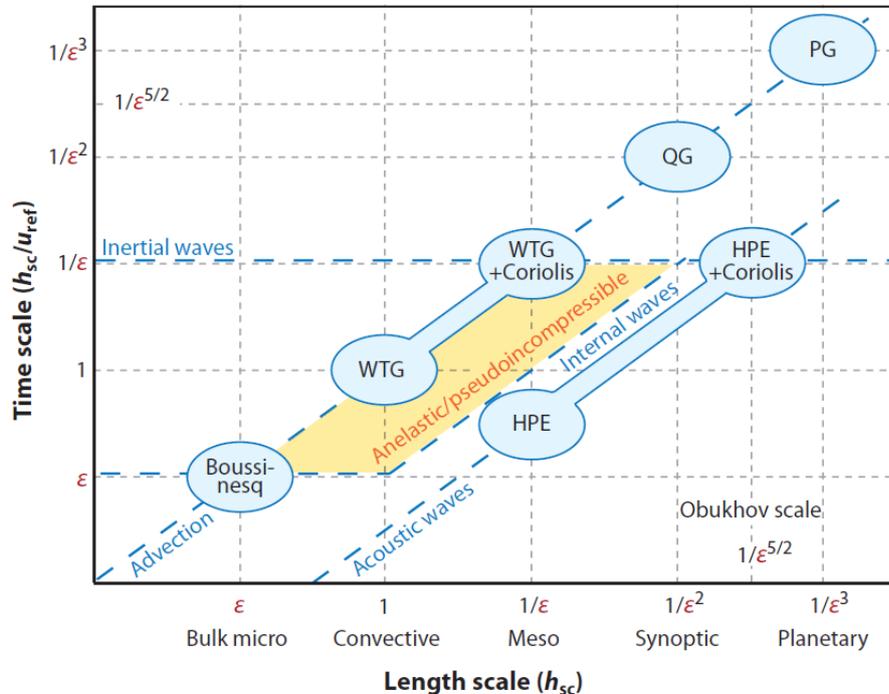


Figure 1.4: Atmospheric scaling regimes and the related model equations. WTG: weak temperature gradient, QG: quasi-geostrophic, PG: planetary geostrophic, HPE: hydrostatic primitive equation. h_{sc} represents the density scale height (~ 11 km) and u_{ref} the thermal wind velocity (~ 12 m/s). $\epsilon \sim 1/9 \sim u_{ref}/c_{int}$, where c_{int} is the internal wave speed (~ 110 m/s). Reproduced from Klein (2010), copyright © 2010 by Annual Reviews.

We have seen in this section that the atmosphere is a highly complex system described by a set of non-linear differential equations, and can be studied based on different interpretations, for example, as a chaotic dynamical system or as a forced-dissipative system in an out-of-equilibrium state. In the next two sections, we describe how can we analyse atmospheric extreme events based on these two perspectives.

1.3 Extreme events and general system properties

A robust theoretical framework for analysing extreme events is provided by *Extreme Value Theory* (EVT). After the early contributions by Fisher and Tippett (1928), EVT was introduced by Gnedenko (1943) who discovered that under rather general conditions the

extreme events associated to stochastic variables can be described by a family of parametric distributions. As a result, the problem of studying the extremes of a stochastic variable can be reduced to estimating the parameters of a known probability distribution. The most important parameter of such a distribution is called the shape parameter, which determines the qualitative properties of the distribution, and, in particular, whether it has a finite or infinite upper point, or, more concretely, whether extremes are bounded by an absolute finite maximum or not. Two of the most widely used methods of EVT are the *Block Maxima* (BM) (Fisher and Tippett, 1928; Gnedenko, 1943) and the *Peak Over Threshold* (POT) (Pickands, 1975; Balkema and de Haan, 1974) approaches, which differ mostly in the selection procedure of extreme events. Whereas the BM method identifies extremes as maxima M_m of independent identically distributed (i.i.d.) random variables (r.v.) separated into sufficiently large blocks of size m as $m \rightarrow \infty$, the POT approach detects extremes considering exceedances of a stochastic variable X above a given threshold u as $u \rightarrow x_*$, where x_* represents the upper end point of the distribution of X . Figure 1.5 illustrates the different selection procedures of extreme events in case of the BM and POT approaches. Both methods are formulated in form of limit laws, and rely on the convergence in distribution of the selected extreme values to one limiting family of distributions, as one considers more and more extreme levels (i.e. for increasing m or u). Thus EVT is often interpreted as an analogy to the Central Limit Theorem (CLT).

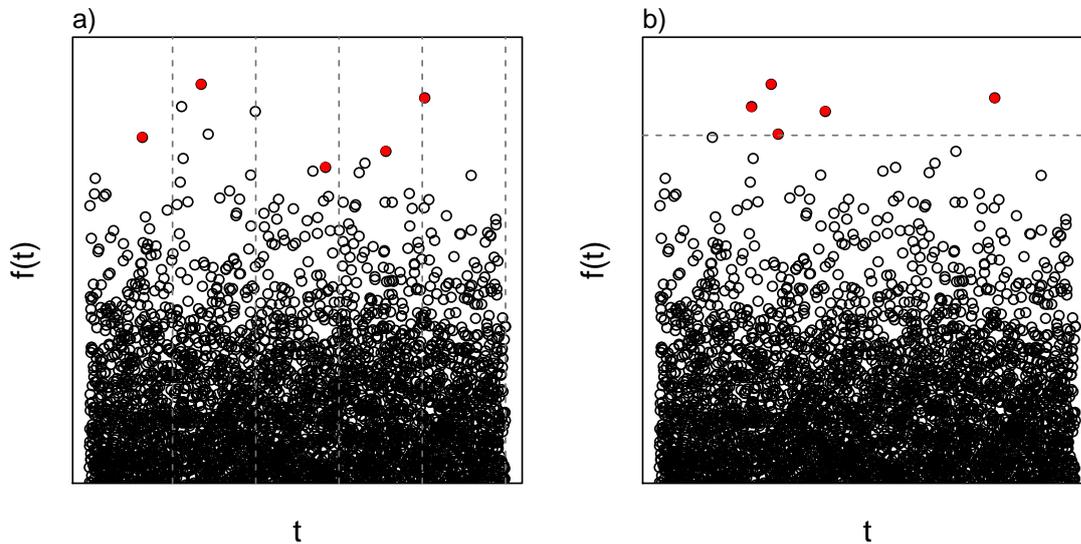


Figure 1.5: Selection of extreme events (red dots). a) BM method, vertical dashed lines separate the blocks; b) POT method, horizontal dashed line represents the threshold.

The limiting family of distributions is the *Generalized Extreme Value* (GEV) distribution in case of the BM method, and the *Generalized Pareto distribution* (GPD) in case of the POT approach. It is remarkable that these two points of view on extremes (based on

different selection procedures and different choice of the events classified as extremes) are, in fact, equivalent. The existence of well-defined functional forms of distributions describing extreme events provides predictive power: one can in principle compute the return time of yet unobserved events. In other words: if the limit law applies, predictive power emerges (in a statistical sense), and is directed towards events with decreasing probability.

The setting behind the construction of EVT can be extended by relaxing the hypothesis of independence of the random variables, as long as correlations are decaying sufficiently fast. This is of clear relevance when trying to use EVT for studying observables of deterministic dynamical systems. In this case, in fact, the underlying dynamics determines the existence of correlations between the values of observables at different times, and one can easily guess that, when the dynamical system is chaotic, there is good hope of deriving EVT for its observables (Lucarini et al., 2016). Consequently, it is possible to establish an EVT for observables of chaotic dynamical systems when one considers Axiom A dynamical systems, due to the convenient characteristics of these systems discussed in Sec. 1.1.

Several studies dealing with EVT for dynamical systems reveal a link between the statistical properties of the extremes and geometric (and possibly in turn global dynamical) characteristics of the system producing these extremes (Faranda et al., 2011; Lucarini et al., 2012a,b; Holland et al., 2012; Lucarini et al., 2014b). The main findings are that when suitable observables are chosen for the dynamical system of interest, it is possible to relate the GEV or GPD parameters describing the extremes to basic properties of the dynamics, and especially to the geometry of the attractor. In particular, depending on the choice of the observable, one can associate the most important parameter of the GEV or GPD distribution to the information dimension of the attractor or to the partial information dimension along the stable and unstable directions of the flow (Lucarini et al., 2014b). These partial dimensions are well-defined everywhere on the chaotic attractor, also for non-uniformly hyperbolic systems (Barreira et al., 1999), beside Axiom A systems. However, Axiom A systems possess an ergodic SRB measure which lends itself to a universality of the shape parameter for all sufficiently smooth observables; the local or point-wise (partial) dimensions taking the same value almost everywhere (Ott, 1993). In this case the uniform shape parameter can be related to the (partial) Kaplan-Yorke dimension(s), defined by the global dynamical characteristic numbers, the Lyapunov exponents. Clearly, this is an asymptotic result, and one must expect that differences emerge on pre-asymptotic level when different observables are studied. Via the connection with fractal dimensions, it can be said that the analysis of extremes acts as a microscope able to assess the fine scale properties of the invariant measures. The reader is referred to Lucarini et al. (2016) for a detailed overview of the field of EVT for dynamical systems.

Some preliminary numerical tests show that the convergence to the asymptotic shape parameter is slow in low-dimensional cases, in systems well known not to be Axiom A (Lucarini et al., 2014b). Bódai (2017) examined the convergence to the GEV distribution in the case of extremes of site variables in the Lorenz 96 model (Lorenz, 1996), investigating separately a range of cases extending from weak to strong chaos. He found that, when considering configurations supporting weak chaos with a low-dimensional attractor, the theoretical results obtained in the context of the Axiom A hypothesis are hard to verify. For lower dimensions, up to a dimension of about 5, shape parameter estimates fluctuate greatly rather than converge, while block maxima data can be shown not to conform to a GEV model; and for somewhat larger dimensions, up to 9 in the study, estimates could diverge from the predicted value while data already conform to a GEV model. Good agreement with the theory was found only in the highly turbulent case possessing a higher-dimensional attractor, about 30, supporting the basic idea behind the chaotic hypothesis. Also in this case, nonetheless, very slow convergence was found.

In previous analysis performed on higher dimensional, intermediate complexity models with $O(10^2 - 10^3)$ degrees of freedom, very slow (if any) convergence to EVT distributions could be found in the case of extremes of local temperature observables (Vannitsem, 2007). In another analysis of a similar model (Felici et al., 2007a), the agreement of the distribution of global energy extremes with a member of the GEV family was indeed good, yet large uncertainty remained on the value of the shape parameter, and no stringent test was used to make sure that the estimate was stable against changes in the block size considered in the BM analysis. Clearly, the specific choice of the observable and the degree of chaoticity of the underlying dynamics is of primary relevance regarding the convergence to the limiting GEV or GPD distribution.

Findings of EVT for dynamical systems have been shown to be applicable not only to simple systems but, recently, even to reanalysis data sets. Although, one has to be aware that, by considering a few observables of reanalysis data, one is far away from obtaining general properties of the very-high dimensional attractor. However, it has been shown that extremes of some dynamical system metrics can be used as proxies for certain atmospheric patterns which can favour extreme weather conditions (Faranda et al., 2017; Messori et al., 2017).

1.4 Large deviations and persistent extreme events

As discussed in Sec. 1.2, there is a non-trivial relationship between temporal and spatial scales of variability in the atmosphere, as well as in the climate system. As a result,

one can guess that persistent climatic extremes, are generally both persistent *in time* and persistent, i.e. extended, *in space*. A heat wave or a cold spell, for example, is not only lasting for a long time (from several days to several weeks, even months) but has also a considerable spatial extension. They are primarily caused by anomalous synoptic conditions, mostly by atmospheric blocking situations, so we talk about space-time persistence on *large synoptic scales* (Vautard et al., 2011; Sillmann et al., 2011; Stefanon et al., 2012). For example, the 2003 and 2010 European heat waves had a temporal and spatial extent of the order of weeks to months and 10^6 km^2 respectively. (Barriopedro et al., 2011).

As explained above, EVT is mathematically very powerful, but in the usual setting is not well suited for studying, in the case of spatio-temporal chaos, anomalously large or small events that are persistent in time and extended in space. This is unfortunate because *persistent climatic extreme events* - like heat waves or cold spells - can have a huge impact: they do not affect only human health, but also ecosystems; they can be a danger for our infrastructures and crops, and have a destabilising effect over entire societies (Easterling et al., 2000; Robinson, 2001; WHO, 2004; IPCC, 2012). Among the most relevant historical examples we would like to mention the mega-drought that played a major role in the collapse of the Maya empire (Kennett et al., 2012), and the recurrent extreme cold spell episodes referred to as Dzug that led to various waves of migration of the nomadic Mongolian populations (Fang and Liu, 1992; Hvistendahl, 2012).

One can analyse persistent events generally in two ways: first, by treating them as a concatenation of successive extreme events and study the properties of clusters of extremes (Ferro and Segers, 2003a; Segers, 2005), or, second, by looking at pdf's of *time-averaged observables*. In this study we follow the second route. Following intuition, if we look at the pdf of finite-size averages of an observable, one expects that the tails of the distribution are mainly populated by averages coming from persistent extremes. A rationale for this is that the averaging window acts like a low-pass filter on the length of the considered persistent event, leading to a connection between extremes of averages and persistent events with a certain length (greater or approximately equal to the chosen averaging window). This will roughly be, in fact, the scenario we will explore below. However, the link between persistence and extremes of finite-size averages is not always true: in case of heavy-tailed r.v., for example, the extremes of averages are dominated by a single very large extreme event within the averaging window (Mikosch and Nagaev, 1998). We remark that, generally, the methods of EVT can be applied the same way also to study extremes of averaged observables. However, the averaging process reduces the number of available data, so that these methods can become more difficult to apply.

A mathematical framework describing properties of pdf's of *averaged* observables is provided by Large Deviation Theory (LDT), introduced by Cramer (1938) and further

developed by other mathematicians, like Donsker and Varadhan (1975a,b, 1976, 1983), Gärtner (1977), and Ellis (1984). The central result of LDT consists of writing the probability of averaged random variables $A_n = \frac{1}{n} \sum_{i=1}^n X_i$ as function of the variables X_i : for $n \rightarrow \infty$, the probability of averages decays exponentially with n , $p(A_n = a) \approx e^{-nI(a)}$. This is called a large deviation principle (LDP). The speed of decay is described by the so called rate function $I(a) \geq 0$. The probability $p(A_n = a)$ decays everywhere with increasing n , except when $I(a) = 0$. Here, $e^{-nI(a)} = 1$. For an i.i.d. r.v. one would have that $\mathbb{E}[A_n] = a^*$, where a^* is such that $I(a^*) = 0$. If the rate function exists, one can estimate the probability of averages for every n . Similarly to EVT, if the limit law applies, we gain predictive power, with the difference that in this case it is directed towards averages with increasing n . This means that one does not have to deal anymore with the problem of decreasing amount of data as n increases. Figure 1.6 shows the first step in the application of LDT, which is obtaining averaged observables over equidistant averaging blocks. For increasing averaging block lengths, the rate function estimates are expected to converge, if a LDP exists.

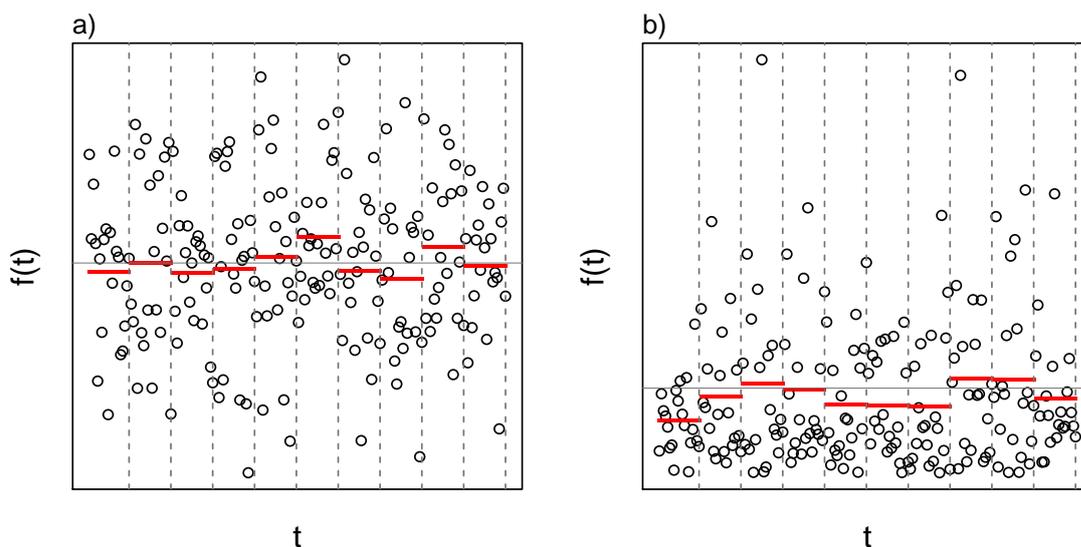


Figure 1.6: First step in the application of LDT: obtaining averages over blocks of equal lengths. a) Normal and b) exponential r.v. (black circles) and their averages over finite blocks (red lines). The vertical dashed lines separate the averaging blocks. The horizontal grey line represents the long-term average.

LDT is used very extensively in physics, mostly in the context of thermodynamics and statistical mechanics. In the case of an equilibrium system, the global minimum of $I(a)$ represents the system state with maximum entropy, which is at the same time the most probable state, according to the maximum entropy principle. In the case of a non-equilibrium system, under suitable assumptions, rate functions are computed taking into account the time evolution and for each observable the minimum of the rate function defines its expectation value.

While they have recently been applied in the context of geophysical flows (see e.g. Bouchet and Venaille 2012; Bouchet et al. 2014; Herbert 2015), techniques of LDT have been used sporadically until now in climate studies, despite the fact that they can be useful whenever the connection between macroscopic or long term observables and microscopic or instantaneous observables is important, and one is interested in persistent and/or extended fluctuations of a climatic field.

One area of climate modelling where techniques of LDT are just beginning to be applied is the sampling of rare events. Rare event computation techniques based on elements of LDT have been developed with the aim to produce reliable statistics of specific rare events of a given model, as an alternative to long direct numerical simulations (Giardina et al., 2016; Wouters and Bouchet, 2016; Lestang et al., 2018). Ragone et al. (2017) describe how model trajectories can be selected, based on a rare event algorithm, by keeping an ensemble realisation of the system in states that are preferentially close to those leading to heat waves. Therefore, one can exponentially oversample events that have ultra long return periods, and thus construct a richer statistics of heat waves than one would get by standard Monte Carlo techniques. The described method provides also the possibility to investigate dynamical properties of the system state (like global circulation patterns and jet stream position) supporting the occurrence of the studied extremes (heat waves).

1.5 This thesis

In this thesis, we analyse extreme or rare events in two atmospheric models based on *Extreme Value Theory* (EVT) and *Large Deviation Theory* (LDT). Beyond the statistical analysis of extremes, we mainly aim at finding out general properties of our dynamical systems (i.e. the numerical models) and characteristics of the system state producing these extremes. The thesis is built up around two main topics. First, we study extreme events of energy observables in a two-layer quasi-geostrophic atmospheric model based on EVT. The main objective here is to see whether we can detect universal properties of extreme events, which can then be related to properties of the attractor of the system, according to a theory developed within EVT for dynamical systems (Holland et al., 2012; Lucarini et al., 2014b). Second, we look at large deviations of surface temperature in a simplified general circulation model of the atmosphere based on LDT. The main objective is to study persistent space-time extreme events of temperature, i.e. heat waves or cold spells, based on their connection to extremes of temperature averages on specific scales.

In practical terms, the main goal behind the study of the extreme events is to understand the properties of the highest quantiles of the variable of interest. A fundamental draw-

back comes from the fact that it is difficult to collect satisfactory statistics from a time series of finite length because, as we know, extreme events are rare. Additionally, in the absence of a strong mathematical framework, it is virtually impossible to make quantitative statements about the probability of occurrence of events larger than observed. Therefore, the analysis in this thesis is based on two rigorous mathematical theories, EVT and LDT, both based on *limit laws*, which provide a *predictive power* to the obtained results, in a statistical sense. Note that statistical inference based on empirical models tends to suffer from the lack of this predictive power. However the predictive power can act only if one is able to *approach* the *theoretical limit*. We use here deliberately the expression “approach” instead of “reach”, because the theoretical limit can, of course, never be reached based on finite-size time series. However, in case of applications, one can *reach* an *empirical limit* which means that the empirical results are indistinguishable from what the theory predicts considering the uncertainty due to the finite size of our data. If we succeed to attain such an empirical limit in this work, we consider to have reached the “*asymptotic level*”, whereas if the empirical limit exists but is not realised yet, it means that we are on a “*pre-asymptotic level*”.

Besides the very important common advantage of having predictive power, we also encounter similar problems - with similar solutions - when we try to apply these methods. These problems, into which we bump several times in this thesis, are mostly related to serial correlations, finite-size of the data, or convergence behaviour of the analysed observables. Although the mentioned similarities between the two methods are very interesting and indeed very important, one should not overlook that the object and the background of the two theories are completely different. While EVT focuses on tails of pdf's by selecting extreme values according to a certain procedure (as maxima over blocks of equal length or as exceedances above a high threshold), by applying LDT, one actually transforms the original observable and finds a limit law for the newly obtained observable. In other words, whereas EVT acts as a magnifying glass able to assess the fine scale properties of the attractor (Lucarini et al., 2016) without changing its structure, when we apply LDT we perform a kind of coarse-graining of the original system.

The atmospheric models considered in this thesis can be seen as being quite simple and not very realistic. However, they are very effective tools for producing long simulations in a relatively short time, which is of utmost importance if one aims to test the applicability of limit laws, as we proceed in this work. Furthermore, they are realistic enough to simulate Earth-like atmospheric processes on, mainly, synoptic scales (order of 10^6 m), and allow also for computing with feasible computational costs some dynamical system properties, like Lyapunov Exponents or Kaplan-Yorke dimensions (Schubert and Lucarini, 2015; De Cruz et al., 2018).

In case of sufficiently high forcing levels, both models exhibit a chaotic behaviour as

described by the *chaotic hypothesis*, stating that high dimensional chaotic systems behave at all practical purposes as so-called *Axiom A* systems (Gallavotti and Cohen, 1995; Gallavotti, 2014). These are special chaotic dynamical systems that exhibit an apparently random dynamics, even though they are completely deterministic. They are also stable against weak stochastic perturbations. Lucarini et al. (2014b) discuss that whenever one performs numerical simulations, one implicitly assumes that the systems is Axiom A or, at least, Axiom A-equivalent. This property of being Axiom A-equivalent is very important for the applicability of the theories considered in this thesis. The connection between statistical properties of extremes and general properties of the system or the existence of large deviation laws holds in the framework of the chaotic hypothesis, which guaranties a strong enough mixing associated with a fast enough decay of correlations - so that two events far away from each other are approximately uncorrelated - necessary for the applicability of these limit laws. In these kind of mixing systems one often encounters on several levels a property of *universality*, which we explore in this thesis from two different perspectives: by searching for universal properties of extremes of energy observables and exploring the universality of temperature values averaged in time and in space (as well as in space-time).

We believe that the concept of universality, as it is studied in this thesis, is relevant for more realistic systems as well, considering that atmospheric or climatic time series based on observations or complex model simulations often exhibit chaotic characteristics and a fast decay of correlations, i.e. they behave on certain scales as Axiom A systems. Thus, the non-universality of extreme value statistics, for example, indicates that asymptotic levels are not reached yet, thus one has to be careful with the estimation of long return periods of extremes based on the available data. The universal property relating temporal to spatial averages, should be also valid in case of realistic data sets satisfying certain conditions (for example considering stationary time series over homogeneous surfaces), and is, in fact, extremely useful because it allows for obtaining characteristics of spatial averages from temporal series and vice-versa.

The structure of the thesis is as follows.

- In *Chapter 2* we provide a theoretical overview of the elements of EVT and LDT used in this thesis.
- *Chapter 3* contains a detailed description of the model physics and methods.
- In *Chapter 4* we present and discuss our results related to extreme events of energy observables in the two-layer quasi-geostrophic model.
- In *Chapter 5* we show and discuss the results regarding large deviations of temperature in the simplified general circulation model PUMA.

- At the end, we provide some general discussion and common conclusions for the two main topics in *Chapter 6*.

After having discussed the general aim of this thesis, we wish to outline the objectives of the two major topics discussed in this work, presented later in Chapters 4 and 5, including their scientific relevance. In Chapter 4, we use a quasi-geostrophic (QG) atmospheric model of intermediate complexity featuring 1056 degrees of freedom, to analyse extremes of different types of observables: local energy (defined at each grid point), zonally-averaged energy, and the average value of energy over the mid-latitudes. Our main objective is to compare the estimated GEV and GPD shape parameters with a shape parameter derived, based on the theory referred to above, from the properties of the attractor along the stable, unstable, and neutral directions. We refer to this as the “theoretical shape parameter”. Thus we explore numerically the link between the purely statistical properties of extreme events based on EVT and the dynamical properties of the system producing these extremes. We perform simulations applying two different levels of forcing: a strong forcing, producing a highly chaotic behaviour of the system, and a weak forcing, producing a less pronounced chaotic behaviour. The dimensionality of the attractor is much larger in the former than in the latter case.

The work presented in Chapter 4 goes beyond the previously mentioned studies based on more simple dynamical systems (Sec. 1.3), in a sense that with our model we can study the convergence for observables being different physical quantities, or, representing different spatial scales/characteristics of the same physical quantity. Additionally, compared to previous studies also performed on intermediate complexity models, we consider longer time series and a variety of observables. Our model is simple compared to a GCM (General Circulation Model), but contains two of the main processes relevant for mid-latitude atmospheric dynamics, i.e. baroclinic and barotropic instabilities. Hence, we contribute to bridging the gap between the analysis of extremes in simple and very high dimensional dynamical systems, as in the case of the GCMs used for atmospheric and climate simulations. The properties of the model have been extensively studied by Schubert and Lucarini (2015, 2016).

Although we use an idealised model, our results are transferable to time series obtained from more realistic model simulations or from measurements. By understanding the differences among the analysed observables, we gain insight into the statistical properties of extremes of geophysical observables with different spatial scales. By using two forcings, we are able to study the convergence to theoretical shape parameters related to different chaotic systems: one exhibiting fast decaying correlations and another one characterised by slower decaying correlations. These aspects are relevant in the case of geophysical applications, where one deals also with time series on several spatial scales and with different degrees of correlations.

In Chapter 5 we adopt LDT to analyse the properties of temporally and/or spatially persistent surface temperature extremes - heat waves or cold spells - generated through simulation performed with the Portable University Model of the Atmosphere (PUMA) (Lunkeit et al., 1998; Fraedrich et al., 2005b). We investigate temperature averages computed in time and/or in space, the spatial averaging being performed along the zonal direction for reasons of symmetry.

PUMA - details given in Sec. 3.2 - describes with a good level of precision the dynamics of the three-dimensional atmosphere as an out-of-equilibrium forced-dissipative system. We analyse the properties of the steady state achieved as a result of time-independent forcing after transient dynamics has been discarded. For a wide range of parameter values, PUMA features high-dimensional chaotic dynamics (De Cruz et al., 2018). By considering the connection between the averaged values and persistent events on suitably defined scales (as explained above), large deviations of temperature can possibly be related to persistent extreme events of temperature.

Following the discussion above, we expect to find a link between spatially extended and temporally persistent events. In order to achieve a large deviation when considering spatial averages in a turbulent system, we need to have occurrence of a spatially extended structure of length say L . In a system possessing a characteristic velocity scale U , one expects such a structure to persist for a typical time of the order L/U . Because persistent events are space-time events, we explore the connection between temporal and spatial large deviations, and we also analyse spatio-temporal large deviations. We seek answers to two main questions:

1. How well does LDT describe persistent in space and/or time temperature fluctuations in PUMA?
2. What is the link between temporal, spatial, and spatio-temporal large deviations?

These questions are potentially relevant, because, if we find experimental proofs that the LDP does hold in the case of our numerical simulations, there is a good chance to calculate the probability of occurrence of arbitrarily long in time and/or extended in space (within the limits allowed by the geometry of the Earth, as seen later) heat waves and cold spells.

In the case we find a link between temporal and spatial large deviations, we can deduce the probability of spatial (or spatio-temporal) averages from the one of temporal averages and vice-versa. This can be very useful in case of applications, when for example only temporal or only spatial series are available. In order to test the quality of predictions of return times based on LDT, we compare the results with what can be obtained using EVT (we use here the POT method).

The common points of the objectives of Chapters 4 and 5 might be not totally clear when formulated in detail. Therefore, we wish to point out again that in both cases we study extreme or rare atmospheric events in numerical models from a theoretical point of view, based on limit laws which allow, under certain conditions, a connection between properties of extremes and general properties dynamical system.

2 Extreme events in chaotic dynamical systems

As discussed already in Chapter 1, we apply for the study of extreme or rare events two asymptotic theories: Extreme Value Theory (EVT) and Large Deviation Theory (LDT), which are described in the following. Both theories are derived in an independent r.v. framework, thus we also discuss the conditions under which they can be applied in case of deterministic dynamical systems. These conditions are generally related to a fast enough decay of correlations, this being usually the case in systems satisfying the chaotic hypotheses. The fact that they rely on limit laws provides predictive power for the obtained results, and represents a clear advantage compared to statistical inference based on empirical models. However, it also requires a rigorous verification of the convergence behaviour from an empirical point of view, under consideration of the uncertainty resulting from the finite data size. This will be discussed in detail also later in Chapters 4 and 5 based on our numerical results .

2.1 Extreme Value Theory for dynamical systems

Let us consider $M_m = \max\{X_1, \dots, X_m\}$, where X_1, \dots, X_m is a sequence of i.i.d. r.v. with common distribution function $F(x)$. The extremal types theorem (Fisher and Tippett, 1928; Gnedenko, 1943) states that if there exist sequences of constants $\{a_m > 0\}$ and $\{b_m\}$, so that the distribution of normalised M_m , i.e., $\Pr\{(M_m - b_m)/a_m \leq z\}$, converges for $m \rightarrow \infty$ to a non-degenerate distribution function $G(z)$, then $G(z)$ is one of three possible types of so-called extreme value distributions, having the cumulative distribution function

$$G(z) = \begin{cases} \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} & \text{for } \xi \neq 0, \\ \exp \left\{ - \exp \left[- \left(\frac{z - \mu}{\sigma} \right) \right] \right\} & \text{for } \xi = 0, \end{cases} \quad (2.1)$$

where $-\infty < \mu < \infty$, $\sigma > 0$, $1 + \xi(z - \mu)/\sigma > 0$ for $\xi \neq 0$ and $-\infty < z < \infty$ for $\xi = 0$ (Coles, 2001).

$G(z)$ represents the GEV family of distributions with three parameters: the location parameter μ , scale parameter σ , and shape parameter ξ . The shape parameter ξ describes the tail behaviour, and determines to which one of the three types of extreme value distributions $G(z)$ belongs. If $\xi = 0$, the tail decays exponentially, $G(z)$ is a type I extreme value distribution or Gumbel distribution. If $\xi > 0$, the tail decays polynomially, and $G(z)$ belongs to the type II or Fréchet distribution. If $\xi < 0$, the domain of the distribution has an upper limit, and is referred to as a type III or Weibull distribution. The three types of GEV distributions are illustrated in Fig. 2.1a.

Under the same conditions, for which the distribution of M_m converges to the GEV distribution, the exceedances $y = X - u$ of a threshold u reaching the upper right point of the distribution of X , given that $X > u$, are asymptotically distributed according to the Generalized Pareto distribution (GPD) family (Coles, 2001)

$$H(y) = \begin{cases} 1 - \left(1 + \frac{\tilde{\xi}y}{\tilde{\sigma}}\right)^{-1/\tilde{\xi}} & \text{for } \tilde{\xi} \neq 0, \\ 1 - \exp\left(-\frac{y}{\tilde{\sigma}}\right) & \text{for } \tilde{\xi} = 0, \end{cases} \quad (2.2)$$

where $1 + \tilde{\xi}y/\tilde{\sigma} > 0$ for $\tilde{\xi} \neq 0$, $y > 0$, and $\tilde{\sigma} > 0$. $H(y)$ has two parameters: the scale parameter $\tilde{\sigma}$ and the shape parameter $\tilde{\xi}$. The shape parameter $\tilde{\xi}$ describes again the tail behaviour (shown in Fig. 2.1b), and determines to which one of the three types of GPD distributions $H(y)$ belongs. If $\tilde{\xi} = 0$, the tail of the distribution decays exponentially; if $\tilde{\xi} > 0$, the tail decays polynomially; and if $\tilde{\xi} < 0$ the distribution is bounded (Pickands, 1975; Balkema and de Haan, 1974; Davison and Smith, 1990). If convergence to the GEV and GPD distributions is realised, $\tilde{\xi} = \xi$ and $\tilde{\sigma} = \sigma + \xi(u - \mu)$. As a result, once we estimate the parameters for the GEV, we can derive the corresponding GPD parameters, and vice versa (Coles, 2001).

From the values of the GEV parameters, which can be inferred using usual proven estimation methods, like maximum likelihood estimation (Coles, 2001) or L-moments (Hosking, 1990), it is possible to infer the expected return levels or extreme quantiles. Return levels z_p are obtained from the GEV distribution by inverting equation (2.1):

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi} [1 - w_p^{-\xi}] & \text{for } \xi \neq 0, \\ \mu - \sigma \log w_p & \text{for } \xi = 0, \end{cases} \quad (2.3)$$

where $w_p = -\log(1 - G(z_p))$, and $1/w_p$ represents the return period. The expected return levels or extreme quantiles are obtained similarly from the values of the GPD parameters. In this case, the level y_m that is exceeded on average once every m -observations is called the m -observation return level and is the solution of $P(Y > y) = \frac{1}{m}$. One obtains $P(Y > y)$ from $H(y) - 1 = P(Y > y|Y > u) = \frac{P(Y > y)}{P(Y > u)}$, and consequently

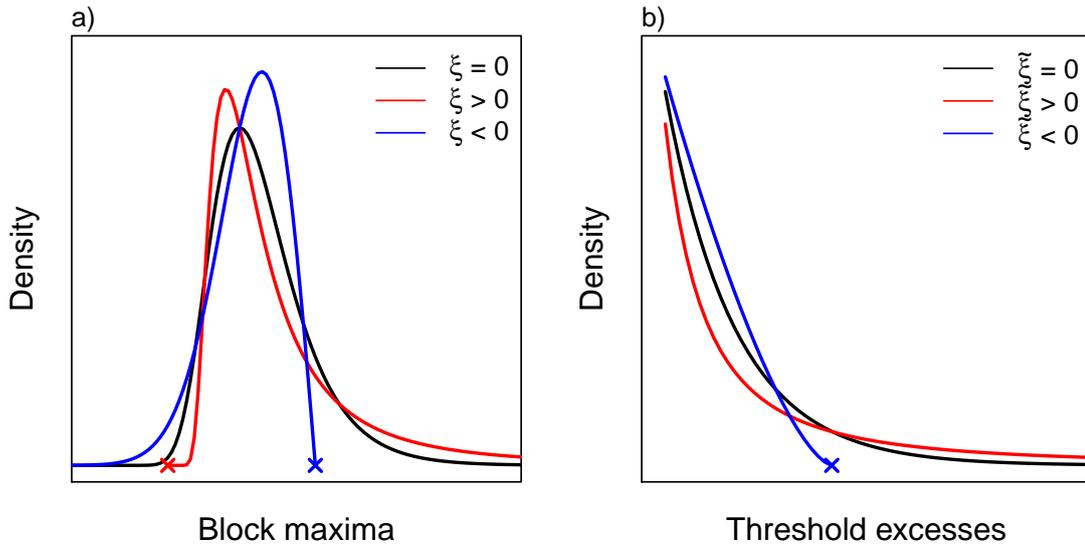


Figure 2.1: a) GEV and b) GPD probability densities for zero (black), positive (red) and negative (blue) shape parameters.

(Coles, 2001):

$$y_m = \begin{cases} u - \frac{\tilde{\sigma}}{\tilde{\xi}} \left[1 - \left(\frac{1}{mP(Y>u)} \right)^{-\tilde{\xi}} \right] & \text{for } \tilde{\xi} \neq 0, \\ u - \tilde{\sigma} \log\left(\frac{1}{mP(Y>u)} \right) & \text{for } \tilde{\xi} = 0. \end{cases} \quad (2.4)$$

By plotting the GEV (GPD) return level z_p (y_m) against the return period $1/w_p$ ($mP(Y > u)$) on a logarithmic scale, the plot is linear if $\xi = 0$ ($\tilde{\xi} = 0$), is convex if $\xi > 0$ ($\tilde{\xi} > 0$), and is concave if $\xi < 0$ ($\tilde{\xi} < 0$).

Problems in applying EVT to actual time series result from the fact that, typically, the observed data feature a certain degree of serial correlations (Ghil et al., 2011). In the case of a correlated stationary stochastic process, the same GEV limit laws apply as for i.i.d. r.v. if certain conditions, regarding the decay of serial correlation, are fulfilled (Leadbetter, 1974; Leadbetter et al., 1989; Lucarini et al., 2016). By stationary, we refer to a sequence of correlated variables whose joint probability distribution is time-invariant. However, an important restriction is that, as an effect of serial correlation, an effective block size can be defined, which is smaller than the number of observations in a block. This can enhance the bias in the parameter estimation, appearing as a slower or delayed convergence of the block maxima distribution to the limiting GEV distribution (Coles, 2001; Rust, 2009). Another possible effect of serial correlation is the appearance of extremes at consecutive time steps (clusters). If an extreme value law does exist in this case, then $G^*(z) = G(z)^\theta$, where θ is called the extremal index and $0 < \theta < 1$ ($G^*(z)$ denotes the limiting distribution of BM from the correlated sequence and $G(z)$ the one from an uncorrelated sequence, having the same marginal distribution). Clusters of extremes represent a problem especially when applying the POT approach. A widely-adopted

method to get rid of the correlated extremes is declustering, which basically consists of identifying the maximum excess within each cluster, and fitting the GPD distribution to the cluster maxima (Leadbetter et al., 1989; Smith, 1989; Ferro and Segers, 2003b).

Obtaining the true limiting EVT can be extremely hard in case of applications. When analysing finite time series, the convergence of the estimated GEV or GPD shape parameters to the asymptotic values can be very slow. The speed of convergence depends on the type of parent distribution (Leadbetter et al., 1983), and can be additionally slowed down by correlations (Rust, 2009; Coles, 2001). Due to the fact that the data size is always limited, there is typically a difference between the asymptotic GEV or GPD parameters and the estimated ones; finite-size estimates are generally biased. For example, the GEV shape parameter of a simple Gaussian process is 0, but, for any finite time series, we would estimate typically a negative shape parameter (Fisher and Tippett, 1928).

When performing statistical inference using the BM or POT method (fitting the GEV or GPD model, respectively, to data), it is crucial to have an appropriate protocol of selection of “good” candidates for extremes (Coles, 2001). On the one hand, if the chosen blocks (for the BM method) are too short or the threshold (for the POT method) is too low, the approximation of the limit model is likely to be inappropriate, leading to false parameter estimates. Hence, the verification of the agreement between the statistical model and the available data is essential. This is often done based on goodness-of-fit tests, like the Kolmogorov-Smirnov (Massey, 1951), Anderson-Darling (Anderson and Darling, 1954) or Pearson’s chi-squared tests (Agresti, 2007). On the other hand, if the blocks are too large or the threshold is too high, the number of extremes may be insufficient for a reliable estimation of the parameters, and uncertainty becomes very high. Coles (2001) shows how to derive an optimal choice for the value of the block size or the threshold, in such a way as to verify that we are close to the asymptotic level as required by EVT but we use the available data as efficiently as possible. We will describe how to do this empirically, when the asymptotic shape parameter is known, later in Chapter 4.

It is possible to establish an EVT also for observables of dynamical systems. However, we wish to remark that this is not trivial at all considering the deterministic nature of dynamical systems, and is reasonable only in case the system and its measure fulfil certain requirements. An ideal framework for a rigorous mathematical derivation of the theoretical results presented below is given by systems with uniformly hyperbolic attractor. This means that every point of the attractor comes with a stable and unstable manifold, which are well separated. The stretching and folding along these manifolds generates a complicated long-term behaviour, leading to apparently random dynamics, similar to a stochastic process, even though the system is completely deterministic. Additionally, the invariant set of these systems and its dynamics are structurally stable,

in the sense that small perturbations preserve the dynamics. Furthermore, the equilibrium measure of a hyperbolic attractor is the SRB measure (Sec. 1.1). As discussed in Sec. 1.1, Axiom A systems satisfy all of these geometrical and dynamical requirements. Despite the very rigorous mathematical conditions necessary for deriving theoretical results, in case of applications, the requirements can be relaxed based on the chaotic hypotheses, stating that high dimensional dynamical systems can be regarded as Axiom A systems for the purpose of computing macroscopic properties (Gallavotti and Cohen, 1995; Gallavotti, 2014). The most important condition for the application of the theory is that the dynamical system is ergodic and chaotic enough.

As mentioned before, several studies on EVT for observables of dynamical systems relate the GEV and GPD shape parameters to certain properties of the attractor. In the case of so-called “distance” observables, one can relate the GEV and GPD parameters to local geometrical properties of the attractor (Faranda et al., 2011; Lucarini et al., 2012a,b). The distance observables $g_i(r)$, with $r = \text{dist}(x(t), \zeta)$, are functions of the Euclidean distance between one point on the attractor ζ and the orbit $x(t)$ (Lucarini et al., 2016):

$$g_1(r) = -\log(r), \quad (2.5)$$

$$g_2(r) = r^{-1/\alpha}, \quad (2.6)$$

$$g_3(r) = -r^{1/\alpha}, \quad (2.7)$$

where α is a positive constant. The function $g_i(r)$ is chosen in a way to have a global maximum for $r = 0$, so that large values of g_i correspond to recurrences of the orbit near ζ . Depending on the choice of the function $g_i(r)$, the extremes of the distance observables can have positive, negative, or vanishing values for the shape parameter (Lucarini et al., 2016):

- g_1 -type observable:

$$\tilde{\sigma} = \frac{1}{D_\zeta}, \quad \tilde{\xi} = 0, \quad (2.8)$$

- g_2 -type observable:

$$\tilde{\sigma} = \frac{u}{\alpha D_\zeta}, \quad \tilde{\xi} = \frac{1}{\alpha D_\zeta}, \quad (2.9)$$

- g_3 -type observable:

$$\tilde{\sigma} = \frac{C - u}{\alpha D_\zeta}, \quad \tilde{\xi} = -\frac{1}{\alpha D_\zeta}, \quad (2.10)$$

where C is a constant. In particular, when $g_i(r)$ is chosen to be a power law, the shape parameter is non-zero, and it is proportional to the inverse of the local information dimension of the attractor. For the relationship between D_ζ and other attractor dimen-

sions, especially the Kaplan-York dimension, see Sec. 1.1. For a detailed treatment of the topic, please consult Lucarini et al. (2016). Note that the above equations are derived for the GPD parameters, however, one can obtain similar - in case of the shape parameters identical - relationships by considering the GEV distribution as well (Faranda et al., 2011; Lucarini et al., 2012a).

While recurrence properties are indeed important for characterising a system, distance observables are not well suited for studying some basic physical properties, such as, in the the case of fluids, energy or enstrophy. Hence, Holland et al. (2012) studied the extremes of smooth functions $S = S(x)$, which take their maximum on the attractor in a point where the corresponding level surface of $S(x)$ is tangential to the unstable manifold, referring to them as “physical” observables. They found a relationship between the GEV shape parameter and some geometrical properties of the attractor dealing with the properties of the unstable and stable directions in the tangent space. The results of Holland et al. (2012) were re-examined by Lucarini et al. (2014b), using the POT approach for physical observables of Axiom A systems. They considered the time-continuous time series of physical observables, and found that for all non-pathological physical observables S the shape parameter can be written as:

$$\xi_\delta = -\frac{1}{\delta}, \quad (2.11)$$

with δ defined as

$$\delta = d_s + (d_u + d_n)/2, \quad (2.12)$$

where d_s , d_u , and d_n are the partial dimensions of the attractor restricted to the stable, unstable, and neutral (i.e., central) directions. As mentioned above in Sec. 1.1, these local or point-wise dimensions take the same value almost everywhere on the attractor if one considers smooth observables of Axiom A systems. d_u is equal to the number of positive Lyapunov exponents (Ott, 1993), d_n is equal to the number of zero Lyapunov exponents, and $d_s = D_{KY} - d_u - d_n$, where D_{KY} is the Kaplan-Yorke dimension obtained according to Eq. (1.7). Figure 2.2 provides a general illustration of the selection of above threshold (vertical lines) events (blow-ups) in phase space, by the example of the two-dimensional Hénon map according to Lucarini et al. (2014b). We remark that a more general point of view, taking into consideration possible geometrical degeneracies in phase space, suggests that $-1/\xi_\delta < d_{KY} < -2/\xi_\delta$, and, additionally, $d_{KY}/2 = (d_s + d_u + d_n)/2 \leq \delta \leq d_s + (d_u + d_n)/2$ (Lucarini et al., 2016).

According to Eq. (2.11) the shape parameter is always negative (due to the compactness of the attractor), and it is close to zero in the case of systems having large Kaplan-Yorke dimension. Furthermore, it shows a universal property of extremes, which does not depend on the chosen observable but only on the geometry of the attractor. In Chapter 4 we will focus on comparing Eq. (2.11) with statistically inferred GEV and GPD shape

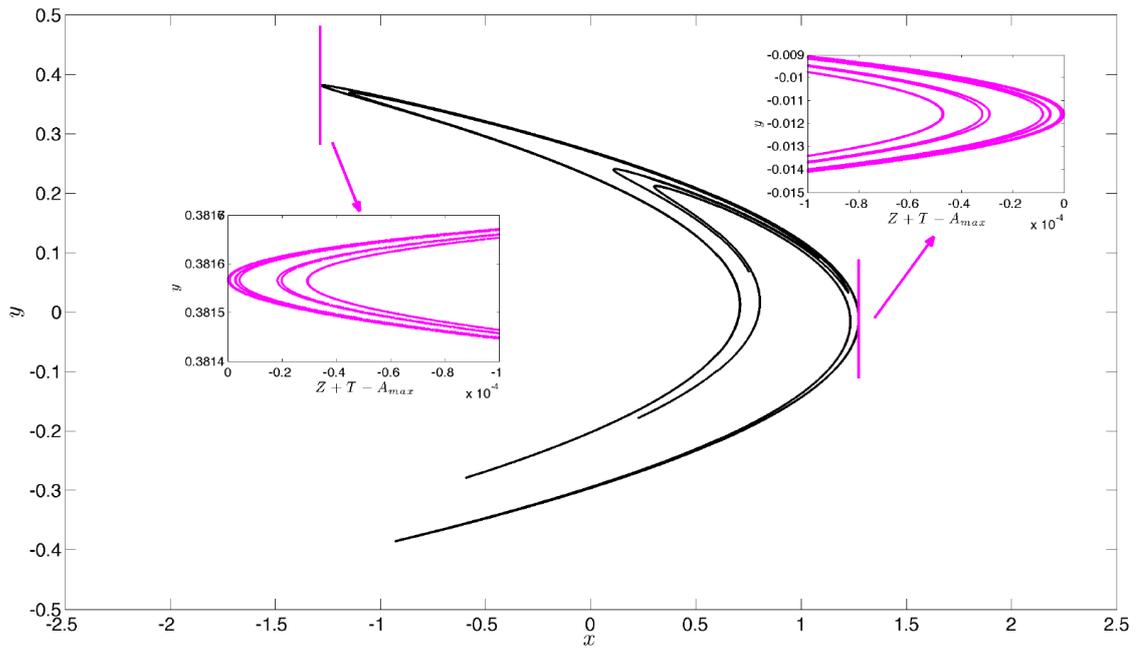


Figure 2.2: Selection of extreme events in phase space, based on the example of the Hénon map. The vertical lines indicate the thresholds, and the blow ups show the portions of the invariant measure corresponding to the extremes. Reproduced from Lucarini et al. (2014b).

parameters in the case of energy extremes in the two-layer QG model.

2.2 Large Deviation Theory

The large deviation theoretical framework can be formulated on three different levels, corresponding to the complexity of the statistical description of the dynamical system. These are, as described by Oono (1989), based on: sample means of observables (level-1), probability distributions on the state space of observables (level-2), and probability distributions on the path or history space, i.e. the entire set of possible orbits or histories of the system (level-3). The below description follows the level-1 approach, according to the scientific purpose of Chapter 5, and is mostly based on the works of Touchette (2009) and Oono (1989). We do not pursue at all a rigorous mathematical formulation here; our aim is rather to recapitulate the main concepts and results, and to introduce our notation. We also give some physical interpretations of the main mathematical concepts.

We say that the r.v. $A_n = \frac{1}{n} \sum_{i=1}^n X_i$, where X_i are identically distributed r.v., satisfies a

large deviation principle (LDP) if the limit

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \ln p(A_n = a) = I(a) \quad (2.13)$$

exists. $I(a) \geq 0$ is the so-called rate function, representing the rate of exponential decay of probabilities. The probability density $p(A_n = a)$ decays exponentially with n for every value of a except the ones for which $I(a) = 0$, where $\lim_{n \rightarrow \infty} p(A_n = a^*) = 1$, and $a^* = \mathbb{E}[A_n]$. Whenever this limit holds and $I(a)$ has a unique global minimum, A_n converges in probability to its mean a^* and obeys the Law of Large Numbers. If then additionally $I(a)$ is quadratic (i.e. twice differentiable) around a^* , the Central Limit Theorem (CLT) holds, meaning that small fluctuations around the mean are normally distributed. The expression “small fluctuations” is very important here, because large fluctuations around the mean are not necessarily normally distributed. Since the rate function describes both small and large deviations, LDT can be considered as a generalisation of CLT. Figure 2.3 shows, as an example, pdf’s of averages for increasing averaging windows and the corresponding rate functions in case of normally and exponentially distributed r.v. (Touchette, 2009).

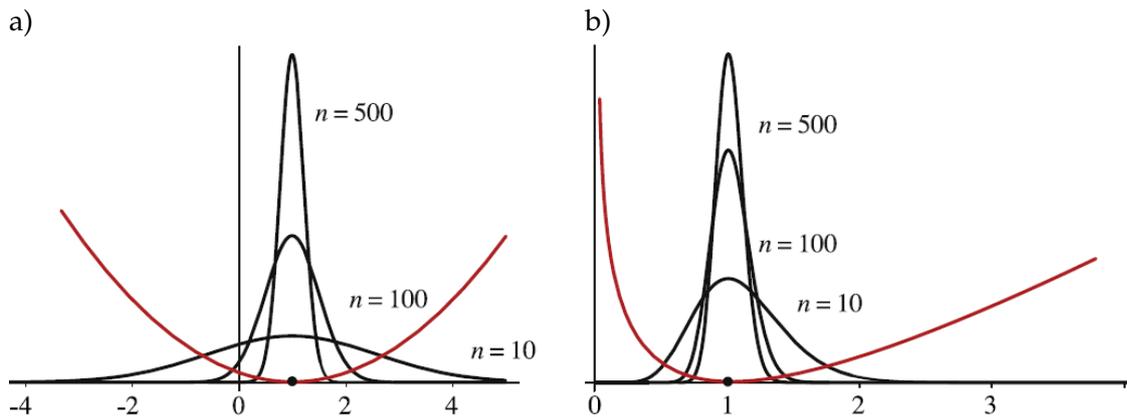


Figure 2.3: Pdf’s of averages (black) for increasing averaging windows and corresponding rate functions (red) for a) normal and b) exponential r.v. (Touchette, 2009).

Now let’s consider, instead of r.v., observables produced by a deterministic dynamical system. If the system is Axiom A, all of its observables obey a LDP (Eckmann and Ruelle, 1985). If we consider a system whose dynamics is characterised by high-dimensional chaos, by invoking the chaotic hypothesis (Gallavotti and Cohen, 1995), one can expect to find large deviation laws, even in systems which are not Axiom A.

The dynamical nature of out-of-equilibrium steady state systems requires, however, a slight modification of our theoretical approach, which mainly implies that time has to be considered in the formulation of the LDP, replacing the parameter n . Due to temporal correlations in these systems the computation of the rate function requires level-2 or

level-3 theory. This has been done for Markov chains and r.v. with a specific form of dependence, and involves mostly the computation of transition matrices or joint pdf's (den Hollander, 2000; Touchette, 2009). In the case of non-Markovian processes and high dimensional systems the computation of analytical rate functions is a hopeless endeavour. Thus, in this work, we adopt another (very simple) strategy in dealing with temporal correlations. Considering weakly correlated observables (i.e. X_j and X_l have an exponentially decreasing correlation if $|j - l|$ is large enough), one can take advantage of the fact that for large enough n the averages A_n become almost uncorrelated. This represents the basis for the block averaging method (Rohwer et al., 2015). We transform the variables X_i into variables $Y_i = \frac{1}{b} \sum_{i=1}^b X_i$, where b represents the size of the averaging block, i.e. $b = n/k$ with the number of blocks k . If Y_i are almost independent and identically distributed (ergodic Markov chain), a large deviation principle can be obtained for:

$$A_n = \frac{1}{k} \sum_{i=1}^k Y_i = \frac{1}{n} \sum_{i=1}^n X_i. \quad (2.14)$$

Intuitively, one can argue that b has to be at least so large that X_{i+1} and X_{i+b} are almost uncorrelated, i.e. $b \geq \rho$ where ρ is a metric of persistence expressed in terms of number of successive correlated data. One usually quantifies persistence in terms of the auto-correlation function. Considering our scientific goal, which is the study of probabilities of averages, it makes sense to choose the integrated auto-correlation as a general measure of patterns in time and space, since this quantity plays a central role in the formulation of the CLT for Markov chains, as described below.

According to a formulation of CLT in case of dependent variables based on Billingsley (1995), suppose that X_1, X_2, \dots is a stationary Markov chain with $\mathbb{E}[X_n] = 0$ and satisfies appropriate mixing conditions, then the variance of the sample mean A_n is

$$n\mathbb{E}[A_n^2] \rightarrow \mathbb{E}[X_1^2] \left(1 + 2 \sum_{k=1}^{\infty} c(k)\right) \quad (2.15)$$

where $c(k) = \frac{C(k)}{C(0)}$ is the auto-correlation, and $C(k)$ denotes the auto-covariance at lag k , $C(k) = \mathbb{E}[X_i X_{i+k}]$. Eq. (2.15) shows that the rescaled variance of the sample mean of the Markov chain converges to the variance of X_1 times the integrated auto-correlation.

Based on the LDP, we can estimate the cumulative distribution function $P(A_n \leq a)$, and obtain the return periods of events exceeding the value a as $\frac{1}{1 - P(A_n \leq a)}$.

3 Atmospheric models

We use for the analysis of extremes two numerical models which simulate atmospheric motions on large scales: the quasi-geostrophic two-layer (QG) model and the simplified global circulation model PUMA. The later one is based on primitive atmospheric equations obtained by adapting the Navier-Stokes equations to a rotating frame of reference where the vertical acceleration of the fluid is constrained to be small with respect to gravity (Klein, 2010; Holton, 2004). These equations constitute a good approximation of the atmospheric dynamics on horizontal spatial scales larger than a few tens of kilometres. The evolution equations of the QG model follow from the primitive equations based on the quasi-geostrophic approximation for the mid-latitudes β -plane. Thus, in this later case, a further approximation is performed valid on even larger horizontal spatial scales of $O(10^3)$ km, and allowing for only small departures from the geostrophic balance. Nonetheless, this model is still able to represent characteristic processes of the mid-latitude atmospheric dynamics, like baroclinic conversion and barotropic stabilisation. Fig. 1.4 (Klein, 2010) above shows the position of the QG and PUMA (corresponding to HPE and HPE+Coriolis in the figure) models embedded in a hierarchy of atmospheric models based on different scaling regimes. The reader is referred to Holton (2004) for a detailed physical and mathematical description of the primitive atmospheric equations and the quasi-geostrophic approximation.

In both models, moist processes are omitted and simple parametrisations are used to account for the effect of friction, diabatic heating, and diffusion, as discussed in details below. However, if the forcing is strong enough, both models represent chaotic dynamical systems that satisfy the chaotic hypothesis and, thus, can be considered as Axiom A-equivalent. This is a very important property for the applicability of the asymptotic theories used in this thesis, as we see later in Chapters 4 and 5. Considering the corresponding temporal and spatial scales, both models simulate atmospheric circulations similar to the real atmosphere, as shown by Fig. 3.2 and Fig. 3.3 below.

3.1 The quasi-geostrophic two-layer atmospheric model

We consider a spectral quasi-geostrophic (QG) 2-layer atmospheric model similar to the one introduced by Phillips (1956). Specifically, our model, including the simulation code, is the same as in Schubert and Lucarini (2015), and is a modified version of the one presented by Frisius (1998). The model represents synoptic scale mid-latitude atmospheric dynamics based on the quasi-geostrophic approximation, which assumes hydrostatic balance and allows only small departures from the geostrophic balance (Holton, 2004). The model features baroclinic conversion and barotropic stabilisation processes, and simulates a turbulent jet-like zonal flow when suitable values are chosen for the parameters of the system.

3.1.1 Model description

The model domain is a rectangular channel with latitudinal and longitudinal coordinates $(x, y) \in [0, L_x] \times [0, L_y]$. $y = 0$ represents the Equator, and $y = L_y$ corresponds to the North Pole. We assume periodicity along the x -direction, so that L_x corresponds to the length of the parallel at 45° N. The vertical structure of the model atmosphere consists of only two discrete layers: this is the minimal vertical resolution needed to represent baroclinic processes (Holton, 2004). As shown in Fig. 3.1, five vertical pressure levels define the two layers with boundaries at: $p_{2.5} = 1000$ hPa (surface level), $p_2 = 750$ hPa, $p_{1.5} = 500$ hPa, $p_1 = 250$ hPa, $p_{0.5} = 0$ hPa (top level). The geostrophic stream function ψ is defined at levels p_1 and p_2 , $\psi(p_1) = \psi_1$ and $\psi(p_2) = \psi_2$, where the quasi-geostrophic vorticity equation for the mid-latitude β -plane (3.1) - (3.2) is applied, while the vertical velocity ω is specified at level $p_{1.5}$, where the thermodynamic energy equation (3.3) is valid.

The model is described by the following equations in terms of the barotropic stream function $\psi_M = (\psi_1 + \psi_2)/2$, baroclinic stream function $\psi_T = (\psi_1 - \psi_2)/2$, and temperature T :

$$\frac{\partial}{\partial t}(\nabla^2 \psi_M) = -J(\psi_M, \nabla^2 \psi_M + \beta y) - J(\psi_T, \nabla^2 \psi_T) - r \nabla^2(\psi_M - \psi_T) + k_h \nabla^4 \psi_M, \quad (3.1)$$

$$\frac{\partial}{\partial t}(\nabla^2 \psi_T) = -J(\psi_T, \nabla^2 \psi_M + \beta y) - J(\psi_M, \nabla^2 \psi_T) + r \nabla^2(\psi_M - \psi_T) + k_h \nabla^4 \psi_T + \frac{f_0}{\Delta p} \omega, \quad (3.2)$$

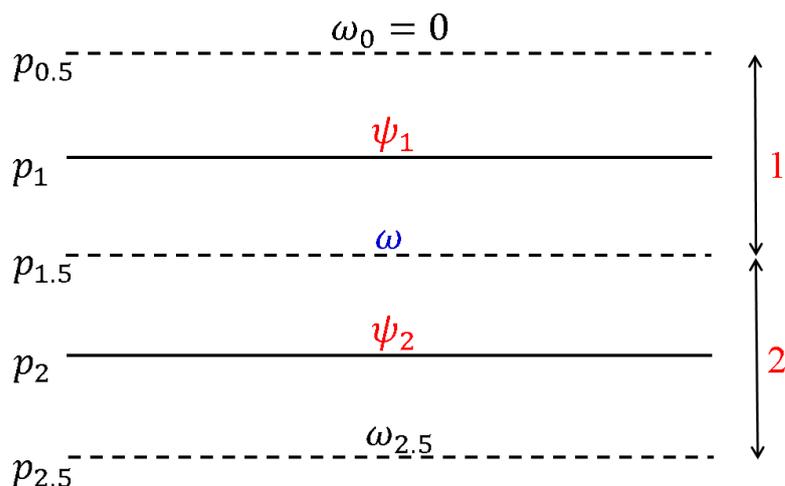


Figure 3.1: Vertical pressure levels of the two-layer QG model. The continuous lines represent model layers, and the dashed lines illustrate their boundaries. ψ represents the geostrophic stream function, and ω is the vertical velocity.

$$\frac{\partial}{\partial t}(T) = -J(\psi_M, T) + S_p \omega + r_R(T_e - T) + \kappa \nabla^2 T \quad (3.3)$$

In the above, we expressed the advection in terms of the Jacobian operator defined as $J(A, B) = \frac{\partial A}{\partial x} \frac{\partial B}{\partial y} - \frac{\partial A}{\partial y} \frac{\partial B}{\partial x}$. S_p represents the static stability parameter (Holton, 2004). We define the stability parameter $S = \frac{R S_p \Delta p}{2 f_0^2} = L_D^2$, where L_D is the Rossby radius of deformation. The name and values of model parameters are listed in Table 3.2.

The vertical velocity is set to 0 at the top level, $\omega_0 = 0$, and is defined through Ekman pumping at the surface level, $\omega_{2.5} = \frac{\Delta p}{f_0} 2r \nabla^2 \psi_2$, which parameterises the dissipative processes occurring in the boundary layer. Subgrid-scale processes are represented by momentum and heat diffusion terms. The system is driven by a Newtonian cooling term that involves the restoration temperature field:

$$T_e = \frac{\Delta T}{2} \cos \frac{\pi y}{L_y}. \quad (3.4)$$

ΔT denotes the forced meridional temperature difference, and quantifies the external forcing in the model. In the performed simulations, no time-dependence of ΔT is assumed, with the aim of creating time series of a deterministic equivalent of a stationary process. If ΔT is sufficiently large, the system reaches a steady state featuring a turbulent atmospheric flow with sensitive dependence on initial conditions. The physical processes responsible for limited predictability are in general the baroclinic and barotropic instability. The Newtonian cooling provides the so-called baroclinic forcing to the system, and activates a set of energy exchanges and transformations summarised by the framework of the Lorenz energy cycle. See discussion in Holton (2004).

According to Fig. 3.2, the QG model captures the basic features of the large scale atmospheric dynamics at mid-latitudes. In the upper layer, the model simulates a baroclinic westerly jet, while, in the lower layer, we find westerly zonal winds in the middle of the channel and easterly winds in boundary regions (Fig. 3.2a). We also notice that by decreasing the forcing level, the strength of the jet decreases considerably (as an effect of the lower meridional temperature difference), and the system becomes weakly chaotic, which can be seen by comparing the snapshots of the stream functions for the two forcing levels (Fig. 3.2b and Fig. 3.2c).

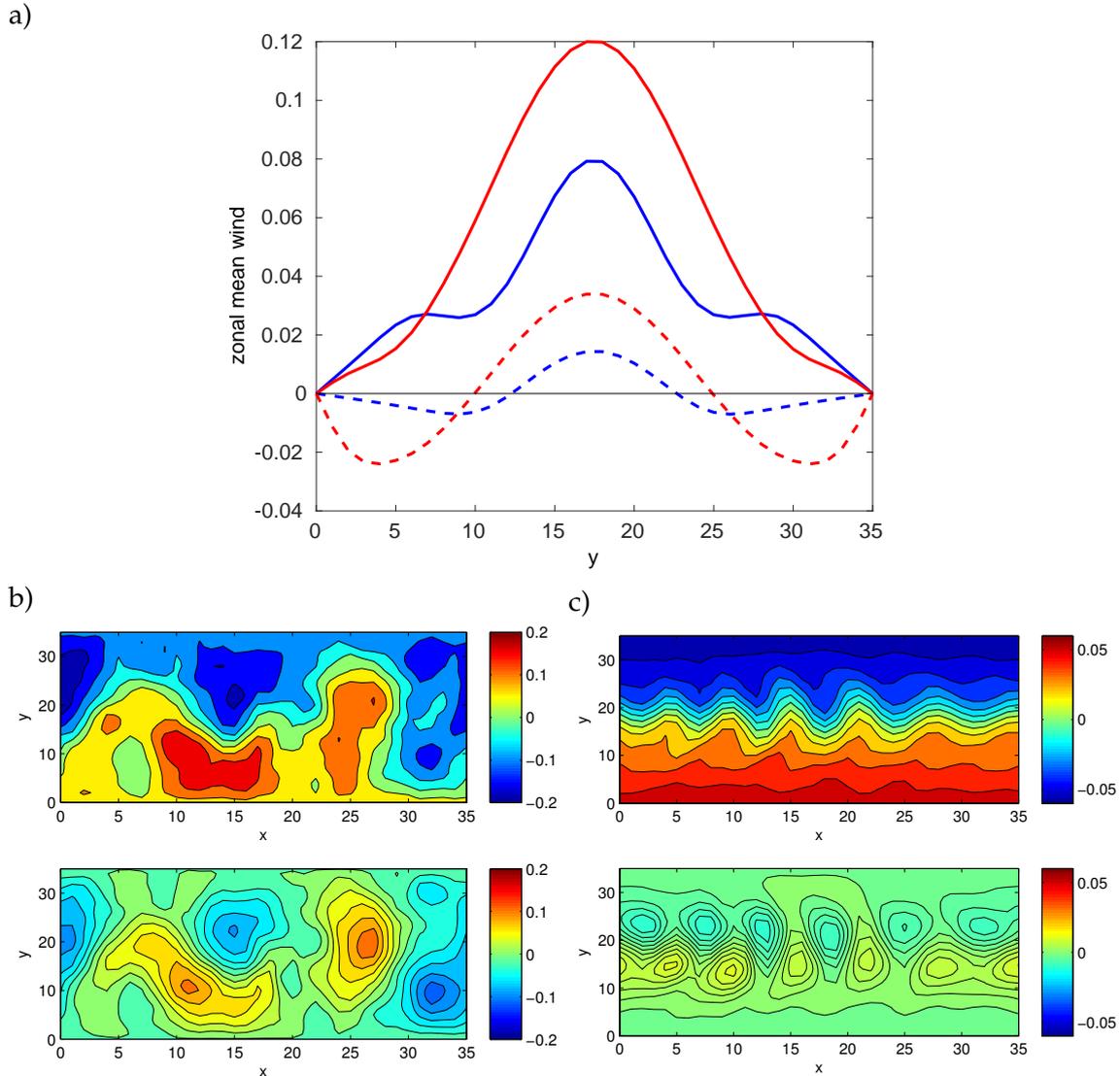


Figure 3.2: Zonal mean wind and stream function in the QG model. a) Zonal mean wind in the upper (continuous lines) and lower (dashed lines) layers for two different levels of forcing: $\Delta T = 133$ K (red lines) and $\Delta T = 40$ K (blue lines). Stream function snapshots in the upper (upper figures) and lower (lower figures) layers for b) $\Delta T = 133$ K and c) $\Delta T = 40$ K. The quantities are displayed in non-dimensional form (see Table 3.2 for scaling factors).

We obtain for our vertical discretization $T = \frac{2f_0}{R}\psi_T$ using hydrostatic approximation (Holton, 2004). Thus, the three model equations (3.1) - (3.3) can be reduced to two equations with two variables ψ_M and ψ_T . The model equations can be transformed (Schubert and Lucarini, 2015) into a non-dimensional form using the scaling factors in Table 3.2. In the following, we use non-dimensional quantities, if not indicated otherwise. As mentioned, the channel is periodic in the x -direction. At the meridional boundaries, we set the meridional velocity and the zonally-integrated zonal velocity to 0, $v = \frac{\partial\psi}{\partial x} = 0$ and $\int_0^{2\pi/a} dx \frac{\partial\psi}{\partial y} \Big|_{y=0,\pi} = 0$. For these boundary conditions, the solution of the model equations is

$$\psi(x, y, t) = \sum_{k,l=1}^{N_x, N_y} (\psi^r(k, l, t)\cos(akx) + \psi^i(k, l, t)\sin(akx)) \sin(ly) + \sum_{l=1}^{N_y} \psi^r(0, l, t)\cos(ly), \quad (3.5)$$

with ψ denoting the non-dimensional stream function, index r represents the real and i the imaginary coefficients. We apply a spectral cut-off at $N_x = N_y = 16$ in both x and y directions. Hence, the total dimension of the model phase space is $2N_y(2N_x + 1) = 1056$.

3.1.2 Methods

We substitute Eq. (3.5) into the evolution equations, perform a Galerkin projection, and eventually integrate numerically the non-dimensional model equations in spectral space using the fourth-order Runge-Kutta scheme. We perform simulations with two different forced meridional temperature differences, $\Delta T = 133$ K and $\Delta T = 40$ K. In the case of strong forcing ($\Delta T = 133$ K), the system has a Kaplan-Yorke dimension $D_{KY} = 585.95$ with 222 positive Lyapunov exponents, so that $d_u = 222$, $d_n = 2$, and $d_s = 361.95$. The computation of these dimensions is explained above in Sec. 2.1. Note that the presence of a second neutral direction is related to the existence of a rotational symmetry in the system and to the fact that we consider a spectral model. This feature is of little relevance for the analysis below. We produce stationary time series of 96,576 years with a time step of 0.7 hours. In the case of weak forcing ($\Delta T = 40$ K), the $D_{KY} = 39.31$ with 17 positive Lyapunov exponents, so that $d_u = 17$, $d_n = 2$, and $d_s = 20.31$. We produce stationary time series of 485,760 years with a time step of 2.8 hours. The spectral coefficients $\psi^{r/i}$ of the stream functions are recorded every 5.5 hours with either forcings. The Lyapunov exponents are obtained by the same simulation code as the one used by Schubert and Lucarini (2015), based on the method of Benettin et al. (1980).

The spectral output of the model is transformed into the grid point space using Fast Fourier Transform resulting $n_x \times n_y$ grid points with $n_x = n_y = 36$ in the x and y directions. We refer to the grid points by indices (i_x, i_y) where $i_x = i_y = 0, \dots, 35$. We

Table 3.1: List of symbols and parameter values for the QG model, Eq. (3.1) – (3.5).

Variable	Symbol	Unit	Scaling factor	
Stream function	ψ	m^2s^{-1}	L^2f_0	
Temperature	T	K	$2f_0^2L^2/R$	
Velocity	\mathbf{v}	ms^{-1}	Lf_0	
Energy	e	Jkg^{-1}	$L^2f_0^2$	
Parameter	Symbol	Dimensional Value	Non-dimensional value	Scaling Factor
Forced meridional temperature difference	ΔT	133 & 40 K	0.188 & 0.0564	$2f_0^2L^2/R$
Ekman friction	r	$2.2016 \times 10^{-6} \text{ s}^{-1}$	0.022	f_0
Eddy-momentum diffusivity	k_h	$10^5 \text{ m}^2\text{s}^{-1}$	9.8696×10^{-5}	L^2f_0
Eddy-heat diffusivity	κ	$10^5 \text{ m}^2\text{s}^{-1}$	9.8696×10^{-5}	L^2f_0
Thermal damping	r_R	$1.157 \times 10^{-6} \text{ s}^{-1}$	0.011	f_0
Stability parameter	S	$3.33 \times 10^{11} \text{ & } 2.52 \times 10^{11} \text{ m}^2$	0.0329 & 0.0247	L^2
Coriolis parameter	f_0	10^{-4} s^{-1}	1	f_0
Beta (df/dt)	β	$1.599 \times 10^{-11} \text{ m}^{-1}\text{s}^{-1}$	0.509	f_0/L
Aspect ratio	a	0.6896	0.6896	-
Meridional length	L_y	10^7 m	π	L
Zonal length	L_x	$2.9 \times 10^7 \text{ m}$	$2\pi/a$	L
Specific gas constant	R	$287.06 \text{ Jkg}^{-1}\text{K}^{-1}$	2	$R/2$
Vertical pressure difference	Δp	500 hPa	1	Δp
Time scale	t	10^4 s	1	$1/f_0$
Length scale	L	$10^7/\pi \text{ m}$	1	$10^7/\pi$

analyse extremes of total energy observables defined in non-dimensional form below. For our extreme value analysis we consider only the “mid-latitudes” of the QG model, which we define as the region between the latitudes $i_y = 9$ and $i_y = 26$, i.e. the latitudinally central 0.5 fraction of the whole domain. The total energy is the sum of the kinetic energy of the lower and upper layers and of the available potential energy:

$$e(i_x, i_y, i_t) = e_{k_1}(i_x, i_y, i_t) + e_{k_2}(i_x, i_y, i_t) + e_p(i_x, i_y, i_t), \quad (3.6)$$

where i_t represents the discrete time coordinate. The components of the right side of (3.6) are defined for each grid point as:

$$e_{k_1}(i_x, i_y, i_t) = \frac{1}{2} (u_1(i_x, i_y, i_t)^2 + v_1(i_x, i_y, i_t)^2), \quad (3.7)$$

$$e_{k_2}(i_x, i_y, i_t) = \frac{1}{2} (u_2(i_x, i_y, i_t)^2 + v_2(i_x, i_y, i_t)^2), \quad (3.8)$$

$$e_p(i_x, i_y, i_t) = 2\lambda^2\psi_T(i_x, i_y, i_t)^2, \quad (3.9)$$

with the zonal component of the horizontal velocity $u_1 = -\frac{\partial\psi_1}{\partial y}$, $u_2 = -\frac{\partial\psi_2}{\partial y}$, the meridional component of the horizontal velocity $v_1 = \frac{\partial\psi_1}{\partial x}$, $v_2 = \frac{\partial\psi_2}{\partial x}$, and $\lambda^2 = 1/(2S)$.

We obtain the zonally-averaged energy by taking the zonal average of the local energy (Eq. (3.6)):

$$e_z(i_y, i_t) = \frac{1}{n_x} \sum_{i_x=0}^{n_x-1} e(i_x, i_y, i_t), \quad (3.10)$$

and the average mid-latitude energy by averaging the local energy over the area corre-

sponding to the mid-latitudes:

$$e_{\text{ml}}(i_t) = \frac{2}{n_x n_y} \sum_{i_x=0}^{n_x-1} \sum_{i_y=9}^{26} e(i_x, i_y, i_t). \quad (3.11)$$

The energy observables are analysed in their non-dimensional form. The physical values expressed in units of J/Kg (J/m^2) can be obtained by multiplying the non-dimensional values by the factor $L^2 f_0^2 = 1.013 \times 10^5$ ($L^2 f_0^2 \Delta p / g = 5.164 \times 10^9$).

Although we record the model output, as stated above, every 5.5 hours, we save only the maximum values over one month in the case of strong forcing and over three months in the case of weak forcing. We estimate the GEV and GPD parameters based on block maxima and threshold exceedances obtained from the monthly, respectively 3-monthly, maxima series. Such an operation has no effect on the subsequent GEV analysis. Instead, it might modestly impact the GPD analysis, as some above threshold events might be lost because they could be masked by a larger event occurring within the same 1-month or 3-months period. Nonetheless, since we consider very high thresholds and an extremely low fraction of events, the risk of losing information is negligible. The GEV and GPD parameters are inferred by maximum likelihood estimation (MLE), as described by Coles (2001). We estimate the GEV and GPD parameters, as well as the confidence intervals, using the MATLAB functions `gevfit` and `gpfit` (MATLAB, 2015). The computed confidence intervals contain the true value of the parameters with a probability of 95%. The auto-correlation coefficients and histograms are obtained based on 1000 years of the “raw” simulated time series.

3.2 The simplified global circulation model PUMA

The Portable University Model of the Atmosphere (PUMA) is a simplified spectral general circulation model (GCM) developed at the University of Hamburg. PUMA has been used for the investigation of several atmospheric phenomena, like storm track dynamics or low frequency variability (Lunkeit et al., 1998; Fraedrich et al., 2005b), and has been even adapted to extra-terrestrial atmospheres (Grieger et al., 2004). A recent study investigates the properties of the Lyapunov spectrum in PUMA, including large deviations of finite time Lyapunov Exponents (De Cruz et al., 2018). PUMA is the dry core of the Planet Simulator (PlaSim), which is a climate model of intermediate complexity (Fraedrich et al., 2005a; Lucarini et al., 2010).

In the following, we summarise the model equations and the applied parametrisations. For a more detailed description of the model, please consult Fraedrich et al. (2009). The

physics of the model is fundamentally described by the primitive equations for the atmosphere, which amount to a modification of the Navier-Stokes equation in a rotating frame of reference where the vertical acceleration of the fluid is constrained to be small compared to gravity (Klein, 2010). These equations provide a good representation of the dynamics of the atmosphere for horizontal spatial scales larger than few tens of kilometres (Holton, 2004). Compared to a full atmospheric GCM, moist processes are omitted, and simple parametrisations are used to account for the effect of friction (Rayleigh friction), diabatic heating (Newtonian cooling), and diffusion. The Newtonian cooling and Rayleigh friction terms are such as that proposed by Held and Suarez (1994) for the comparison of dynamical cores of GCMs. The model equations allow for the conservation of momentum, mass, and energy. The prognostic equations for absolute vorticity ($\zeta + f$), where $\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$ represents the relative vorticity (u and v are the zonal and meridional wind components) and $f = 2\Omega \sin \phi$ the planetary vorticity (Ω : angular velocity of the Earth, ϕ : latitude), divergence $D = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$, temperature T , and surface pressure p_s can be written by using spherical coordinates and the vertical σ -system as follows:

$$\frac{\partial(\zeta + f)}{\partial t} = \frac{1}{1 - \mu^2} \frac{\partial F_v}{\partial \lambda} - \frac{\partial F_u}{\partial \mu} - \frac{\zeta}{\tau_F} - \kappa \nabla^8 \zeta \quad (3.12)$$

$$\frac{\partial D}{\partial t} = \frac{1}{1 - \mu^2} \frac{\partial F_u}{\partial \lambda} + \frac{\partial F_v}{\partial \mu} - \nabla^2 \left(\frac{U^2 + V^2}{2(1 - \mu^2)} + \Phi + T_0 \ln p_s \right) - \frac{D}{\tau_F} - \kappa \nabla^8 D \quad (3.13)$$

$$\frac{\partial T'}{\partial t} = -\frac{1}{1 - \mu^2} \frac{\partial(U T')}{\partial \lambda} - \frac{\partial(V T')}{\partial \mu} + D T' - \dot{\sigma} \frac{\partial T}{\partial \sigma} + \kappa \frac{T}{p} \omega + \frac{T_R - T}{\tau_R} - \kappa \nabla^8 T \quad (3.14)$$

$$\frac{\partial \ln p_s}{\partial t} = - \int_0^1 D + \vec{V} \cdot \nabla \ln p_s d\sigma \quad (3.15)$$

with

$$F_u = V(\zeta + f) - \dot{\sigma} \frac{\partial U}{\partial \sigma} - T' \frac{\partial \ln p_s}{\partial \lambda}$$

$$F_v = -U(\zeta + f) - \dot{\sigma} \frac{\partial V}{\partial \sigma} - T'(1 - \mu^2) \frac{\partial \ln p_s}{\partial \mu}.$$

The variables and parameters used in Eq. (3.12) – (3.15) are listed in Table 3.2.

The horizontal representation of the prognostic model variables is given by a series of spherical harmonics, which are integrated in time by a semi-implicit time differencing scheme (Hoskins and Simons, 1975). The linear contributions in the prognostic equations are computed in spectral space, the non-linear contributions in grid point space.

Table 3.2: List of variables and parameters in PUMA, eq. (3.12) – (3.15).

symbol	value	description
$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$		relative vorticity
$f = 2\Omega \sin \phi$		Coriolis parameter
Ω	$7.292 \cdot 10^5 \text{ s}^{-1}$	angular velocity of the Earth
$D = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$		horizontal divergence
T		temperature
T_0	250 K	reference temperature
$T' = T - T_0$		temperature deviation from T_0
p		pressure
p_s		surface pressure
$\sigma = p/p_s$		vertical coordinate
$U = u \cos \phi$		zonal velocity in spherical coordinates
$V = v \cos \phi$		meridional velocity in spherical coordinates
\vec{V}		horizontal velocity with components U and V
t		time
ϕ		latitude
μ		$\sin \phi$
λ		longitude
Φ		geopotential
$\omega = dp/dt$		vertical velocity in p-system
$\dot{\sigma} = d\sigma/dt$		vertical velocity in σ -system
τ_F		time scale for Rayleigh friction
K		hyperdiffusion coefficient
τ_R		time scale for Newtonian cooling
T_R		restoration temperature
κ	0.286	adiabatic coefficient

The horizontal resolution is defined by triangular truncation. The vertical discretization is based on finite differences on equally spaced σ -levels. The vertical velocity is set to 0 at the upper ($\sigma = 0$) and lower ($\sigma = 1$) boundaries.

A Rayleigh damping of horizontal velocities with time scale τ_F accounts for the effect of boundary layer friction in the lowest levels. $\tau_F = 0.6$ days at $\sigma = 0.95$ (the vertical level nearest to the surface), and $\tau_F = 1.65$ days at $\sigma = 0.85$. No friction is considered for higher levels, i.e. $\tau_F = \infty$. The effect of non resolved processes on the energy and enstrophy cascade is represented by hyperdiffusion ($\sim \nabla^{2h}$). The hyperdiffusion coefficient K is such that provides a maximal damping of the shortest waves and has no effect on the mean state (wave number 0). The integer exponent $h = 4$ leads to an additional damping of short waves. The diffusion time scale for the shortest wave is 1/4 days. The diabatic heating (cooling) is parameterized by a Newtonian cooling term. This forces the relaxation of the model temperature to a so-called radiative-convective *equilibrium* state specified by the restoration temperature T_R , which depends only on vertical level and latitude.

$$T_R(\phi, \sigma) = T_R(\sigma) + f(\sigma)T_R(\phi) \quad (3.16)$$

$T_R(\phi)$ describes the meridional form of the restoration temperature, whereas $f(\sigma)$ ac-

counts for the vertical changes in this meridional profile:

$$T_R(\phi) = (\Delta T_R)_{NS} \frac{\sin \phi}{2} - (\Delta T_R)_{EP} \left(\sin^2 \phi - \frac{1}{3} \right), \quad (3.17)$$

where $(\Delta T_R)_{NS}$ is the temperature difference between the North and South poles, and $(\Delta T_R)_{EP}$ represents the equator-to-pole temperature difference. The meridional temperature gradient decreases with height in the troposphere, $f(\sigma) = \sin(0.5 \pi(\sigma - \sigma_{tp})/(1 - \sigma_{tp}))$ for $\sigma \geq \sigma_{tp}$, and vanishes at the tropopause, $f(\sigma) = 0$ for $\sigma < \sigma_{tp}$, where σ_{tp} is the height of the tropopause. $T_R(\sigma)$ describes the vertical changes in the restoration temperature, representing a radiative-convective profile, i.e. accounts for the vertical effects of radiation and convection, in a general and simplified way, by connecting the vertical temperature profile to the moist adiabatic lapse rate:

$$T_R(\sigma) = (T_R)_s - Lz_{tp} + \sqrt{\left[\frac{L}{2} (z_{tp} - z(\sigma)) \right]^2 + S^2} + \frac{L}{2} (z_{tp} - z(\sigma)), \quad (3.18)$$

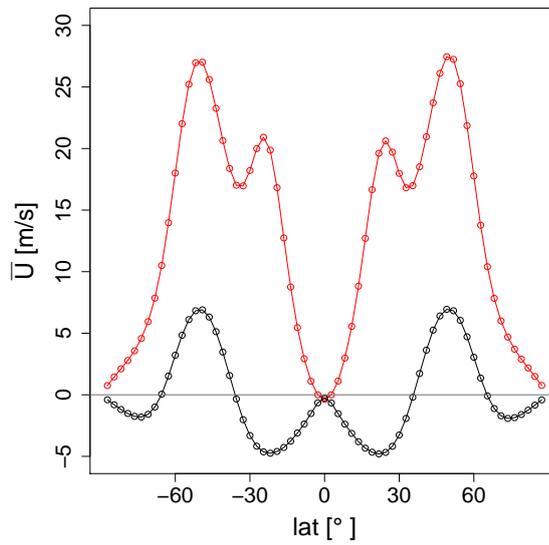
with: restoration temperature at the surface, $(T_R)_s = 288$ K; moist adiabatic lapse rate, $L = 6.5$ K/km; global constant height of the tropopause, $z_{tp} = 12$ km; geometric height z . S allows for a smoothing of the temperature profile at the tropopause. In case of 10 vertical levels l , the time scale of the Newtonian cooling τ_R is 2.5 days in the lowest level at $l = 10$, and 7.5 days at $l = 9$. τ_R continues to increase monotonically with height until the upper 3 levels, where it is set to 30 days.

We run the model in a simple symmetric setting (usually referred to as aqua-planet), i.e without orography. We remove the annual and diurnal cycle, and use a symmetric forcing with respect to the Equator, $(\Delta T_R)_{NS} = 0$. We set the equator-to-pole temperature difference $(\Delta T)_{EP}$ to 90 K, thus creating a baroclinically more unstable atmospheric state than in the standard setting with $(\Delta T)_{EP} = 70$ K. We run the model with constant forcing in time using a time step of 30 minutes. The horizontal resolution is T42 (triangular spectral truncation with 42 zonal waves), and the vertical resolution consists of 10 levels. The length of the simulations is 10^4 years, not counting the first 5 years, which are discarded to reach steady state. We consider for our analysis the air temperature in the lowest vertical level at 960 hPa, with daily output. The spectral temperature variable is transformed during the post-processing into grid point space consisting of a 65×128 equidistant latitude-longitude grid.

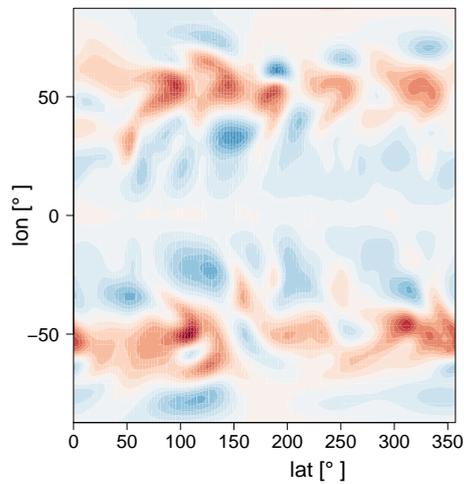
Figure 3.3a shows the long-term average of the zonal wind at 960 hPa (near-surface) and at 450 hPa. Near to the surface, we observe a westerly zonal flow at mid-latitudes as well as easterly winds in the subtropics (suggesting the existence of the Hadley cell) and near to the Poles. At 450 hPa, a westerly jet stream is dominant. Figure 3.3b and Fig. 3.3c present snapshots of the zonal and meridional wind components. One can identify the effect of eddies at mid-latitudes, which cause the main zonal and meridional wind

activity. These are responsible also for the meridional heat transport directed from the Equator to the Poles, acting against the mean meridional temperature gradient. Figure 3.3 suggests that the model simulates an atmospheric flow similar to the real atmosphere at large scales. Furthermore, the snapshots confirm the high level of chaos in the system.

a)



b)



c)

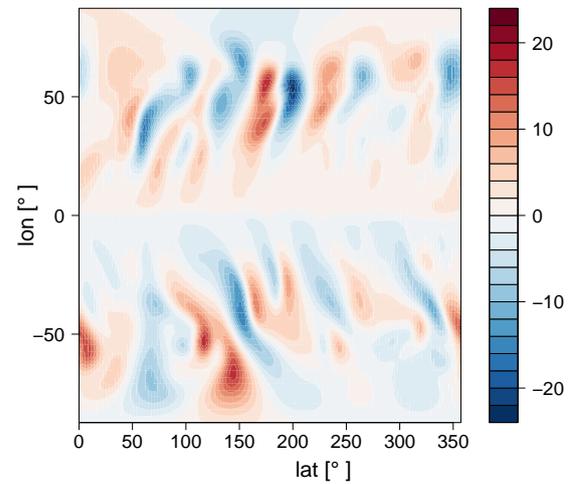


Figure 3.3: Zonal mean wind and wind components in the PUMA model (horizontal resolution: T42, vertical resolution: 10 levels, $(\Delta T)_{EP} = 90$ K). a) Zonal mean wind at 960 hPa (black) and at 450 hPa (red). Snapshots of b) zonal and c) meridional wind components at 960 hPa.

Using the same model settings as above, but with a lower equator-to-pole temperature difference, De Cruz et al. (2018) estimate a Kaplan-York dimension D_{KY} of 187 and a number of positive Lyapunov exponents of 68 for $(\Delta T)_{EP} = 60$ K. In this study, $(\Delta T)_{EP} = 90$ K, hence the model atmosphere is baroclinically substantially more unstable than in

the mentioned study. Thus, to provide a rough estimation, $D_{KY} > 200$ and the number of positive Lyapunov exponents > 80 in our system. Consequently, we expect for this set-up a very high dimensional chaos, which fulfils the chaotic hypothesis, as shown also by the fast decay of auto-correlations in Fig. 5.1c,d below. As a result, we expect that the outputs of our model can be analysed using EVT and LDT, as discussed above. Nonetheless, it is *a priori* unclear whether the asymptotic result can be clearly detected at finite size given the length of our numerical integrations. Note that in De Cruz et al. (2018) it was shown that the finite time Lyapunov exponents obey a large deviation law¹.

¹This is not always fulfilled in models which incorporate multiscale phenomena, as in the case of coupling atmospheric and oceanic dynamics. In the same study, the authors could not detect large deviation laws in case of a quasi-geostrophic coupled ocean-atmosphere model. Vannitsem and Lucarini (2016) analysed the large deviations of finite time Lyapunov exponents as well in a low-order version of the above mentioned coupled model, and found a LDP only in case of nonzero Lyapunov exponents, whereas the convergence was considerably slower or even absent in case of near-zero Lyapunov exponents.

4 Extreme events of energy observables in the two-layer quasi-geostrophic atmospheric model

As mentioned before, the simulations with the QG model are performed using two different configurations, where the value of the parameter ΔT describing the baroclinic forcing, is set to 133 K and 40 K respectively. As we see below, in the case of strong forcing we find good agreement between the results of the statistical inference based on the GEV or GPD distribution and the theory presented in Sec. 2.1, for local observables at least, even if the speed of convergence of the estimated shape parameters (not predicted by the theory), to the value that is predicted by the theory is rather diverse among the considered observables. In the case of weak forcing and the resulting weakly turbulent behaviour, the results of the statistical inference analysis are in worse agreement with the theory, and we find that for the different observables the shape parameter estimates have various non-monotonic dependence on the block size. We will investigate possible reasons for such a behaviour.

4.1 Strong forcing

Before presenting the results related to the statistics of extreme events, we outline some general statistical properties of the analysed observables. As emphasised in Sec. 2.1, correlations have an effect on the convergence of the distribution of block maxima (threshold exceedances) to the GEV (GPD) distribution.

By taking the ergodic hypothesis, we estimate the auto-correlation coefficients for the local energy, as explained in Sec. 2.2, by taking the ratio between the auto-covariance $C(l)$ at lag l and the variance $C(0)$, $\rho_e = C(l)/C(0)$, and obtain an estimate for each grid point $\rho_e = \rho_e(i_x, i_y, l)$. We calculate the integrated auto-correlation time scale (Franzke et al., 2005) $\tau_e = \tau_e(i_x, i_y)$ according to $\tau_e = \sum_{l=0}^{n_l} |\rho_e|^l$. We set $n_l = 604$

¹If the decay of the auto-correlation is exponential, the integrated auto-correlation time scale is equal to the e-folding time.

(corresponding to about 140 days) as an upper limit for the integration in order to avoid the noisy tail of the auto-correlation coefficient. We proceed the same way in the case of the zonally-averaged and average mid-latitude energy to obtain $\tau_{e_z}(i_y) = \sum_{l=0}^{n_l} |\rho_{e_z}(i_y, l)|$ and $\tau_{e_{ml}} = \sum_{l=0}^{n_l} |\rho_{e_{ml}}(l)|$. The integrated auto-correlation time scales, expressed in days, are shown in Fig. 4.1a. As expected, the weakest auto-correlations are recorded for the local observables, yielding about 1-2 days. Because of the information propagation along parallels due to the prevailing zonal winds, the zonal average time series are impacted by a low-pass filtering as result of averaging along a latitudinal band, thus the correlations become stronger. For these zonally-averaged observables, as opposed to the local ones, the integrated auto-correlation changes substantially with latitude. We observe a minimum in the middle of the channel (≈ 3.5 days) and an increase outwards to the boundaries (≈ 15 days). Through averaging over the area of mid-latitudes, the zonally-averaged time series with different properties are merged together. The resulting time series has an integrated auto-correlation time scale of about 3.6 days. Note that if in a time series of length N with reasonably fast (e.g. similar to exponential) decay of correlations the integrated auto-correlation time scale is τ , then one can deduce that the time series has approximately $N \times \Delta t / \tau$ effectively independent entries, where Δt is the time interval. This can have important effects in determining when the asymptotic behaviour of the EVT statistics is valid.

Figures 4.1b,c,d illustrate the histograms and the approximated probability density functions (pdf's) of our observables. Although, the dynamical properties of the flow as a function of latitudes are not exactly symmetric with respect to the meridional middle of the channel due to the β -effect, our estimations of statistical quantities and density functions exhibit approximate meridional symmetries. Therefore, only half of the channel's meridional extension (at every second latitude) is shown in the case of local and zonally-averaged observables. The strongest skewness and the longest right tails are observed in the case of pdf's of local observables. After spatial averaging, the pdf's become more symmetric and almost similar to a Gaussian distribution (which would look like a parabola on a semi-logarithmic scale), according to the central limit theorem.

In what follows, we present the results of the extreme value analysis starting with the local observables. We first discuss the convergence of the shape parameter for GEV and GPD, then the convergence of the GPD modified scale parameter (to be introduced below), and, at the end, the convergence of return levels. Taking advantage of the fact that statistics are uniform in the zonal direction, we concatenate the monthly maxima series for every second longitude one after the other in the x -direction, thus increasing the data length to about 1.7×10^6 (from about 9.6×10^4) years. Therefore, we can estimate the GEV and GPD shape parameters for larger block sizes and higher thresholds than in the case of zonally-averaged or average mid-latitude observables. Although the time series at every second longitude are correlated with each other, the correlation almost

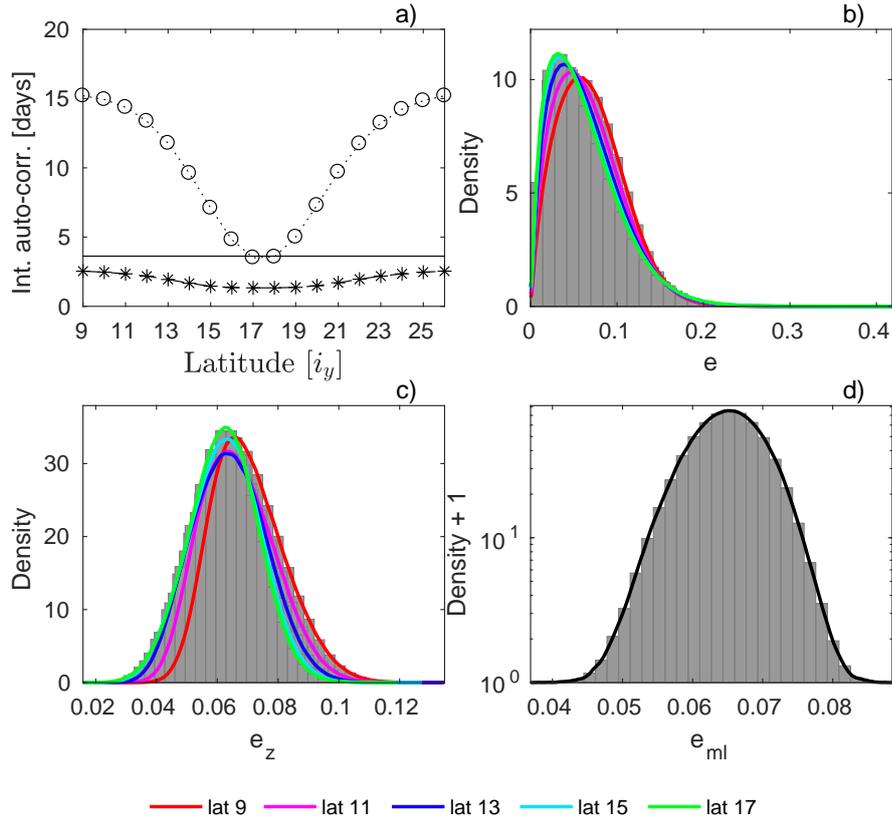


Figure 4.1: Statistical properties of the total energy for $\Delta T = 133$ K. a) Integrated auto-correlation time scales: zonal average of τ_e (dashed line with star markers), τ_{e_z} (dotted line with circle markers), $\tau_{e_{ml}}$ (continuous line); histograms of the b) local, c) zonally-averaged, and d) average mid-latitude observables. In the case of b) – d), the continuous lines show the approximation of the pdf's by kernel smoothing (*ksdensity* function in MATLAB), the colours mark different latitudes according to the legend.

vanishes at the block size of 8 years, being below 0.15 at every latitude. In other words, correlations are very weak at extreme levels, which is according to Coles (2001) the only important condition for the GEV limit laws to apply in the case of a stationary process. Block sizes smaller than 8 years are not relevant for our analysis, since (as presented below) much larger ones are needed to approach the theoretical shape parameter. In the case of the POT approach, we use the same argument of choosing very high thresholds, above which the correlations are extremely weak.

The theory discussed in Sec. 2.1 indicates that the true (asymptotic) GEV shape parameter is given by ξ_δ , as expressed by Eq. (2.11), which corresponds to approximately -0.002, and is indicated by the straight line in Fig. 4.2. Note that the range of the theoretical values derived taking into consideration possible geometrical degeneracies, according to what is described in Sec. 2.1, is too small to be visible in this case. We define the *precision* $P(m)$ of estimation by considering half of the width of the 95 % maximum likelihood

confidence interval. Additionally to the precision, we define the *trueness* of a *single* estimate by the distance between ξ_δ and $\xi(m)$: $T(m) = |\xi_\delta - \xi(m)|$. Note that the latter is different from the usual definition in that the reference from which we measure the distance is ξ_δ , not the true value of the distribution from which the BM data is drawn. In fact, strictly, the BM data is not drawn from a GEV distribution that we are fitting, and hence we cannot even talk about the true value of an underlying GEV distribution. We emphasise that our interest is the convergence to the asymptotic value, which is why we take a reference value in our definition other than customary. Accordingly, we shall refer to the ‘bias’ of the estimator, again different from customary, as the expected trueness. We remark that, since our estimates are obtained based on one realisation instead of several realisations yielding a distribution of estimates, our trueness T approximates the bias of the estimates as long as $T \gg P$. We emphasise that we are able to calculate T here because we know the true ξ_δ ; this is not the case in practice when facing just a measured time series. Obviously, we aim at obtaining a joint optimisation by having a bias and a precision as small as possible. Clearly, optimality requires a compromise between these two requirements. When we apply the BM method and increase the block size m , the number of blocks and of BM decreases, thus the estimation of $\xi(m)$ becomes more and more uncertain, and P increases monotonically. At the same time, for increasing m we expect a (not necessarily monotonic) convergence of our estimated shape parameter to the true value, so that the actual bias should (on the long run) decrease with m . Clearly, instead, our approximation T decreases only until a certain block size, above which it becomes more uncertain with increasing values of m , because less BM are available. We choose as optimal block size $m = m^*$ the smallest block size for which the estimate of the bias is lower than the estimate of the precision $m^* = \min(m; T(m) < P(m))$. On the scale of variation ranges of $P(m)$ and $T(m)$ we have $T(m^*) \approx P(m^*)$. With this we obtain a single number that can quantify the *accuracy* of estimation. This measure of accuracy provides here a basis for comparing different observables with regard to the speed of convergence, or a basis for assessing the degree of non-uniformity of estimates of various observables of interest (in terms of the range of accuracy values), as a finite-data-size deviation from the uniformity predicted by theory. We have verified that the optimal choice for m^* is virtually unchanged when we use an alternative definition of the accuracy such as $T^2 + P^2$, borrowing an idea concerning the optimality of MLE estimators (not shown).

First we assess the uniformity for the local observables. Figure 4.2a shows the GEV shape parameter estimates against exponentially increasing block sizes of $m = 2^i$ years ($i = -2, -1, \dots, 13$), for different latitudes. The estimated GEV shape parameters $\xi(m)$ seem to converge monotonically for every latitude to ξ_δ . The monotonic convergence is pointed out also in Fig. 4.2b in terms of $T(m)$. In this diagram we display $P(m)$ too, by which we can determine the optimal m^* and the accuracies of estimation. These

accuracies, depending on the latitude, have a range of $5 \times 10^{-3} - 2 \times 10^{-2}$. At the same time, the value of m^* ranges from several tens of years to a few hundreds of years depending on the latitude we are considering. This is unsurprising, because the speed of convergence to the asymptotic level is not universal. As a consequence, when finite block sizes are considered, extremes of different observables can feature rather distinct properties. The slow convergence suggests that customary choices like yearly maxima are not always good enough for an accurate modelling of extremes. Figure 4.2c, giving a different view of the same data seen in Fig. 4.2a, illustrates the estimated GEV shape parameter as a function of latitudes for various block sizes. For small block sizes, we observe a slight latitudinal dependence of the shape parameter. This latitudinal structure flattens as one increases the block size, and the estimated shape parameters get closer to the theoretical value. According to Fig. 4.2c, universality emerges as we approach the asymptotic level.

To assess the goodness of fit, we perform a one-sample Kolmogorov-Smirnov-test (KS-test) (Massey, 1951) at 5 % significance level using the MATLAB function `kstest` (MATLAB, 2015). We remark that the KS-test is performed in case of each block size based on the whole BM data, meaning that this amount of data decreases as we increase the block size. The shape parameter values for which the KS-test p-value p is above 0.05 (i.e. the hypothesis that the distribution of BM is a GEV distribution cannot be rejected) are marked by circle markers in Fig. 4.2a. We define m_{KS} as the smallest block size for which $p > 0.05$, $m_{KS} = \min(m; p > 0.05)$, and $\xi_{KS} = \xi(m_{KS})$. Figure 4.2 points out that the KS-test suggests a good fit already at smaller block sizes than the optimal block size, $m_{KS} < m^*$, and for lower shape parameter values than the best estimate, $\xi_{KS} < \xi(m^*)$. Thus, a very important conclusion is that the p-value of the KS-test is not an appropriate measure for the convergence to the limiting distribution. More precisely, it indicates that we have indeed agreement with a member of the GEV family of distributions, but we cannot say what is the error from the asymptotic value of the parameters. We emphasise that m_{KS} , just like m^* , depends on the time series length, and it would be even smaller if shorter time series were considered. This implies that in the case of applications with less data, the results of the KS-test are even less reliable. The misleading property of p-values was also shown by Bóday (2017), who studied the convergence to the GEV distribution of extremes of site variables in the Lorenz 96 model, and found p-values above the significance level in cases where the theoretical prediction did not even apply, and the shape parameter did not converge. The goodness-of-fit test was in the mentioned study a Pearson's chi-squared test. Misleading p-values based on the KS-test were pointed out also by Faranda et al. (2011) in the case of the BM approach in simple systems. A slow convergence of the estimated GEV shape parameters and a poor quality of diagnostic tools (return level and quantile plots) in case of small block sizes were also found by Vannitsem (2007) in case of local temperature extremes in a

three-layer QG model with orography.

Figure 4.2d illustrates the GPD shape parameter estimates as a function of decreasing exceedance ratio (the fraction of above-threshold data) r , which is equivalent with an increasing threshold. To ensure direct comparability between the BM and POT approaches of EVT, sample values of the threshold are chosen corresponding to the sample values of the block size, in such a way that $r = \frac{1}{m \times m_y}$, where m_y is the data amount in a year. Thus the number of threshold exceedances is equal to the number of block maxima. By comparing the GPD shape parameter (Fig. 4.2d,e,f) with the GEV shape parameter (Fig. 4.2a,b,c) we generally observe the same characteristics. More precisely, the changes of the GPD shape parameter as a function of exponentially decreasing exceedance ratio are very similar to the variation of the GEV shape parameter according to an exponentially increasing block size. Both GEV and GPD shape parameters seem to converge to ξ_δ . This is also consistent with theoretical results according to which the two distributions are asymptotically equivalent (Coles, 2001; Lucarini et al., 2014b). However, we expect that in case of finite block sizes (i.e. in case of every practical application) differences might emerge in the estimates of the GEV and GPD shape parameters. Although, in the case of consistent estimations one would expect that at large block sizes, corresponding to low exceedance ratios, the difference between them should be small, as it is the case for our estimators. Besides the mentioned similarities, we observe some differences between the estimates of the GEV and GPD shape parameters. These differences concern, for example, their latitudinal dependence (less pronounced in case of the GPD shape parameter) or the width of the confidence intervals (larger in case of the GPD shape parameter, indicating larger estimation uncertainty). The most relevant difference is, however, that the GPD shape parameter seems to converge faster to ξ_δ . This is unsurprising as in many applications it is usually suggested to use the POT over the BM method as the former is less data-hungry and provides (usually) a faster convergence (Lucarini et al., 2016).

We perform another test to check whether the GPD distribution is a good approximation for the distribution of threshold exceedances based on our data, and consider the GPD modified scale parameter. The GPD scale parameter depends on the chosen threshold according to $\sigma_u = \sigma_{u_0} + \xi(u - u_0)$ (Coles, 2001), where $\xi = \xi_\delta$ represents the asymptotic shape parameter, u_0 is the lowest threshold at which the GPD distribution is a reasonable model for exceedances, and u represents any other threshold $u > u_0$. The scale parameter can be reparameterized yielding the modified scale $\hat{\sigma} = \sigma_{u_0} - \xi u_0 = \sigma_u - \xi u$, which should converge to a nonzero value. Figure 4.3 illustrates the modified scale parameter estimates (calculated based on the finite-size GPD parameter estimates, i.e., taking threshold dependent GPD shape parameter estimates instead of ξ_δ) as a function of the exceedance ratio r . We observe estimates of $\hat{\sigma}$ relatively stable to further decreases of r (for $r < r^*$, $r^* = \max(r; T(r) < P(r))$), which supports the conclusions

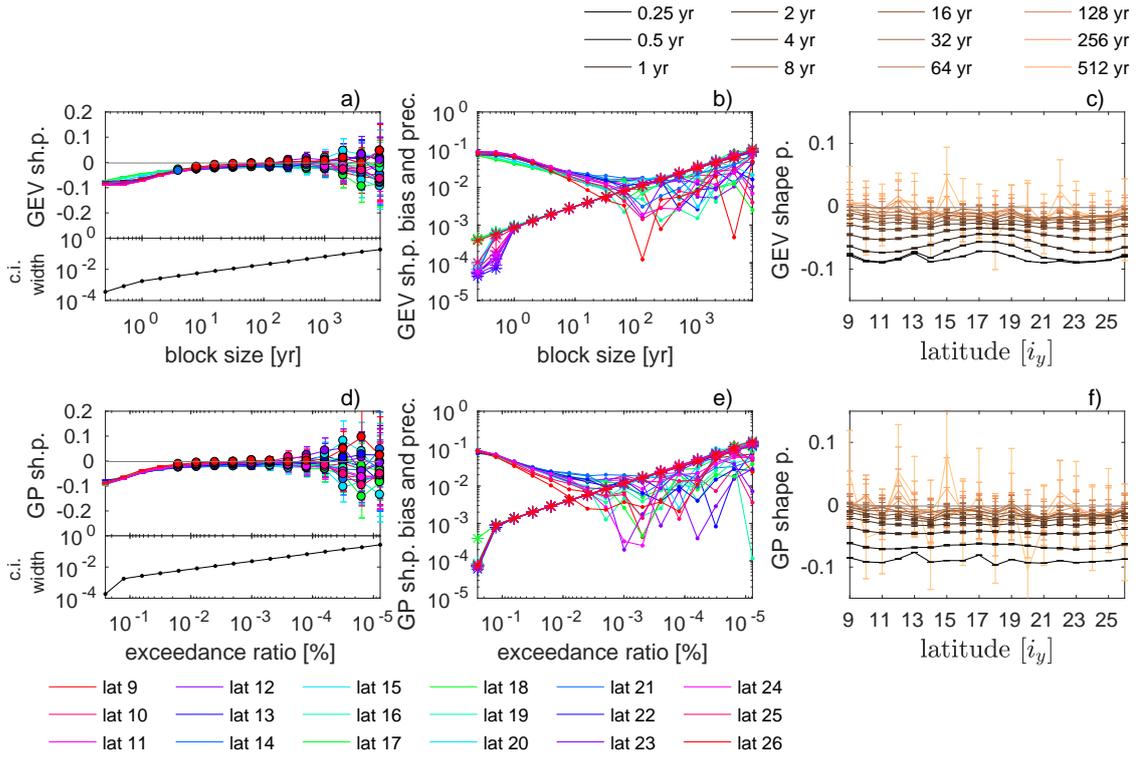


Figure 4.2: GEV and GPD shape parameters as well as bias and precision estimates in case of the local observables, for $\Delta T = 133$ K. a) GEV and d) GPD shape parameter estimates as functions of the block size and exceedance ratio. The circle markers point out shape parameter values, for which the p-value of the KS-test is above 0.05. Lower y-axis: zonal mean of the maximum likelihood 95 % confidence interval widths. b) and e): Estimates for the bias (dot markers) and precision (star markers) of the shape parameter. c) GEV and f) GPD shape parameter estimates as functions of the latitude. The grey, horizontal line illustrates the theoretical shape parameter of -0.002. The range of theoretical values resulting from taking into consideration possible geometrical degeneracies is invisible in this case. The error bars show the 95 % confidence intervals of the MLE. Different colours represent different latitudes (a, b, d, e) or different block sizes (c) or exceedance ratios (f).

drawn before that we are indeed close to asymptotic levels as required by EVT. Note that in this case there is no universality in the value of the modified scale parameter, as for stochastic variables one has that the upper right endpoint of the distribution is given by $\Lambda_{\max} = -\sigma_{u_0}/\xi + u_0 = -\hat{\sigma}/\xi$. Such an endpoint is clearly observable-specific.

Having practical applications in mind, the BM and POT methods aim at obtaining statistical estimates of either return levels or expected return periods, for even unobserved extreme events. Figures 4.4a,b show GEV and GPD return level plots for the local observables based on a fixed block size, $m = 128$ years, and corresponding (as explained) $r = 5 \times 10^{-4}$ % respectively, at five different latitudes (every second latitude from the

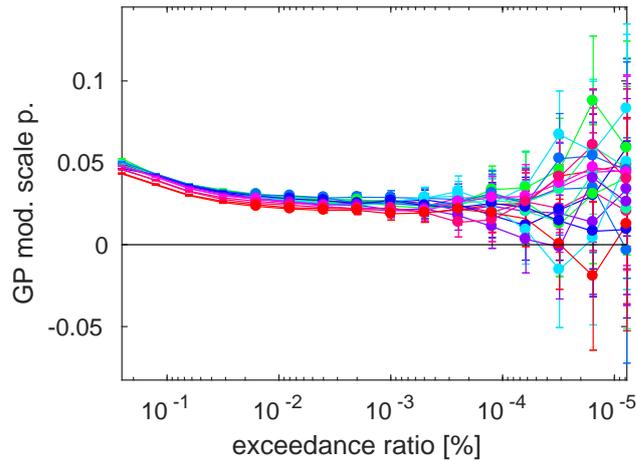


Figure 4.3: GPD modified scale parameter estimates in case of the local observables, for $\Delta T = 133$ K. The error bars show the 95 % confidence intervals of the maximum likelihood estimation. Different colours represent different latitudes according to Fig. 4.2.

southern meridional boundary to the channel centre). We compute the GEV return levels according to Eq. (2.3), the GPD return levels based on Eq. (2.4), and estimate the 95 % confidence intervals using the delta method described by Coles (2001). The GEV and GPD return level plots look very similar, except two minor differences. One emerges simply from the different equations for the GEV and GPD distributions, leading to slightly different definitions of return levels (as described in Sec. 2.1 and in more details in Coles, 2001) and affects short return periods; and the other one comes from the larger uncertainty in the estimation of the GPD parameters compared to the GEV parameters, and results slightly wider confidence intervals in the case of GPD return levels. The main message of Fig. 4.4a,b is, however, that the GEV and GPD return level estimates using the chosen m and r fit the empirical data quite well, which is in agreement with the results of the KS-test reported above. The 95 % confidence intervals of the estimated return levels (continuous lines) contain the empirical return levels (dot markers) or are very near to them, except a few very high extremes at some latitudes. The return level is almost linear to the logarithm of the return period, showing the effect of a shape parameter very close to 0 (see Eq. (2.3) and (2.4)).

If the GEV distribution is an adequate model for extreme events for a certain block size, one expects return levels with a certain return period not to change much any more with increasing block size. Figure 4.4c,d,e shows indeed that, above a certain block size, the estimated return levels for three different return periods (10^3 , 10^4 , and 10^5 years) are stable against further increase of m . But it also shows that the longer the return period, the slower the convergence. While in the case of the 10^3 -years return period we obtain stable return level estimates already at m_{KS} , in case of 10^5 -years the return

level estimates are still increasing for $m > m_{KS}$. Here we experience the practical effect of the issue mentioned above, namely, that the KS-test suggests a good fit even for $\xi_{KS} < \xi(m^*)$. This implies that the estimation of return levels with long return periods can be erroneous even if the KS-test does not reject the GEV distribution. We also notice that the return levels are underestimated if the block size is too small, and this underestimation is more severe in the case of return levels with longer return periods. We come to the same conclusion by considering the convergence of the GPD return levels (not shown), as suggested already by the similarity between Fig. 4.4a and Fig. 4.4b.

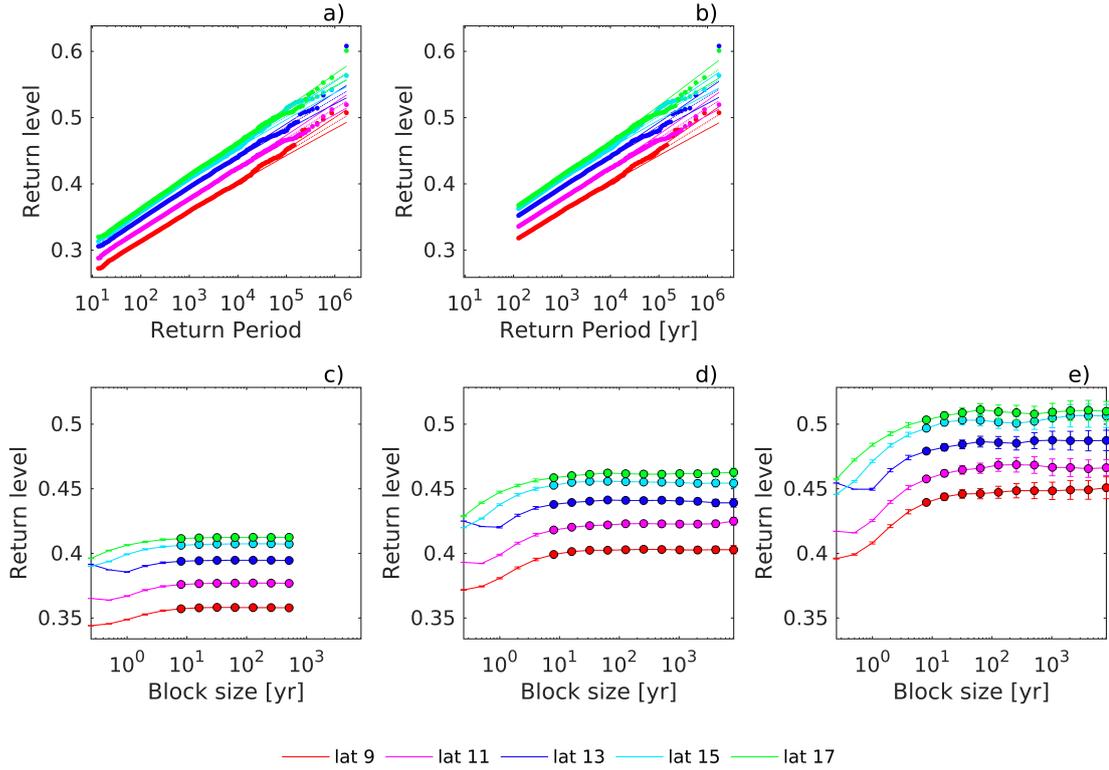


Figure 4.4: Return levels for $\Delta T = 133$ K. a) Return levels vs. return periods based on GEV parameters using a block size of 128 years and b) based on GPD parameters using an exceedance ratio of 5×10^{-4} %. (Dotted lines: estimated return levels; continuous lines: 95 % maximum likelihood confidence interval limits of the return level estimates; dot markers: empirical return levels.) GEV Return levels for c) 10^3 -years, d) 10^4 -years, and e) 10^5 -years return periods as functions of the block size. The error bars show the 95 % confidence intervals of the MLE. The circle markers point out estimates for which the p-value of the KS-test is above 0.05. The colours mark different latitudes according to the legend.

After having discussed in detail the convergence in case of the local observables, we proceed with the results for the zonally-averaged observables. Figure 4.5 illustrates the GEV and GPD shape parameters for the zonally-averaged observables (a, c, d, f) as well as the estimated bias and precision of the inferred shape parameters (b, e). As mentioned

above, in the case of zonally-averaged observables we have shorter time series (9.6×10^4 instead of 1.7×10^6 years). Because of this, results for the accuracies of estimates cannot be ‘fairly’ compared to the accuracies found for local observables. Nevertheless, we produce the same type of diagrams suitable to determine the accuracies and show it in Fig. 4.5b,e. Clearly, the range of accuracy values depending on the latitude and the maximal value of the accuracies (i.e of the bias at the optimal block size) are both considerably larger than those for the local observables. What is fair to compare, however, is the range of biases for a certain block size where the confidence of the estimates is high, $P \ll T$, and the amount of data does not affect significantly the parameter estimate. In this regard, the zonal observables display a much larger non-uniformity regarding the shape parameter estimates. Otherwise, the estimates feature typically a monotonic change towards the theoretical value (up to at least the optimal block size), what can be seen as convergence.

Our observation that the estimated shape parameters depend strongly on the considered latitude has to do with the effect of serial correlation on the convergence to the limiting distribution. We obtain weak auto-correlations, fast convergence to ξ_δ , and low bias in the middle of the channel, versus strong auto-correlations, slow convergence, and large bias at the margins of the channel. As already mentioned in Sec. 2.1, the stronger the serial correlation the less the number of uncorrelated data in a block, and the larger block sizes are needed in order to approach asymptotic levels (see also Coles, 2001). Thus the latitudinal structure of the GEV shape parameter estimates (Fig. 4.5c) is related to the one of the integrated auto-correlation time scale (Fig. 4.1a, dotted line with circle markers). By increasing the block size, this latitudinal structure flattens, the estimated shape parameters seem to approach ξ_δ , and the confidence intervals contain ξ_δ in the case of several latitudes, especially the central ones. Nonetheless, we note that, due to the presence of (relatively) large statistical uncertainty on the shape parameter, we cannot make more precise statements on the success of the analysis.

We present now the analysis of extremes of the average mid-latitude observable. Figure 4.6 shows the GEV and GPD shape parameter estimates for the average mid-latitude observable and their estimated bias and precision as a function of the block size and exceedance ratio respectively. In the case of average mid-latitude energy, we have the same amount of data as in the case of zonally-averaged energy. Similarly to the zonally-averaged observables, the estimated GEV and GPD shape parameters seem to approach the theoretical shape parameter, but, when more stringent definitions for selecting the extremes are used, the bias is relatively large, being about 4×10^{-2} at the optimal block size in the case of the GEV and about 7×10^{-2} at the optimal exceedance ratio in the case of the GPD shape parameter. Again, also in this case, our analysis is limited by the amount of available data.

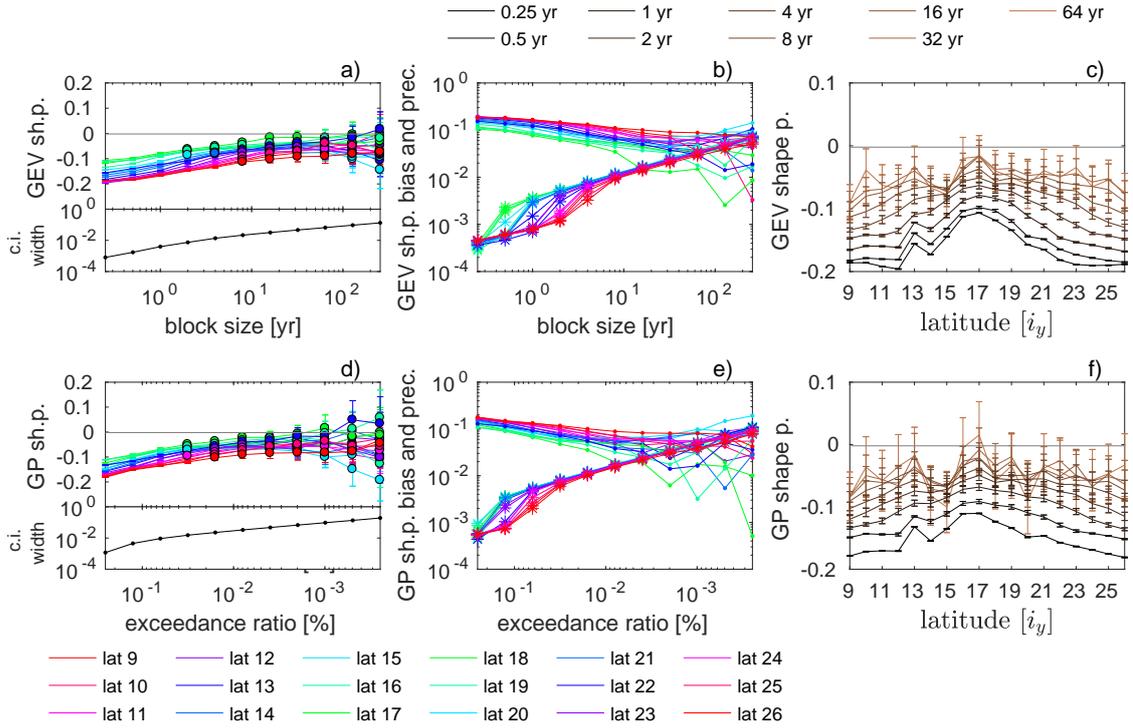


Figure 4.5: Same as Fig. 4.2, but for the zonally-averaged observables.

In short, our numerical results do allow for conclusions regarding the universality of extremes, as predicted by the theory presented in Sec. 2.1. However, considering the most various observables one would typically see a non-uniformity in the finite-size shape parameter estimates simply because of their distinct convergence properties (not predicted by the theory). The observables that we found in our study to have the fastest converging shape parameter estimates are the local observables at every latitude and the zonally-averaged observables at central latitudes, where the auto-correlation has a minimum. However, convergence is very slow, and is additionally slowed down by the presence of serial correlations in the time series. Thus, the estimated shape parameters are relatively far from the theoretical value in case of several latitudes of the zonally-averaged observables (especially marginal latitude exhibiting strong auto-correlations) and in case of the average mid-latitude observable. This slow convergence in combination with the finite size of the data makes the actual observation of the theoretical limit extremely difficult.

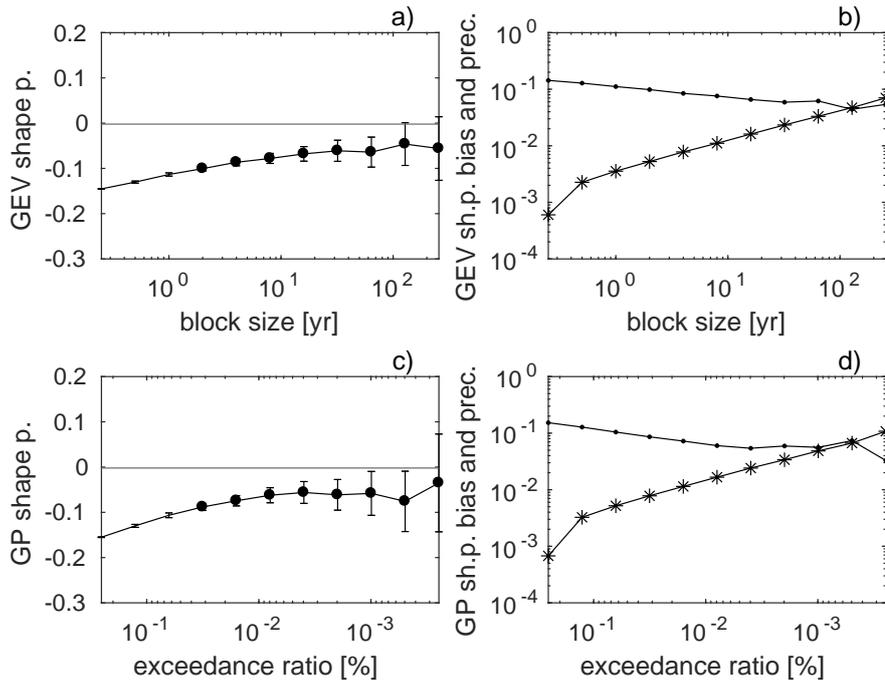


Figure 4.6: GEV and GPD shape parameter as well as bias and precision estimates in case of the average mid-latitude observable, for $\Delta T = 133$ K. a) GEV and c) GPD shape parameters as functions of block size and exceedance ratio. The circle markers point out shape parameter values for which the p-value of the KS-test is above 0.05. The grey, horizontal line illustrates the theoretical shape parameter of -0.002. The range of theoretical values resulting from taking into consideration possible geometrical degeneracies is invisible in this case. The error bars show the 95 % confidence intervals of the MLE. b), d) Estimates of the bias (dot markers) and precision (star markers) of the shape parameter.

4.2 Weak forcing

Before analysing the extreme events for weak forcing, we discuss some statistical (and dynamical) properties of our observables, which influence directly the statistics of extremes. Figure 4.7a shows the integrated auto-correlation time scales for the three observables: local, zonally-averaged, and average mid-latitude energy. We compute the integrated auto-correlation time scale according to the method described in Sec. 4.1 for the strong forcing. In the case of weak forcing, however, we set the time lag $n_1 = 1728$ (corresponding to about 400 days) as upper limit for the integration, according to the slow decay of the auto-correlation (especially in the case of zonally-averaged and average mid-latitude observables). The integrated auto-correlation time scales are substantially higher than for strong forcing: around 10 days in the case of local, about 30 – 48 days in the case of zonally-averaged observables, and approximately 45 days for the average mid-latitude observable. Figure 4.7b shows the time series of the local observables at

the central latitude $i_y = 17$ (at two different longitudes $i_x = 4$ and $i_x = 19$) and suggests two alternating states of our system: one with strong fluctuations and another one with reduced fluctuations. Thus, it seems that our system exhibits a regime behaviour, which definitely supports the presence of strong correlations.

In contrast to the case of strong forcing, the zonal averages of the local energy observables show remarkable deviations from a Gaussian behaviour, even more than the pdf's of the local energy observables (Fig. 4.7c,d,e). One has that the pdf's of the zonally-averaged observables typically have a marked skewness and very strong large kurtosis, and often contain rather pronounced "shoulders", where smoothness is basically lost. The presence of large kurtosis indicates that there is significant positive spatial correlation of the energy along a longitude. The presence of skewness indicates that there is asymmetry between the occurrence of anomalies of either sign. Another particular property of the spatial energy field for weak forcing is the strong anti-correlation (especially in the case of zonally-averaged observables) between time series at central and marginal latitudes (not shown). Accordingly, the "shoulders" appear in different parts of the pdf's at different latitudes: on the left in the case of central latitudes and on the right in the case of marginal latitudes. We conclude that the regime behaviour is connected to non-trivial spatial structures, with the system living in a transitional range where one can still distinguish long-lived unstable waves amidst chaos. We note that such conditions are different from what is foreseen by the chaotic hypothesis, and, therefore, the statistics of extremes might not converge (according to our finite-sized data set) to what is predicted by the theory developed for Axiom A systems.

For the analysis of extreme events, we use a similar procedure as in the case of strong forcing ($\Delta T = 133$ K), and concatenate the three-monthly maxima series for every second longitude one after the other in x -direction. Thus, we increase the length of available data for the local observables to about 8.7×10^6 (from about 4.8×10^5) years. Although the time series at every second longitude are correlated with each other, the correlation almost vanishes at extreme levels, being below 0.1 for every latitude in the case of the 8-years BM. We define the GPD exceedance ratios so that the number of threshold exceedances corresponds to the number of block maxima, as described in Sec. 4.1.

In the case of weak forcing, the theoretical shape parameter is -0.03, shown by the grey horizontal line in Fig. 4.8. The grey shading represents the range of theoretical values resulting from taking into consideration possible geometrical degeneracies according to the limits described in Sec. 2.1. We plot the GEV shape parameter against exponentially increasing block sizes of $m = 2^i$ years, where $i = 0, \dots, 15$ for the local observables and $i = 0, \dots, 11$ for the zonally-averaged and average mid-latitude observables. Focusing first on the local observables, we notice a non-monotonic change of the shape parameter according to increasing block sizes. For block sizes smaller than 30 years, the shape

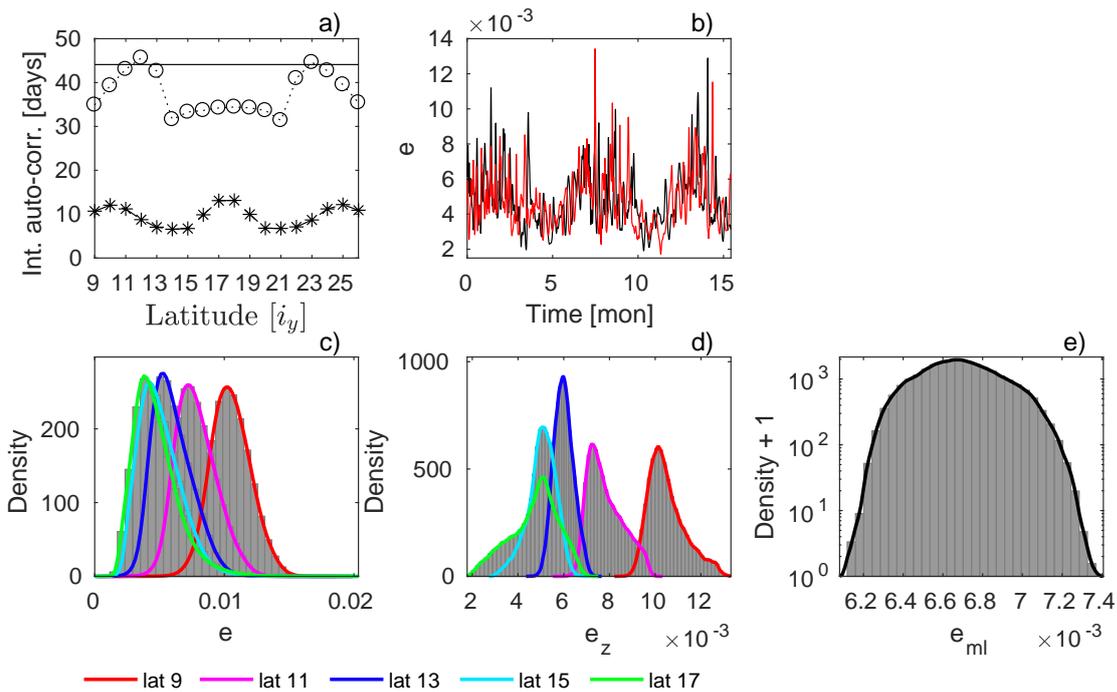


Figure 4.7: Statistical properties of the total energy for $\Delta T = 40$ K. a) Integrated auto-correlation time scales for the local (dotted lines with star markers), zonally-averaged (dashed lines with circle markers), and average mid-latitude (continuous line) observables. b) Time series of the local energy at latitude $i_y = 17$ and at two different longitudes: $i_x = 4$ (red line) and $i_x = 19$ (black line). Histograms of the c) local, d) zonally-averaged, and e) average mid-latitude observables. In the case of c) – e), the continuous lines show the approximation of the pdf’s by kernel smoothing (*ksdensity* function in MATLAB); the colours mark different latitudes according to the legend.

parameter even reaches non-physical, positive values for certain latitudes. This change of sign of the estimated shape parameters is similar to what has been observed by Vanitsem (2007) in case of local temperature extremes in a more realistic QG model with orography. The non-monotonic changes and the positive shape parameter estimates have to do with the fact that, if the block size is not large enough, we select events from both regimes (more and less fluctuating) thus “contaminating” the statistics of extremes; whereas if the block size is large enough, only extremes from the more unstable regime are selected. Figure 4.8a also shows that the estimated shape parameter seems to converge at almost every latitude to a value which is lower than the theoretical shape parameter, yet near to the range of values obtained taking into consideration possible geometrical degeneracies, see Sec. 2.1. As discussed above, this is in fact unsurprising given the qualitative properties of the system in the low forcing regime.

In case of the zonally-averaged and average mid-latitude observables we cannot detect any convergence. This is an expected result, considering the statistical and dynamical

characteristics of our data and the fact that the length of the time series is in this case even shorter than for the local observables. As an effect of the “shoulders” in the pdf’s, we obtain very uncertain estimates even for large block sizes, and the KS-tests reject the hypothesis of a GEV model in these cases. The shape parameter estimates have a large latitudinal spread due to the varying form of pdf’s according to latitudes. Except the differences between the GEV (Fig. 4.8a,b,c) and GPD (4.8d,e,f) shape parameters at small block sizes and high exceedance ratios, both methods show us basically the same picture. The misleading property of the KS-test p-values p is underlined by Fig. 4.8. Even in case of the zonally-averaged and average mid-latitude observables, where we cannot detect any convergence at all, we find $p > 0.05$ for a wide range of block sizes and exceedance ratios (circle markers).

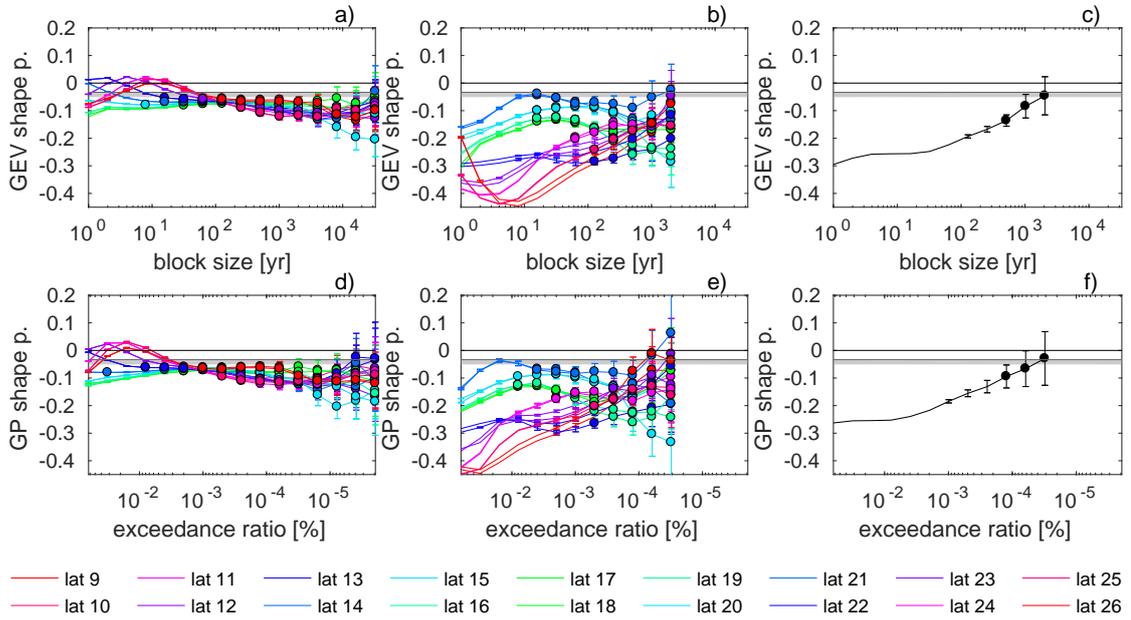


Figure 4.8: Shape parameter for $\Delta T = 40$ K. GEV shape parameter for a) local, b) zonally-averaged, and c) average mid-latitude energy. GPD shape parameter for d) local, e) zonally-averaged, and f) average mid-latitude energy. The circle markers point out shape parameter values, for which the p-value of the KS-test is above 0.05. Grey, horizontal line: theoretical shape parameter. The grey shading represents a possible range of the theoretical shape parameter according to the limits described in Sec. 2.1. The error bars show 95 % confidence intervals of the maximum likelihood estimation. Different colours represent different latitudes

4.3 Summary and discussion

In this chapter we have studied the convergence of statistically estimated GEV and GPD shape parameters to a theoretical shape parameter. The latter is calculated based on properties of the attractor (Holland et al., 2012; Lucarini et al., 2014b). We analyse a quasi-geostrophic 2-layer atmospheric model. We study extremes of different types of energy observables: local, zonally-averaged, and average mid-latitude energy. We perform simulations with two different forcing levels: a strong forcing ($\Delta T = 133$ K), producing a highly chaotic behaviour of the system, and a weak forcing ($\Delta T = 40$ K), producing a less pronounced chaotic behaviour. In the case of strong (weak) forcing, we produce time series of about 9.6×10^4 (4.8×10^5) years, representing a deterministic equivalent to a stationary process. We estimate the GEV and GPD shape parameters for exponentially increasing block sizes and exponentially decreasing exceedance ratios (fractions of above-threshold events), i.e. increasing thresholds, by performing maximum likelihood estimation. For comparability, we choose the GPD thresholds so that the number of threshold exceedances corresponds to the number of block maxima. We take advantage of the fact that statistics are uniform in the zonal direction, and use the data from every second longitude for the analysis of extreme events, thus increasing the length of available data for the local observables to about 1.7×10^6 (8.7×10^6) years in the case of strong (weak) forcing.

We start the discussion of our results with the strong forcing regime. In this case, we observe a roughly monotonic increase of the estimated GEV (GPD) shape parameters towards the theoretical value $\xi_\delta = -0.002$. The estimated shape parameters seem to converge to ξ_δ in the case of local observables at every latitude and in the case of zonally-averaged observables at central latitudes. Thus, our numerical results allow for robust conclusions regarding the universality of extremes, according to the theory presented in Sec. 2.1. However, in the case of several (especially marginal) latitudes of the zonally-averaged observables, as well as for the average mid-latitude observable, the estimated shape parameter is relatively far from the theoretical one. For these observables the amount of data seems to be not enough to approach asymptotic levels, thus we cannot make more precise statements on the success of the analysis. Even in this extremely chaotic case, the convergence is very slow, suggesting that customary choices like yearly maxima are not always the best option for an accurate modelling of extremes.

Despite the predicted universal asymptotic properties of extremes, if we consider a certain block size (threshold), we find that the shape parameter estimates are different among the observables and latitudes. Thus, on pre-asymptotic level, extremes show rather diverse properties. The speed of convergence to the asymptotic level is not universal. The local observables exhibit high-frequency fluctuations, as an effect of boundary fluxes, and at the same time, the fastest convergence of the shape parameter estimates to

the theoretical value. Since the energy is transported mostly along the zonal direction by the zonal mean flow, by averaging along a latitudinal band, the highest frequencies are filtered out, and fluctuations with lower frequencies become stronger. In the case of the zonally-averaged observables, we obtain weak auto-correlations and fast convergence to ξ_δ in the middle of the channel where the baroclinicity is the strongest, versus high auto-correlations and slow convergence at the margins of the channel where instead the baroclinicity is weak. The stronger the serial correlation, the less the number of uncorrelated data in a block, and the larger block sizes are needed in order to approach asymptotic levels (see also Coles, 2001). By averaging over the mid-latitude area, one merges zonally-averaged time series exhibiting different auto-correlations. Thereby, the convergence to ξ_δ is faster than in the case of zonally-averaged observables at marginal latitudes. To sum up, a very important conclusion of our study is the existence of latitude-dependent pre-asymptotic differences, as a counterpart to the universal asymptotic properties.

We assume that the extremely slow convergence has to do mainly with the fact that ξ_δ is negative but very close to 0. Based on ξ_δ and on the estimated GPD modified scale parameter, one is able to estimate according to Lucarini et al. (2014b) the absolute maximum, which is the upper end point of the GPD distribution, as mentioned in Sec. 4.1. By performing a very rough estimation (and neglecting the weak latitude-dependence of the GPD modified scale parameter), the absolute maximum in case of the local observables $A_{\max} \approx 12.5$, which is about 200 times the mean local energy value (see Fig. 4.1) and 20 times larger than some of the largest estimated return levels obtained for the largest return times considered here (see Fig. 4.4). This means that extremes are bounded, and an absolute maximum does exist, but the tail is extremely stretched out, and ultra long simulations are needed to explore this absolute maximum. Our results point out the discrepancy between the existence of a mathematical limit and the actual possibility of observing it. Note that if the asymptotic shape parameter is lower, the absolute maximum will be much closer to the maximum observed within a long, yet finite, time series, as it is shown in a recent study on temperature extremes in Southern Pakistan (Zahid et al., 2017).

Our conclusions regarding the convergence of the estimated shape parameter to ξ_δ are confirmed by results based on the GPD modified scale and return level estimates, in the case of local observables. We point out, however, that the longer the return period, the slower the convergence of the estimated return levels to their asymptotic values, and the larger the underestimation of the asymptotic return levels if we consider small block sizes (low thresholds).

In the case of weak forcing, temporal and spatial correlations are very strong due to a regime behaviour of our system, which exhibits two regimes: a more unstable one

with stronger fluctuations and a less unstable one with reduced fluctuations. Due to this regime behaviour the statistics of extreme events are “contaminated”: if the block size (threshold) is not large (high) enough, we select events from both regimes, whereas if it is large (high) enough, only extremes from the more unstable regime are selected. This induces non-monotonic changes of the estimated shape parameters by increasing the block size (threshold), and leads to the appearance of positive, i.e., non-physical, or very low shape parameter estimates. In the case of the local observables, the estimated shape parameters seem to converge at almost every latitude to a value which is lower (≈ -0.06) than the theoretical shape parameter ($\xi_\delta = -0.03$). Furthermore, in the case of the zonally-averaged and average mid-latitude observables, we cannot detect any convergence at all. The inconsistency of our numerical results with the theory is, in fact, unsurprising given the qualitative properties of the system in the low forcing regime, which do not resemble characteristics of Axiom A systems (at least on the finite time scales we are able to explore based on the available data).

Our results show that with increasing block size or threshold the shape parameters of the GEV and GPD distributions are becoming more and more similar, according to the asymptotic equivalence of the two models (Coles, 2001; Lucarini et al., 2014b). Both methods show us basically the same picture regarding the statistical properties of extreme events. Despite the mentioned similarities, we observe also some differences between the two approaches. The convergence to the limiting distribution seems to be somewhat faster in the case of the POT approach. This is in agreement with the well-established fact that the POT approach produces often more accurate predictions in the case of applications (Davison and Smith, 1990; Coles, 2001). Despite the faster convergence, however, the best GPD shape parameter estimates (defined in Sec. 4.1) do not approximate ξ_δ more accurately than the best GEV shape parameter estimates. Therefore, the advantage of the POT approach compared the BM approach is irrelevant in the case of very long time series.

We use the Kolmogorov-Smirnov test (KS-test) to verify the fit of the GEV (GPD) distribution to the distribution of extremes, selected as block maxima (threshold exceedances). Our results show that the KS-test is merely an indicator of the fit quality, and does not show whether the convergence to the correct GEV (GPD) distribution is reached or not. The KS test suggests a good fit to the GEV (GPD) distribution even in cases when the distance between the estimated and the asymptotic shape parameter is substantial and even if no convergence can be detected. The misleading property of p-values of the KS and Pearson’s chi-squared tests was also pointed out in previous studies in the case of more simple systems (Faranda et al., 2011; Bódai, 2017). In this work, we estimate the GEV and GPD parameters performing maximum likelihood estimation (Coles, 2001), but it would be relevant to find out to what extent other estimation procedures, like the L-moments (Hosking, 1990) or probability-weighted moments methods (Hosking et al.,

1985), would change the results.

Concluding, we would like to emphasise some key messages one can get from our results:

- Indeed, we have been able to find the signature of the universal properties of the extremes of physical observables in strongly chaotic dynamical systems, as predicted in the case of Axiom A systems. Nonetheless, given the availability of very long, yet finite, time series, we have been able to find more convincing results (yet with a relatively large uncertainty) only for specific observables, because in the case of observables featuring serial correlations it is extremely hard to collect robust statistics of extremes.
- We have observed that in the case of strong forcing the estimate of the shape parameter increases monotonically towards its asymptotic value for stricter and stricter criteria of selection of extremes. This corresponds to the fact that we manage to collect more detailed information on the local properties of the attractor near the point of absolute maximum of the observable, and thus explore all the dimensions of the attractor.
- We also remark that agreement of the results with the theory of extremes of observables of dynamical systems developed in the context of Axiom A flows cannot be found in the case of the weakly chaotic flow featuring regime behaviour and strong spatial and temporal correlations, as these features suggest strong deviations from the conditions behind the chaotic hypothesis. Note that conceptually analogous results had been found in a simple model in Bódai (2017).
- We note that the predicted and estimated shape parameters are extremely small so that the statistics of extremes is virtually indistinguishable, up to ultra long return periods, from what would be predicted by a Gumbel distribution ($\xi = 0$), which emerges as the statistical model of reference for physical extremes in high dimensional chaotic systems, and suggests in the case of fluids the existence of a well-developed turbulent state.
- We conclude by noting that in some cases of great practical relevance one finds results in contradiction with the basic tenets of the theory of extremes of dynamical system, suggesting that one should never find block maxima distributed according to the Fréchet distribution, which allows for arbitrary large extremes. The precipitation, as opposed to geophysical fields like temperature or pressure, is a non-smooth intermittent field with multifractal properties in space and time (Deidda, 2000; Bernardara et al., 2007), as a result of the complex chain of multiscale physical processes ranging from large scale water vapour transport on scales of 10^6 m and days, convection occurring on scales of 10^4 m and hours, and phase tran-

sitions occurring at microscopic level and on ultra short time scales. As a result of the intermittency of the rainfall, a very large (yet finite) amount of water can precipitate at a specific location, with little or no precipitation occurring nearby, as in the case of localised intense thunderstorms. Instead, extremely large anomalies of temperature or pressure cannot be reached as the climate has efficient mechanisms to dissipate them via, e.g. waves. Indeed, the analysis of block maxima of rain gauge readings shows many cases where the Fréchet distribution appears as the optimal model (Koutsoyiannis, 2004a,b). This is a result of the fact that in order to observe the actual physical limit of rainfall one should observe the system for an impossibly long time, and that closed physical budgets exist locally (and on a finite spatial domain) for the water balance (involving evaporation from surface and horizontal convergence of water transport), not in the precipitation per se. One can expect that such an anomalous behaviour is reduced if one chooses a smoother, better defined observable, such the spatial average of precipitation over a region.

5 Large deviations of temperature in the general circulation model of the atmosphere PUMA

After analysing extremes in the QG model based on EVT and connecting their statistics to universal properties of the system, we continue with the study of persistent extreme events in the simplified general circulation model of the atmosphere PUMA. For this, we use LDT and we rely on the connection between extremes of finite size averages and persistent events, which is true for certain temporal or spatial scales, as shown below. We see furthermore that the LDP in one direction, i.e. in space or in time, is not suited for the analysis of persistent events because it becomes valid on large scales, where persistence vanishes and universality prevails. To capture persistent spatio-temporal events, we need to perform averaging in two dimensions: first we average in space on intermediate spatial scales (defined later in this chapter), and then we obtain a LDP in time. We find that on these intermediate scales the system allows for a higher degree of spatio-temporal organisation than on very small or very large ones. We also search for the connection between temporal and spatial, as well as spatio-temporal averaging. We compare the results obtained by LDT based on return level plots with the empirical data and with return levels obtained by the POT approach of EVT.

Before discussing our main results related to the large deviations of temperature, it is useful to have a general picture about the properties of the simulated temperature field at 960 hPa (i.e. close to the surface). For the analysis of temporal, zonal, and spatio-temporal large deviations, we select three latitudes: 60° , 46° , and 30° (dashed horizontal lines in Fig. 5.1a). We focus on the mid-latitudes because it is the region of the atmosphere with the strongest turbulence, so that we expect that the corresponding observables should behave in coherence with the chaotic hypothesis. The two hemispheres are symmetric, thus it is not important which one we consider. In the following, we provide first a qualitative comparison of temporal and spatial characteristics of the temperature field, and then we quantify the persistence in time and space based on the integrated auto-correlation.

Figure 5.1a illustrates the temperature field $T(x, y, t^*)$ as function of longitude x and

latitude y at one selected time point t^* , whereas Fig. 5.1b represents the temperature field $T(x^*, y, t)$ as function of latitude y and of time t at one selected longitude x^* . Qualitatively similar figures would be obtained for different values of t^* and of x^* , respectively. Note that, to facilitate the comparison between space and time, the x -axis in Fig. 5.1b is backward in time according to the prevailing eastward zonal wind at mid-latitudes. Additionally, the range of the x -axis in Fig. 5.1b is the same as in Fig. 5.1a once we rescale the time axis according to the scale velocity U_τ introduced below (computed for 46°), which weights the decay of correlation in space at fixed time and in time at fixed location. Comparing these two figures we realise that by cutting across time or across longitudes we obtain very similar wavy patterns, which is non-surprising since the forcing is invariant in time and along a latitudinal band.

While this result would be trivial when observing a periodic or quasi-periodic signal, we need to consider here that the dynamics of the atmosphere features a nontrivial mixture of wave, turbulence, and particles (Ghil and Roberston, 2002), so that we need to look at this space-time similarity from a statistical point of view. According to this, we have that, at a given latitude y^* , the temporal series $T(x^*, y^*, t)$ and the zonal series $T(x, y^*, t^*)$ are sampled from two similarly distributed random processes, given the condition of steady state and the discrete symmetry with respect to translation along latitudes.

The main difference is related to distinct temporal and spatial characteristic scales, i.e. to temporal or spatial correlations. At mid-latitudes, cyclones have a typical temporal scale of ≈ 1 day and a characteristic spatial scale of about 1000 km (Holton, 2004). Obviously, these scales are relevant when we try to obtain a LDP, thus it is very important to find an adequate metric to describe them. We quantify the typical temporal and zonal scales based on the integrated auto-correlation, as explained in Sec. 2.2.

We calculate the auto-correlations of the temporal and zonal series at a selected latitude y^* , based on which we obtain later the integrated temporal and zonal auto-correlations. For this, we use 1000 years of our simulation out of a total of 10000 years, as this proves to be more than enough to reach robust estimates. As explained in Sec. 2.2, the auto-correlation is defined as the ratio between the auto-covariance $C(l)$ at lag l and the variance σ^2 : $c(l) = C(l)/C(0) = C(l)/\sigma^2$. To obtain better auto-correlation estimates, we calculate the spatio-temporal mean and variance at each y^* , and use these estimates for the computation of both temporal and zonal auto-correlations:

$$\mu = \frac{1}{N_t N_x} \sum_{j=1}^{N_t} \sum_{i=1}^{N_x} T(i, y^*, j)$$

and

$$\sigma^2 = \frac{1}{N_t N_x - 1} \sum_{j=1}^{N_t} \sum_{i=1}^{N_x} (T(i, y^*, j) - \mu)^2,$$

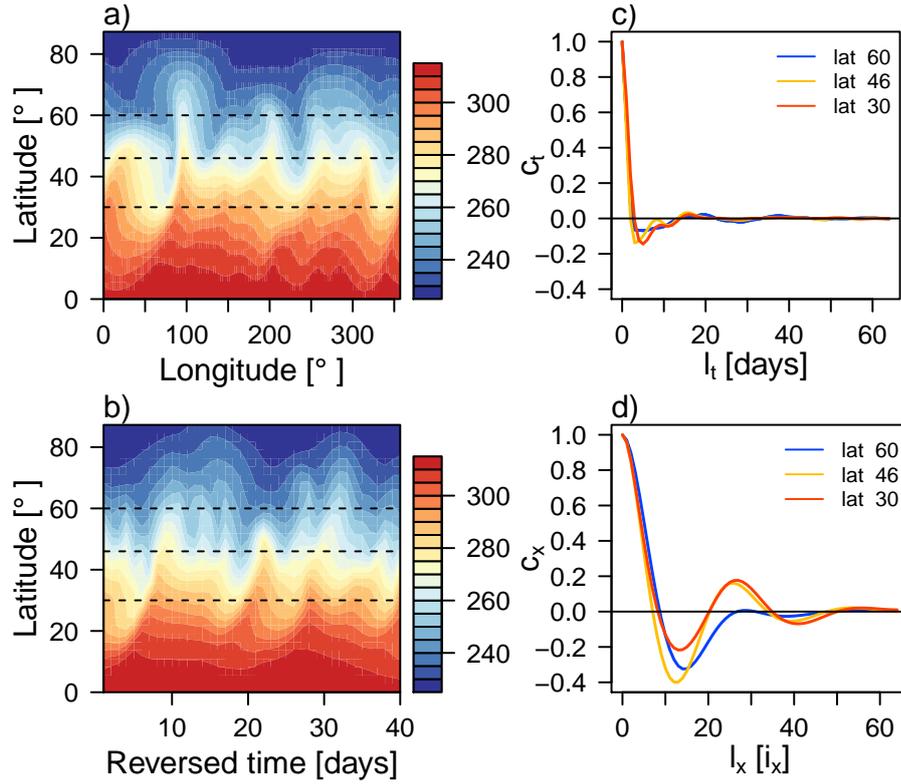


Figure 5.1: General properties of the temperature field at 960 hPa. a) Temperature values $T(x, y, t^*)$ as function of longitude x and latitude y at one selected time point t^* . b) Temperature values $T(x^*, y, t)$ as function of latitude y and of time t at one selected longitude x^* (the x -axis is backward in time). The dashed lines in a) and b) mark the selected latitudes. c) Temporal and d) zonal auto-correlation functions according to (5.1) and (5.2) for the selected latitudes (different colours according to the legend).

where $N_t = 3.6 \times 10^5$ is the number of considered points in time (daily data), and $N_x = 128$ is the number of grid points in the zonal direction. This is reasonable considering the symmetries in our system in time and along latitudinal circles. The subscripts t and x refer to time and to the zonal dimension, also in what follows.

In case of the temporal series $T(x^*, y^*, t)$, we calculate the auto-covariance at one selected longitude x^* . This estimate is independent of x^* , thus it is unimportant which longitude we choose. We have:

$$c_t(l_t) = \frac{1}{\sigma^2} \frac{1}{N_t} \sum_{i=1}^{N_t-l_t} (T(x^*, y^*, i) - \mu) (T(x^*, y^*, i + l_t) - \mu) \quad (5.1)$$

The length of the zonal series $T(x, y^*, t^*)$, however, is too short to obtain reliable auto-correlation estimates. The number of grid points along the zonal dimension is only 128. Together with such a restriction related to the *size* of the Earth, there is another

one related to the *shape* of the Earth. In fact, we have to reduce the maximum lag to $N_x/2 = 64$ because at larger lags the correlations start to increase again due to the periodicity along a latitudinal circle. To increase the robustness of our estimate, we first calculate the lagged zonal auto-correlation coefficients at each time point and then we take the average over time:

$$c_x(l_x) = \frac{1}{\sigma^2} \frac{1}{N_x N_t} \sum_{j=1}^{N_t} \sum_{i=1}^{N_x-l_x} (T(y^*, i, j) - \mu) (T(y^*, i + l_x, j) - \mu). \quad (5.2)$$

Figure 5.1c shows the temporal auto-correlation coefficient as function of the temporal lag in units of days, whereas Fig. 5.1d illustrates the zonal auto-correlation coefficient as function of the spatial lag expressed as longitude indexes $i_x = 0, 1, 2, \dots$. Both temporal and spatial auto-correlations decay to zero, meaning that two temperature values, which are far away from each other in time or in space are almost independent. We finally estimate the integrated temporal and zonal auto-correlations by taking the sum of the auto-correlations coefficients until the maximum lag $l_t = l_x = 64$. Note that we use the same temporal and zonal maximum lags for consistency reasons. The temporal integrated auto-correlation can be obtained also for larger maximum lags, but this changes only negligibly the estimate value because the decay to 0 is relatively fast. We define:

$$\tau_t = 1 + 2 \sum_{l_t=1}^{64} c_t(l_t) \quad (5.3a)$$

$$\tau_x = 1 + 2 \sum_{l_x=1}^{64} c_x(l_x) \quad (5.3b)$$

τ_t is 1.32 at 60° , 1.05 at 46° , and 1.61 at 30° , whereas τ_x is 3.26 at 60° , 3.54 at 46° , and 7.68 at 30° . We define τ_t and τ_x in a non-dimensional form (i.e. as number of time units or zonal data points) to facilitate the comparison of temporal and spatial persistence based on the resolution of our data, and because in this form we can use them directly for scaling the rate function, as we show below.

In order to express the temporal and zonal persistence in physical units, one just needs to multiply τ_t by the time step $\delta_t = 1$ day, and τ_x by the zonal grid spacing δ_x , which is latitude dependent. Thus, the temporal persistence expressed in units of days is equivalent to τ_t , and the zonal persistence in units of km amounts to 391 at 60° , 732 at 46° , and 1292 at 30° . We define the scale velocity $U_\tau := \frac{\tau_x \delta_x}{\tau_t \delta_t}$. From a statistical point of view, U_τ represents the ratio between spatial and temporal persistence. From a geometrical/dynamical point of view, U_τ represents the ratio between spatial and temporal typical scales. Thus, U_τ is a measure for the anisotropy between space and

time. At 60° $U_\tau = 4.25 \text{ ms}^{-1}$ and at 46° $U_\tau = 8.47 \text{ ms}^{-1}$. U_τ is not necessarily the same as the temporally averaged zonal mean wind $\overline{[U]}$ (at the same vertical level of 960 hPa), although, clearly, a dynamical relationship exists, because the turbulent structures are advanced, to a first approximation, by the mean flow. In fact, we find that $\overline{[U]}$ is 3.6 ms^{-1} at 60° , and 6 ms^{-1} at 46° , which bear a good agreement with the scale velocities at the same latitudes.

The agreement is lost when looking at 30° , which is at the boundary of the mid-latitude baroclinic zone where the qualitative description given above applies. As we approach the equator, the atmospheric dynamics has a much lower degree of chaoticity with respect to the mid-latitudes, unless we look at the convective scales, which are not resolved at all in this model, and spatial persistence is very enhanced (see also Fig. 5.1a,b). In this case we find $U_\tau = 14.95 \text{ ms}^{-1}$, while $\overline{[U]}$ is -3.4 ms^{-1} , which indicates prevailing easterly flow, a clear signature of tropical dynamics.

Before continuing with the description of the temporal and spatial large deviations, we briefly discuss the connection between high values of coarse grained temperatures and long individual events where the temperature readings are persistently above the long-term average, discussed already in Sec. 1.4. Figures 5.2a,b,c show three short temporal series at latitude 46° together with the corresponding series of the coarse grained quantities where averages are computed using block lengths of $20\tau_t$, $10\tau_t$, and $5\tau_t$, respectively. The three short time series have been specifically chosen because they feature a large fluctuation in the coarse grained quantity. Figures 5.2d,e,f show the same in case of the zonal fields. The main finding is that up to moderately long averaging windows of about $10\tau_t$ (or $5\tau_x$ for spatial averages) it is possible to link large fluctuations with individual persistent events. When a coarser graining is considered, using a window of $20\tau_t$ for time averages and $10\tau_x$ or $20\tau_x$ for spatial ones, thus going in the direction of the regime of the large deviations discussed below, we do not have such an one-to-one identification. Instead, large ultralong fluctuations are related to the occurrence of subsequent moderately long persistent features.

5.1 The link between temporal and zonal large deviations

At this point, we turn our attention to the estimation of the temporal and zonal rate functions. For this, we first have to obtain sequences of temporal and zonal averages for increasing lengths of averaging blocks n_t and n_x , for which we use the total length of our simulation of 10000 years.

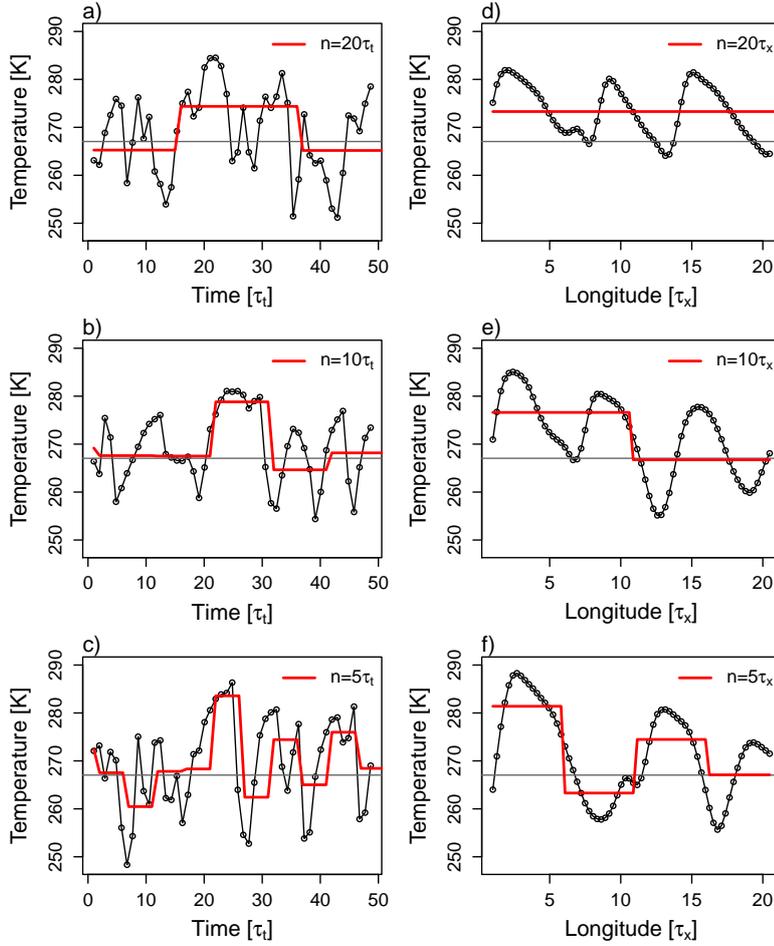


Figure 5.2: Relationship between persistent events and large fluctuations of the coarse-grained fields. a) Time series (black line) of near-surface temperature at 46° in the case of a large event of the coarse grained time series (red line) with averaging window of $20\tau_t$; x-axis in units of $\tau_t = 1.05$ days. b) Same as a), for averaging window of $10\tau_t$. c) Same as a), for averaging window of $5\tau_t$. d) Zonal series (black line) of surface temperature at 46° in the case of a large event of the coarse grained zonal series (red line) with averaging window of $20\tau_x$; x-axis in units of $\tau_x = 3.54$ grid points. e) Same as f), for averaging window of $10\tau_x$. f) Same as d), for averaging window of $5\tau_x$. In all panels the grey horizontal line represents the long term and longitudinal average.

$$A_{n_t} = \frac{1}{n_t} \sum_{i=1}^{n_t} T(x^*, y^*, t = i) \quad (5.4a)$$

$$A_{n_x} = \frac{1}{n_x} \sum_{i=1}^{n_x} T(x = i, y^*, t^*) \quad (5.4b)$$

The lengths of temporal averaging blocks are chosen to be multiples of τ_t : $n_t =$

$5\tau_t, 10\tau_t, \dots, 40\tau_t$. Similarly, the lengths of zonal averaging blocks are multiples of τ_x , but in this case the largest possible multiple m is limited due to the size and shape of the Earth, as mentioned above: $n_x = 5\tau_x, 10\tau_x, \dots, m\tau_x$. $m = 20$ in the case of latitudes 60° and 46° , whereas $m = 10$ in the case of latitude 30° . To increase the number of averaged values for the computation of the temporal rate functions, we lump together the temporal averages from every 25th longitude along a latitudinal circle. Since $\tau_x \ll 25$, these temporal sequences can be treated as independent repetitions. In case of zonal averaging, we take one averaged value in space from every 10th point along the time axis, which we consider to be independent repetitions as well. This assumption is reasonable because the integrated temporal auto-correlation of zonal averages is much lower than 10, even for the largest n_x (as shown later in Fig. 5.4). We obtain for each value of n_t and n_x estimates of the rate functions, after using the re-normalising factors given by $1/\tau_t$ or $1/\tau_x$, respectively:

$$\tilde{I}_{n_t}(a) = -\frac{\ln p(A_{n_t} = a)}{n_t} \tau_t, \quad (5.5a)$$

$$\tilde{I}_{n_x}(a) = -\frac{\ln p(A_{n_x} = a)}{n_x} \tau_x, \quad (5.5b)$$

where $p(A_{n_t} = a)$ and $p(A_{n_x} = a)$ represent empirical estimates of the pdf's of the temporally and zonally averaged sequences. Due to the re-normalisation, the logarithm of the probabilities is scaled by n_t/τ_t or n_x/τ_x , i.e. by the number of uncorrelated instead by the total number of data in an averaging block. Thus, we eliminate the effect of correlations.

Figure 5.3 displays \tilde{I}_{n_t} (a-c) and \tilde{I}_{n_x} (d-f) for every n_t and n_x as function of temperature fluctuations $T' = T - \mu$. As a side note, we remark that in every figure below the shown re-normalised rate function estimates are shifted vertically so that their minimum is at 0. In case of the temporal rate functions, it is clear that for $n_t \geq 20\tau_t$ the estimates \tilde{I}_{n_t} do not change in shape by further increase in n_t , meaning that we obtain stable and reliable estimates, i.e. there is a proof in our data for a LDP in time. We also notice that the range of A_{n_t} values becomes narrower as n_t increases as an effect of averaging, which reduces the amount of available data. Thus, we obtain our best estimate at an optimal averaging block length n_t^* , which is large enough to allow for the convergence of rate function estimates, but is in the same time small enough so that the range of A_{n_t} is not too narrow, i.e. $n_t^* = \min(n_t; \tilde{I}_{n_t} \approx I_{n_t})$. We choose the same optimal averaging length for all three latitudes $n_t^* = 20\tau_t$, although in the case of latitudes 60° and 30° , $\tilde{I}_{n_t=10\tau_t}$ seems to be already a good estimate for the asymptotic I_{n_t} . In the case of zonal rate functions, we first notice that the largest n_x seems to be too small for a clear convergence. In other words, the length of a latitudinal circle is not long enough to

clearly obtain a large deviation limit. However, the dependence of \tilde{I}_{n_x} on n_x seems to decrease as n_x is increasing, thus we choose the largest possible n_x as the optimal zonal averaging length $n_x^* = \max(n_x)$. $n_x^* = 20\tau_x$ in the case of latitudes 60° and 46° , whereas in the case of latitude 30° , n_x^* is only $10\tau_x$ because of stronger zonal auto-correlations.

The best estimates of the temporal and zonal re-normalised rate functions $\tilde{I}_{n_t}^* = \tilde{I}_{n_t=n_t^*}$ and $\tilde{I}_{n_x}^* = \tilde{I}_{n_x=n_x^*}$ are shown again in Fig. 5.3g,h,i. The shading represents the 95% confidence intervals of 2000 nonparametric ordinary bootstrap estimates based on the normal distribution (functions `boot` and `boot.ci` of the R package `boot`, Davison and Hinkley, 1997; Canty and Ripley, 2017). We notice that $\tilde{I}_{n_t}^* \approx \tilde{I}_{n_x}^*$. The equivalence is very good in the case of latitude 60° and in the case of negative anomalies at latitudes 46° and 30° . We also notice some differences regarding positive anomalies at latitudes 46° and 30° , with larger differences at 30° . At this later latitude, however, it has to be considered that the maximum possible zonal averaging length is $10\tau_x$, whereas in the other cases it is $20\tau_x$. We assume that the differences between the temporal and zonal re-normalised rate function estimates have to do with the fact that n_x^* is not large enough to estimate the rate function properly. Larger values of n_x are needed to overcome the enhanced skewness in the distribution of zonal averages as effect of spatial correlations, however this is impossible due to limitations coming from the size and shape of the Earth. These findings have correspondence with the large value of U_τ at this latitude, defining the anisotropy between space and time. While the temporal rate function can be estimated reliably at a relatively small n_t , the estimation of the zonal one is a much more difficult task.

However, the main message of Fig. 5.3 is that the temporal and zonal re-normalised rate functions seem to be equal, $I_{n_t} = I_{n_x}$, if the probability of averages is scaled by the number of uncorrelated data in an averaging block n_t/τ_t or n_x/τ_x , as explained above. In other words, there is a link connecting temporal and spatial large deviations or averages, due to the existence of a *universal function* I_n ; universal in the sense that it represents large deviations in both dimensions: time and space.

Obviously, based on the LDP in time or in space, one cannot characterise persistent temporal or spatial events, because the limit law starts to act on larges scales, where persistence is lost and universality emerges. However, one can capture persistent space-time events by averaging in both dimension space and time. To achieve this, it is important that the spatial averaging length is not too small but not too large either, as we show in the following.

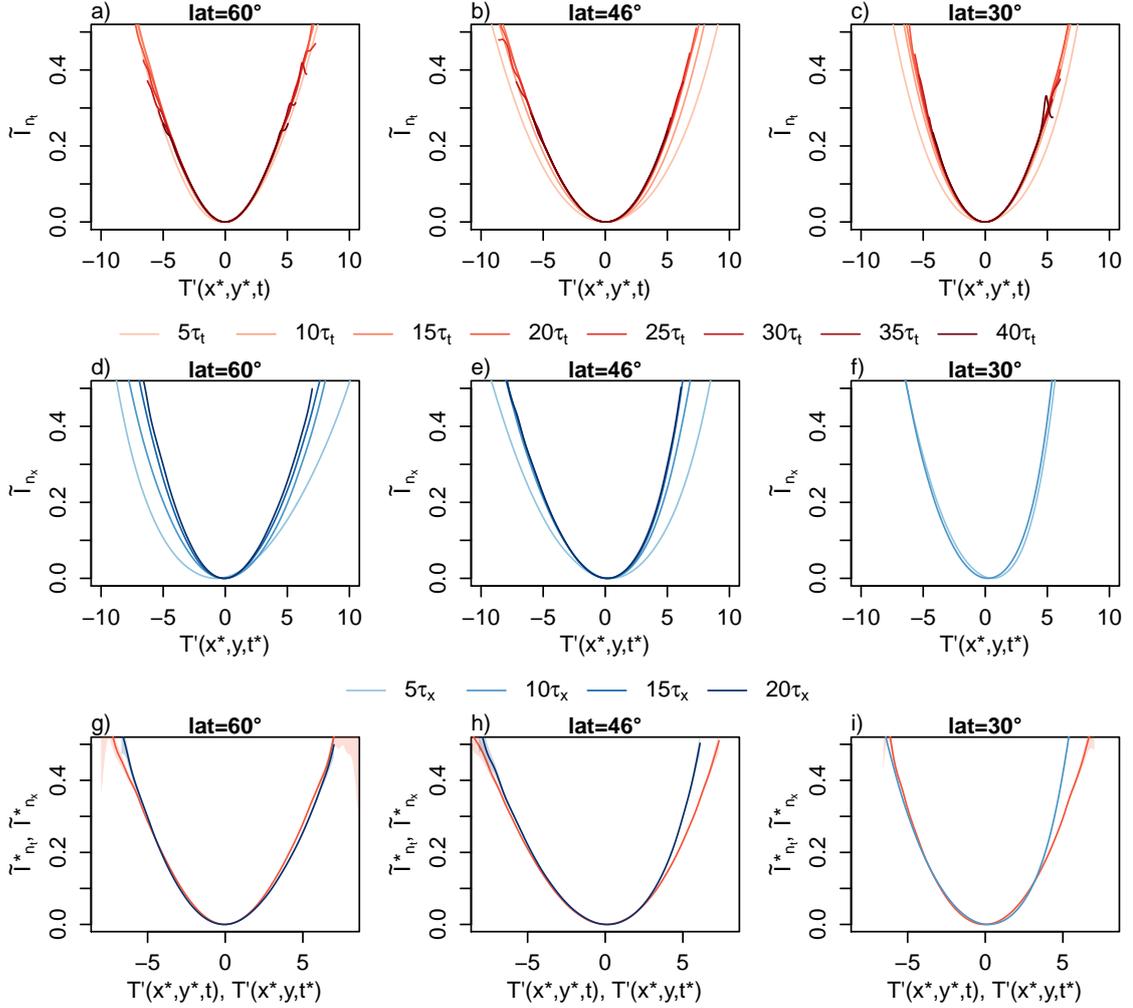


Figure 5.3: a) – c) Temporal re-normalised rate function estimates $\tilde{\Gamma}_{n_t}$ and d) – f) zonal re-normalised rate function estimates $\tilde{\Gamma}_{n_x}$ for the three considered latitudes and for increasing averaging lengths n_t and n_x according to the different colours (see legend). g) – i) Best estimates of the temporal (red) and zonal (blue) re-normalised rate functions. All estimates are shifted vertically so that their minimum is at 0. $T' = T - \mu$ represents temperature fluctuations around the mean.

5.2 Spatio-temporal large deviations

We consider temporal sequences of zonally averaged observables over averaging lengths $n_x = 1\tau_x, 5\tau_x, 10\tau_x, 20\tau_x$, and then average each sequence in time for increasing averaging lengths $\hat{n}_t = 1\hat{\tau}_t, 5\hat{\tau}_t, 10\hat{\tau}_t, 15\hat{\tau}_t, \dots, 40\hat{\tau}_t$. The notation $\hat{\cdot}$ is meant to indicate that we average in space and *additionally* in time, and $\hat{\tau}_t$ is the decorrelation time of the spatially averaged observable. By considering several n_x values, we choose the spatial scale at which we analyse the large deviations in time. The spatio-temporal averages are computed as:

$$A_{n_x, \hat{n}_t} = \frac{1}{\hat{n}_t} \sum_{j=1}^{\hat{n}_t} \frac{1}{n_x} \sum_{i=1}^{n_x} T(i, y^*, j) = \frac{1}{\hat{n}_t} \sum_{j=1}^{\hat{n}_t} A_{n_x}(j). \quad (5.6)$$

Similarly to the previous cases, also in the case of spatio-temporal averages, we have to take into account the strength of auto-correlations if we pursue to compare the spatio-temporal rate functions with the temporal and zonal ones. We estimate the integrated temporal auto-correlation $\hat{\tau}_t$ of spatio-temporal averages analogously to τ_t or τ_x , but to assure the stability of $\hat{\tau}_t$ we choose a higher maximum lag of 120 days, because the auto-correlation in time of zonal averages has a slower decay compared to the one of unaveraged temporal or zonal observables. Figure 5.4 shows $\hat{\tau}_t$ as function of zonal averaging length n_x and temporal averaging length \hat{n}_t . The temporal auto-correlations of the spatio-temporal observables are increasing with n_x and decreasing with \hat{n}_t . The increase with n_x , on the one hand, can be explained by the connection between temporal and spatial scales. Large events in space are long lasting events in time. The decrease of the temporal integrated auto-correlation with \hat{n}_t , on the other hand, can be explained by the increase of the number of uncorrelated events with respect to the number of correlated events in an averaging block as a consequence of increasing the block length. This is automatically the case for large averaging blocks when correlations are finite, and is crucial for the applicability of the block averaging method. The different behaviour with n_x and \hat{n}_t , however, has to do mainly with the discrepancies in the temporal resolution of the newly obtained averages. While, in the case of zonal averaging, the temporal resolution remains one day, in the case of additional averaging in time, the temporal resolution decreases with \hat{n}_t , thus the temporal auto-correlation lag increases. However, this is not a problem for our analysis since we are interested in the correlations of the averaged observables measured in number of averaged data. A stronger increase of $\hat{\tau}_t$ at the “end” of the channel underlines the above discussed effect of averaging along a latitudinal circle. At the zonal “end” of the channel, the temperature values are strongly correlated with the ones at the “beginning” of the channel.

We observe the largest increase of $\hat{\tau}_t$ with n_x in case of latitude 60° . However, taking into account the different typical zonal scales at different latitudes and comparing the auto-correlations at same zonal averaging lengths in units of τ_x (for example along the vertical lines with same colours in Fig. 5.4), things look more similar. Here, we deliberately use the expression “more similar” instead of “the same” because the differences in the zonal tendencies are, of course, not only a matter of scale. The changing dynamical properties of the zonal mean temperature are determined by the latitudinal dependence of diabatic forcing, meridional heat transport due to eddy activity, zonal mean wind, static stability, etc.

Additionally, we also have to consider that the distance between the longitudes is de-

creasing from equator to pole due to the geometry of the sphere. This, of course, contributes to the fact that the temperature values at neighbouring longitudes are stronger correlated at latitude 60° than at 46° or 30° .

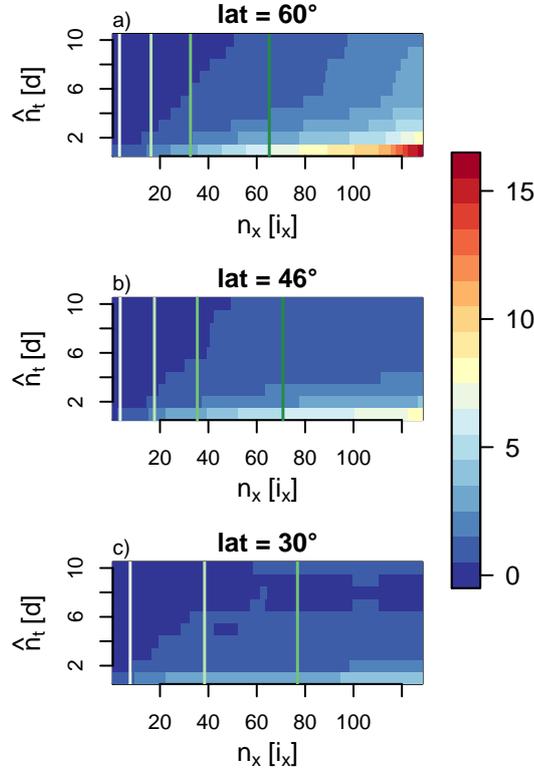


Figure 5.4: Integrated temporal auto-correlation of spatio-temporal averages for the selected latitudes as function of zonal and temporal averaging lengths. The vertical lines mark longitudes representing upper limits of zonal averaging (lower limit is 1), corresponding to the multiples of τ_x : $1\tau_x, 5\tau_x, 10\tau_x, 20\tau_x$ (from white to green).

Estimates of spatio-temporal re-normalised rate functions are then computed for each n_x and \hat{n}_t as:

$$\tilde{I}_{n_x, \hat{n}_t}(a) = -\frac{\ln \tilde{p}(A_{n_x, \hat{n}_t} = a)}{\hat{n}_t n_x} \hat{\tau}_t \tau_x \quad (5.7)$$

We remark that Eq. (5.7) accounts for both zonal and temporal auto-correlations by multiplication with $\hat{\tau}_t \tau_x$, similarly to the case of temporal and zonal rate functions. The spatio-temporal re-normalised rate function estimates are displayed by Fig. 5.5 (coloured lines). For comparison reasons, we also show the best temporal and zonal estimates $\tilde{I}_{n_t}^*$ (continuous black lines) and $\tilde{I}_{n_x}^*$ (short-dashed black lines), together with the estimate of the zonal re-normalised rate function at the selected zonal averaging length \tilde{I}_{n_x} (long-dashed black lines). The main message here is that:

- The spatio-temporal re-normalised rate function seems to be equal to the universal function I_n for *small* and *large* zonal averaging lengths.
- We suppose that, in the case of *small* zonal averaging lengths $n_x \ll n_x^*$, like $n_x = 1\tau_x$, the zonally averaged observable is not significantly different from the spatially localised observable, so that $\tilde{I}_{n_x, \hat{n}_t}$ converges to the universal function I_n .
- In the case of *large* zonal averaging lengths $n_x \geq n_x^*$, like $n_x = 20\tau_x$, the zonal averages already exhibit universal characteristics, which are not altered by the additional temporal averaging, thus $\tilde{I}_{n_x, \hat{n}_t}$ corresponds again with the universal function I_n .
- On *intermediate* levels however, i.e. $\tau_x < n_x < n_x^*$, due to the non-trivial zonal correlations one obtains after zonal averaging a totally different observable from the original one.

The spatio-temporal re-normalised rate functions are different from the universal function I_n at $n_x = 5\tau_x$ in the case of latitudes 60° and 46° , as well as at $n_x = 10\tau_x$ in the case of latitude 60° , whereas in this last case is worth mentioning that the spatio-temporal rate function corresponds with the zonal rate function estimate at $n_x = 10\tau_x$. In all these cases, the spatio-temporal re-normalised rate functions are flatter than the universal function, pointing out a higher probability of large deviations, which are associated to the presence of organised structures. Thus, these scales are exactly the ones at which one can analyse persistent space-time events, like heat waves or cold spells, based on the LDP. Figure 5.6 represents schematically the ranges of temporal and zonal averaging lengths, at which universality emerges (blue) or is hindered (light blue) due to zonal correlations. Pre-asymptotic regions, where the LDP is not valid yet, are left white.

As a side note, the horizontal shifts of the rate function estimates at small averaging lengths (n_x or \hat{n}_t) emphasise that these estimates are not reliable because the averaging length is too small for the law of large numbers to hold. We also wish to remark that differences emerge when looking at temperature data from latitude 30° . Here, the spatio-temporal re-normalised rate function $\tilde{I}_{n_x, \hat{n}_t}$ at $n_x = 1\tau_x$ is not equivalent with the universal function I_n . One possible reason for this is that when averaging over a length $n_x = 1\tau_x$ the newly defined observable has already significantly different properties from the temporal observable. The universality of the spatio-temporal rate function cannot be checked properly due to the limit in zonal averaging length of $10\tau_x$. What we see, however, is that at $n_x = 10\tau_x$ spatio-temporal re-normalised rate function is quite similar to - yet distinct from - the universal function.

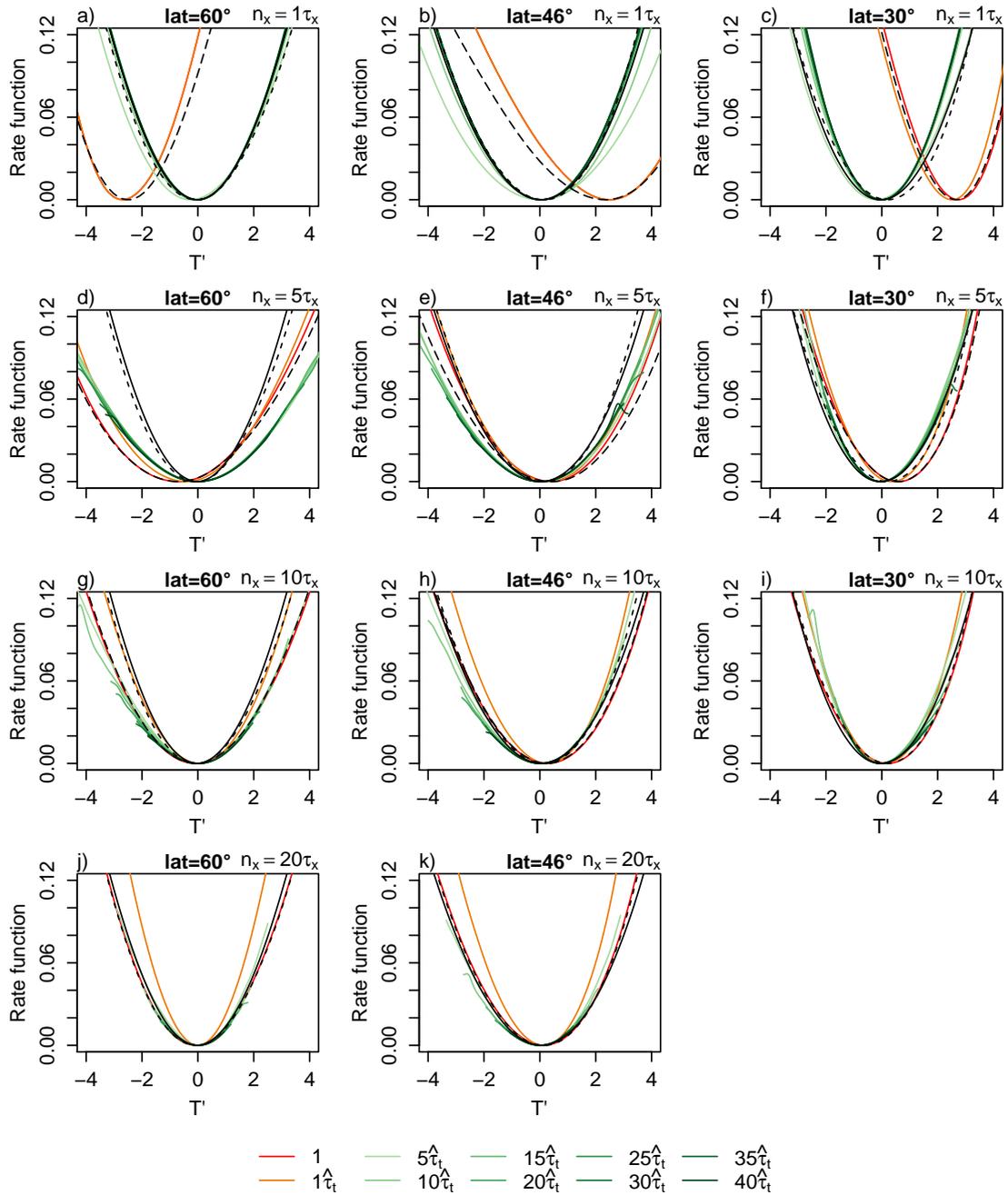


Figure 5.5: Rate functions of spatio-temporal averages for the selected latitudes and different zonal averaging lengths: a) – c) $n_x = 1\tau_x$, d) – f) $n_x = 5\tau_x$, g) – i) $n_x = 10\tau_x$, j) – k) $n_x = 20\tau_x$. The coloured lines represent spatio-temporal rate functions for different temporal averaging lengths \hat{t}_t according to the legend. The black continuous line is the best temporal rate function estimate, the black short-dashed line is the best zonal rate function estimate, and the black long-dashed line is the zonal rate function estimate at the selected n_x . The rate function estimates are shifted vertically so that their minimum is at 0. $T' = T - \mu$ represents temperature fluctuations around the mean.

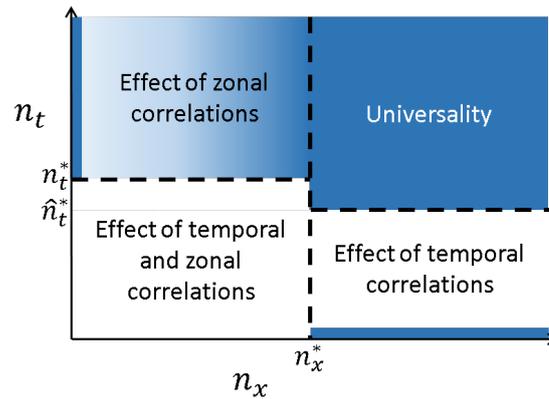


Figure 5.6: Schematic representation of universality and effect of correlations depending on the zonal and temporal averaging lengths. The dark blue colour marks the region where universality emerges. The light blue colour represents the region with non-universal spatio-temporal rate functions as an effect of zonal correlations. Pre-asymptotic regions, i.e. where the LDP is not valid yet, are white.

5.3 Return levels of large deviations

We summarise shortly our main findings presented until now:

1. The estimates of the rate functions seem to converge to an asymptotic function, and we obtain the best estimate of the rate function at an optimal averaging block length n^* . This is mostly clear in case of the temporal rate functions, showing that there is a large deviation principle, i.e. a *universal law* that allows us to estimate the probabilities of occurrence of averages over $n \geq n^*$, without having to actually perform the averaging.
2. We have found an *equivalence* between the temporal, spatial and spatio-temporal re-normalised rate functions, in the later case for very small and very large zonal averaging lengths. This means that we can deduce statistical properties of temporal averages from the ones of spatial averages and vice-versa. We can do the same in the case of spatio-temporal averages and temporal or spatial averages.

Now, the question is how can we use these information in a practical way? One possibility of application, which we present in this subsection by the example of latitude 60° , arises in the context of computing return periods of large events. Figure 5.7 shows return level plots, i.e. return levels as function of return periods obtained in three different ways, based on: empirical data (circle markers), large deviation principle (continuous lines), and the Generalized Pareto distribution (dashed lines). For the estimation of return periods based on the LDP, we first obtain kernel density estimates (function density of the R package stats, R Core Team, 2016) of the pdf's $p(A_n = a)$ at fixed

equidistant return levels $A_{n,1}, \dots, A_{n,256}$, based on which we estimate the cumulative distribution function $P(A_n \leq a)$, and then compute the return periods for $A_n \geq a$ as $\frac{1}{1-P(A_n \leq a)}$ and for $A_n \leq a$ as $\frac{1}{P(A_n \leq a)}$. Thus we obtain the return periods of both positive (Fig. 5.7a,c,e) and negative (Fig. 5.7b,d,f) large deviations. The shading around the continuous lines in Fig. 5.7 represents the 95% confidence intervals of 2000 nonparametric ordinary bootstrap return period estimates based on the normal distribution.

We compute the GPD return levels based on (2.4) using the maximum likelihood estimates of GPD parameters (functions `gpd.fit` and `gpd.rl` of the R package `ismev` functions written by Janet E. Heffernan with R port and documentation provided by Alec G. Stephenson., 2016). We analyse return levels of high temperature values exceeding a threshold equal to the 99.9% quantile of the averaged series, as well as return levels of low temperature values below the 0.1% quantile. To verify the applicability of the GPD method, the stability of return levels was checked also for a higher (lower) quantile of 99.99% (0.01%). The return level estimates seem to be stable even if the threshold is increased (not shown). Note that, although the very slow convergence of the GPD shape parameter is well known in some cases, the stability of return level estimates still holds if the change in the shape parameter is relatively small as the threshold increases, as shown in Sec. 4.1. The shading around the dashed lines in Fig. 5.7 represents 95% maximum likelihood confidence intervals of return level estimates. As a side note, in the case of GPD, the estimation concerns the return levels while the return periods are fixed, whereas we proceed the other way around in the case of LDT. This is necessary because we estimate the rate function $I(a)$ at fixed equidistant values a .

In Fig. 5.7a,b the return levels of temporal averages are shown for three different averaging windows $20\tau_t, 30\tau_t, 40\tau_t$. Here we use point 1. from above, and obtain the return periods based on the LDP for every averaging window from $p(A_{n_t^*=20\tau_t} = a)$. We notice a very good agreement with the empirical data and the GPD return levels, not only for $20\tau_t$ but also in case of $30\tau_t$ and $40\tau_t$, for both high (Fig. 5.7a) and low (Fig. 5.7b) extremes of averages. In case of $n_t = 20\tau_t$, the confidence intervals of the largest return periods based on the LDP become very unstable, the lower limits reach even negative values, thus they cannot be displayed on this semi-logarithmic scale.

The return periods from the LDP have an upper (or lower) limit because the estimation relies on empirical pdf's. This is not the case for the GPD return periods since these can be extrapolated to even unobserved events. The LDP, however, is a limit law that gives us return periods for every averaging length $n > n^*$, whereas the GPD return periods have to be computed separately for every n . This becomes more and more difficult with increasing n due to the decreasing data amount as effect of averaging. In other words, Fig. 5.7 points out the different dimensions in which the two limit laws act, as mentioned already in Chapter 1. The predictability of GPD is directed towards larger

and larger events, i.e. towards unobserved ones, whereas the predictability of LDT is directed towards larger and larger averaging lengths, i.e. towards observables that, by construction, reduce dramatically the amount of data available for statistical analysis.

Point 2. presented above is illustrated by Fig. 5.7c,d and Fig. 5.7e,f. In the first case, return periods of temporal averages are computed based on the LDP obtained for zonal averages ($n_x^* = 20\tau_x$), and, in the second case, return periods of spatio-temporal averages (with a spatial averaging length of $20\tau_x$) are obtained from the LDP for temporal averages ($n_t^* = 20\tau_t$). In both cases, but especially for the spatio-temporal averages, the agreement with the empirical data and the GPD return levels is good. The differences between the return levels based on the LDP and the empirical data (also GPD) are related to the discrepancies in the estimation of the temporal and zonal, as well as temporal and spatio-temporal re-normalised rate functions. For example, the underestimation of low extremes of temporal averages based on the zonal rate function has to do with higher re-normalised zonal rate function values compared to the temporal ones in their left tails (see Fig. 5.3g).

We remark that the possibility of *commuting* between averages of different dimensions (time and space) is due to the fact that, by eliminating the effect of serial correlations, the large deviations of these different dimensions follow a *universal function*. This universal function, however, is inadequate to describe persistent space-time events, like heat waves or cold spells, because it acts on large scales, where persistence becomes irrelevant. To characterise the *persistent space-time events*, one needs to compute the temporal rate function of zonal averages *on intermediate scales*.

5.4 Summary and discussion

We have analysed the properties of temporal and spatial near-surface (960 hPa) temperature averages in the PUMA simplified global atmospheric circulation model based on LDT. Extremes of averages on specific scales are related to persistent extreme events, like heat waves or cold spells. We run the model for 10000 years with a constant (only latitude dependent) forcing, creating non-equilibrium (due to the forced-dissipative nature of the model) steady state simulations without orography, annual or daily cycle. The forcing is symmetric for the two hemispheres. The horizontal resolution is T42 with 10 vertical levels, and the temperature values are recorded daily. We compute and compare re-normalised rate functions based on the integrated auto-correlation for temporal and zonal temperature sequences at selected latitudes (60° , 46° , and 30°), focusing in the mid-latitudes region, where turbulence is best developed. The spatial averaging is performed only in zonal direction, because this is a geometrical direction along which

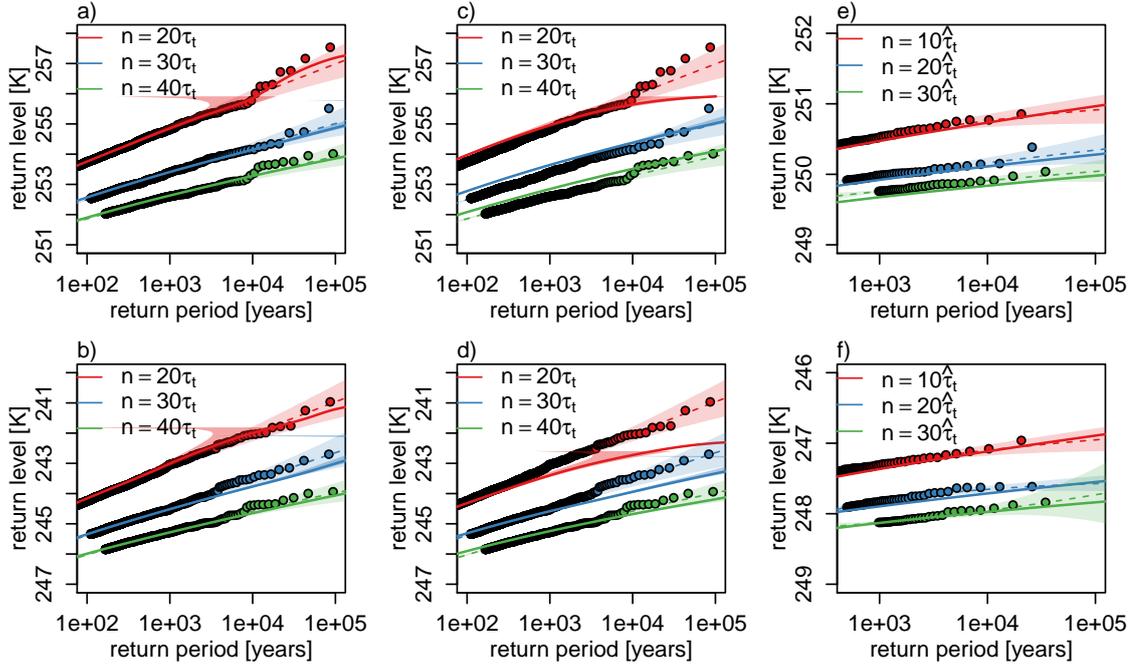


Figure 5.7: Return levels and return periods of positive (upper row) and negative (lower row), a) – d) temporal and e) – f) spatio-temporal large deviations of temperature at latitude 60° . Circle markers: empirical data; continuous line with shading: estimates based on LDT with 95% confidence intervals of 2000 non-parametric bootstrap samples based on the normal distribution; dashed line with shading: GPD estimates with 95% confidence intervals based on Maximum Likelihood Estimation. The different colours represent different averaging lengths. The LD-estimates are obtained based on a), b), e), f) temporal averages at $n_t^* = 20\tau_t$, and c), d) zonal averages at $n_x^* = 20\tau_x$.

the system has a symmetry. We also analyse the case of two-dimensional, i.e. spatio-temporal averaging. We verify the correctness of our results by comparing the return periods based on the rate functions with return periods from the empirical data and from the POT method. Before discussing them in detail, we summarise first our main findings:

1. The temperature averages in PUMA follow a LDP.
2. The temporal and zonal re-normalised rate functions are equal if we compute them by eliminating the effect of temporal and zonal correlations. Thus, we can define a *universal function*, describing temporal as well as spatial large deviations.
3. The spatio-temporal re-normalised rate functions are *equal* to the universal function for *small and large* spatial averaging lengths. *On intermediate levels*, as an effect of non-trivial spatial correlation, the spatio-temporal re-normalised rate functions *differ* from the universal one.

Our results show that the temperature averages in PUMA follow a LDP. The estimated rate functions clearly converge in case of temporal averages. We obtain reliable estimates at an optimal averaging length n_t^* , which is about $20\tau_t$, where τ_t represents the temporal integrated auto-correlation. The fact that the temperature averages follow a LDP might seem unsurprising, but actually it has extremely important consequences on a practical level. Based on the LDP, we can estimate the probabilities of averages, and thus for the practical use very important return periods, for every averaging length $n_t \geq n_t^*$. All we need to know is the probability of averages $A_{n_t^*}$, which we can estimate empirically. In contrast to the temporal averages, in case of zonal averaging, the spatial averaging length n_x is substantially limited by the size and shape of a latitudinal circle. The temporal averaging is performed on a theoretically infinite (and practically very long) line, whereas the zonal averaging takes place on a circle. Thus, the convergence of the estimated rate functions is not that clear as for temporal averages. However, the comparison of the zonal results with the temporal re-normalised rate function estimates shows that the averaging length $n_x^* = 20\tau_x$ seems to provide a reasonable rate function estimate, thus we choose this one as the optimal zonal averaging length. In the case of latitude 30° , $20\tau_x$ cannot be reached due to stronger zonal correlations. Here, the maximum averaging length is $10\tau_x$.

We find that the temporal and spatial re-normalised rate functions seem to be equal if we eliminate the effect of correlations according to Eq. (5.5), where we basically scale the rate functions by the number of uncorrelated data instead of the whole number of data in an averaging block. Based on this equivalence, one finds a universal function $I_n = I_{n_t} = I_{n_x}$, in the sense that it describes both temporal and spatial large deviations. From a statistical-mechanical point of view, this means that fluctuations (around the steady state) in time and in space exhibit same basic properties, but differ as an effect of temporal and spatial persistence. From a practical point of view, this implies that one can commute between space and time: we can deduce statistical properties of spatial averages (including return level estimates) from a single time series, and this is, of course, true the other way round too.

Obviously, based on the LDP in one dimension - time or space - we cannot describe persistent events, because the limit law is acting on very large scales, where spatial or temporal organisation is lost and universality emerges. However, persistent space-temporal events can be studied based on LDT if one performs the averaging in both dimension - time and space.

Therefore, we extend our analysis also to spatio-temporal large deviations. Here, we average first in zonal direction taking different averaging lengths $n_x = 1\tau_x, 5\tau_x, 10\tau_x, 20\tau_x$, and then we search for a LDP in time of the zonally averaged observables. We find that the spatio-temporal re-normalised rate function, computed again by eliminating

the correlations according to (5.7), is equal to the universal function I_n in two cases: 1) for small zonal averaging lengths $n_x \approx \tau_x$, and 2) for large ones $n_x \geq n_x^*$. We suppose that in the first case, due to the small n_x , the zonally averaged observable is not significantly different from the temporal observable, and thus the rate function converges to the universal function. In the second case, the zonal averages already exhibit universal characteristics because the large n_x allows for enough mixing in the series of zonal averages. These universal characteristics are not altered by the additional temporal averaging. On intermediate scales however, i.e. $\tau_x < n_x < n_x^*$, due to the non-trivial zonal correlations, one obtains after zonal averaging a totally different observable, whose large deviations follow a different re-normalised rate function than the universal one. Consequently, by computing large deviations in time of zonal averages, we get rid of temporal persistence if the temporal averaging length is large enough, but we cannot eliminate the effect of zonal persistence on intermediate scales, which then leads to a non-universal re-normalised rate function. This also means that in this way we can study persistent extreme events based on LDT. These intermediate scales of about $5 - 10 \tau_x$ or $\approx 2000 - 4000$ km are approximately equal with the scale of persistent synoptic disturbances, like the ones causing severe heat waves. According to this points of view, long lasting synoptic scale disturbances are large deviations from the steady state, which allow for a higher degree of spatio-temporal organisation and, in a loose sense, a lower entropy compared to disturbances at any other scales.

Additionally, we compare the two frameworks for investigating rare events, i.e. LDT and the POT approach of EVT, from a practical point of view, based on return level and return period estimates. Both methods are based on limit laws, but they differ in the way the limit is obtained, and thus also in the direction in which the limit acts. The POT approach deals with the conditional probabilities of averages exceeding a high threshold. The limit law is obtained as one considers larger and larger extremes, thus it is directed towards large, even unobserved events. In case of LDT, we approach the limit as we consider averages with increasing averaging length n , thus the limit is directed towards $n \rightarrow \infty$. Our results point out these differences. On the one hand, the return time estimates based on LDT are limited from above at small averaging lengths because they are obtained based on empirical distributions, whereas the estimates based on the POT approach can be extrapolated to unobserved events. On the other hand, the return times based on LDT can be obtained for every $n \geq n^*$ based on the probabilities of A_{n^*} , whereas in the case of the POT approach they have to be estimated for every n separately. We also have to remark that the convergence to the limit law seems to be easier to achieve in case of LDT than in case of EVT.

As mentioned above, we eliminate the effect of correlations in the computation of the rate functions by multiplication with the integrated auto-correlation. We estimate both temporal and zonal integrated auto-correlations, τ_t and τ_x . By computing the ratio

between spatial and temporal persistence, we define a scale velocity $U_\tau = \tau_x/\tau_t$, which is a measure for the anisotropy between space and time. If the anisotropy between space and time is strong, it becomes more difficult to show the existence of a universal rate function. In the case of latitude 30° , for example, due to a strong zonal persistence, the largest zonal averaging length is limited to $10\tau_x$, which is not large enough to obtain reliable estimates of the spatial rate function. We remark that the scale velocity we find by such asymptotic procedure could be viewed in connection with the research lines aiming at identifying the multifractal nature of the weather and climate fields (Lovejoy and Schertzer, 2013) and, in particular, of precipitative fields (Deidda, 2000).

While nature and society do not typically conform to the hypotheses of the theorems one needs to establish universal laws, such asymptotic results can nonetheless be extremely useful for studying observational data, just as in the widely case of EVT. Therefore, this work should be seen as a first step towards the use of LDT for the analysis of actual climatic data and the outputs of state-of-the art climate models. The perspective is to find new ways to estimate efficiently the probability of occurrence of extremely rare events associated to persistent climatic conditions. In this work, we have focused on time scales which are long compared to those typical for the atmosphere, but one can adopt the same methods for studying persistent events of multiannual scales, where the oceanic variability is, instead, essential. This has, potentially, great relevance for addressing the problem of assessing human and environmental resilience to the low-frequency variability of the climate system.

6 Conclusions and outlook

In this thesis we have analysed extreme or rare events in two numerical atmospheric models, having a twofold objective. First, we study extremes of energy observables in a two-layer quasi-geostrophic (QG) atmospheric model and compare our results with theoretical ones from the field of extreme value theory (EVT) for dynamical systems. These energy observables can be considered as so-called “physical observables”, thus we search for a relationship, derived theoretically for Axiom-A systems, between extremes of physical observables and the dimension of the attractor of the system, considered to be a global property. We apply the two popular methods of EVT, the Block Maxima (BM) method, and the Peak over Threshold (POT) approach. Second, we study properties of temporal, spatial, as well as spatio-temporal averages of near-surface temperature in the simplified global circulation model PUMA. The main purpose here is to study persistent extreme events of temperature, like heat waves or cold spells, based on their connection with extremes of averages, from the perspective of Large Deviation Theory (LDT) for non-equilibrium steady states. In both cases, we are interested in the relation between statistical properties of extremes and dynamical properties of the system generating them.

Both theories are asymptotic, i.e. their statements are based on limit laws, which provide predictive power to the obtained results if the limit is reached empirically. As discussed in Sec. 1.5 and shown by our results in Sec. 4.1, this means that, beyond this empirical limit, the results are indistinguishable from the theoretical predictions taking into account the uncertainty due to the finite data size. If this is the case, we consider to have reached the asymptotic regime or level. From the dynamical systems point of view, when applying EVT we select one point on the attractor which we try to explore based on close recurrences in phase space, i.e. we look at a certain point through a kind of magnifying glass and try to explore the fine scale structure of the attractor (Lucarini et al., 2016). Thus, we still consider the original system, but instead of a global view, we switch to a local, in-depth view around the point of absolute maximum of the considered observable. When applying LDT, however, we adopt a completely different approach. In this case, the original observable is transformed through the averaging process, thus one considers a coarse grained observable with different, probably less heterogeneous, phase space properties. We adopt a global view, and try to approximate the full distribution of the new observable. Obviously, since the two different theories

operate in different dimensions, the predictive power is directed towards different dimensions as well: towards unobserved events, i.e. events with decreasing probability, in case of EVT; and towards averages obtained by increasing the averaging block length in case of LDT.

Both limit laws have been derived under certain conditions, which means that in anticipation of reliable results, one has to choose a system that is able to satisfy these conditions. Our models can generate strongly chaotic systems if the applied forcing is strong enough. These systems fulfil the so-called chaotic hypothesis, which means that they behave at all practical purposes as Axiom A systems. These systems are characterised by a fast decay of correlations, which provides the necessary mixing conditions to reach asymptotic levels even in a finite-sized data set. We have seen that the QG model with the strong forcing and the PUMA model under the applied forcing both represent systems where this is true. In these highly chaotic systems with well-developed turbulence, universality emerges as one approaches asymptotic regimes. Consequently, to find universality in a deterministic chaotic system, two conditions have to be fulfilled:

1. the dynamics of the system has to feature a sufficient degree of chaoticity to fulfil the mathematical conditions required for deriving the limit laws for its observables;
2. one has to reach asymptotic levels: the block size (threshold) has to be large (low) enough when applying the BM (POT) approach, and the averaging block length has to be long enough in case of LDT.

We summarise the core results of this thesis based on the conditions above:

- If both conditions are met, *asymptotic universality* emerges. This is the case of energy extremes in the QG model under the strong forcing. We have seen that the GEV or GPD estimated shape parameter is the same for different observables and corresponds in several cases with the theoretical one related to the partial dimensions of the attractor. In case of large deviations of temperature in PUMA, we find a universal function which describes both spatial, temporal, and spatio-temporal averages, for very small and very large zonal averaging lengths in the later case. These are certainly different types of universality: in the first case we find an equivalence between properties of different observables and general system characteristics, whereas in the second case we discover an equivalence between averaging in space and in time (and in space-time). However, there is a connection between them: in chaotic system with strong enough mixing or sufficiently fast decaying correlations, there exist levels where the effect of correlations, memory, or persistence vanishes, thus allowing for the emergence of universal properties described by asymptotic laws.

- If only condition 1. is fulfilled, one experiences *pre-asymptotic non-universality*. The pre-asymptotic differences are mostly caused by serial correlations in our systems. Hence, they seem to be characteristics of deterministic chaotic systems and represent a counterpart to asymptotic universality. The strength of correlations influence the effective block size and thus the values of the pre-asymptotic shape parameter estimate. In case of strong correlations, the number of uncorrelated data in a block is lower than in weakly correlated cases, and the difference between the estimated shape parameter and the asymptotic value (bias) is higher, at least considering monotonic convergence. Correlations also have an effect on the pre-asymptotic rate function estimates, which are broader in case of strongly correlated data compared to weakly correlated data, hinting to a higher probability of large deviations. Even the asymptotic rate function is affected by correlations, if one does not re-normalise it using the integrated auto-correlation.
- If both conditions are violated, one cannot apply the theories because the *asymptotic regime does not exist* or it is *impossible to be reached* on realistic time scales. This is the case when the forcing is not strong enough, and thus the system under investigation is not chaotic enough. The QG model under the weak forcing exhibits a regime behaviour, which “contaminates” the statistics of extreme events leading to non-monotonic shape parameter estimates, which do not converge to the theoretical value, at least on the explorable finite scales.
- In case of two dimensional, i.e. spatio-temporal, averaging of the temperature field in PUMA, we find deviations from asymptotic universality, although both conditions are met. Here, the spatial averaging is performed on intermediate zonal scales, where as an effect of non-trivial spatial correlations the coarse grained observable is totally different than the ones with universal properties. The re-normalised large deviation rate function is lower than the universal function, hinting to a higher probability of large fluctuations of the system, related to a higher degree of spatio-temporal organisation. This is a very interesting and important case of *asymptotic non-universality* due to persistent events, considering the high impact of events like heat waves or cold spells.

We have seen that by applying EVT and LDT we encounter similar problems or difficulties, which are mostly caused by correlations, finite data size, and convergence behaviour. Accordingly, we use similar strategies to deal with them. As mentioned above, in case of sufficiently fast decaying correlations, and if our system fulfils the chaotic hypothesis, the asymptotic limit exists, but it is more difficult to be reached. In the case of the BM method, we need longer blocks to guarantee for the independence of block maxima, in the case of LDT, we have to choose larger averaging block lengths to have approximately uncorrelated averages. When applying the POT method, things can be-

come even more complicated: declustering might be necessary to get rid of correlated threshold exceedances.

When applying asymptotic theories, it is of utmost importance to verify whether the asymptotic regime has been reached, otherwise the results can be erroneous. In the case of the BM method, this is equivalent with searching for an optimal block size, at which the shape parameter estimate corresponds to the theoretical value, it is stable against further increase of block size, and the estimation uncertainty is as small as possible. In the case of the POT method, one proceeds in an analogous way, but here the shape parameter is investigated as function of increasing threshold values. In our study, the theoretical shape parameter is known, thus we can estimate the actual bias. If this is not the case, one can rely only on verifying the stability of the shape parameter. Thus, the optimal block size (threshold) has to be large (high) enough, to allow for the asymptotic limit to hold, but at the same time small (low) enough, to limit uncertainty. Similarly, also in case of applying LDT, the optimal averaging block length has to be large enough, for the LDP to hold, but also small enough, so that the range of averaged values does not become too narrow as an effect of increasing averaging block lengths. Summing up, the *optimal level of estimation* when applying an asymptotic theory, is the *onset of the asymptotic regime*.

The speed of convergence to the asymptotic regime depends usually on the selected observable and the properties of the system. We have shown that, in case of a high dimensional system with a high degree of chaoticity, the convergence of the GEV and GPD shape parameters is monotonic, which can be related to the fact that we manage to collect more detailed information on the local properties of the attractor near the point of absolute maximum of the observable, and thus explore all the dimensions of the attractor. However, the convergence is very slow, which has to do probably with the high number of dimensions one has to explore. In the framework of EVT, theoretical studies show that the speed of convergence to the limiting distribution depends on the type of parent distribution (Leadbetter et al., 1983), and, as we have seen, it is additionally slowed down by correlations. In this work we experience similar effects in the case of LDT, where correlations and the skewness of the original distribution seem to slow down the convergence. Although we consider here different systems and different observables, it is still worth to mention the different scales at which the two limit laws converge in our work. Whereas the averages seem to converge for averaging block lengths of order of 10τ , with τ representing the integrated auto-correlation, the block maxima converge for block sizes larger than $O(10^4)\tau$. We assume that the slower convergence in case of EVT has to do, on the one hand, with a more agitated, more noisy, phase space structure considering the original observables, which makes it very difficult to explore every dimension, on the other hand, with the focus on very close and thus extremely rare recurrences in phase space.

We would like to point out that, although the above presented results and the related conclusions are quite theoretical, they are still relevant for applications related, for example, to climate risk assessment. We summarise the most important results from a practical perspective:

- The convergence to the limiting GEV and GPD distributions is slow, suggesting that customary choices like yearly maxima are not always the best option for an accurate modelling of extremes. If the estimated shape parameter is different from the asymptotic value, one has to be extremely careful especially in case of estimating events with long return times, because these can be over- or underestimated. We show that return levels with relatively short return periods can be stable, even though the shape parameter is still slightly changing with increasing block size (or threshold). However, the longer the return period, the more difficult it is to obtain stable return level estimates.
- Our results also show that significance tests, like the Kolmogorov-Smirnov test, are merely an indicator of the fit quality, but they do not show whether the convergence to the asymptotic GEV or GPD distributions is reached or not. Thus, once again, care is needed with the estimation of events with long return periods.
- In the case of LDT, we observe a significantly faster convergence than for EVT. Therefore, we believe that LDT can be an adequate and powerful method to analyse return times of persistent extreme events - like heat waves, cold spells, or droughts - and the eventual modification of their statistics due to climate change. However, this work is among the first ones in this direction, thus future studies are needed to improve the methodology, especially regarding applications to realistic data.
- We find a connection between properties of temporal and spatial (as well as spatio-temporal) averages by eliminating the effect of correlations in the computation of the rate functions. This relation is extremely important also from a practical perspective, because it allows for obtaining spatial characteristics from temporal data, and vice-versa. Additionally, we define a measure, which expresses how difficult it actually is to obtain the space-time connection. This measure, which we call scale velocity U_τ , is a fraction of the spatial and temporal integrated auto-correlation, thus it is related to the anisotropy between space and time.

Since the essence of scientific development is asking (and answering) questions, we finish this thesis with a few questions, which hint at possible directions for future studies. These questions concentrate around two main ones: 1) Can we find similar results in other, more realistic systems? 2) What is the connection between pre-asymptotic non-universality and asymptotic universality?

It would be very appealing to find out whether the results found in this thesis apply also in case of more realistic data, for example based on realistic climate model simulations or observations. Especially, the theory of LD could be easier to apply due to the apparently faster convergence. Thus, the next possible step would be to check whether one finds similar space-time asymptotic universality of the re-normalised rate functions together with the non-universality on intermediate spatial or temporal averaging scales in Large Eddy Simulations.

While asymptotic theories provide powerful results if the asymptotic regime is reached, they do not offer information about pre-asymptotic levels, where the effect of correlations and memory is still important. However, pre-asymptotic scales are the ones where temporal evolution takes place, and thus they are more comprehensible and considered to be more relevant for human individuals or societies. Pre-asymptotic non-universality is a big challenge for theoretical approaches, although some attempts have been made, for example, in the context of penultimate EVT (Cohen, 1982; Gomes and de Haan, 1999). In this work, we have seen that one can capture persistent events by averaging in two dimensions, space and time. By choosing a spatial averaging length on intermediate scales and then obtaining a LDP in time, one can analyse properties of persistent events. This strategy, i.e. computing spatial averages over intermediate scales and then analysing their extremes over time, is similarly applicable also in case of EVT. However, some opened questions still remain, for example: Can we describe the path leading from the phase space of a correlated observable to the one of the coarse grained observable in the asymptotic regime, and, even more relevant, can we follow this path also in the opposite direction in some cases? The answer to this question is extremely important to several sub-fields of climate modelling, like parametrisations or downscaling, thus depicts an exciting challenge for future studies.

References

- Agresti, A., 2007: *An Introduction to Categorical Data Analysis*. John Wiley and Sons, 2 edition.
- Anderson, T. W. and D. A. Darling, 1954: A test of goodness of fit. *Journal of the American Statistical Association*, **49**, 765–769.
- Balkema, A. A. and L. de Haan, 1974: Residual life time at great age. *The Annals of Probability*, **2**, 792–804.
- Barreira, L., Y. Pesin, and J. Schmeling, 1999: Dimension and product structure of hyperbolic measures. *Annals of Mathematics*, **149**, 755–783.
- Barriopedro, D., E. Fischer, J. Luterbacher, R. Trigo, and R. Garcia-Herrera, 2011: The hot summer of 2010: Redrawing the temperature record map of europe. *Science*, **332**, 220–224.
- Benettin, G., L. Galgani, A. Giorgilli, and J.-M. Strelcyn, 1980: Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory. *Meccanica*, **15**, 9–20.
- Bernardara, P., M. Lang, E. Sauquet, D. Schertzer, and I. Tchiriguyskaia, 2007: *Multifractal Analysis in Hydrology: Application to Time Series*. Quae.
- Billingsley, P., 1995: *Probability and Measure*. Wiley, 3 edition.
- Bódai, T., 2017: Extreme value analysis in dynamical systems: Two case studies. *Non-linear and Stochastic Climate Dynamics*, C. Franzke and T. O’Kane, eds., Cambridge University Press, Cambridge.
- Bouchet, F., C. Nardini, and T. Tangarife, 2014: Stochastic averaging, large deviations and random transitions for the dynamics of 2d and geostrophic turbulent vortices. *Fluid Dynamics Research*, **46**, 061416.
- Bouchet, F. and A. Venaille, 2012: Statistical mechanics of two-dimensional and geophysical flows. *Phys. Reports*, **515**, 227–295.
- Canty, A. and B. D. Ripley, 2017: *boot: Bootstrap R (S-Plus) Functions*. R package version 1.3-20.

- Cohen, J., 1982: Convergence rates for the ultimate and penultimate approximations in Extreme-Value Theory. *Adv. Appl. Prob.*, **14**, 833–854.
- Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer Verlag.
- Cramer, H., 1938: Sur un nouveau théorème limite dans la théorie des probabilités. *Colloque consacré à la théorie des probabilités*, Hermann, Paris, volume 3, 2–23.
- Davison, A. C. and D. V. Hinkley, 1997: *Bootstrap Methods and Their Applications*. Cambridge University Press, Cambridge, ISBN 0-521-57391-2.
URL <http://statwww.epfl.ch/davison/BMA/>
- Davison, A. C. and R. L. Smith, 1990: Models for exceedances over high thresholds. *J. R. Statist. Soc. B*, **52**, 393–442.
- De Cruz, L., S. Schubert, J. Demaeyer, V. Lucarini, and S. Vannitsem, 2018: Exploring the Lyapunov instability properties of high-dimensional atmospheric and climate models. *Nonlin. Processes Geophys.*, **2018**, 387412.
- Deidda, R., 2000: Rainfall downscaling in a space-time multifractal framework. *Water Resources Research*, **36**, 1779–1794.
- den Hollander, F., 2000: *Large Deviations*. American Mathematical Society.
- Donsker, M. and S. Varadhan, 1975a: Asymptotic evaluation of certain Markov process expectations for large time. I. *Comm. Pure Appl. Math.*, **28**, 1–47.
- 1975b: Asymptotic evaluation of certain Markov process expectations for large time. II. *Comm. Pure Appl. Math.*, **28**, 279–301.
- 1976: Asymptotic evaluation of certain Markov process expectations for large time. III. *Comm. Pure Appl. Math.*, **29**, 389–461.
- 1983: Asymptotic evaluation of certain Markov process expectations for large time. IV. *Comm. Pure Appl. Math.*, **36**, 183–212.
- Easterling, D., G. Meehl, C. Parmesan, S. Changnon, T. Karl, and L. Mearns, 2000: Climate extremes: Observations, modeling, and impacts. *Science*, **289**, 2068–2074.
- Eckmann, J.-P. and D. Ruelle, 1985: Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, **57**, 617–656.
- Ellis, R., 1984: Large deviations for a general class of random vectors. *Ann. Probab.*, **12**, 1–12.
- Fang, J.-Q. and G. Liu, 1992: Relationship between climatic change and the nomadic southward migrations in eastern asia during historical times. *Climatic Change*, **22**, 151–

168, doi:10.1007/BF00142964.

URL <https://doi.org/10.1007/BF00142964>

- Faranda, D., V. Lucarini, G. Turchetti, and S. Vaienti, 2011: Numerical convergence of the block-maxima approach to the Generalized Extreme Value distribution. *J. Stat. Phys.*, **145**, 1156–1180.
- Faranda, D., G. Messori, and P. Yiou, 2017: Dynamical proxies of North Atlantic predictability and extremes. *Scientific Reports*, **7**, 41278.
- Felici, M., V. Lucarini, A. Speranza, and R. Vitolo, 2007a: Extreme value statistics of the total energy in an intermediate-complexity model of the midlatitude atmospheric jet. Part I: Stationary case. *Journal of Atmospheric Sciences*, **64**, 2137–2158.
- Ferro, C. A. T. and J. Segers, 2003a: Inference for clusters of extreme values. *J. R. Statist. Soc. B*, **65**, 545–556.
- 2003b: Inference for clusters of extreme values. *J. R. Statist. Soc. B*, **65**, 545–556.
- Fisher, R. A. and L. H. C. Tippett, 1928: Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, **24**, 180–190.
- Fraedrich, K., H. Jansen, E. Kirk, U. Luksch, and F. Lunkeit, 2005a: The planet simulator: Towards a user friendly model. *Meteorologische Zeitschrift*, **44**, 299–304.
- Fraedrich, K., E. Kirk, U. Luksch, and F. Lunkeit, 2005b: The portable university model of the atmosphere (PUMA): Storm track dynamics and low-frequency variability. *Meteorologische Zeitschrift*, **14**, 735–745.
- Fraedrich, K., E. Kirk, and F. Lunkeit, 2009: PUMA Portable University Model of the Atmosphere. Technical report, World Data Center for Climate (WDCC) at DKRZ. URL https://doi.org/10.2312/WDCC/DKRZ_Report_No16
- Franzke, C., A. J. Majda, and E. Vanden-Eijnden, 2005: Low-order stochastic mode reduction for a realistic barotropic model climate. *Journal of the Atmospheric Sciences*, **62**, 1722–1745.
- Frisius, T., 1998: A mechanism for the barotropic equilibration of baroclinic waves. *Journal of Atmospheric Sciences*, **55**, 2918–2936.
- functions written by Janet E. Heffernan with R port, O. S. and R. documentation provided by Alec G. Stephenson., 2016: *ismev: An Introduction to Statistical Modeling of Extreme Values*. R package version 1.41. URL <https://CRAN.R-project.org/package=ismev>

- Gallavotti, G., 2014: *Nonequilibrium and irreversibility*. Springer.
- Gallavotti, G. and E. Cohen, 1995: Dynamical ensembles in stationary states. *Journal of Statistical Physics*, **80**, 931–970.
- Gärtner, J., 1977: On large deviations from the invariant measure. *Theory Probab. Appl.*, **22**, 24–39.
- Ghil, M. and W. Roberston, 2002: Waves vs. particles in the atmospheres phase space: A pathway to long-range forecasting? *PNAS*, **99**, 2493–2500.
- Ghil, M., P. Yiou, S. Hallegatte, B. D. Malamud, P. Naveau, A. Soloviev, P. Friederichs, V. Keilis-Borok, D. Kondrashov, V. Kossobokov, O. Mestre, C. Nicolis, H. W. Rust, P. Shebalin, M. Vrac, A. Witt, and I. Zaliapin, 2011: Extreme events: dynamics, statistics and prediction. *Nonlin. Processes Geophys.*, **18**, 295–350.
- Giardina, C., J. Kurchan, V. Lecomte, and J. Tailleur, 2016: Simulating rare events in dynamical processes. *J. Stat. Phys.*, **145**, 787–811.
- Gnedenko, B., 1943: Sur la distribution limite du terme maximum d’une serie aleatoire. *Annals of Mathematics*, **44**, 423–453.
- Gomes, M. and L. de Haan, 1999: Approximation by penultimate extreme value distributions. *Extremes* **2**, **1**, 71–85.
- Grieger, B., J. Segschneider, H. Keller, A. Rodin, F. Lunkeit, E. Kirk, and K. Fraedrich, 2004: Simulating Titan’s tropospheric circulation with the Portable University Model of the Atmosphere. *Advances in Space Research*, **34**, 1650–1654.
- Held, I. and M. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Amer. Meteor. Soc.*, **75**, 1825–1830.
- Herbert, C., 2015: An introduction to large deviations and equilibrium statistical mechanics for turbulent flows. *Stochastic Equations for Complex Systems: Theoretical and Computational Topics*, S. Heinz and H. Bessaih, eds., Springer, Switzerland.
- Holland, M. P., R. Vitolo, P. Rabassa, A. E. Sterk, and H. W. Broer, 2012: Extreme value laws in dynamical systems under physical observables. *Physica D*, **241**, 497–513.
- Holton, J. R., 2004: *An introduction to dynamic meteorology*. Elsevier Academic Press.
- Hosking, R. M., 1990: L-moments: Analysis and estimation of distributions using linear combination of order statistics. *J. R. Statist. Soc. B*, **52**, 105–124.
- Hosking, R. M., J. R. Wallis, and E. F. Wood, 1985: Estimation of the Generalized Extreme-Value distribution by the method of Probability-Weighted Moments. *Technometrics*, **27**, 251–261.

- Hoskins, B. and A. Simons, 1975: A multi-layer spectral model and the semi-implicit method. *Q. J. R. Meteorol. Soc.*, **101**, 637–655.
- Hvistendahl, M., 2012: Roots of empire. *Science*, **337**, 1596–1599, doi:10.1126/science.337.6102.1596.
URL <http://science.sciencemag.org/content/337/6102/1596>
- IPCC, 2012: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*. Cambridge University Presse, [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change.
- 2013: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, [Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (eds.)].
- Kennett, D. J., S. F. M. Breitenbach, V. V. Aquino, Y. Asmerom, J. Awe, J. U. Baldini, P. Bartlein, B. J. Culleton, C. Ebert, C. Jazwa, M. J. Macri, N. Marwan, V. Polyak, K. M. Prufer, H. E. Ridley, H. Sodemann, B. Winterhalder, and G. H. Haug, 2012: Development and Disintegration of Maya Political Systems in Response to Climate Change. *Science*, **338**, 788–791, doi:10.1126/science.1226299.
URL <http://science.sciencemag.org/content/338/6108/788>
- Kleidon, A., 2010: Non-equilibrium thermodynamics, maximum entropy production and Earth-system evolution. *Phil. Trans. R. Soc. A*, **368**, 181–196.
- Klein, R., 2010: Scale-dependent models for atmospheric flows. *Annu. Rev. Fluid Mech.*, **42**, 249–274.
- Koutsoyiannis, D., 2004a: Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation / Statistiques de valeurs extremes et estimation de precipitations extremes: I. Recherche theorique. *Hydrological Sciences Journal*, **49**, 575–590.
- 2004b: Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records / Statistiques de valeurs extremes et estimation de precipitations extremes: II. Recherche empirique sur de longues series de precipitations. *Hydrological Sciences Journal*, **49**, 591–610.
- Leadbetter, M. R., 1974: On extreme values in stationary sequences. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, **28**, 289–303.

-
- Leadbetter, M. R., G. Lindgren, and H. Rootzén, 1983: *Extremes and Related Properties of Random Sequences and Processes*. Springer Verlag.
- Leadbetter, M. R., I. Weissman, L. de Haan, and H. Rootzen, 1989: On clustering of high values in statistically stationary series. *Proceedings of the 4th International Meeting on Statistical Climatology*, J. Sansom, ed., New Zealand Meteorological Service, Wellington, 217–222.
- Lestang, T., F. Ragone, C. Brehier, C. Herbert, and F. Bouchet, 2018: Computing return times or return periods with rare event algorithms. *J. Stat. Mech.*, **2018**, 043213.
- Lobell, D., W. Schlenker, and J. Costa-Roberts, 2011: Climate trends and global crop production since 1980. *Science*, **333**, 616–620.
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, **20**, 130–141.
- 1984: Irregularity: a fundamental property of the atmosphere. *Tellus*, **36A**, 98–110.
- 1996: Predictability – a problem partly solved. *Seminar on Predictability*, ECMWF, volume I.
- Lovejoy, S. and D. Schertzer, 2013: *The weather and Climate: emergent laws and multifractal cascades*. Cambridge U. Press.
- Lucarini, V., R. Blender, C. Herbert, F. Ragone, S. Pascale, and J. Wouters, 2014a: Mathematical and physical ideas for climate science. *Rev. Geophys.*, **52**, 809859.
- Lucarini, V., D. Faranda, A. C. M. Freitas, J. M. Freitas, M. Holland, T. Kuna, M. Nicol, M. Todd, and S. Vaienti, 2016: *Extremes and Recurrence in Dynamical Systems*. Wiley.
- Lucarini, V., D. Faranda, G. Turchetti, and S. Vaienti, 2012b: Extreme value theory for singular measures. *Chaos*, **22**, 023135.
- Lucarini, V., D. Faranda, and J. Wouters, 2012a: Universal behaviour of extreme value statistics for selected observables of dynamical systems. *J. Stat. Phys.*, **147**, 63–73.
- Lucarini, V., D. Faranda, J. Wouters, and T. Kuna, 2014b: Towards a general theory of extremes for observables of chaotic dynamical systems. *J. Stat. Phys.*, **154**, 723–750, licensed under CC BY 4.0, <http://creativecommons.org/licenses/by/4.0/>.
- Lucarini, V., K. Fraedrich, and F. Lunkeit, 2010: Thermodynamic analysis of snowball Earth hysteresis experiment: Efficiency, entropy production and irreversibility. *Quarterly Journal of the Royal Meteorological Society*, **136**, 2–11, doi:10.1002/qj.543.
- Lunkeit, F., K. Fraedrich, and S. Bauer, 1998: Storm tracks in a warmer climate: sensitivity studies with a simplified global circulation model. *Clim. Dynamics*, **14**, 813–826.
-

- Massey, F. J. J., 1951: The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, **46**, 68–78.
- MATLAB, 2015: *MATLAB and Statistics Toolbox Release 2015b*. The MathWorks Inc., Natick, Massachusetts, United States.
- Messori, G., D. Faranda, and R. Caballero, 2017: A dynamical systems approach to studying midlatitude weather extremes. *Geophys. Res. Lett.*, **44**, 3346–3354.
- Mikosch, T. and A. Nagaev, 1998: Large deviations of heavy-tailed sums with applications in insurance. *Extremes* **1**, 1, 81–110.
- Oono, Y., 1989: Large deviation and statistical physics. *Progr. Theoret. Phys. Suppl.*, **99**, 165–205.
- Oseledets, V., 1968: A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, **19**, 197–231.
- Ott, E., 1993: *Chaos in dynamical systems*. Cambridge University Press.
- Phillips, N. A., 1956: The general circulation of the atmosphere: a numerical experiment. *Q. J. R. Meteorol. Soc.*, **82**, 123–164.
- Pickands, J., 1975: Statistical inference using extreme order statistics. *The Annals of Statistics*, **3**, 119–131.
- Prokofiev, 2010: Ikeda attractor for parameters $a=1$, $b=0.9$, $k=0.4$ and $p=6$. Accessed on: 07-08-2018, licensed under CC BY-SA 3.0, <http://creativecommons.org/licenses/by-sa/3.0/deed.en>.
URL https://commons.wikimedia.org/wiki/File:Ikeda_map_a%3D1_b%3D0.9_k%3D0.4_p%3D6.jpg#file
- R Core Team, 2016: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
URL <https://www.R-project.org/>
- Ragone, F., J. Wouters, and F. Bouchet, 2017: Computation of extreme heat waves in climate models using a large deviation algorithm. *PNAS*, **115**, 24–29.
- Robinson, P., 2001: On the definition of a heat wave. *J. Appl. Met.*, **40**, 762–775.
- Rohwer, C., F. Angeletti, and H. Touchette, 2015: Convergence of large-deviation estimators. *Phys. Rev. E*, **92**, 052104.
- Ruelle, D., 1989: *Chaotic Evolution and Strange Attractors*. Cambridge University Press.

-
- Rust, H. W., 2009: The effect of long-range dependence on modelling extremes with the generalised extreme value distribution. *Eur. Phys. J. Special Topics*, **174**, 91–97.
- Schubert, S. and V. Lucarini, 2015: Covariant Lyapunov vectors of a quasi-geostrophic baroclinic model: analysis of instabilities and feedbacks. *Quarterly Journal of the Royal Meteorological Society*, **141**, 3040–3055.
- 2016: Dynamical analysis of blocking events: spatial and temporal fluctuations of covariant Lyapunov vectors. *Quarterly Journal of the Royal Meteorological Society*, **142**, 2143–2158.
- Segers, J., 2005: Approximate distributions of clusters of extremes. *Stat. and Prob. Letters*, **74**, 330–336.
- Sillmann, J., M. Croci-Maspolli, M. Kallache, and R. Katz, 2011: Extreme Cold Winter Temperatures in Europe under the Influence of North Atlantic Atmospheric Blocking. *J. Climate*, **24**, 5899–5913.
- Smith, R. L., 1989: Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Statistical Science*, **4**, 367–393.
- Stefanon, M., F. D’Andrea, and P. Drobinski, 2012: Heatwave classification over Europe and the Mediterranean region. *Environ. Res. Lett.*, **7**, 014023.
- Stull, R., 2015-2016: *Practical Meteorology: An Algebra-based Survey of Atmospheric Science*. The University of British Columbia, licensed under CC BY-NC-SA 4.0, <http://creativecommons.org/licenses/by-nc-sa/4.0/>.
URL http://www.eos.ubc.ca/books/Practical_Meteorology/
- Touchette, H., 2009: The large deviation approach to statistical mechanics. *Physics Reports*, **478**, 1–69.
- Vannitsem, S., 2007: Statistical properties of the temperature maxima in an intermediate order Quasi-Geostrophic model. *Tellus*, **59A**, 80–95.
- Vannitsem, S. and V. Lucarini, 2016: Statistical and Dynamical Properties of Covariant Lyapunov Vectors in a Coupled Atmosphere-Ocean Model multiscale effects, geometric degeneracy, and error dynamics. *J. Phys. A-Math. Theor.*, **49**, 224001.
- Vautard, R., P. Yiou, F. D’Andrea, N. de Noblet, N. Viovy, C. Cassou, J. Polcher, P. Ciais, M. Kageyama, and Y. Fan, 2011: Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit. *Geophys. Res. Lett.*, **34**, L07711.
- WHO, 2004: Heat-waves: risks and responses. *Health and Global Environmental Change*, **2**.

- Wouters, J. and F. Bouchet, 2016: Rare event computation in deterministic chaotic systems using genealogical particle analysis. *J. Phys. A: Math. Theor.*, **49**, 374002.
- Zahid, M., R. Blender, V. Lucarini, and M. C. Bramati, 2017: Return levels of temperature extremes in Southern Pakistan. *Earth Syst. Dynam. Discuss.*, **8**, 1263–1278.

Acronyms

EVT Extreme Value Theory

LDT Large Deviation Theory

LDP Large Deviation Principle

QG Quasi-Geostrophic

PUMA Portable University Model of the Atmosphere

BM Block Maxima

POT Peak over Threshold

GEV Generalised Extreme Value

GPD Generalised Pareto Distribution

CLT Central Limit Theorem

MLE Maximum Likelihood Estimation

KS Kolmogorov-Smirnov

List of important symbols in Chapter 4

ΔT	meridional Temperature difference
x	longitude
y	latitude
t	time
e	local energy
e_z	zonally averaged energy
e_{ml}	averaged mid-latitude energy
τ	integrated auto-correlation
m	block size
m^*	optimal block size
ξ	shape parameter
ξ_δ	theoretical shape parameter
$\xi(m)$	estimated shape parameter
T	trueness
P	precision
r	exceedance ratio
r^*	optimal exceedance ratio
u	GPD threshold
$\hat{\sigma}$	GPD modified scale parameter
A_{max}	absolute maximum

List of important symbols in Chapter 5

x	longitude
y	latitude
t	time
T	temperature
T'	temperature fluctuation
τ	integrated auto-correlation
n	averaging block length
n^*	optimal averaging block length
A	averaged temperature
I	rate function
\tilde{I}	rate function estimate

Acknowledgements

First of all, I would like to thank to my supervisor Valerio Lucarini for his guidance during my PhD, for his kind support, great ideas, but also for giving me the necessary freedom to make my own decisions. I would like to thank to my co-supervisor Christian Franzke as well for his caring support and help in every matter. I am also thankful to Johanna Baehr and Cathy Hohenegger for taking care that everything is going well.

I am grateful to the International Max Planck Research School on Earth System Modelling for the financial support during my PhD. My special gratitude goes to the IMPRS-ESM office, to Antje Weitz, Cornelia Kampmann, Michaela Born, and Wiebke Böhm, for being always supportive and for finding solutions in case of problems.

My sincere thanks goes also to the co-authors of our papers, Tamás Bódai and Jeroen Wouters, for answering my questions and providing help anytime it was needed.

Special thanks goes to Edilbert Kirk, Frank Lunkeit, and Frank Sielmann for help with the model simulations. I am grateful to Richard Blender for taking time and explaining concepts of Extreme Value Theory and statistics in general.

I would like to thank to all of my colleagues for their support and for the cheerful lunch, coffee, and ice-cream breaks. I am indebted to Sebastian Schubert for answering my questions about the QG model and dynamical systems in general. I would particularly like to thank to Maida Zahid for her encouragement during the whole PhD, for the fruitful discussions about EVT, for being so modest and at the same time so strong. I am extremely thankful to Federica Gugole for being supportive and listening to the usual problems and worries. I would like to sincerely thank also to Valerio Lembo, Denny Gohlke, Gabriele Vissio, Markus Kilian, and Guannan Hu for their support and pleasant company.

I am extremely thankful to my parents for their love and care, and for always believing in me. I am deeply grateful to Tassilo for his immense support in everything I do, for being so optimistic, and giving me the strength to solve any problems.

Aus dieser Dissertation hervorgegangene Vorveröffentlichungen und Einsendungen

List of publications and submissions

1) **Galfi, V.M.**, T. Bodai, and V. Lucarini, 2017: Convergence of extreme value statistics in a two-layer quasi-geostrophic atmospheric model. *Complexity*, 2017, 5340858, <https://doi.org/10.1155/2017/5340858>.

Own contributions: performing the numerical simulations and the analysis, plotting figures, developing the concept of the paper, writing the paper.

Co-authors' contributions: TB helped with performing the analysis. VL and TB contributed to develop the concept of the paper and to write the paper.

2) **Galfi, V.M.**, V. Lucarini, and J. Wouters, 2018: A Large Deviation Theory-based Analysis of Heat Waves and Cold Spells in a Simplified Model of the General Circulation of the Atmosphere, submitted to *J. Stat. Mech.*.

Own contributions: performing the numerical simulations and the analysis, plotting figures, developing the concept of the paper, writing the paper.

Co-authors' contributions: VL and JW contributed to develop the concept of the paper and to write the paper.

Eidesstattliche Versicherung

Declaration on oath

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation mit dem Titel: "Mathematics of extreme events in atmospheric models" selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Dissertation oder Teile davon vorher weder im In- noch im Ausland in einem anderen Prüfungsverfahren eingereicht habe und die eingereichte schriftliche Fassung der auf dem elektronischen Speichermedium entspricht.

I hereby declare, on oath, that I have written the present dissertation by myself and have not used other than the acknowledged resources and aids. The submitted written version corresponds to the electronic one. I declare that the present dissertation has not been submitted in any other previous doctorate examination procedure.

Hamburg,

