Institut für Physik · Arbeitsgruppe Nichtlineare Dynamik

What Can We Learn from Climate Data? Methods for Fluctuation, Time/Scale and Phase Analysis

Dissertation zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
in der Wissenschaftsdisziplin Nichtlineare Dynamik

Eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

von
Douglas Maraun

Potsdam, im Juni 2006

The following contributions have been published or submitted for publication as part of this work or related to it:

D. Maraun, H.W. Rust, and J. Timmer. Tempting long-memory. On the interpretation of DFA results. *Nonlin. Proc. Geoph.*, 11(4):495–503, 2004.

D. Maraun and J. Kurths. Cross wavelet analysis. Significance testing and pitfalls. *Nonlin. Proc. Geoph.*, 11(4):505–514, 2004.

S. Rahmstorf, D. Archer, D.S. Ebel, O. Eugster, J. Jouzel, D. Maraun, U. Neu, G.A. Schmidt, J. Severinghaus, A.J. Weaver, and J. Zachos. Cosmic rays, carbon dioxide, and climate. *Eos, Trans. AGU*, 85(4):38, 2004.

S. Rahmstorf, D. Archer, D.S. Ebel, O. Eugster, J. Jouzel, D. Maraun, U. Neu, G.A. Schmidt, J. Severinghaus, A.J. Weaver, and J. Zachos. Reply to comment on ”cosmic rays, carbon dioxide, and climate”. *Eos, Trans. AGU*, 85(48):511, 2004.

D. Maraun and J. Kurths. Epochs of phase coherence between El Niño/ Southern Oscillation and Indian monsoon. *Geophys. Res. Lett.*, 32:L15709, doi:10.1029/2005GL023225, 2005.

J. Saynisch, J. Kurths, and D. Maraun. A Conceptual ENSO Modell under Realistic Noise Forcing. *accepted for publication in Nonlin. Proc. Geophys.*, 2006.

D. Maraun, J. Kurths, and M. Holschneider. Nonstationary Gaussian processes in wavelet domain: Definitions, estimation and significance testing. *submitted to Phys. Rev. E*, 2006.

D. Maraun, B. Schaefli and M. Holschneider. Wavelet Spectral and Cross Spectral Analysis in Hydrology. submitted for publication in the proceedings of the workshop “New Developments in Trend- and Extreme Value Analysis of Hydrometeorological Time Series”, Springer, 2006.

P.M. Crowley, D. Maraun and D. Mayes. How hard is the euro area core? A wavelet analysis of growth cycles in Germany, France and Italy. *submitted to Econ. J.*, 2006.

To Antje, the friend of my years in Berlin

Abbreviations often Used in this Thesis

ACF	Auto correlation function
AIR	All India Rainfall Index
CWT	Continuous Wavelet Transformation
CDP	Curvature defined phase
DFA	Detrended Fluctuation Analysis
DWT	Discrete Wavelet Transform
ENSO	El Niño/Southern Oscillation
HPA	Hilbert phase analysis
ISM	Indian summer monsoon
MODWT	Maximum Overlap Discrete Wavelet Transformation
NAO	North Atlantic Oscillation
SST	Sea surface temperature

Notation

a	Scale (in wavelet domain), $a = 1/f$
$A(t)$	Amplitude of an oscillation
$A(.)$	Averaging operator
b	Time (in wavelet domain)
$C(s)$, $C(r)$	Autocorrelation function
$COH^2(b,a)$	Squared wavelet coherency
$CS(b,a)$	Wavelet cross spectrum
Cov	Covariance
f	Frequency
$f(\omega)$	Fourier multiplier
$g(t)$	Wavelet
$h(t)$	Reconstruction wavelet
$H(.)$	Hilbert transformation
H	Hurst exponent
$\mathbb{H}$	Positive halfplane
$\mathbb{I}$	Identity
$m(a,b)$	Wavelet multiplier
M_h	Inverse continuous wavelet transformation with reconstruction wavelet h
$P_X(\lambda)$	Probability distribution
$\mathcal{P}$	Probability
$\rho_X(x)$	Probability density function of the random variable X
$s(t)$	Time series
s	Scale (in fluctuation analysis)
SE, SP	Sensitivity, specificity
SNR	Signal to noise ratio
$S(b,a)$	Wavelet spectrum
Var	Variance
W_g	Continuous wavelet transformation with wavelet g
$x(t)$, x_i	Time series
γ	Scaling exponent of the autocorrelation function
η	noise
$\Phi(t)$	Phase of an oscillation
τ_D	Characteristic time scale
ω	Frequency, $\omega = 2\pi f$
ω_0	Eigen-frequency of an oscillator, parameter of the Morlet wavelet
$\bar{.}$	Complex conjugate
$\hat{.}$	Estimator, Fourier transformation (depending on context)
$\langle . \rangle$	Expectation value

Contents

Introduction **xi**

1 General Concepts **1**
- 1.1 Random Variables . . . 1
  - 1.1.1 Random Variables . . . 1
  - 1.1.2 Multivariate Random Variables . . . 2
  - 1.1.3 Moments . . . 2
  - 1.1.4 Quantiles . . . 3
- 1.2 Stochastic Processes . . . 3
  - 1.2.1 Properties of Stochastic Processes . . . 4
  - 1.2.2 Gaussian Processes . . . 4
- 1.3 Estimators . . . 6
  - 1.3.1 Bias . . . 6
  - 1.3.2 Variance . . . 6
  - 1.3.3 Consistency . . . 6
- 1.4 Significance Testing . . . 7
  - 1.4.1 Statistical Errors and Power of the Test . . . 7
  - 1.4.2 Sensitivity vs. Specificity . . . 7
  - 1.4.3 Multiple Testing . . . 8
- 1.5 Inverse Problems . . . 8

2 Tempting Long Memory **9**
- 2.1 Introduction . . . 9
- 2.2 Autocorrelation Function of Stochastic Processes . . . 10
  - 2.2.1 Short Range Correlations . . . 10
  - 2.2.2 Long Range Correlations . . . 11
  - 2.2.3 Discrimination between Short and Long Memory . . . 11
- 2.3 Detrended Fluctuation Analysis (DFA) . . . 11
  - 2.3.1 The Algorithm . . . 12
  - 2.3.2 Estimating the Strength of Long Range Correlations . . . 12
- 2.4 Inference of Long Memory . . . 13
  - 2.4.1 Two Example Processes . . . 13
  - 2.4.2 Establish Scaling of the Fluctuation Function . . . 14
- 2.5 Inference of Short Memory . . . 18
  - 2.5.1 The Fluctuation Function of AR[1]-Processes . . . 19
  - 2.5.2 Finite Scaling of Short-Memory Processes . . . 19
  - 2.5.3 Shift of the finite scaling region . . . 20

2.6 Memory of Temperature Records 21
2.7 Summary and Conclusions 23

3 Continuous Wavelet Spectral Analysis 25
3.1 Introduction 25
3.2 Mathematical Background 26
  3.2.1 From Fourier to Wavelet Transformation 26
  3.2.2 Wavelets 27
  3.2.3 Continuous Wavelet Transformation 28
  3.2.4 Inverse Continuous Wavelet Transformation 28
  3.2.5 Reproducing Kernel 29
  3.2.6 Choice of the Wavelet 29
  3.2.7 Discrete Wavelet Transformation 29
3.3 Motivation - The State of the Art 29
  3.3.1 Conventional Definitions of Wavelet Spectral Measures 30
  3.3.2 Conventional Smoothing and Global Wavelet Spectrum 31
  3.3.3 Conventional Pointwise Significance Testing 31
  3.3.4 Questions arising 32
3.4 Stochastic Processes defined in Wavelet Domain 34
  3.4.1 Stationary Gaussian Processes in Fourier Domain 34
  3.4.2 Definitions 34
  3.4.3 Spectral Measures 36
  3.4.4 Example 37
3.5 Estimating Wavelet Spectra 37
  3.5.1 Spectral Estimators 38
  3.5.2 Distribution, Variance and Bias of the Estimator 39
  3.5.3 Example 42
3.6 Significance Testing 43
  3.6.1 Pointwise Testing of the Wavelet Spectrum 43
  3.6.2 Areawise Testing of the Wavelet Spectrum 44
  3.6.3 Sensitivity and Specificity of the Areawise Test 46
  3.6.4 Testing of Covarying Power 52
3.7 Applications 54
  3.7.1 Wavelet Analysis of ENSO and NAO 54
  3.7.2 Wavelet Analysis in Hydrology 55
3.8 Summary and Conclusions 57

4 Phase Coherence between ENSO and Monsoon 61
4.1 Introduction 61
4.2 Sketching the Physics of ENSO and Monsoon 62
  4.2.1 ENSO 62
  4.2.2 Indian Summer Monsoon 63
  4.2.3 Coupling between ENSO and Monsoon 63
4.3 Questions 64
4.4 Data 65
4.5 Concepts and Methods 66
  4.5.1 Phase Synchronization 66

4.5.2 Measures of Synchronization . . . 66
4.5.3 Inference of Coupling Direction . . . 67
4.5.4 Hilbert Phase Estimation . . . 67
4.6 Phase of ENSO and Indian Monsoon . . . 72
4.7 Results . . . 73
4.7.1 Phase Relation . . . 73
4.7.2 Significance Testing . . . 76
4.7.3 Coupling Direction . . . 78
4.8 Climatological Discussion . . . 78
4.8.1 Interdecadal Variability . . . 78
4.8.2 Temporarily Weakening Relationship with Precedents . . . 79
4.8.3 Possible Volcanic Forcing . . . 79
4.9 Summary and Conclusions . . . 80

5 Discussion, Conclusions and Outlook **81**
5.1 Discussion and Conclusions . . . 81
5.2 Outlook . . . 83

A Data Sets **85**
A.1 Large Scale Climate Indeces . . . 85
A.1.1 El Niño/Southern Oscillation . . . 85
A.1.2 North Atlantic Oscillation . . . 85
A.1.3 Indian Monsoon . . . 85
A.1.4 Volcanic Radiative Forcing Index . . . 85
A.2 Regional Climate data from the Swiss Alps . . . 85
A.2.1 Precipitation at Bourg St. Pierre . . . 85
A.2.2 Temperature at Grand St. Bernard . . . 88
A.2.3 Run-Off of River Drance de Bagnes . . . 88

B Continuous Wavelet Spectral Analysis **89**
B.1 Properties of the Transformation . . . 89
B.1.1 Covariances . . . 89
B.1.2 Projection Property . . . 89
B.1.3 Reproducing Kernel . . . 89
B.1.4 Example Wavelets . . . 90
B.1.5 The Morlet Wavelet . . . 91
B.1.6 Wavelet Transformation of Discrete Sampled Data . . . 91
B.1.7 Cone of Influence . . . 92
B.2 Locally Stationary Wavelet Processes . . . 93
B.3 Properties of Gaussian Processes in Wavelet Domain . . . 93
B.3.1 The Dependency on the Wavelet . . . 93
B.3.2 The Relation to Fourier Spectra . . . 94
B.3.3 Distribution of the Wavelet Sample Spectrum . . . 95
B.3.4 Bias of the Spectral Estimator . . . 96
B.4 Estimating the Patch-size . . . 97
B.4.1 Stochastic Root Finding . . . 97

C Phase Coherence **99**
C.1 A Stochastically Perturbed ENSO Delay Oscillator 99
C.2 The Roessler Oscillator 99

D Software **101**
D.1 Wavelet Analysis 101
D.2 DFA and Phase Analysis 101

Introduction

Those among us who are unwilling to expose their ideas
to the hazard of refutation do not take part in the scientific game.
Sir Karl Popper (1902-1994)

Since early human history mankind has been eager to understand weather and climate. The ancient civilizations along the great rivers Euphrates and Tigris, Nile or Indus, in Mesoamerica or in the Andes heavily relied on the annual rainfall and floods. Long lasting dry periods arguably caused the decline of the Maya civilization in the 9th century [38] and the collapse of the old kingdom of Egypt around 4200 years ago [115], a story that is told of in the bible. Being at the mercy of seemingly unpredictable forces, humans worshiped the gods of rain or rivers, agriculture and fertility as Varuna in India, Chaac in the Maya kingdom or Hapi in ancient Egypt.

The observational field of meteorology was first defined by Aristotle 350 years BC. Nevertheless, a modern scientific investigation did not emerge before the time of Renaissance, when Galileo Galilei invented the first thermometer. Parallel to the stunning development of physics, researchers formulated the principles of the atmospheric and the ocean sciences.

However, even though meteorology and climatology - as other geosciences - are based on physical and chemical theories, fundamental differences to the classical scientific disciplines do exist. Popper [88] defined the key standards for a scientific theory and experimental testing by the terms of refutability or falsifiability and reproducibility: Since the truth of a theoretical concept can never be inferred empirically, progress in science arises from testing and either falsifying or corroborating competing hypotheses. A theory that is not falsifiable is tautological and unscientific. Hence, experiments have to be designed to test and possibly refute a theoretical framework. As any experiment might be subject to measurement errors, inherent stochasticity or deviations in the initial conditions, the essential statements to be drawn have to be reproducible. In geosciences, one often cannot design experiments at all. Although laboratory research might investigate certain material parameters, small scale behavior or basic principles, one will never be able to experimentally scan the whole parameter space of, for instance, the El Niño/Southern oscillation (ENSO) system. In general, climatologists are restricted to observing a single realization of a phenomenon. If a scientist has access to several realizations they often differ strongly in the initial and boundary conditions and are not easily comparable. Another typical limit of geosciences is the shortage of observational data in relation to the complexity of the processes. The development of modern computers partly provides a way out: Numerical models allow for the design of simulated experiments to test and falsify hypotheses. However, a self-consistent model will never falsify itself. Thus, without a connection to the real world, model experiments are rather a useful meta science. The desired connection to observations finally leads to the concept of time series analysis to corroborate the numerical models and to propose new hypotheses reaching beyond the established theory.

Time series analysis aims to infer characteristic properties of an underlying process, for instance its parameters or even the process structure itself, from observational data. This approach constitutes an inverse problem [76]. Corresponding to the complexity and often stochastic nature of the observed systems, testing and falsification are limited to statistical statements characterized by confidence intervals and significance testing. In fact, the concept of falsification is turned upside down in the concept of significance testing: An observation is attributed being statistically significant, when it with high probability is not compatible with a trivial null hypothesis. Thus, one does not try to falsify the hypothesis itself but the null hypothesis. However, for a sufficient amount of data every null hypothesis will be rejected. Furthermore, a researcher has to accurately understand the null hypothesis to reliably interpret the meaning of its rejection. Beside these well known inherent problems of significance testing, one has to consider sensitivity and specificity: On the one hand, a test has to be capable of extracting the interesting statements - this demand is characterized by the sensitivity of the test. On the other hand, the null hypothesis should be rejected in case of a real deviation only - this demand is characterized by the specificity of the test. In the statistical context, the corroboration of a hypothesis by trying to falsify it requires a significance test of high specificity.

The above discussion sets the stage for the work at hand: My thesis mainly aims to contribute to methodological aspects of time series analysis, but also to provide applications to important climatological problems. In the framework of the collaborative research center SFB 555 of the German research council (DFG), I exemplarily study, enhance and apply three statistical methods of interest for climatology. Here I focus on the inverse problem perspective, i.e. a reliable and unambiguous inference by means of significance testing. In this context, the discussion of sensitivity and specificity takes up a central role.

The first chapter provides a brief overview of the basic concepts of random numbers, stochastic processes, parameter estimation and significance testing. It contains no new results and is rather dedicated to readers not too familiar with these concepts.

In the second chapter, I present work done in collaboration with Prof. Dr. Jens Timmer and Henning Rust. We contribute to a recent discussion about the autocorrelation structure of temperature time series. Prominent publications suggest a universal law of long memory that is not accounted for in state-of-the-art climate models. The authors conclude that global warming derived from trend estimates and model runs has been exaggerated. However, we studied the method used for the inference in terms of sensitivity and specificity and showed that the mentioned conclusions cannot be drawn from the given amount of data. This work has been published in [70].

In chapter three, I study a very popular method for time/scale resolved analysis. Continuous wavelet spectral analysis is widely used in climatology. However, I realized that the suggested methodology is far away from being a reliable and closed concept. After initial studies [67] and a draft of an areawise significance test in mind, I started to collaborate with Prof. Dr. Matthias Holschneider, who had the idea to define stochastic processes in the wavelet domain. Hereupon we developed a framework of nonstationary Gaussian processes in wavelet domain. Starting from this concept, I formulated estimators for wavelet spectra and developed a new significance test with a considerably improved specificity [71]. These works provide a theory for the continuous wavelet spectral analysis. Together with Dr. Bettina Schaefli I applied the method to a set of hydrological problems from the Swiss Alps [69].

Chapter four deals with a time series analysis method originating from nonlinear dynamics. To study the coupling between ENSO and the Indian summer monsoon (ISM) rainfall, I estimated the oscillating phases of the underlying processes and studied their time dependent difference. To account for the typical variability of climate phenomena, I combined the phase reconstruction by means of Hilbert transformation with a recent approach of a geometry based filter. In order to study the performance of the method, I constructed a simple toy model and estimated the bias and variance of the phase reconstruction. For the teleconnection between ENSO and monsoon, I inferred two epochs of phase coherence that reproduce earlier findings with improved accuracy. To corroborate these results, I developed a significance test of high specificity. Furthermore, I detect intervals of coupling invisible to linear methods. These findings mark a considerable contribution to the ongoing discussion, whether the coupling is weakening due to global warming. Finally, I outline a possible volcanic influence on the ENSO monsoon coupling. This work has been published in [68].

In the last chapter, I summarize and discuss the results and achievements from a climatological and data analysis point of view and discuss limits and challenges for future research. In order to give consideration to the collaborative character of parts of the publications underlying this thesis, I decided to present the main part in plural form. Finally I would like to mention further work I have done with Prof. Dr. Stefan Rahmstorf and others [95, 96], Prof. Dr. Crowley [15] and Jan Saynisch [110], which has not entered this thesis.

Chapter 1

General Concepts

This first chapter aims to provide the conceptual background for the studies in the thesis at hand. Time series analysis investigates observational data to infer characteristic properties of the underlying processes. The underlying process is either assumed to be purely stochastic, a deterministic process driven by noise or a deterministic process, superimposed by measurement noise, the latter being again modeled as a (often white) stochastic process. Interesting characteristics can, for instance, be the correlation structure of the stochastic process, its marginal distribution, or the parameters of the underlying dynamical system. In this thesis, we will study the following properties: In chapter 2, we will investigate the correlation structure of the Prague daily temperature dynamics. Chapter 3 deals with the analysis of time dependent correlation structures, and in chapter 4, we infer the time dependent phase relation between ENSO and the Indian monsoon.

Stochastic processes are time dependent random variables with a certain dependency structure. In this context, inference means to construct an estimator, that maps the data onto an estimation of a particular process property. Because of a limited amount of data, the estimation is always subject to random variability. Thus, the quality of the estimate has to be quantified, either by means of confidence intervals or significance testing. The approach of inferring characteristic properties of an underlying process constitutes an inverse problem.

In the first section, we present the concept of random variables and their probability density distributions. Sec. 1.2 extends this concept to stochastic processes. The estimation of the properties of random variables is presented in Sec. 1.3, followed by an introduction to significance testing. In the last section, the concept of inverse problems and its relevance for time series analysis are outlined.

1.1 Random Variables

1.1.1 Random Variables

A *discrete random variable* X defined on the *probability space* $(\Omega, \mathcal{P})$ is a function that maps a *space of events* Ω to the real axis $X : \Omega \to \mathbb{R}$ by $\omega \to X(\omega)$, where $\omega \in \Omega$ is a particular *elementary event*. The *probabilities* $\mathcal{P}_X(\omega)$ map each ω to values between zero and one. A *realization* of X yields a particular element ω with a probability $\mathcal{P}_X(\omega)$. Repeating a realization N times one obtains the *relative frequency* $\hat{\mathcal{P}} = N(\omega)/N$ as the ratio between realizations of ω and the total number of realizations N.

A *continuous random variable* X is defined on a continuous space $\mathcal{R}$ with a corresponding *probability density function* $\rho_X(x)$ with $x \in \mathcal{R}$. The integrated probability density function is called *probability distribution*:

$$P_X(\lambda) = \int_{-\infty}^{\lambda} dx\, \rho_X(x) \tag{1.1}$$

A realization of X yields a value $x < \lambda$ with probability $\mathcal{P}\{x|x < \lambda\}$ according to the probability distribution:

$$\mathcal{P}\{x|x < \lambda\} = P_x(\lambda), \tag{1.2}$$

For a general introduction to random numbers and probability and more rigid definitions, please refer to [e.g. 41, 80].

A typical example for a continuous random number is one exhibiting a *Gaussian distribution* with density function

$$\rho_X(x) = \mathcal{N}(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{1.3}$$

1.1.2 Multivariate Random Variables

A *multivariate random variable* is a vector $\mathbf{X}$ of n random variables X_i with $i = 1...n$. The corresponding *joint probability distribution* is then a function defined on $\mathbb{R}^n$, $\rho_{\mathbf{X}} \equiv \rho_{\mathbf{X}}(x_1, ..., x_n)$ exhibiting values between zero and one. As $\rho_{\mathbf{X}}$ is a probability density, $\int d^n x\, \rho_{\mathbf{X}}(x_1, ..., x_n) = 1$.

Marginal Distribution

The probability density for a subset of variables in the vector irrespective of the values of the other variables is called *marginal distribution* and can be calculated by integrating the joint probability distribution over the other variables. The marginal distribution of X_1, e.g., is given as

$$\rho_{X_1}(x_1) = \int dx_2...dx_n\, \rho_{\mathbf{X}}(x_1, ..., x_n). \tag{1.4}$$

Conditional Probabilities

Given a two-dimensional random vector $\mathbf{X} = (X_1, X_2)$, then the *conditional probability density function* for X_1 given x_2 is written as $\rho(x_1|x_2)$. It is related to the joint probability density function by

$$\rho_{\mathbf{X}}(x_1, x_2) = \rho(x_1|x_2)\rho_{X_2}(x_2). \tag{1.5}$$

If $\rho(x_1|x_2) = \rho_{X_1}(x_1)$, i.e. $\rho_{\mathbf{X}}(x_1, x_2) = \rho_{X_1}(x_1)\rho_{X_2}(x_2)$, two random variables X_1 and X_2 are called *independent*.

1.1.3 Moments

The *expectation value* of a function $H(X)$ of the random variable X is given as

$$\langle H(X) \rangle = \int dx\, H(x)\, \rho(x) \tag{1.6}$$

Of particular interest is the expectation value of X itself:

$$\mu = \langle X \rangle = \int dx\, x\, \rho(x). \tag{1.7}$$

In general, one can define the *m-th moment* of X as

$$\langle X^m \rangle = \int dx\, x^m\, \rho(x), \tag{1.8}$$

with μ being the first moment of X. An important measure is the *variance*, being a combination of the first and second moment:

$$\mathrm{Var}(X) = \langle (X - \langle X \rangle)^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2. \tag{1.9}$$

For a Gaussian distribution Eq. (1.3), mean and variance are given by μ and σ^2, respectively. It is defined by the first and second moment only.

All these measures can also be extended to multivariate random variables $\mathbf{X}$ defined in $\mathbb{R}^n$. In particular, one can define the *covariance matrix*:

$$\mathrm{Cov}(X_i, X_j) \equiv \sigma_{ij}^2 \equiv \langle (X_i - \mu_i)(X_j - \mu_j) \rangle \tag{1.10}$$

1.1.4 Quantiles

Given a random variable X with probability density function $p(x)$, the *α-quantile* is defined as the value x_α for which

$$P(x_\alpha) = \int_{-\infty}^{x_\alpha} dx\, \rho(x) = \alpha. \tag{1.11}$$

For a Gaussian distribution Eq. (1.3), the quantiles defining 95% of the mass symmetric around the mean μ are given by $\approx \mu \pm 1.96\sigma$.

1.2 Stochastic Processes

A *stochastic process* $X(t)$ is a time dependent random variable. If time is discrete, then X_i is an indexed random variable with $i = 1...$ representing times $t_1....$ The process is defined by the joint probability density function $\rho(x_1, ...)$. The linear interrelations between two points in time t_1 and t_2 are captured by the *autocovariance function*

$$\mathrm{Cov}(t_1, t_2) = \langle (X_1 - \langle X_1 \rangle)(X_2 - \langle X_2 \rangle) \rangle \tag{1.12}$$

A *Markov process of order p* is defined as a (discrete) stochastic process fulfilling

$$\rho_n(x_n | x_1, ..., x_{n-1}) = \rho_p(x_n | x_{n-p}, ..., x_{n-1}), \tag{1.13}$$

i.e. the conditional probability density distribution of X_n given all previous values $x_1, ..., x_{n-1}$ depends on the last p values only. A first order Markov process is often simply referred to as Markov process. A Markov process is called *homogeneous* if the conditional probability density depends on time differences only, but not on absolute times.

1.2.1 Properties of Stochastic Processes

Stationarity

A stochastic process is *strictly stationary* if for any $t_i, t_j, ...$, the joint probability distribution $\rho(x_i, x_j, ...)$ stays invariant under time shifts k, i.e. $\rho(x_i, x_j, ...) = \rho(x_{i+k}, x_{j+k}, ...)$. Since the inference of such a property from real data is in principle impossible, one introduces the notion of weak stationarity, which refers to the moments of the process distribution. A process is *first order stationary*, if it exhibits a time independent mean. If a process additionally exhibits a time independent variance and if its covariance function is a function of time differences only, the process is called *second order stationary* or weakly stationary:

$$\begin{array}{rl} i. & \langle X_i \rangle \equiv \mu = const. \\ ii. & \mathrm{Var}(X_i) \equiv \sigma^2 = const. \\ iii. & \mathrm{Cov}(t_i, t_j) \text{ is a function of } (t_j - t_i) \text{ only.} \end{array} \tag{1.14}$$

Autocovariance Function and Spectrum

As stated in the previous section, a stationary process exhibits an autocovariance function of the form

$$\mathrm{Cov}(k) = \langle (X_{i+k} - \langle X \rangle)(X_i - \langle X \rangle) \rangle \tag{1.15}$$

Normalized to the variance (i.e. the autocovariance function at lag $k = 0$) one gets the *autocorrelation function*:

$$C(k) = \frac{\mathrm{Cov}(k)}{\mathrm{Cov}(0)} \tag{1.16}$$

As the frequency domain counterpart of the autocovariance function of a stationary process, one can define the *spectrum* as

$$S(\omega) = \widehat{\mathrm{Cov}}(k), \tag{1.17}$$

where the hat denotes the Fourier transformation.

Ergodicity

Ergodicity is a notion from statistical physics and dynamical systems theory. Loosely speaking, in time series analysis one calls a process *ergodic*, if the estimation of the mean and auto covariance function based on the time average converges to the estimation based on the ensemble average[1] for $N \rightarrow \infty$. As one often has access to only one realization of a process, this property is important for the estimation.

1.2.2 Gaussian Processes

A time discrete stationary process X_t is called *Gaussian* if the joint probability distribution of $\{X_{t_1}, ..., X_{t_N}\}$ for any n and any $t_1, ..., t_n$ is multivariate Gaussian [90]. This is equivalent to the condition that X_t can be modeled as Gaussian white noise ϵ_t, $\epsilon_t \sim \mathcal{N}(0, \sigma)$, filtered by a time independent linear filter $L(.)$ [91]:

$$X_t = \sum_{-\infty}^{\infty} L_i \epsilon_i \tag{1.18}$$

---

[1] Ensemble refers to a set of realizations of the process.

Gaussian processes are completely defined by their autocovariance function or equivalently by their spectrum.

ARMA processes

An important class of stationary Gaussian processes are causal *autoregressive moving average process* of order [p,q] (ARMA[p,q]-process) [9, 90],

$$X_t = \epsilon_t + \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=1}^{q} b_j \epsilon_{t-j}, \tag{1.19}$$

with AR coefficients a_i and MA coefficients b_j. With $b_j = 0$ for all j, the process is called AR[p] process, with $a_i = 0$ for all i it is called MA[p] process. Any Gaussian process Eq. (1.18) can in principle[2] be represented by an ARMA[p,q] process.

The ACF of general ARMA[p,q] models can be calculated using the Yule-Walker equations [9, 90] and decays exponentially for large time lags. For an AR[1] process $X_t = \epsilon_t + aX_{t-1}$ with $|a| < 1$, the ACF reads

$$C(r) = a^{|r|}, \qquad r = 0, \pm 1, \pm 2, ... \tag{1.20}$$

The spectrum of an ARMA[p,q] process is given as

$$|f_X(\omega)|^2 = \frac{\sigma^2}{2\pi} \frac{|\beta(e^{-i\omega})|^2}{|\alpha(e^{-i\omega})|^2} \quad \text{with} \quad \omega \in [-\pi, \pi], \tag{1.21}$$

where $\alpha(z)$ and $\beta(z)$ are the autoregressive and moving average polynomials defined as $\alpha(z) = 1 - a_1 z - ... - a_p z^p$ and $\beta(z) = 1 + b_1 z + ... + b_q z^q$ [9, 90].

Fractal Processes

A special class of Gaussian processes are fractal processes. These cannot be modeled by ARMA[p,q] processes of finite order. *Fractional Brownian motion* (FBM) is defined by the following autocovariance function:

$$\mathrm{Cov}(t_1, t_2) \equiv \langle X(t_1) X(t_2) \rangle = \frac{\sigma^2}{2} \left(|t_1|^{2H} + |t_2|^{2H} - |t_2 - t_1|^{2H} \right) \tag{1.22}$$

The exponent H is called *Hurst parameter*. For $H = 1/2$, one gets the classical (non-fractal) Brownian motion. FBM is called self similar, because a scaling of the time axis can be compensated by a scaling of the amplitude. The process defined by the increments of FBM

$$Y(t) = X(t+1) - X(t) \tag{1.23}$$

is called *fractional Gaussian noise* (FGN). It is a stationary Gaussian process with an algebraically decaying ACF

$$\mathrm{Cov}(r) \equiv \langle X(t) X(t+r) \rangle = \frac{\sigma^2}{2} \left(|r+1|^{2H} + |j-1|^{2H} - 2|j|^{2H} \right), \qquad r = 0, \pm 1, \pm 2, ... \tag{1.24}$$

---

[2] in general of possibly infinite order

The power spectrum of FGN reads

$$S(\omega) = \frac{(2\pi)^\alpha \sigma^2}{2|\omega|^\alpha} \tag{1.25}$$

with $\alpha = 2H + 1$. More general linear processes consist of ARMA components, which are additionally integrated (ARIMA) with a possibly fractional integration (FARIMA or ARFIMA) [5].

Realizations of Gaussian Processes

Realizations of ARMA[p,q] processes can be simulated in time domain by applying Eq. (1.19) to a realization of a white noise time series ϵ_t. This procedure, however, does not apply to e.g. fractal processes, where no straight forward model in the time domain exists. An alternative solution applicable to any Gaussian process makes use of the spectrum: One might transform the white noise time series ϵ_t to the Fourier domain, multiply it with the desired spectrum and transform it back to the time domain [e.g. 119].

1.3 Estimators

Given a random variable X, one can utilize a number N of realizations of X to reconstruct the parameters θ defining the underlying probability density function $\rho_X(x)$ such as the moments or even the density itself. This procedure is referred to as estimation. The *estimator* of the parameter θ is denoted by a hat, $\hat{\Theta}$, and is a function of the N copies of the random variable $\hat{\Theta} \equiv \hat{\Theta}(X_1, ...X_N)$. Hence, the estimator is a random variable itself with a particular realization $\hat{\theta}(x_1, ..., x_N)$ [41].

1.3.1 Bias

The *bias* of an estimator $\hat{\Theta}$ is defined as the difference between the expectation value of the estimator and the underlying parameter,

$$\mathrm{Bias}(\hat{\Theta}) = \langle \hat{\Theta} \rangle - \theta. \tag{1.26}$$

An estimator is said to be *unbiased*, if $\mathrm{Bias}(\hat{\Theta}) = 0$.

1.3.2 Variance

According to Eq. (1.9), the variance of an estimator $\hat{\Theta}$ is given as

$$\mathrm{Var}(\hat{\Theta}) = \langle (\hat{\Theta} - \langle \hat{\Theta} \rangle)^2 \rangle = \langle \hat{\Theta}^2 \rangle - \langle \hat{\Theta} \rangle^2. \tag{1.27}$$

1.3.3 Consistency

An estimator is said to be *consistent* if the probability that the difference between a realization $\hat{\theta}(x_1, ..., x_N)$ and the underlying true value θ is larger than an ϵ tends to zero for large N:

$$\mathcal{P}\left(|\hat{\theta}(x_1, ..., x_N) - \theta)| > \epsilon \right) \to 0 \qquad \text{for} \qquad N \to \infty. \tag{1.28}$$

This requires asymptotically vanishing bias and variance for $N \to \infty$.

1.4 Significance Testing

Given an observation x, one is often interested if the measured value (or more general, the measured sample) is compatible with some trivial assumption or if it shows some interesting behavior. Translated to statistics, the question is if x is a realization of a (trivial) random variable X with a specific distribution or if it differs. If the support of X ranges from $-\infty$ to ∞ (e.g. if X is Gaussian distributed), any x can be a realization of X and it is in principle impossible to answer this question. However, one can ask how *likely* it is that x is a realization of X. This approach leads to the concept of significance testing.

One assumes a certain (trivial) *null hypothesis* H_0 about X (e.g. X is Gaussian) and defines a *significance level* α, usually[3] 0.1, 0.05 or 0.01. If x exceeds the $1-\alpha$ quantile of the distribution under the Null hypothesis[4], one considers the random variable underlying x as being significantly different from the Null hypothesis on the α-level.

1.4.1 Statistical Errors and Power of the Test

Even if the null hypothesis H_0 is right, one observes on average α percent rejections of H_0. This result is called *false positive* or an α-error. This error is in principle not evitable and occurs by construction. On the other hand, a significance test might fail to reject H_0, even if it was wrong. This result is called *false negative* or a β-error. The probability, that a test rejects H_0 if it is wrong, i.e. the probability not to produce a β-error is called power of the test.

1.4.2 Sensitivity vs. Specificity

Accounting for its purpose, every significance test has to be designed individually: A medical test should always warn the patient in case of a severe disease; on the other hand, establishing a result by means of statistics, a test should preferably reject false positive results. These antithetic demands are represented by the terms of *sensitivity SE* and *specificity SP*:

$$\begin{aligned} SE &= \mathcal{P}(H_0 \text{ rejected} \,|\, H_0 \text{ wrong}) \\ SP &= \mathcal{P}(H_0 \text{ accepted} \,|\, H_0 \text{ right}) \end{aligned} \tag{1.29}$$

A sensitive test rejects H_0 in preferably every case it is wrong (low β-error), whereas a specific test preferably only rejects H_0 when it is definitely wrong (low α-error). For finite data, no test can be perfectly sensitive and specific simultaneously.

Assume a population N, where a null hypothesis H_0 is right in N_R cases and wrong in $N_W = N - N_R$ cases. Applying a significance test for H_0 to measurements of every element, the numbers of true negative and false positive results are denoted as N_{TN} and N_{FP}, with $N_{TN} + N_{FP} = N_R$. The numbers of false negative and true positive results are given as N_{FN} and N_{TP}, with $N_{FN} + N_{TP} = N_W$. Then one can estimate sensitivity and specificity as

$$\begin{aligned} \widehat{SE} &= \frac{N_{TP}}{N_W} = \frac{N_{TP}}{N_{TP} + N_{FN}}, \\ \widehat{SP} &= \frac{N_{TN}}{N_R} = \frac{N_{TN}}{N_{FP} + N_{TN}}. \end{aligned} \tag{1.30}$$

---

[3] In geosciences, one often calls $1-\alpha$ the significance level, e.g. the 90%, 95% or 99% level.

[4] In case of a two sided test, one considers the $\alpha/2$ and the $1-\alpha/2$ quantiles.

The sensitivity relates the number N_{TP} of true rejections of H_0 to the total number of wrong H_0, N_W. On the other hand, the specificity measures the number N_{TN} of true acceptances of H_0 in relation to the total number of right H_0, N_R. The notions of sensitivity and specificity are often used more general without referring to a quantitative analysis.

1.4.3 Multiple Testing

Assume a test being independently repeated on the α level. Even if H_0 is true, it will by definition be rejected in on average α percent of the realizations. The probability, that H_0 will at least once be falsely rejected after N realizations is $1 - (1 - \alpha)^N$, which is essentially one for large N. This effect is called multiple testing and is inherent to every repeatedly realized statistical test. A conservative approach to handle this problem is the Bonferroni-correction: One divides the desired over all significance level $\tilde{\alpha}$ by the number of realized tests N to get a corrected significance level $\alpha_B = 1 - (1 - \tilde{\alpha})^{1/N} \approx \tilde{\alpha}/N$ for the single tests. A very pointed example that illustrates multiple testing and also its consequences for scientific progress is the problem of publication bias: Assume 100 researchers performing the same experiment with the same test without knowing of each other. Hence, even if H_0 were right, on average α-percent of these researchers will obtain a false positive result. These results - but probably not the 95 others - might be published causing a bias of scientific discovery [e.g. 20].

1.5 Inverse Problems

Given a model M with a set of parameters P (including possible input, boundary conditions and initial values), the generation of data D from this model is called *direct problem* [76]:

$$M(P) \Longrightarrow D \tag{1.31}$$

Examples are any numerical simulations, e.g. using a simple dynamical system as the Lorenz equations, a model for a biochemical reaction or a general ocean atmosphere circulation model, or experiments where the underlying dynamics and the parameters are well known.

However, in many scientific problems, one faces the opposite problem: One measures a set of data and aims to infer the parameters of the underlying model or even the model itself. This task is referred to as an *inverse problem*[41, 76]:

$$M(P) \Longleftarrow D \tag{1.32}$$

The task of inferring the parameters, given the model, is called inverse problem of first kind. The inference of a model, given a set of parameters, is called inverse problem of second kind. An inverse problem is called *well posed* [35] if the solution

- exists,
- is unique,
- depends continuously on the data.

Usually, inverse problems are *ill-posed* and need to be regularized by additional constraints. Examples for inverse problems of first kind are inverse linear transformations, determination of physical constants by a chosen experimental setting, or tomography [76]. Inverse problems of second kind are e.g. the inference of a physical law from observations, significance testing and directly related, the verification and falsification of hypothesis [76]. In this respect, all the time series analyses presented in this thesis constitute inverse problems.

Chapter 2

Tempting Long Memory

2.1 Introduction

Understanding the variability of the climate system on scales from days to decades or even longer is a fundamental issue of climatology. In 1976, Hasselmann [37] introduced the concept of stochastic climate models and showed that climate variability in principle can be modeled by rather simple AR-processes. The time dependency of linear stochastic processes is completely captured by the autocorrelation function $C(s)$, describing the average linear relation of two points in time with lag s. The sum of $C(s)$ over all lags s is often called memory or persistence of the process [124]. Typical properties of AR-processes are an exponential decay of the autocorrelation function for large s and thus a finite memory which is equal to the existence of a characteristic time scale. Processes with these attributes are called short-range correlated or short-memory processes.

Later, Pelletier and Turcotte [81] as well as Koscielny-Bunde et al. [55] reported long-range correlations in temperature data. This process property is characterized by infinite memory and the absence of a typical time scale. For large time lags r, the autocorrelation function of such long-memory processes decays according to a power-law and hence exhibits scaling with a characteristic exponent γ.

Govindan et al. [31] reported, that widely used general coupled climate models do not exhibit the behavior found by Koscielny-Bunde et al. [55] and raised the question whether estimates of the warming-trend in global mean temperature based on these models could be exaggerated. However, Fraedrich and Blender [22] extended the analysis to various stations around the globe and showed that state-of-the-art climate models are able to reproduce the empirical findings.

Except for Pelletier and Turcotte [81], the authors applied Detrended Fluctuation Analysis (DFA, Peng et al. [82]) which investigates the average variability on a scale[1] s. For long-memory processes in the limit of large scales, this fluctuation function $F(s)$ can directly be linked to the autocorrelation function. However, in all works known to the authors the properties of DFA have only been studied for certain long-memory processes. The behavior for other linear processes has not been investigated and it is not clear if the results can be easily transferred.

In this chapter we interpret the inference of long-memory from empirical data as an inverse problem. Following Popper [88], we test competing hypothesis, long memory but also short

---

[1] scale refers to the length of a time interval

memory, and try to falsify them. The main insight is, that power-law scaling of the fluctuation function and thus long-memory may not be assumed *a priori* but have to be established. We argue that this requires the investigation of the local slopes of the fluctuation function. Furthermore one has to account for the variability characteristic for stochastic processes. Associated with the inference of long-range correlations, we investigate the sensitivity and specificity of DFA: When long-memory is present, does DFA detect it? When no long-memory is present, does DFA reject it? We illustrate our discussion by comparing a long-memory with a short-memory model. Finally, we review the results of Koscielny-Bunde et al. [55] for the Prague temperature data set and show that their conclusions are not unambiguous.

In Sect. 2.2, we introduce the processes considered and discuss the inference of long-range correlations. The method of DFA is explained in Sect. 2.3. In Sect. 2.4, we interprete the inference of long-range correlations from empirical data as an inverse problem. To exemplify the discussion, we present two competing models. Establishing power-law scaling of the fluctuation function is discussed with respect to local slopes and natural variability. Section 2.5 gives an overview over typical pitfalls for the case that the unknown underlying process is short range correlated. As a consequence of the foregoing discussions, we review the DFA results of the Prague daily temperature record in the last section.

2.2 Autocorrelation Function of Stochastic Processes

Many processes in nature are of such a high complexity that a description by deterministic models is difficult or not desirable. However, their characteristic behavior often is effectively captured by stochastic processes. In 1976, Hasselmann [37] introduced the concept of stochastic climate models and showed that the variability of the climate system in principle can be modeled by rather simple and linear AR-processes (For an overview over Gaussian and ARMA processes refer to Sec. 1.2.2). The time dependent structure of stationary linear stochastic processes is captured by the autocorrelation function $C(s)$ of two points in time with lag s or equivalently by the spectrum $S(\omega)$. The integral of $C(s)$ over all lags s is often called memory or persistence [124]. Different classes of autocorrelation structures can be distinguished with respect to the form of their decay for large time lags s:

2.2.1 Short Range Correlations

Many stochastic processes in nature exhibit *short-range correlations*, which decay exponentially:

$$C(s) \propto e^{-s/\tau}, \, \text{for} \, s \to \infty. \tag{2.1}$$

These processes exhibit a typical time scale τ. This decay is fast enough to let the sum $\sum C(s)$ converge resulting in a finite *decorrelation time* [128], e.g. for $C(s) = e^{-s/\tau}$

$$\begin{aligned} \tau_D &= 1 + 2\sum_{s=1}^{\infty} C(s) = 1 + 2\sum_{s=1}^{\infty} e^{-s/\tau} \\ &\approx 2\tau \, \text{for} \, \tau \gg 1. \end{aligned} \tag{2.2}$$

We employ the decorrelation time as a measure for the memory or persistence of a process. Correlations on scales large compared to the decorrelation time are negligible due to the fast exponential decay. Thus, one also refers to these processes as having short-range or finite memory. Typical examples are AR-processes (see Sec. 1.2.2).

2.2.2 Long Range Correlations

A second class are *long-range correlated* processes characterized by algebraically decaying correlations:

$$C(s) \propto s^{-\gamma}, \text{ for } s \to \infty, \quad \text{with} \quad 0 < \gamma < 1. \tag{2.3}$$

A characteristic time scale as defined above does not exist. The calculation of the decorrelation time, e.g. for $C(s) = s^{-\gamma}$ for $s > 0$, $C(0) = 1$

$$\tau_D = 1 + 2\sum_{s=1}^{\infty} C(s) = 1 + 2\sum_{s=1}^{\infty} s^{-\gamma} = \infty \tag{2.4}$$

results in infinite memory.

A physical example for such a process can be found in Kolmogorov's theory of turbulence [23]. A mathematical model exhibiting long-range correlation is e.g. fractional Gaussian noise (FGN) [66] (see Sec. 1.2). For this process the asymptotic relation Eq. 2.3 is already well satisfied for finite sample sizes because it is the increment of the self-similar process fractional Brownian motion. A more flexible long-range correlated process is given by fractional ARIMA$[p, d, q]$ (FARIMA) that contains additional short range correlated ARMA components [5, 32, 43]. However, for non-trivial autoregressive (AR) or moving average (MA) components ($p + q > 0$) it is possible that the asymptotic relation Eq. 2.3 is not satisfied for a finite sample.

The concept of long-memory refers to non-periodic processes. Thus, the recurrence due to periodicities like the Milankovic-cycles in the climate system are not to be considered as long-range correlations, even if their (deterministic) behavior causes correlations for infinite time lags.

2.2.3 Discrimination between Short and Long Memory

An important question to characterize a given non-periodic process is now to investigate, if its autocorrelation decays exponentially or according to a power law. However, for observational data, one cannot investigate the decay of the autocorrelation function on arbitrarily large scales. Measurement limitations always restrict the analysis and interpretation to a range between two time scales independent of the nature of the process: First, the sampling interval Δt defines a minimum time scale. Second, the finite length of the record T defines a maximum time scale. Intuitively, one might easily be tempted to refer to records with characteristic time scales close to Δt as short-range and close to T as long-range correlated. This, however, ignores the fundamental difference of short-range and long-range correlated processes concerning the form of the decay of their autocorrelation as given in Eqs. (2.1) and (2.3). This difference is essential e.g. for estimation of the variance of the mean, prediction [5], trend assessment [51, 29] and extreme value statistics [21].

2.3 Detrended Fluctuation Analysis (DFA)

Estimating the autocorrelation function $C(s)$ from empirical data is limited to rather small time lags s and is affected by observational noise and instationarities like trends. The same restrictions hold for the sample spectrum at low frequencies. Peng et al. [83] suggested *Detrended Fluctuation Analysis* (DFA) to indirectly gain information about the correlation structure imposed on a time series.

2.3.1 The Algorithm

The method works as follows: Given a realization[2] $x(i)$ of length N of a stochastic process, one calculates the aggregated time series or profile $y(j)$:

$$y(j) = \sum_{i=1}^{j}[x(i) - \bar{x}], \quad \text{with} \quad \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x(i) \tag{2.5}$$

The profile $y(i)$ is divided into M non-overlapping segments of length s. For DFAn, in each segment m a best fit polynomial trend $p^n_{s,m}$ of order n is subtracted from the profile:

$$Y_{s,m}(i) = y(i) - p^n_{s,m}(i) \tag{2.6}$$

For each segment m the squared fluctuation is calculated:

$$F^2_m(s) = \frac{1}{s}\sum_{i=(m-1)s+1}^{ms} Y_{s,m}(i)^2 \tag{2.7}$$

The squared fluctuation function of the process is estimated by averaging over all segments:

$$F^2(s) = \frac{1}{M}\sum_{m=1}^{M} F^2_m(s) \tag{2.8}$$

This procedure is repeated for several s. The minimum scale is given by the sampling interval Δt and the order n of DFA, a reasonable choice for the maximum scale is about 1/10 of the total record length, due to natural variability as will be discussed in Sect. 2.4.2. For a more comprehensive description of the algorithm and the influence of trends refer to Kantelhardt et al. [52].

For an uncorrelated series $x(i)$ we get a squared fluctuation function $F^2(s) \propto s$ (i.e. $F(s) \propto s^{0.5}$) which reflects the linear increase of the variance of the profile. For time series $x(i)$ with algebraically decaying autocorrelations (Eq. (2.3)), it can be shown that in the limit of large s the fluctuation function increases according to a power law [117]

$$F(s) \propto s^{\alpha}\,, \quad \alpha = 1 - \gamma/2\,, \quad \text{with} \quad 0.5 < \alpha < 1\,. \tag{2.9}$$

Thus, analyzing the fluctuation function of long-range correlated processes for large s reveals the decay exponent of the corresponding correlation function for large s.

2.3.2 Estimating the Strength of Long Range Correlations

For FGN, one can estimate the scaling exponent α of the fluctuation function by fitting a straight line to the result in a double logarithmic plot (for FARIMA processes, the fit has to be restricted to large scales where short memory components can be neglected). Then, according to Eq. (2.9), one can easily derive the power-law decay of the autocorrelation function with exponent γ or equivalently the Hurst coefficient H (see Sec. 1.2.2) with $\gamma = 2 - 2H$. Taqqu et al. [117] have shown that this procedure is an effective estimator of the scaling exponent. The ability of DFA to estimate α from realizations of long-memory processes corrupted by different instationarities like trends, spikes, harmonics and gaps has been studied by Hu et al. [44], Chen et al. [13] and Kantelhardt et al. [52].

---

[2] For clarity reasons during the following calculations, here we write the discrete time index in brackets, not as subscript.

2.4 Inference of Long Memory

The straight line fit discussed in the previous section was often used to estimate scaling exponents from observational data [e.g. 31, 55, 22]. However, this procedure is valid only in the case, that a scaling parameter exists at all, i.e. only when the underlying process in fact is long range correlated. But in general, the correlation structure of a complex real world process is not known *a priori* and first has to be identified. Thus, one faces an inverse problem as introduced in Sec. 1.5.

Given the set of all possible correlation structures, the inference of long-range correlations means

1. showing the compatibility of the given data set with a realization of a long-memory process,
2. excluding other possible correlation structures.

These two conditions - the necessary and the sufficient - are often discussed with respect to the terms of sensitivity and specificity (see Sec. 1.4): In this case, a procedure, that with a high probability detects compatibility with long-range correlations, whensoever they are present, would be sensitive. An algorithm that with a high probability rejects long-range correlations, when they are not present, were specific. The optimal algorithm would be sensitive and specific. A sensitive but unspecific algorithm, however, would frequently detect long-range correlations. This algorithm would not be suitable for a reliable inference. On the other hand, an un-sensitive but specific algorithm would be very conservative and would often reject the existence of long-range correlations.

To our knowledge it has not been studied, if DFA can be used to infer long-memory from realizations of a process when it is not *a priori* clear, if this process is long-range correlated. I.e. it is still an open question, how sensitive and specific DFA behaves when investigating processes of unknown correlation structure for long memory. Closely connected to this discussion is the question, how DFA behaves analyzing short memory processes.

To summarize, for the inference of long memory we mainly address the following questions:

1. How to conclude scaling from the DFA fluctuation function?
2. Does scaling necessarily imply long-range correlations?

2.4.1 Two Example Processes

To illustrate our line of argumentation and to investigate the questions listed above, we consider a short-memory as well as a long-memory process and apply DFA to both. The choice of the model parameters is motivated by the example of the Prague temperature record studied in Sect. 2.6 and will become clear during the discussion. As an example for a long-range correlated process we simulate fractional Gaussian noise according to the method given in Timmer and König [119] with $\alpha = 0.6$ (see also Sec. 1.2.2 for an outline of the method). This process shows power-law scaling in the autocorrelation function for a wide range of scales. For the short-range correlated process we choose a superposition of three AR[1]-processes,

$$x(i) = \sum_{j=1}^{3} A_j y_j(i)\,, \quad y_j(i) = a_j y_j(i-1) + \eta_j(i)\,, \tag{2.10}$$

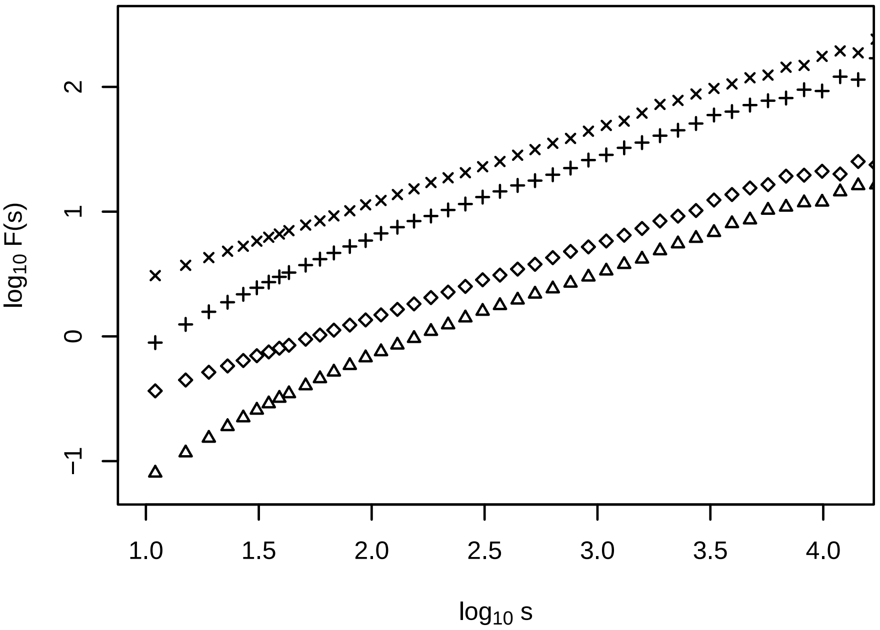

Figure 2.1: Fluctuation functions calculated for an artificial long-range correlated process with exponent $\alpha = 0.6$ (× DFA1, ◇ DFA2) and a superposition of three AR-processes (+ DFA1, △ DFA2) as defined in Sect. 2.4.1. For each order of magnitude, approx. 50 values are calculated. To enhance clarity, only every third value is plotted.

with $\eta_j(i)$ being Gaussian white noise of zero mean and variance $1 - a_j^2$. The latter insures $\mathrm{var}(y_j) = 1$. We choose $A_1 = 0.913$, $A_2 = 0.396$, $A_3 = 0.098$, $a_1 = 0.717$, $a_2 = 0.953$ and $a_3 = 0.998$. Using $a_j = e^{-1/\tau_j}$ we find the following characteristic time scales for the individual AR[1] processes: $\tau_1 = 3$d, $\tau_2 = 21$d and $\tau_3 \approx 1.5$ yrs.

2.4.2 Establish Scaling of the Fluctuation Function

Figure 2.1 shows the fluctuation functions for a realization of each of the two example processes defined in the previous section with $N = 70,492$ and $\Delta t = 1$. For each order of magnitude 50 values of equal distance in logarithmic scale are calculated. For clarity reasons, we plotted only every third value.

The Straight Line Fit

Investigating only the double logarithmic plot of the fluctuation function, one is tempted to rashly conclude for long-range correlations. Due to properties of the logarithm, fluctuations are suppressed in a log-log plot and the deviation from a straight line is not easily visible [125]. Also, restricting the analysis to a straight line in the log-log plot forces $F(s)$ in the *procrustean bed* of power-laws. It will always yield some value for the slope but the suitability of the linear description is not evaluated. For the inference of long-range correlations, this procedure would be sensitive but not specific in the sense, that long range correlations would be attributed to all processes with $\alpha \neq 0.5$ for the largest scale observed. Such a result would trivialize the concept of long-range correlations and provide no insight into the process. Even if scaling is present, it is difficult to determine the beginning and ending of the scaling region in the log-log plot. However, the resulting value for α derived from a straight line fit strongly depends on the fit boundaries if the realization does not stem from a scale free process.

Estimating Local Slopes

Thus, to reliably infer a power-law, a straight line in the log-log plot may not be assumed *a priori* but has to be established. Since a straight line is tantamount to a constant slope, the *local slopes* $\alpha(s)$ of $\log_{10} F(s)$ vs $\log_{10} s$ have to be evaluated for constancy in a sufficient range [53, 125, 120]. The extend of a sufficient range is still a matter of debate (see e.g. Avnir et al. [3] and references therein). This concept has been introduced in the context of estimating correlation dimensions [11, 125] and, in a different setting, has also been suggested for DFA [82].

For a finite amount of data the estimation of the local slopes brings along a certain variability and even for a long-memory process like FGN, the local slopes of the empirical fluctuation function show variations around a constant α. This has two consequences for the calculation and interpretation of the local slopes: First, estimating the local slopes by finite differences results in a large variability. This can be reduced fitting a straight line to $\log(F(s))$ vs. $\log(s)$ within a small window. The window is then shifted successively over all calculated scales s. Figure 2.2 shows the local slopes of a realization of the short-memory model for different window sizes. Choosing the optimal window size, one has to trade bias for variance: For small windows, the bias is small, but the variability renders the interpretation difficult, whereas for large windows, the variance is reduced at the cost of a biased estimate of α. Thus, the extreme case of a single straight line fit to the whole range of scales considered is maximally biased. Since only one value of α is calculated, this does not allow to evaluate constancy.

As a second consequence of the finite amount of data, one has to quantify the variability for a given length of the record. Since vicinal local slopes are not independent, confidence regions cannot be estimated easily from the procedure described in Sect. 2.4.2 [18]. Instead, we perform Monte Carlo simulations: For the two example processes, we simulate 1,000 realizations to estimate mean and standard deviation of α for the scales considered. For a fixed scale s, the distribution of $\alpha(s)$ is approximately gaussian. Thus, we employ $\bar{\alpha}(s) \pm 1.96\sigma(s)$ as estimates of the 95% confidence bands.

Figure 2.3 displays the local slopes of the DFA1 (a) and DFA2 (b) fluctuation functions, estimated from one realization of each of the example models using a window of 21 points. Additionally, the corresponding 1.96σ intervals around the mean of each model are plotted. The realization of the long-memory process shows fluctuations around a constant α within the corresponding 1.96σ interval, increasing like $\sigma \propto \sqrt{s}$ [82]. The local slope $\alpha(s)$ of the short-memory realization, however, decreases constantly in the beginning and basically follows the local slope of the long-memory realization for scales larger than $\log_{10} s \approx 2.5$. Thus, for a certain choice of parameters, a short-range correlated model can mimic scaling in a finite range. Due to the principle of variance superposition for DFA [44], a suitable superposition of three AR[1] processes produces this effect in the fluctuation function analogously to the same effect in the spectral domain described in Hausdorff and Peng [39].

Analyzing the long-memory properties one studies primarily the behavior on large scales s assuming that influences from short-range components are negligible and do not bias the

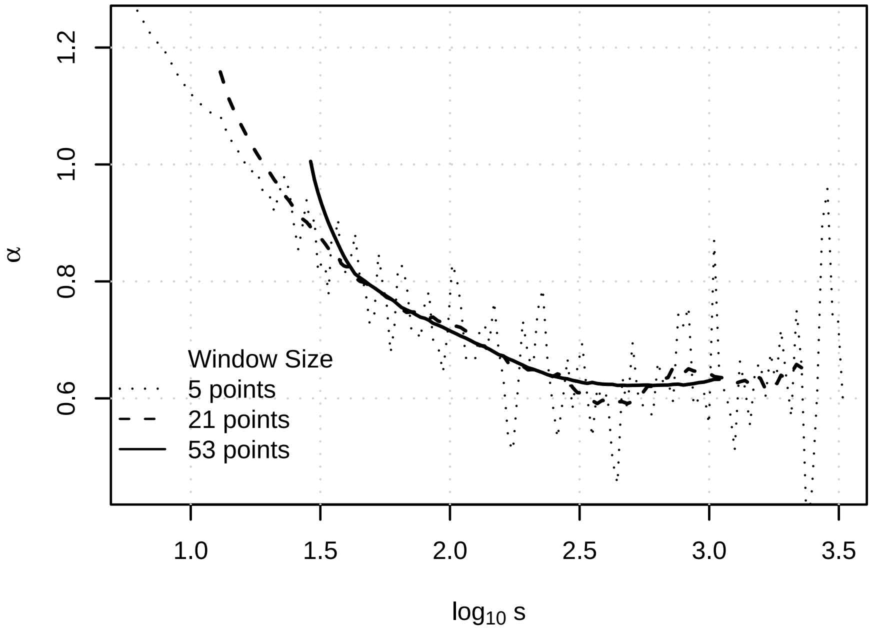

Figure 2.2: Local slopes of a realization of the short-memory model for different window sizes. For F(s), approx. 50 points per order of magnitude are calculated. For small windows, the bias is very low, but the variability renders the interpretation difficult, whereas for large windows, the variance is reduced at the cost of a biased estimate of α.

estimation of the long-range dependence parameter. In our example, the 1.96σ-cones are virtually indistinguishable in this range. Thus, based on the given record length and only considering large s, one cannot distinguish the realizations of the two models by means of DFA. For longer time series, the cones would shrink and the region of overlapping would become smaller.

At this point, a general dilemma related to the inference of long-memory emerges: For a finite time series, one will always find a short-range correlated model to describe the data [5]. Thus, the inference of long memory from finite data is in principle impossible, or in other words: Considering the inference of long range correlations, DFA is sensitive but not specific. The capability of short-memory models to reproduce findings which are associated with long-memory has also been considered with respect to the Hurst phenomenon [73, 74].

However, it is not always meaningful to model finite data with a short-memory model: the longer the scaling region of a short-range model shall be, the more parameters and tuning are required. It may be advantageous to describe a data set exhibiting several orders of magnitude of power-law scaling with a long-range correlated model with few parameters rather than with a short-range correlated model with a large number of parameters. To decide which model to prefer with respect to parameter parsimony, one could e.g. employ a likelihood approach combined with an Akaike-type model selection criterion [6]. A framework for this model based approach is under development [107].

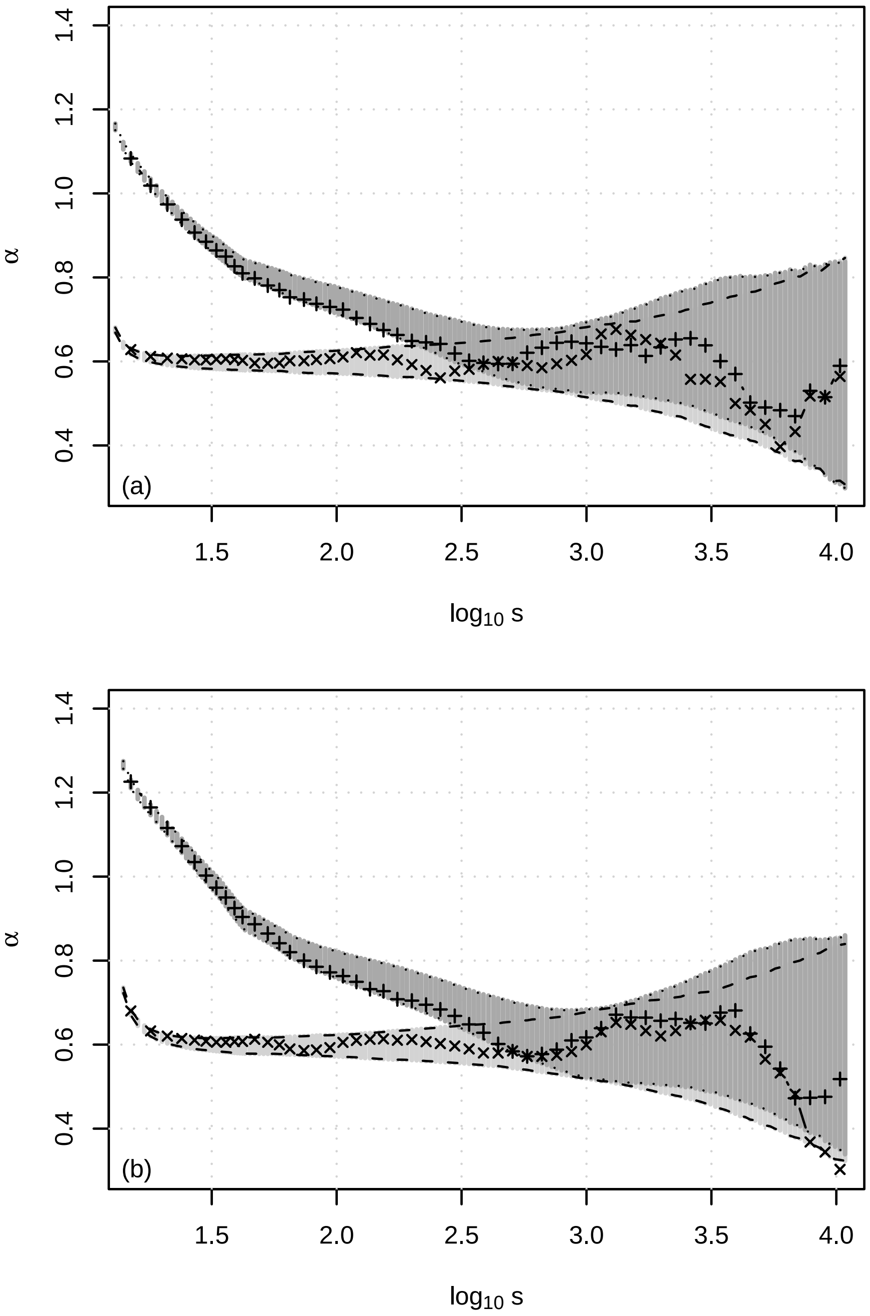

Figure 2.3: Local slopes of the (a) DFA1 and (b) DFA2 fluctuation function calculated for an artificial long-range correlated process with exponent $\alpha = 0.6$ (×) and a superposition of three AR-processes (+) as defined in Sect. 2.4.1. The dashed and the dotted lines border the shadowed 1.96σ intervals obtained from 1,000 realizations of the two processes, respectively.

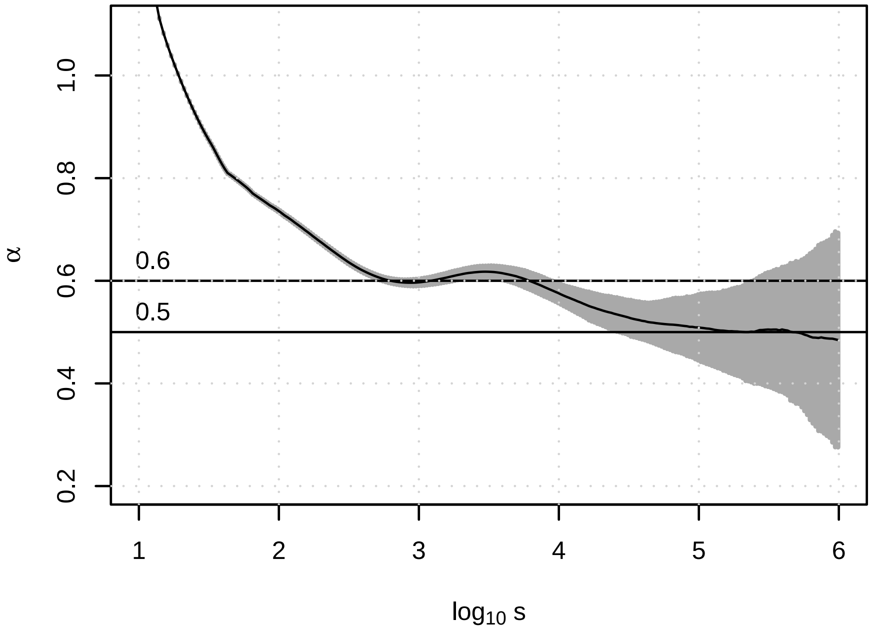

Figure 2.4: Empirical fluctuation function of the short-memory model Eq. (2.10), estimated from 200 realizations of length N=1,000,000 (solid line). A region of approximatively constant slope occurs between $\log_{10} s \approx 2.8$ ($s \approx 600$) and $\log_{10} s \approx 3.8$ ($s \approx 6,000$, $\simeq 16$ years). On larger scales, the slope reduces to $\alpha = 0.5$ characterizing the short-memory nature of the model.

2.5 Inference of Short Memory

To circumvent the principle problem outlined in the previous section, one can alternatively investigate if the underlying process is short-range correlated. Transferring the discussion from above, this requires:

1. To show compatibility with a short-range correlated model.

2. To exclude possible long-range correlated models.

The first condition is always fulfilled, since one will always find a short-range correlated model to describe a finite data set. Thus, for the inference of short-memory, DFA is sensitive. The second condition is not fulfilled for the given example, because the record length is not sufficient to detect the short-memory character $\alpha = 0.5$ for large s of the AR-model by means of DFA. For longer time series as shown in Fig. 2.4, when a plateau of $\alpha = 0.5$ is identifiable, long-memory can be excluded and the specificity of DFA to infer short-range correlations increases.

However, setting aside the realizations of the two example models being two short for the discrimination between long and short memory, one might investigate the performance of DFA with respect to analyzing short memory processes in a more general context.

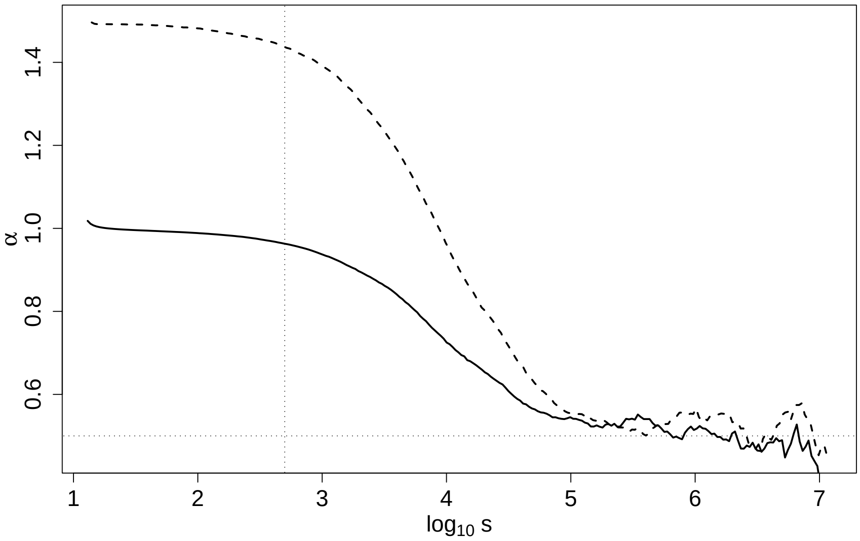

Figure 2.5: Local Slopes of the empirical fluctuation function of an AR[1]-process ($\tau = 500$) estimated from one realization of length N=10,000,000. Solid line: DFA1, dashed line: DFA2. The horizontal dotted line marks $\alpha = 0.5$, the vertical dotted line marks $\tau = 500$. For low scales, α exhibits values larger than one. Only for scales approx. 1.8 orders of magnitude larger than τ, α decrease to 0.5.

2.5.1 The Fluctuation Function of AR[1]-Processes

According to Sect. 2.2, the autocorrelations of short memory processes decay exponentially for large s and are negligible on scales large compared to the decorrelation time τ_D (Eq. (2.2)). Consequently, for large scales, the slope of the fluctuation function of such a process converges to $\alpha = 0.5$. Figure 2.5 shows the local slopes of the empirical fluctuation function calculated with DFA1 (solid line) and DFA2 (dashed line) for a realization (N=10,000,000) of an AR[1] process with $\tau = 500$.

The expected convergence to $\alpha = 0.5$ occurs only for $\log s \approx 4.5$, i.e. $s \approx 32,000$, which is about 1.8 orders of magnitude larger than τ. Consequently, one might be tempted to conclude a much longer characteristic time scale of a short memory process when investigating DFA fluctuation functions.

2.5.2 Finite Scaling of Short-Memory Processes

As stated in the previous section, one expects the slope of the fluctuation function of a short memory process to converge to $\alpha = 0.5$ for scales large enough. However, for a finite set of data one cannot be *a priori* sure that the series is long enough to observe this. For a record of the short-memory model defined in Sect. 2.4.1 of length 70,492 points the local slopes of the fluctuation function of the largest observed scales is compatible with power-law scaling. A plateau with $\alpha = 0.5$ is not observed (Fig. 2.3). Thus, one might be tempted to conclude long-memory. However, analyzing a much longer record (1,000,000 points) of the same model yields a plateau with $\alpha = 0.5$ for large s as can be seen from Fig. 2.4. Therefore, for a

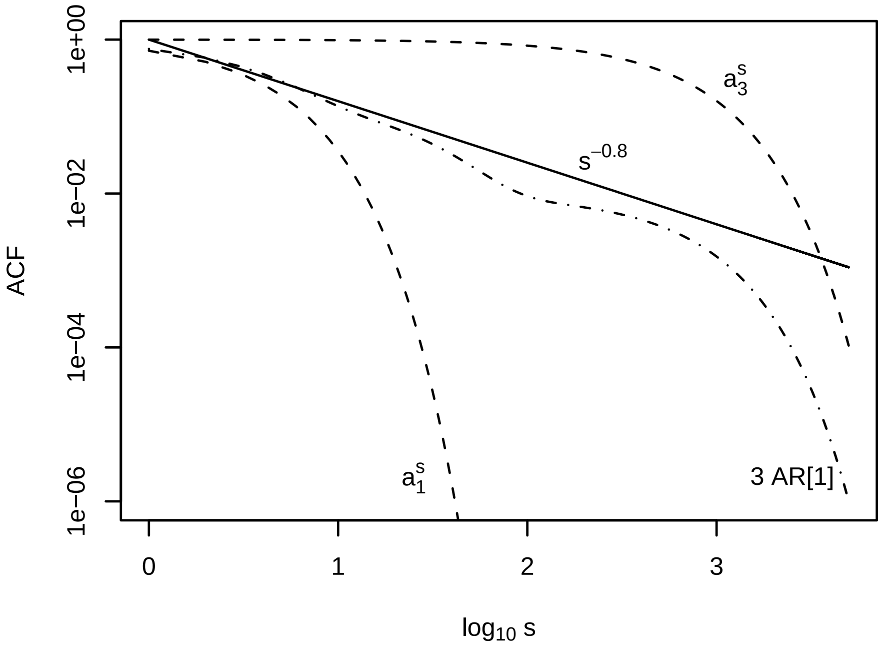

Figure 2.6: Analytical autocorrelation function of the short-memory model Eq. (2.10) (dashed dotted line). The dashed lines depict the autocorrelation functions of the single AR[1] processes with shortest and longest time constant respectively. In a scale range from $\log_{10} s \approx 0$ ($s \approx 1$) to maximally $\log_{10} s \approx 3$ ($s \approx 1,000$) the autocorrelation function approximately follows the power law with $\gamma = 0.8$ (solid line). For larger scales, it turns into an exponential decay determined by the AR[1] process with the largest time constant $\tau \approx 1.5$years.

process with unknown correlation structure it is misleading to use $\alpha > 0.5$ as evidence for long-memory. It might very well be that the record is too short to observe a plateau with $\alpha = 0.5$.

2.5.3 Shift of the finite scaling region

As shown in Sect. 2.4.2, under certain conditions also short-memory processes can exhibit a finite "scaling" region. Thus, the question arises, if such a scaling region derived from the fluctuation function corresponds to the same region in the auto correlation function. To address this question, we relate the fluctuation function shown in Fig. 2.4 to the analytical autocorrelation function as shown in Fig. 2.6. The dashed lines depict the autocorrelation functions of the single AR[1] processes with the largest and the smallest time scale, the autocorrelation function of the superposition of the three AR[1] processes is given by the dashed-dotted line. The solid line represents a power-law with exponent $\gamma = 0.8$ as expected, when applying Eq. (2.9) to the exponent $\alpha = 0.6$ as derived from the fluctuation function. We find that the region of almost constant slope of the autocorrelation function is located on smaller scales between $s \approx 1$ and maximally $s \approx 1,000$ ($\simeq 3$ years). Thus, based on a finite scaling region found in the fluctuation function of a short-memory process, it is not valid to conclude that an equal scaling region exists also for the autocorrelation function, as assumed by Koscielny-Bunde et al. [55].

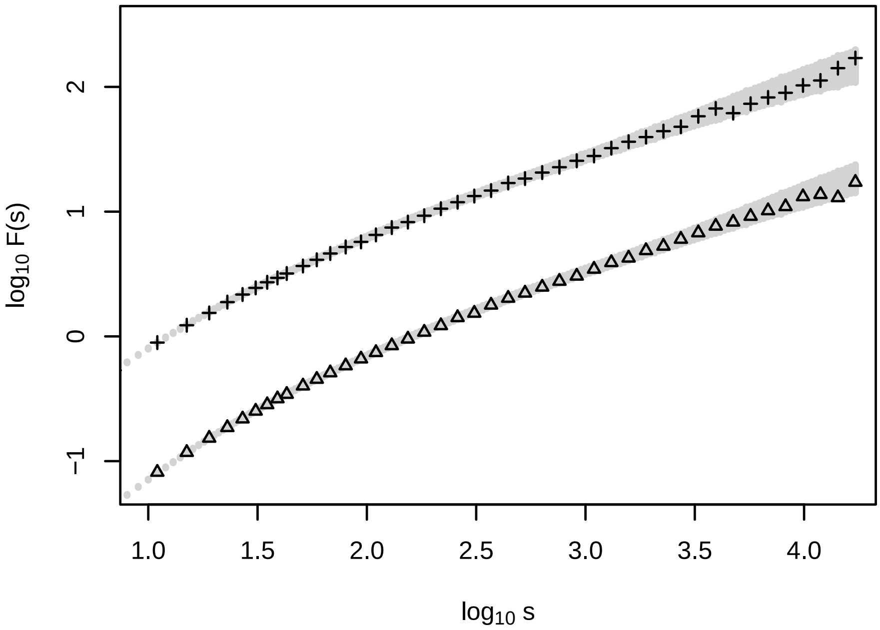

Figure 2.7: DFA1 (+) and DFA2 (△) fluctuation function of Prague daily air temperature data calculated for approximately 50 points per order of magnitude. Only every 4th point is shown to enhanced clarity. The shadows mark the 1.96σ confidence regions derived from 1,000 runs of the AR model.

2.6 Memory of Temperature Records

As a consequence of the previous discussion, we review a prominent DFA result. It has been stated by Koscielny-Bunde et al. [55], that temperature records follow a universal scaling law. This behavior was said to be found in numerous data sets. For a discussion of the existence of universal scaling exponents see [126]. Here, we want to check these findings exemplarily in the daily air temperature dataset of Prague starting from 01/01/1800 up to 12/31/1992 ($N = 70,492$). To test, whether the underlying process is long-range or short-range correlated, we adjusted the long-range correlated model and the AR-model given in Sect. 2.4.1 to the data set.

We apply DFA1 and DFA2 to the temperature anomalies $\Delta T_i = T_i - \bar{T}_i$ giving the deviation of a days temperature from the average over several years for this day of the year. An investigation of higher orders of DFA does not significantly affect the discussion presented while for DFA1 the effect of a trend might be suspected. $F(s)$ is calculated for approximately 50 points per order of magnitude upto $s_{\max} = N/4$ and is shown in double logarithmic representation in Fig. 2.7 for DFA1 and DFA2.

The behavior is qualitatively different from white noise. However, following the discussion in Sect. 2.4.2, we have to estimate the local slopes to investigate for power-law scaling. From the fluctuation function, we estimate the local slopes using a straight line fit in a small window of 21 points.

According to Sect. 2.4.2, DFA is not specific when investigating for long-memory. However, we alternatively study if long-memory can be excluded. Figures 2.8 (a) and (b) show the

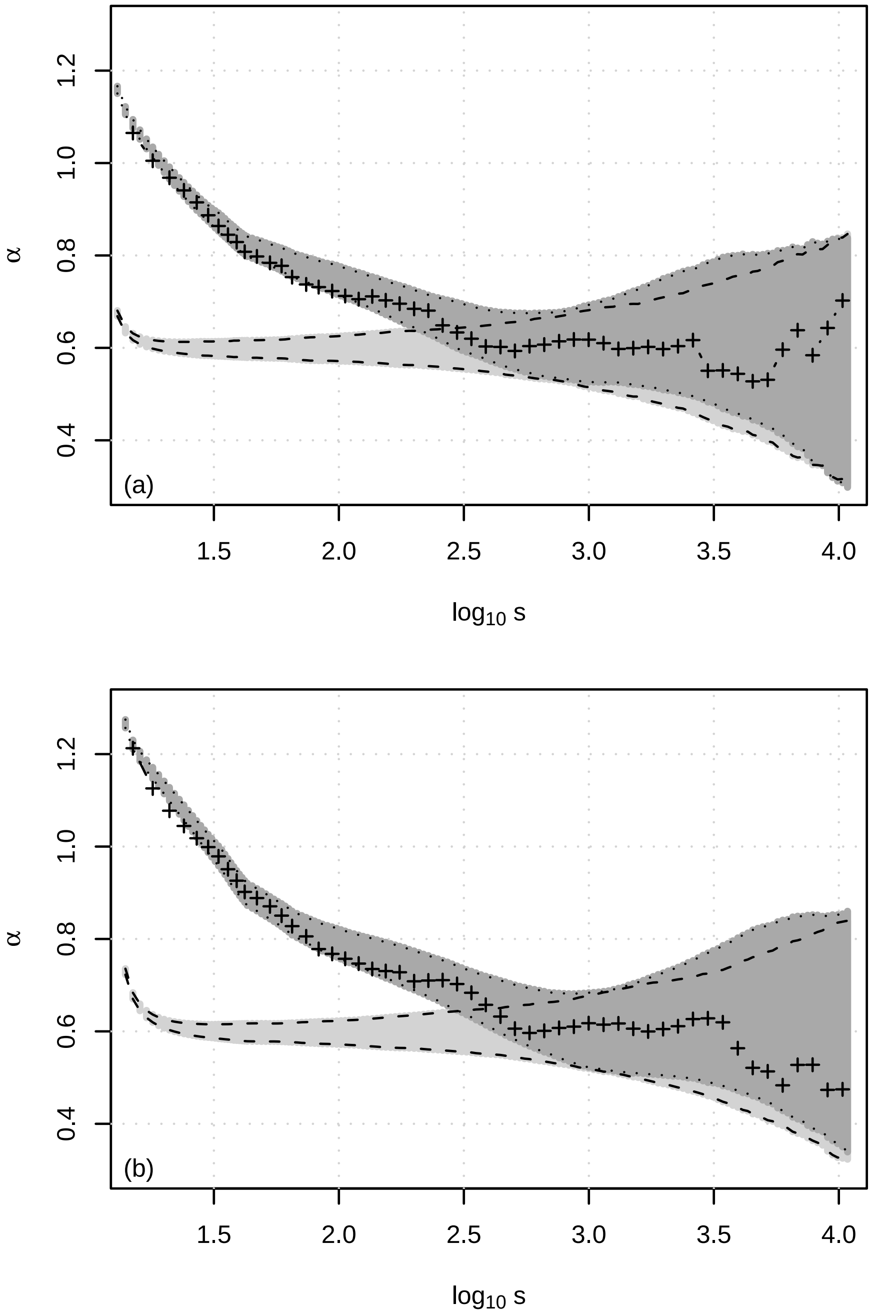

Figure 2.8: Local slopes of the fluctuation functions plotted in Fig. 2.7 for (a) DFA1 and (b) DFA 2 of the Prague daily temperature data. The dotted lines border the 1.96σ confidence regions of the short-range correlated model Eq. (2.10) (dark shadow), the dashed lines those of the long-memory model with $\alpha = 0.6$ (light shadow).

local slopes for DFA1 and DFA2 of the Prague daily temperature record and additionally the 1.96σ confidence intervals derived from the two models. For both orders of DFA the result for the empirical data is almost completely within the 1.96σ interval of the short-memory process. In the range of large s with high variability, we find the local slopes also within the 1.96σ interval corresponding to the long-memory process. Thus, from the given data, one cannot decide whether the Prague temperature time series is a realization of a short-memory or a long-memory process. However, considering all scales observed, the short- memory model describes the data better than the long-memory model with respect to DFA. If the underlying process was short-range correlated, data sets of sufficient length would provide estimates of sufficient accuracy to identify the short-memory character. According to Sect. 2.5.2, we expect such a time series to be much longer as the recorded ones.

2.7 Summary and Conclusions

In this chapter we studied the inference of long-range correlations by means of DFA with respect to the notions of sensitivity and specificity. In the sense of falsification, we argue that the inference of a long-range correlated process underlying a given time series requires not only to show compatibility of the data with a long-range correlated process. Furthermore, other possible correlation structures, especially short-range correlations, have to be excluded.

Power-law scaling of the DFA fluctuation function is frequently taken as evidence for long-range correlations. To reliably infer power-law scaling, it must not be assumed but has to be established. Thus, fitting a straight line to the log-log representation of the fluctuation function is not sufficient. In fact, scaling can be established by estimating local slopes and investigating them for constancy in a sufficient range. However, finite datasets bring along natural variability. To decide, if a fluctuating estimation of the slope has to be considered as being constant, we calculated empirical confidence intervals for a long-range and a simple short-range correlated model.

As a main insight, the inference of long-range correlations from a finite amount of data turned out to be not specific by means of DFA. Alternatively, one can investigate, if short-range correlations can be inferred for the underlying process. For this setting, the specificity increases with the length of the data record. In terms of falsifiability, these findings have to be interpreted as follows: First, to even attempt to falsify the long memory hypothesis, one has to test it and compare it with alternative hypothesis, in this case short memory. It turned out, that for the given amount of data, both hypothesis are not refutable. For longer time series, the long memory hypothesis becomes in principle falsifiable. This, however, requires much longer time series.

Discussing typical difficulties of interpreting DFA results from short memory processes, we remark that the characteristic time scale of a short memory process cannot be directly inferred from the fluctuation function by means of DFA. Additionally, we show that a local slope larger than $\alpha = 0.5$ for large scales does not necessarily imply long-memory. If the length of the time series is not sufficiently large compared to the time scales involved, also for short-memory processes $\alpha = 0.5$ may not be reached. Finally, we demonstrated, that it is not valid to conclude from a finite scaling region of the fluctuation function to an equivalent scaling region of the autocorrelation function. In fact, a finite scaling region of the autocorrelation function is shifted to much longer time scales in the fluctuation function, so that a much longer memory is mimicked.

Keeping these findings in mind, we reviewed the DFA results for the Prague daily air temperature record and compare them to a long-range and a short-range correlated model constructed using AR[1]-processes. Considering only large scales, for the given data set, one cannot reject the one or the other model. For the short-range correlated model, a much longer series would be necessary to identify a local slope of $\alpha = 0.5$. The short-range correlated model we used involves time scales up to ≈ 1.5 years. This might be regarded as long compared to the sampling time of one day but it is considerably small compared to the decades of persistence which are proposed in e.g. [55] and [31].

However, it is not our intention to advocate for AR-processes being suitable to model temperature variation. Following the idea of Occam's razor, for a certain range of scales a power-law model can be formally suitable to describe characteristics of even short-memory processes. This, however, does not give evidence for long-memory. In this context, we suggest to forbear from applying the nonparametric DFA-procedure[3] to investigate the correlation structure of stochastic processes in favor of a parametric analysis by means of FARIMA models. In this context, one can fit FARIMA models of different complexity, including or excluding long memory, to the data and compare the results employing certain information criteria. This approach is currently under development by Rust [107].

---

[3] The straight line fit in fact represents the simplest possible model, namely fractional gaussian noise, which is definitely too simple to describe the majority of real world processes.

Chapter 3

Continuous Wavelet Spectral Analysis

3.1 Introduction

In climatology, observational data often represent a superposition of different underlying processes with variability on a wide range of scales. For directly observed instrumental data, the processes of interest range from local atmospheric weather phenomena with scales of days up to the interdecadal variability of the oceans. Paleoclimatic proxy data from ice cores eventually show variability from centuries up to tens of thousands of years imprinted by the Milancovic-cycles.

The method of choice to analyze processes with a wide frequency band is Fourier analysis [e.g. 9, 90]. However, geophysical processes are furthermore often nonstationary. For instance, the statistical properties of the ENSO system are not constant in time but change considerably from decade to decade because of internal changes and varying boundary conditions as, e.g., a changing global climate. To handle such a class of processes, Gabor [24] proposed a windowed Fourier transformation that assumes local stationarity in a time window shifted over the data set. The time resolution is given by the window and is constant for all frequencies. This, however, is a drawback since for high frequencies one desires a higher time resolution as for low frequencies. In this regard, the optimal method is continuous wavelet analysis as it intrinsically adjusts the time resolution to the analyzed scale [e.g. 12, 16, 40].

The continuous wavelet spectra of paradigmatic processes as Gaussian white noise [40] or fractional Gaussian noise [65] have been studied. Continuous wavelet spectral analysis has been applied to real world problems in physics, climatology [e.g. 34], life sciences [e.g. 33] and other fields of research. Hudgins et al. [46] defined the wavelet cross spectrum to investigate scale and time dependent linear relations between different processes. This measure has been applied in the analysis of atmospheric turbulence [46], time varying relations between El Niño/Southern Oscillation (ENSO) and the Indian monsoon [122], as well as interrelations of business cycles from different national economies [15].

Except for analytically derived wavelet spectra of simple processes as, for instance, white noise or stationary increment processes, the wavelet spectrum hitherto has been defined as the expectation value of the wavelet scalogram of a given time series [65, 67]. However, this concept brings along certain difficulties: These process measures depend on realizations and the wavelets chosen for the analysis, but not on a priori defined spectra in the wavelet domain. Consequently, one can neither simulate data with a given nonstationary wavelet spectrum, i.e. solve the direct problem, nor evaluate the quality of the estimator in terms of bias, i.e. study the inverse problem.

A second question concerns the significance testing of wavelet spectra. Torrence and Compo [121] formulated pointwise significance tests against reasonable background spectra. However, Maraun and Kurths [67] pointed out a serious deficiency of pointwise significance testing: Given a realization of white noise, large patches of spurious significance are detected, making it - without further insight - impossible to judge which features of an estimated wavelet spectrum differ from background noise and which are just artefacts of multiple testing. Under these conditions, a reliable corroboration of a given hypothesis is impossible. This demonstrates the necessity to study the significance testing of continuous wavelet spectra in terms of sensitivity and specificity.

A third problem is the interpretation and significance testing of cross wavelet measures. Huang et al. [45] suggested to test the wavelet cross spectrum against red noise. They applied this method to study the interrelation between ENSO and the North Atlantic oscillation (NAO) and found an instantaneous influence of strong El Niños on North Atlantic climate. However, we argue that the results might be an artefact of the wavelet cross spectrum [67].

In this chapter we suggest a framework of nonstationary Gaussian processes defined in the wavelet domain. These processes are characterized by their time dependent spectral properties. They can be used to simulate data with a specific nonstationary wavelet spectrum and to investigate the behavior of wavelet spectral estimators. To overcome the problems arising from pointwise significance testing, we develop an areawise significance test, taking advantage of basic properties of the continuous wavelet transformation. We evaluate this test in terms of sensitivity and specificity within the framework suggested.

In Sec. 3.2 we present an overview of the mathematical background of continuous wavelet transformation to introduce all properties being important for the following study. In Sec. 3.3, we intensively discuss the problems regarding continuous wavelet spectral analysis and formulate a set of questions as a reference for the subsequent investigations. The framework of nonstationary Gaussian processes in wavelet domain is formulated and studied in Sec. 3.4. The next Sec. 3.5 addresses the estimation of wavelet spectra, followed by the study and development of significance tests in Sec. 3.6. This section also discusses the significance testing and interpretation of cross wavelet measures. We close the chapter about wavelet spectral analysis with a set of climatological applications in Sec. 3.7 and summarizing conclusions in Sec. 3.8. To improve the legibility, most derivations have been moved to App. B.

3.2 Mathematical Background

This section is intended to provide the required basics of continuous wavelet transformation that have been extensively studied. For a detailed presentation, please refer to the many standard text books [e.g. 12, 16, 40].

3.2.1 From Fourier to Wavelet Transformation

The calculation of the Fourier transformation $f(\omega)$ of a time series $s(t)$ often reveals valuable information about the underlying process:

$$f(\omega) = \int_{-\infty}^{\infty} dt\, s(t) e^{i\omega t}, \tag{3.1}$$

where ω denotes frequency. The Fourier periodogram $|f(\omega)|^2$ estimates the spectrum of the underlying process at frequency ω. However, real world processes are often nonstationary, i.e. their statistical properties change over time. Thus a time independent description of the spectrum might be no longer suitable. A possible solution is the application of the Gabor transformation [24]: The Fourier transformation is calculated locally around the desired point in time b by multiplying the time series with an envelope function $g(t)$ centered around b to damp distant contributions,

$$G(b,\omega) = \int_{-\infty}^{\infty} dt\, s(t) e^{-i\omega t} g(t-b), \tag{3.2}$$

where the localizing function $g(t)$ might be chosen as a Gaussian, $g(t) = \exp(-t^2/(2\sigma^2))$. This method assumes local stationarity and provides a slowly changing Fourier spectrum with a time resolution given by the width of the Gaussian σ.

As a drawback of this idea, the time resolution is fixed for every frequency ω. On the one hand, if one chooses a narrow window width σ to obtain a good time resolution for high frequencies, the window width will be short compared to the oscillation length on low frequencies. Consequently, the frequency resolution for low frequencies will be very poor. On the other hand, if one chooses a wide window to obtain a good frequency resolution for low frequencies, the window width will be very wide compared to oscillations on high frequencies. Consequently, the time resolution for high frequencies will be unnecessarily poor.

Thus one might want to adjust the time resolution, i.e. the width of the Gaussian function σ to the frequency analyzed, $\sigma = \omega_0/\omega$. The parameter ω_0 determines the relative width of the Gaussian, i.e. the localization. Considering scales instead of frequencies $a = 2\pi/\omega$ (i.e. $\sigma = a\omega_0/(2\pi)$), one gets

$$M(b,a) = \int_{-\infty}^{\infty} dt\, s(t) e^{-2\pi i t/a} e^{-\frac{1}{2}\left(\frac{t-b}{a\omega_0/(2\pi)}\right)^2}. \tag{3.3}$$

Except for a (rather arbitrary) phase factor $\exp(i\omega_0 b/a)$, this is exactly the wavelet transformation of $s(t)$ using the Morlet wavelet $m(\theta) = \exp(i\omega_0\theta)\exp(-\theta^2/2)$, with $\theta = (t-b)/a$ and $\omega_0 = 2\pi$:

$$W_m(b,a) = \int_{-\infty}^{\infty} dt\, \bar{m}\left(\frac{t-b}{a}\right) s(t). \tag{3.4}$$

Varying ω_0 changes the width of the envelope and thus adjusts the time/scale resolution. The better the scale resolution, the poorer the time resolution and vice versa. The choice of a Gaussian envelope is rather arbitrary; also not only complex exponentials might be considered as basis functions for the frequency decomposition. Thus, one can choose a more general function $g(t)$ instead of the Morlet wavelet $m(t)$.

3.2.2 Wavelets

A *wavelet* $g(t)$ is a function fulfilling the following conditions:

$$\begin{aligned} \int_{-\infty}^{\infty} dt\, g(t) &= 0 \\ \int_{-\infty}^{\infty} dt\, |g(t)|^2 &= 1 \end{aligned} \tag{3.5}$$

The first condition implies a vanishing zero frequency component, $\hat{g}(0) = 0$, and is called the admissibility condition. The second ensures a sufficient localization of the wavelet in the time domain. A wavelet is called *progressive*, if it contains only positive frequencies [40]:

$$g(t) = \frac{1}{2\pi} \int_0^\infty d\omega \hat{g}(\omega) e^{i\omega t} \tag{3.6}$$

For details refer to [e.g. 16, 50, 40].

3.2.3 Continuous Wavelet Transformation

Given a time series $s(t)$, its *continuous wavelet transformation* (CWT) $W_g s(t)[b,a]$ at time b and scale a with respect to the chosen wavelet $g(t)$ is given as

$$W_g s(t)[b,a] = \int dt \, \frac{1}{\sqrt{a}} \, \bar{g}\left(\frac{t-b}{a}\right) s(t). \tag{3.7}$$

The wavelet transformation along one fixed scale, $a = $ *const.*, is called *voice*. The normalization of the transformation can in principle be chosen arbitrarily. However, for the definition of our estimators (Sec. 3.5.1), the L^2-normalization $1/\sqrt{a}$ turns out to be the naturally consistent choice. The CWT from one dimension to two dimensions does not produce any new information, hence any wavelet transformation contains a great quantity of redundant information that will be discussed in Sec. 3.2.5.

Being important for the further discussion, the covariances of CWT are recalled in App. B.1.1. For a detailed discussion of CWT basics please refer to the comprehensive literature [e.g. 12, 16, 40].

3.2.4 Inverse Continuous Wavelet Transformation

For every wavelet $g(t)$ in a strict sense, i.e. fulfilling Eq. (3.5), a *reconstruction wavelet* $h(t)$ can be found [40]. For a progressive wavelet $g(t)$, the reconstruction wavelet $g(t)$ is defined by the equation

$$c_{g,h} = \int_0^\infty \frac{d\omega}{\omega} \bar{\hat{g}}(\omega) \hat{h}(\omega), \qquad 0 < |c_{g,h}| < \infty. \tag{3.8}$$

Using the reconstruction wavelet, one can define an *inverse transformation* of a function $r(b,a)$ from the positive half plane $\mathbb{H}$ to the real axis, i.e. the time domain.

$$M_h r(b,a)[t] = \int_{\mathbb{H}} \frac{db\, da}{a^2} \, r(b,a) \, \frac{1}{\sqrt{a}} \, h\left(\frac{t-b}{a}\right) \tag{3.9}$$

This operation from two dimensions to one dimension is not the unique possible inverse transformation.

Successively applying the wavelet transformation and its inverse to a time series $s(t)$ will reconstruct $s(t)$: $s(t) \equiv M_h W_g s(t)$. However, because of the redundancies characterizing any wavelet transformation not every function on the positive half-plane is a wavelet transformation. Thus the successive transformation of an arbitrary function $r(b,a)$ in time and scale to the time domain and back to the time/scale domain in general does not reconstruct $r(b,a)$, i.e. in general $r(b,a) \neq W_g M_h r(b,a)$. In fact, this operation projects $r(b,a)$ onto the subspace of all wavelet transformations (App. B.1.2).

3.2.5 Reproducing Kernel

The redundancies of CWT (and any time-frequency resolved analysis) are reflected in inherent correlations of any wavelet transformation $r(b,a)$. For Gaussian white noise $\eta(t)$, these correlations are given exactly by the *reproducing kernel* $K_{g,h}((b-b')/a', a/a') = W_g h((b-b')/a')$ (For details refer to App. B.1.3 or [e.g. 12, 40, 67]):

$$\begin{aligned} C(b,a;b',a') &= \langle W_g\eta[t](b,a), W_g\eta[t](b',a')\rangle \\ &\sim K_{g,h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right). \end{aligned} \tag{3.10}$$

As an example, the reproducing kernel of the Morlet wavelet is given in App. B.1.5. These inherent correlations constitute a fundamental difference of any time-frequency (or scale) resolved analysis to time independent Fourier analysis, where neighboring frequencies are asymptotically uncorrelated. The reproducing kernel can also be interpreted as a time-scale uncertainty as e.g. the time/energy uncertainty relation in quantum mechanics.

3.2.6 Choice of the Wavelet

A special set of progressive wavelets are those exhibiting a real valued Fourier transform. Their representation in time domain is complex and thus they are capable of detecting modulus and phase of the analyzed time series. This property makes them especially useful for the spectral analysis of real oscillations: The analysis of a sine function with a real (in time domain) wavelet would yield a real oscillating function in the wavelet domain. The analysis with a complex (in time domain) wavelet, however, yields a constant modulus and a time varying phase which is exactly what is desired. Also, as one is interested in decomposing a signal with respect to time and scale, one should utilize wavelets which are well localized in the time/scale domain. For some illustrative examples, please refer to the appendices B.1.4 and B.1.5.

3.2.7 Discrete Wavelet Transformation

Parallel to CWT, one can define a discrete wavelet transformation (DWT). This transformation decomposes a signal into contributions on dyadic scales $s_i \sim 2^i$. The time resolution is down-sampled, i.e. bisected for every scale. Without down-sampling, this transformation is called maximum overlap discrete wavelet transformation (MODWT). The latter transformation is the CWT applied to a subset of voices only. These techniques reduce the redundancies given by the reproducing kernel and decrease the computational cost compared to CWT. However, oscillations on scales not captured by the choice given above are either redistributed to neighboring scales or possibly totally overseen if the reproducing kernel is too localized in scale. These transformations open another wide field of research which is not the subject of this work. For details about DWT and MODWT, please refer to Percival and Walden [84].

3.3 Motivation - The State of the Art

Although the mathematical properties of CWT are well studied, the formulation of continuous wavelet spectral analysis has got a rather tentative character. This section sketches the status quo and introduces the problems concerning the current formulation of spectral measures in the wavelet domain and the significance testing of wavelet spectra. This discussion is intended to motivate the further developments and studies. As a guideline, we formulate a set of questions in the end of the section.

3.3.1 Conventional Definitions of Wavelet Spectral Measures

For Gaussian white noise and self similar stationary increment processes, the continuous wavelet spectrum has been calculated analytically [40, 65]. For nonstationary processes, continuous wavelet spectral measures have been defined as the expectation value of the corresponding estimator [65, 67]: Given a process with a realization $s(t)$, a wavelet spectrum denoting the variance at a certain scale a and time b might be defined as

$$WSP_g(b,a) = \langle |W_g s(t)|^2 \rangle . \tag{3.11}$$

Given two processes with realizations $s_1(t)$ and $s_2(t)$, a wavelet cross spectrum expressing the time scale resolved covariance might be defined as

$$WCS_g(b,a) = \langle W_g s_1(t) \bar{W}_g s_2(t) \rangle \tag{3.12}$$

Normalizing the modulus of the cross spectrum to the single spectra, one obtains the squared coherence

$$WCO_g^2(b,a) = \frac{|\, WCS_g|^2}{WSP_{g,1}\, WSP_{g,2}}, \tag{3.13}$$

reflecting the scale resolved linear correlation. These measures will be discussed later in Sec. 3.4.3. Here, we just aim to point out several problems:

- In general, the given expectation values $\langle . \rangle$ denote ensemble averages. However, as one in practice often observes only one realization of a particular process, one has no access to the expectation value as an ensemble average. In time independent Fourier analysis, one assumes ergodic processes (see Sec. 1.2) and replaces the ensemble expectation value by the time expectation value. Then a consistent spectral estimate can be obtained in the limit of $N \to \infty$ by averaging in the Fourier domain with a smoothing window growing slower than the frequency resolution [9, 90]. For the local analysis of nonstationary processes, the time expectation value does not exist and the replacement is not valid. Thus, averaging in time (as well as in scale) direction produces a bias and prevents a consistent estimate.

- These measures depend on the wavelet chosen for the analysis, i.e. the definition of the underlying process depends on the technique chosen for its analysis.

- These definitions are based on given realizations $s(t)$. Thus, the underlying process has no a priori mathematical definition. Consequently, it is impossible to study the direct problem of defining a desired wavelet spectrum and producing corresponding time series. However, this also means that it is impossible to investigate the inverse problem: Without a well defined wavelet spectrum, it is impossible to even formulate properties as the bias of a certain estimate (compare Sec. 1.1). Also the relation of a stationary wavelet spectrum to the Fourier spectrum is not clear.

3.3.2 Conventional Smoothing and Global Wavelet Spectrum

Torrence and Compo [121] suggested to smooth the wavelet sample spectrum to reduce the variance. In scale direction, they have chosen a window width being constant in the typical logarithmic scale axis of the discrete wavelet spectral matrix (see App. B.1.6). In time direction, they suggested a window width constant for all scales; extending the width over the length of the whole time series yields the global wavelet spectrum. The idea of a constant time window, however, contradicts the essence of wavelet analysis: The motivation to define a wavelet transformation instead of using the Gabor transformation was just to provide a scale dependent time resolution. This advantage is foiled by applying a constant time window for smoothing. The global wavelet spectrum has been reported to be a consistent estimator for the Fourier Spectrum [121]. However, because of the reproducing kernel, the estimate should in general be biased.

3.3.3 Conventional Pointwise Significance Testing

To our knowledge, Torrence and Compo [121] were the first to define a significance test for continuous wavelet spectral and cross spectral analysis. Given a wavelet sample spectrum, they assumed reasonable background spectra. Separately for every point in time and scale (i.e. pointwise), they checked whether the particular spectral coefficient exceeded a certain critical value corresponding to the chosen significance level.

For testing wavelet spectra, they used the spectrum of a red noise process fitted to the time series as background spectrum. In case of testing wavelet cross spectra and wavelet coherence, they have estimated the background spectrum from a pair of two red noise processes fitted to the corresponding two time series. The method of pointwise significance testing will be explained in detail in Sec. 3.6. Here, we just want to point out some important problems arising from the presented concept of significance testing.

Wavelet Spectrum

Figure 3.1(a) shows the wavelet sample spectrum of the NINO3 index (for a description of the data set, see App. A). The black contour lines depict patches significant on the 0.05% level after applying the pointwise significance test against a red noise spectrum as outlined above. Many of these patches coincide with major El Niños and reveal additional frequency information of the particular event. Also patches on lower scales (i.e. higher frequencies) show up suggesting further superimposed dynamics. However, taking into account Fig. 3.1(b) casts doubt onto these presumptions. The lower panel shows the wavelet sample spectrum of a Gaussian white noise realization. As white noise is a special case of red noise, this sample spectrum contains by construction no significant patches, hence all the observed patches have to be false positive results. Some of these patches are remarkably long compared to the scale they are to be found on (e.g. the patch in Fig. 3.1(b) marked (A)). Also complex structures (e.g. that one marked (B)) seem to appear randomly. The reason for these spurious results is the effect of multiple testing (see Sec. 1.4): Testing every point separately by definition leads to false positive results (see Sec. 1.4). However, without further insight it is impossible to distinguish a real deviation from red noise (as given by the El Niño events) from spurious patches.

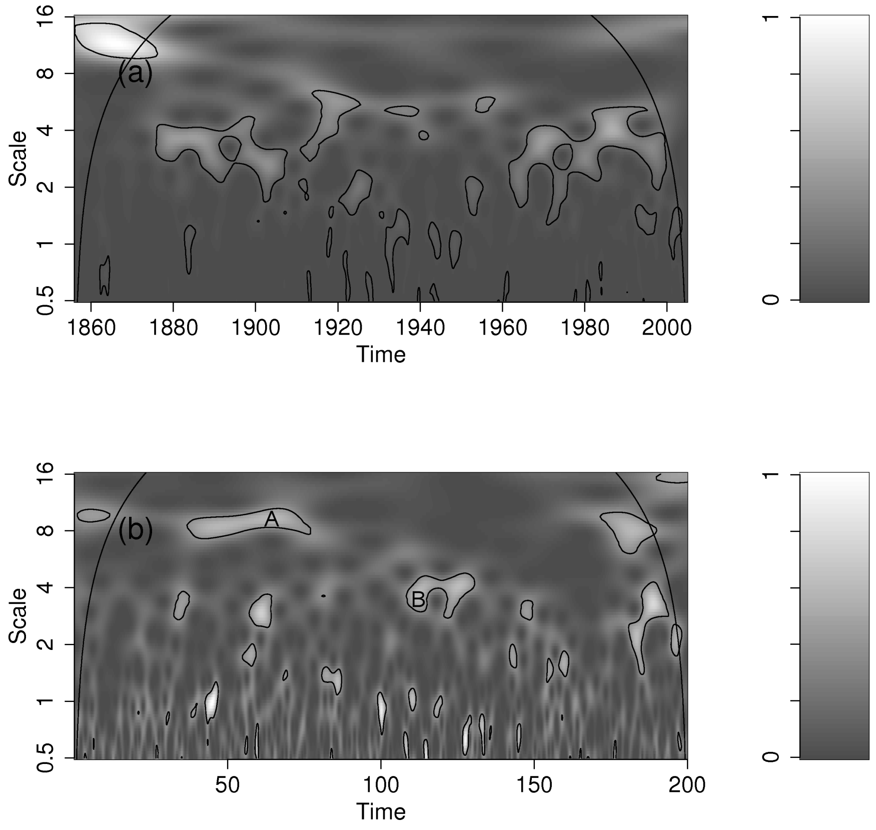

Figure 3.1: Wavelet sample spectra calculated with a Morlet wavelet, $\omega_0 = 6$, without smoothing, including a pointwise significance test against red noise on the $\alpha = 0.05$-level. (a) NINO3 index, (b) a realization of Gaussian white noise.

Wavelet Cross Spectrum

Figure 3.2(a) depicts the wavelet sample cross spectrum of the NINO3 index vs. the NAO index (see App. A). This plot was first published by Huang et al. [45] including the significance test against two red noise spectra as introduced above. Many prominent peaks appear, most of them coinciding with major El Niño events. From these results, the authors concluded that every strong El Niño exhibits a significant influence on North Atlantic climate. Figure 3.2(b), however, shows the wavelet sample cross spectrum between a Gaussian white noise realization and a sine function. Both "processes" are by construction independent of each other. Nevertheless, large patches positioned in the local extrema of the sine wave appear. This finding demonstrates that the wavelet sample cross spectrum produces spurious results and raises the question, whether the findings of Huang et al. [45] could be mere artefacts.

3.3.4 Questions arising

In the previous paragraphs we presented an overview of the state of the art of continuous wavelet spectral analysis. During this discussion, we pointed out shortcomings of the un-

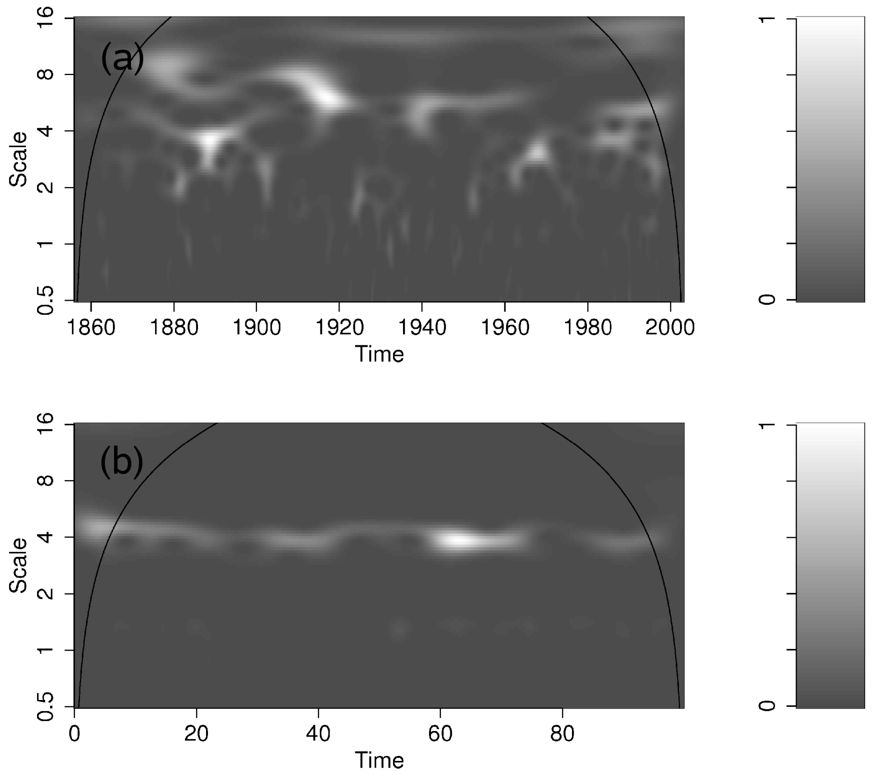

Figure 3.2: Wavelet sample cross spectra calculated with a Morlet wavelet, $\omega_0 = 6$, without averaging. (a) NINO3 index vs. NAO index (b) realization of Gaussian white noise vs. sine function with $f = 1/4$.

derlying theory and illustrated several pitfalls. Based on these problems, one might raise the following questions regarding the direct problem of simulating nonstationary Gaussian processes and the inverse problem of wavelet analysis:

- How can the true underlying wavelet spectrum of a nonstationary gaussian process be defined? How far does this spectrum depend on the chosen wavelet? How can one formulate the direct problem of simulating realizations of processes with given continuous wavelet spectrum? How are the wavelet spectrum and the Fourier spectrum related?

- What are bias and variance of a wavelet sample spectrum and their dependence of the wavelet chosen?

- How does one have to smooth wavelet sample spectra in order to keep the advantage of a scale adjusted time resolution? What is the relation between the global wavelet spectrum and the Fourier spectrum?

- How can multiple testing effects in significance testing of wavelet spectra be taken into account?

- How can one test for covarying oscillations of two processes?

It turned out that continuous wavelet analysis is far away from being a closed theory providing reliable and unambiguous results for the application in time series analysis. The following sections mainly deal with the development of a framework to study and answer the listed questions.

3.4 Stochastic Processes defined in Wavelet Domain

To solve the direct problem of simulating arbitrary nonstationary Gaussian processes, one needs to formulate a mathematical framework to represent the corresponding processes. Given a concept for the direct problem, one can furthermore investigate the inverse problem of estimating the underlying nonstationary Gaussian process from an observed time series. Percival and Walden [84] developed a framework to model stochastic processes in wavelet domain based on MODWT. Another concept also confined to dyadic scales is the approach of Nason and von Sachs [77], that is outlined in App. B.2. Here, we present a concept using CWT that allows for the analysis of arbitrary scales.

3.4.1 Stationary Gaussian Processes in Fourier Domain

Stationary Gaussian processes are completely defined by their ACF or equivalently by their spectrum (see Sec. 1.2.2). Realizations of Gaussian processes with a spectrum $|f(\omega)|^2$ can be simulated in the following way: One simulates a white noise time series ϵ_t, transforms it to the Fourier domain, multiplies it with the desired spectrum and transforms it back to the time domain (see Sec. 1.2.2). The desired spectrum $|f(\omega)|^2$ can either be interpreted parametrically, e.g. as the spectrum of an ARMA[p,q] process, or non-parametrically, e.g. when it is estimated from a given time series. In this generalized context, the function $f(\omega)$ is called a Fourier multiplier. We extend this concept to define nonstationary Gaussian processes using wavelet multipliers $m(b, a)$ as a function of time and scale.

3.4.2 Definitions

We define an equivalence class of nonstationary Gaussian processes in the wavelet domain by the *wavelet multipliers* $m(b, a)$ as a function of time b and scale a. The multipliers can be given as a parametric function or nonparametrically, e.g. estimated from data. Realizations $s(t)$ are given as

$$s(t) = M_h m(b, a) W_g \eta(\tau), \tag{3.14}$$

i.e. a driving Gaussian white noise $\eta(t) \sim \mathcal{N}(0, 1)$ with $\langle \eta(t_1)\eta(t_2)\rangle = \delta(t_1 - t_2)$ is transformed to the wavelet domain, multiplied with $m(b, a)$ and transformed back to the time domain.

We assume the following natural behavior for the wavelet multipliers $m(b, a)$:

$$\begin{aligned} \partial_a m(b, a) &< O(1/a) \\ \partial_b m(b, a) &< O(1/a) \end{aligned} \tag{3.15}$$

These relations reflect the covariances of wavelet transformation (see App. B.1.1) and thus prevent structures small compared to the reproducing kernel (see App. B.1.3) that are in principle undetectable.

Dependency of the Realization on the Wavelet and Asymptotic Behavior

Because every wavelet transformation is subject to redundancies, the realization $m(b,a)W_g\eta(\tau)$ in the wavelet domain[1] is in general not a wavelet transformation (see App. B.1.2). Thus, realizations $s(t)$ in the time domain depend (usually weakly, see Sec. 3.2.6 and App. B.3.1, Fig. B.4) on the chosen wavelet $g(t)$ and the reconstruction wavelet $h(t)$ respectively. Hence, the individual process is completely defined by $m(b,a)$, $g(t)$ and $h(t)$.

To ensure at least asymptotic independence of the wavelet $g(t)$ and the reconstruction wavelet $h(t)$, one has to demand a certain asymptotic behavior of the process $m(b,a)$. As wavelet analysis is a local analysis, the behavior for long time series is not relevant. Alternatively, one has to consider one of the following two cases: Either the limit of small scales or the limit of stationary white noise processes.

Considering the first case, the process $m(b,a)$ gets asymptotically independent of the wavelet, if for small scales the following relation holds:

$$\begin{aligned} \partial_a m(b,a) &< O(a^{-1+\epsilon}) \\ \partial_b m(b,a) &< O(a^{-1+\epsilon}), \end{aligned} \tag{3.16}$$

with $\epsilon > 0$. Looking with a microscope into finer and finer scales, the derivatives of $m(b,a)$ grow slower and slower such that the process looks more and more stationary and white.

For the second case, we consider a class of processes $m_\lambda(b,a) = m(\lambda b, a^\lambda)$ being a function of a parameter λ. Here, the process $m_\lambda(b,a)$ gets asymptotically independent of the wavelet, if $m_\lambda(b,a)$ gets stationary and white in the limit of $\lambda \to \infty$: Using Eq. (3.15), one gets immediately

$$\begin{aligned} \partial_a m_\lambda(b,a) &= \lambda a^{\lambda-1}\partial_a m(\lambda b, a^\lambda) \overset{(3.15)}{<} O(\frac{\lambda}{a}) \\ \partial_b m_\lambda(b,a) &= \qquad \lambda\partial_b m(\lambda b, a^\lambda) \overset{(3.15)}{<} O(\frac{\lambda}{a^\lambda}) \end{aligned} \tag{3.17}$$

The asymptotic dependencies on the wavelet for the two cases are derived in App. B.3.1. Both cases are equivalent in the sense that in both limits, more and more reproducing kernels fit into local structures of the process[2] (see App. B.3.1).

Relation to the Fourier Spectrum

$m(b,a) \equiv m(a)$ defines a stationary Gaussian process in wavelet domain. In this special case, the stationary Fourier spectrum $|f(\omega)|^2$ exists:

$$f(\omega) \approx m\left(\frac{2\pi}{\omega}\right)C_1 + \frac{2\pi}{\omega}m'\left(\frac{2\pi}{\omega}\right)C_2, \tag{3.18}$$

with $a = 2\pi/\omega$ and C_1 and C_2 being constants depending on the localization of the used wavelets. This equation is derived in App. B.3.2. As expected, the Fourier spectrum is given by the wavelet spectrum plus a correction term. The latter depends on the slope of the wavelet spectrum $m'(b,a)$ (Details are explained in App. B.3.2). For the asymptotic behavior defined in Eqs. 3.16 and 3.17, the correction term vanishes for high frequencies.

---

[1] which is a wavelet transformation multiplied with an arbitrary function $m(b,a)$

[2] Note that stationarity alone is not sufficient, as the reproducing kernel is extended not only in time direction, but also in scale direction, compare Sec. 3.4.2.

3.4.3 Spectral Measures

As stated in Sec. 3.3.1, continuous wavelet spectral measures have been defined as the expectation value of the corresponding estimator. This definition turned out to have several drawbacks as mentioned in Sec. 3.3.1. Using wavelet multipliers one can define time dependent spectral measures that elegantly overcome these difficulties. Following Sec. 3.4.2, the wavelet dependent measures defined in Sec. 3.3.1 converge to these wavelet independent measures in the limit of small scales or stationary and white processes.

Spectrum

Given wavelet multipliers $m(b,a)$, one can define a wavelet independent[3] *spectrum*

$$S(b,a) = |m(b,a)|^2 \tag{3.19}$$

The spectrum denotes the variance of the process at a certain time b and scale a. With the chosen normalization of the wavelet transformation Eq. (3.7), white noise is given by $S(b,a) = |m(b,a)|^2 = const.$

Cross Spectrum

Consider two interacting processes $m_{1c}(b,a)$ and $m_{2c}(b,a)$, i.e. both are driven by the same noise realization: $s_1(t) = M_h m_{1c}[b,a] W_g \eta_c(t)$ and $s_{2c}(t) = M_h m_2[b,a] W_g \eta_c(t)$ respectively. Then the *cross spectrum* reads

$$CS(b,a) = m_{1c}(b,a)\bar{m}_{2c}(b,a) \tag{3.20}$$

In general, this is a complex function and may be decomposed into amplitude and phase:

$$CS(b,a) = |CS(b,a)| \exp(i\Phi(b,a)) \tag{3.21}$$

The cross spectrum denotes the covarying power of two processes, i.e. the predictive information between each other. Possibly superimposed independent noise only appears in the single spectra but not in the cross spectra; this also implies that the cross spectrum vanishes for two independent processes.

Coherence

The *coherence* is defined as the modulus of the cross spectrum, normalized to the single spectra. Exhibiting values between zero and one, it denotes the linear relationship between two processes. In general, one rarely finds perfect linear dependence; the single processes $m_1(b,a)$ and $m_2(b,a)$ rather consist of covarying parts $m_{1c}(b,a)$ and $m_{2c}(b,a)$ and superimposed independent contributions m_{1i} and m_{2i}: $s_1(t) = M_h m_{1c}[b,a] W_g \eta_c(t) + M_h m_{1i}[b,a] W_g \eta_{1i}(t)$ and $s_2(t) = M_h m_{2c}[b,a] W_g \eta_c(t) + M_h m_{2i}[b,a] W_g \eta_{2i}(t)$. Then the squared coherence reads

$$COH^2(b,a) = \frac{|CS(b,a)|^2}{S_1(b,a)S_2(b,a)} = \frac{|m_{1c}(b,a)\bar{m}_{2c}(b,a)|^2}{|m_1(b,a)|^2\,|m_2(b,a)|^2}. \tag{3.22}$$

with $m_1(b,a) = m_{1c}(b,a) + m_{1i}(b,a)$ and $m_2(b,a) = m_{2c}(b,a) + m_{2i}(b,a)$.

---

[3] Note that the individual process and its realizations depend on the wavelet chosen for the construction.

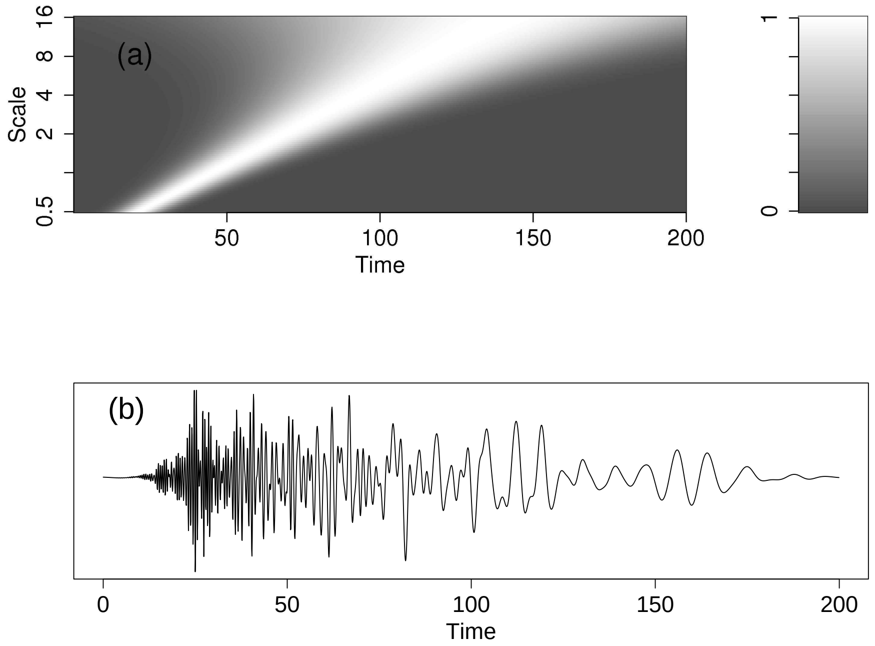

Figure 3.3: Stochastic chirp with $\epsilon = 0.3$. (a) The spectrum $|m(b,a)|^2$. (b) A typical realization in the time domain, calculated with a Morlet Wavelet, $\omega_0 = 6$.

3.4.4 Example

To illustrate the concept, we define a linear stochastic chirp given by

$$m(b,a) = \exp\left(-\frac{(b-b_0(a))^2}{2\sigma^2(a)}\right) \tag{3.23}$$

with $b_0(a) = \beta_0 + c\log(a)$ and $\sigma(a) = \sigma_0 a^{1-\epsilon}$, i.e. every voice is given by a Gaussian with time position and width varying with scale. The center of the Gaussian at time a is given by $b_0(a)$, the width as $\sigma(a)$, determined by the constants β_0, c and σ_0. The power of $1-\epsilon$ ensures the process exhibiting the desired asymptotical behavior Eq. (3.16). Figures 3.3(a) and (b) show the spectrum $S(b,a) = |m(b,a)|^2$ and a typical realization in the time domain, respectively.

3.5 Estimating Wavelet Spectra

Given a time series, inferring the wavelet spectrum of the underlying process is an estimation problem, i.e. one has to formulate estimators for the spectral measures presented in the previous section. We study these estimators with respect to variance and bias, especially in the presence of averaging.

3.5.1 Spectral Estimators

Wavelet Sample Spectrum

Given a realization $s(t)$ of a nonstationary process defined by the wavelet multipliers $m(b,a)$ and the wavelet pair $g(t)$,$h(t)$, one can estimate its spectrum (i.e. calculate the *wavelet sample spectrum*) using a wavelet $g(t)$ by

$$\widehat{S}_g(b,a) = A(|W_g s(t)|^2), \tag{3.24}$$

where $A(.)$ denotes an averaging operator defined later in Sec. 3.5.2. Following the terminology of Fourier analysis, the wavelet sample spectrum without averaging is either called *scalogram* or *wavelet periodogram*.

Wavelet Sample Cross Spectrum

Given realizations $s_1(t)$ and $s_2(t)$ of two processes, the cross spectrum can be estimated as

$$\widehat{CS}_g(b,a) = A(W_g s_1(t)\overline{W}_g s_2(t)), \tag{3.25}$$

or decomposed into amplitude and phase,

$$\widehat{CS}_g(b,a) = |\widehat{CS}_g(b,a)| \exp(i \arg(\widehat{CS}_g(b,a))). \tag{3.26}$$

Wavelet Sample Coherence

The squared coherence is estimated as

$$\widehat{COH}^2_g(b,a) = \frac{|\widehat{CS}_g(b,a)|^2}{\widehat{S}_{g,1}(b,a)\widehat{S}_{g,2}(b,a)}. \tag{3.27}$$

For the sample coherence, averaging is essential. Already in Fourier coherency, when considering only one frequency bin the statistical concept of coherency breaks down: Two pure harmonic oscillations of identical frequency are always perfectly coherent. For wavelet coherency, the problem is even more obvious: Investigating power in a single point in time and scale prevents any information about the oscillation and thus covariance. Consequently, nominator and denominator become equal and one obtains a trivial value of one for any two processes.

Global Wavelet Spectrum and Fourier Spectrum

Given a time series $s(t)$ of a stationary process with a Fourier periodogram $|f(\omega)|^2 = |\hat{s}(\omega)|^2$, the wavelet transformation of $s(t)$ in the Fourier domain reads [40]:

$$W_g(b,a) = \frac{\sqrt{a}}{2\pi} \int_{-\infty}^{\infty} \bar{\hat{g}}(a\omega) e^{ib\omega} f(\omega) d\omega \tag{3.28}$$

Here, $\hat{.}$ denotes the Fourier transform. The essential message of this equation is the following: A (sample) spectrum $|f(\omega)|^2$ with a well localized peak (e.g. the spectrum of a sine function) always gets smeared out to the width of the wavelet $\bar{\hat{g}}(a\omega)$ in Fourier domain. The better the wavelet is localized in the frequency domain, the lower is the broadening effect.

The *global wavelet sample spectrum* $\hat{G}_g(a)$ of a signal $s(t)$ with respect to the wavelet $g(t)$ is defined as the integral of the wavelet sample spectrum $\widehat{S}_g(b, a)$ over time

$$\hat{G}_g(a) = \int_{-\infty}^{\infty} \widehat{S}_g(b, a) db \tag{3.29}$$

Following Eq. (3.28), the global wavelet spectrum always smears out peaks to the width of the wavelet in frequency domain on the particular scale, $\hat{g}(a\omega)$. Thus, contradicting earlier propositions [121], the global wavelet spectrum is in general not an unbiased estimator of the Fourier spectrum.

3.5.2 Distribution, Variance and Bias of the Estimator

Qiu and Er [92] have studied the variance and bias of wavelet spectral estimates for deterministic periodic oscillations corrupted by white noise. For practical applications, however, these investigations are of rather theoretical interest: Many real world systems are barely represented by periodic but rather mixing or stochastic processes.

Distribution of the Wavelet Sample Spectrum

The Fourier periodogram $|f(\omega)|^2$ and also the Fourier cross periodogram $f_1(\omega)\bar{f}_2(\omega)$ asymptotically obey a χ^2-distribution with two degrees of freedom[4]; the variance equals two times the corresponding mean, $\text{Var}_S = 2 < |f(\omega)|^2 >$ [9, 90]. These relations also hold (at least approximately) for the localized wavelet analysis: The wavelet transformation of a Gaussian process sums up Gaussian random numbers. Hence, the wavelet scalogram $|W_g s(t)|^2$ and cross scalogram $W_g s_1(t)\bar{W}_g s_2(t)$ approximately follow a χ^2-distribution with two degrees of freedom[5], with variance $\text{Var}_S = 2 < |W_g s(t)|^2 >$.

Since the Fourier periodogram is asymptotically uncorrelated, averaging neighboring frequencies results in a χ^2 distribution with a number of degrees of freedom determined by the length of the averaging window. This simple relation is e.g. important for the straight forward analytical calculation of a significance test for the Fourier coherency [9, 90]. However, for the localized wavelet analysis, this result does not hold: As neighboring wavelet coefficients are correlated due to the reproducing kernel (see Sec. 3.2.5), reducing the variance by kernel averaging with the operator $A(.)$ destroys the simple χ^2-distribution [67] (see App. B.3.3). This drawback of wavelet analysis is the reason, why all quantiles for significance testing will have to be estimated by Monte Carlo simulations.

Asymptotic Variance of the Wavelet Sample Spectrum

The variance $\text{Var}_A(a)$ of the averaged wavelet sample spectrum scales with the number of effective data points $N_{\mathit{eff}}(a)$ that are averaged on a scale a:

$$\text{Var}_A(a) \sim \frac{1}{N_{\mathit{eff}}(a)}. \tag{3.30}$$

---

[4] This holds for $f(\omega)$ being complex. For $f \in \mathbb{R}$, the periodogram would obey a χ^2-distribution with one degree of freedom.

[5] Here, analyzing a real $s(t)$ with a real wavelet would lead to a χ^2-distribution with one degree of freedom.

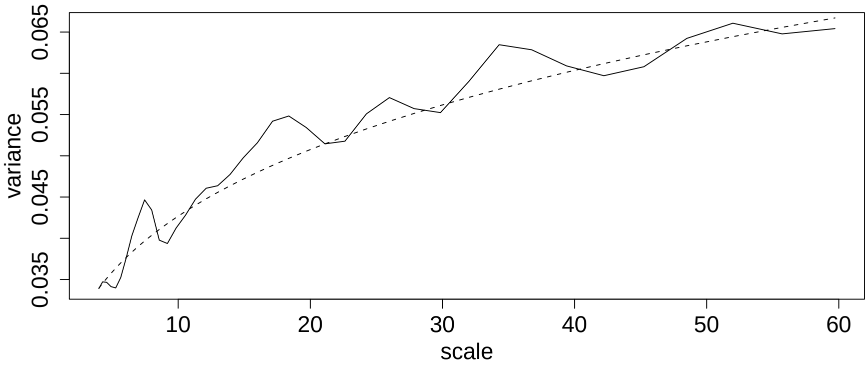

Figure 3.4: Asymptotic behavior of the variance of the averaged wavelet sample spectrum of Gaussian white noise. Solid line: estimation based on 1,000 realizations, dashed line: theoretically expected behavior $\sim a^{1-0.75}$.

The number of effective data points in turn is determined by the width of the averaging window $w(a)$ relative to the effective width of the reproducing kernel $w_{RK}(a)$,

$$N_{\mathit{eff}}(a) \sim \frac{w(a)}{w_{RK}(a)}. \tag{3.31}$$

The width of the reproducing kernel is given by the wavelet covariances (see App. B.1.1), i.e. $w_{RK}(a) \sim a$. If, firstly, the length of the scaling window is chosen to decrease slower than the width of the reproducing kernel, the variance vanishes for small scales, as more and more reproducing kernels (and thus effective data points) are averaged. If, secondly, the length of the scaling window additionally decreases faster than local structures of the process ($\sim a^{-1+\epsilon}$ according to Eq. (3.16)), the section of the process covered by the averaging window becomes more and more stationary and white, i.e. no bias is produced. Hence, if the width of the averaging window is chosen as $w(a) \sim a^{\alpha}$ with an exponent α with $1 > \alpha > 1 - \epsilon$, then the variance of the averaged wavelet sample spectrum decreases as

$$\mathrm{Var}_A(a) \sim a^{1-\alpha} \tag{3.32}$$

for small scales. Figure 3.4 shows the variance of the averaged wavelet sample spectrum of white noise ($\alpha = 0.75$): The solid line depicts the variance at one point in time estimated from an ensemble of 1,000 Gaussian white noise time series, the theoretically expected behavior is plotted as a dashed line. A similar picture would emerge when plotting the variance as a function of λ for the limit of stationary white noise according to Eq. (3.16).

Averaging in Practical Applications

The prior discussion appears to be rather academic: In general, one does not know, a priori, the asymptotical behavior for small scales of an observed process. Also, finite sampling restrains the investigation of small scales. Finally, one is often interested in nonstationary behavior on scales large compared to the sampling time. For practical applications, retaining a scale independent variance appears to be a reasonable choice. This might be accomplished

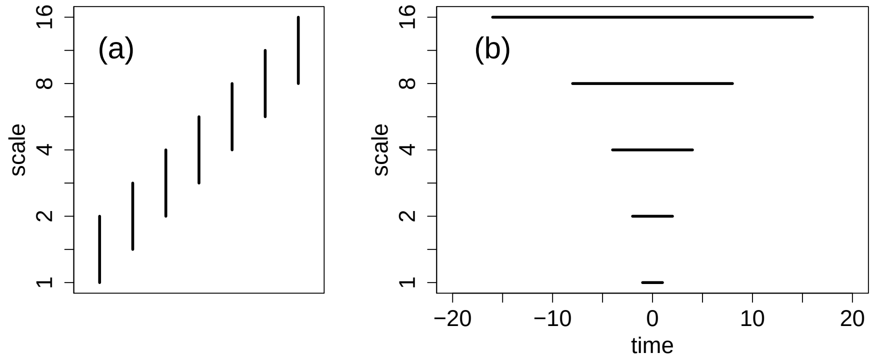

Figure 3.5: Smoothing according to the reproducing kernel to provide a constant variance for all scales. (a) In scale direction, the length of the smoothing window stays constant (for a logarithmic scale axis), $w_a = const.$ (b) In time direction, the length of the smoothing window increases linearly with scale, $w_b \cdot a$.

by averaging the same amount of independent information on every scale, i.e. by choosing the length of the averaging kernel according to the reproducing kernel. Following App. B.1.1 and Sec. 3.2.5 this means [67]:

- Averaging in scale direction should be done with a window exhibiting constant length on logarithmic scales (see App. B.1.6) as shown in Fig. 3.5(a). Here, w_a denotes the half window length in the same units as N_{voice}.
- Averaging in time direction should be done with a window exhibiting a length proportional to scale as shown in Fig. 3.5(b). Here, $w_b \cdot a$ denotes the half window length in units of time.

The (scale independent) variance of the wavelet sample spectrum of a Gaussian white noise realization as a function of the width of a rectangular averaging window is shown in Fig. 3.6. The graphs for averaging in scale as well as in time direction resemble the shape of the reproducing kernel. An averaging window short compared to the effective width of the reproducing kernel includes only a minor part of independent information and thus fails to notably reduce the variance. Figure 3.6 provides guidance for the choice of the averaging window length.

Bias of the Estimator

Given realizations of a Gaussian process defined by $m(b, a)$ and constructed with the wavelet pair $g(t)$ and $h(t)$, one can estimate the wavelet sample spectrum using a wavelet $k(t)$ and an averaging operator $A(.)$. The bias at scale a and time b of the wavelet sample spectrum then reads

$$\text{Bias}(\widehat{S}_g(b,a)) = \langle\, A(\,|\,\underbrace{W_k M_h}_{P_{h\to k}}\, m(b,a)\, W_g\eta(t)\,|^2\,)\,\rangle \;-\; |m(b,a)|^2. \tag{3.33}$$

$P_{h\to k}$ denotes the projector defined in App. B.1.2. The bias consists of two contributions: The averaging with the operator $A(.)$ produces an *averaging bias* of the smoothed wavelet

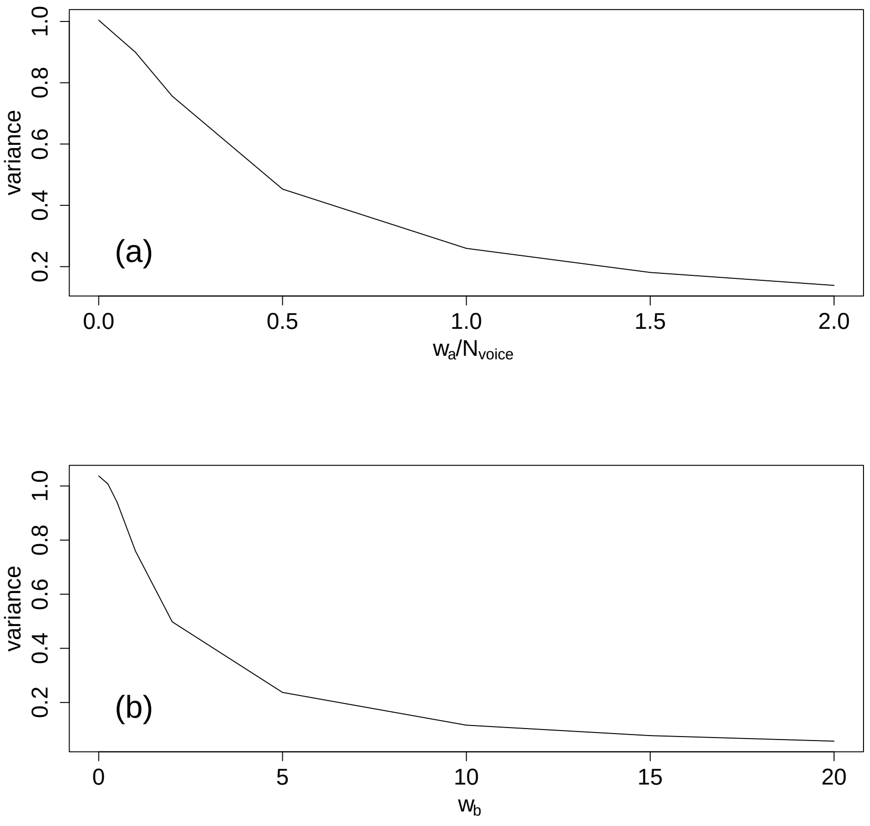

Figure 3.6: Scale independent variance of the wavelet sample spectrum of a Gaussian white noise realization as a function of the lengths of an rectangular averaging window. (a) Averaging in scale direction with half window length w_a. (b) Averaging in time direction with half window length $w_b \cdot a$. The graphs resemble the shape of the reproducing kernel.

sample spectrum in comparison to the wavelet periodogram. Furthermore, not even the wavelet periodogram is an unbiased estimator: The projection property App. B.1.2 results in an *inherent bias* of the wavelet periodogram in relation to the underlying spectrum $m(b, a)$[6]. For processes following Eqs. 3.16 or 3.17, the bias vanishes for small scales (See App. B.3.4); for averaging on finite scales, one has to consider the trade off between bias and variance. Both, the averaging bias and the inherent bias cause that the wavelet sample spectrum is no consistent estimator even in the limit of an infinite number of realizations. An alternative notion of consistency is only fulfilled in the rather theoretical asymptotic behavior given by Eqs. 3.16 or 3.17 in the limit of small scales or stationary white noise, respectively.

3.5.3 Example

We recall the example from Sec. 3.4.4. Figure 3.7(a) depicts the wavelet scalogram of the realization shown in Fig. 3.3(b). It is easy to see that a single realization without averaging

---

[6] Note, that this bias also emerges when the same wavelet $g(t)$ is used for the construction of $s(t)$ and the estimation

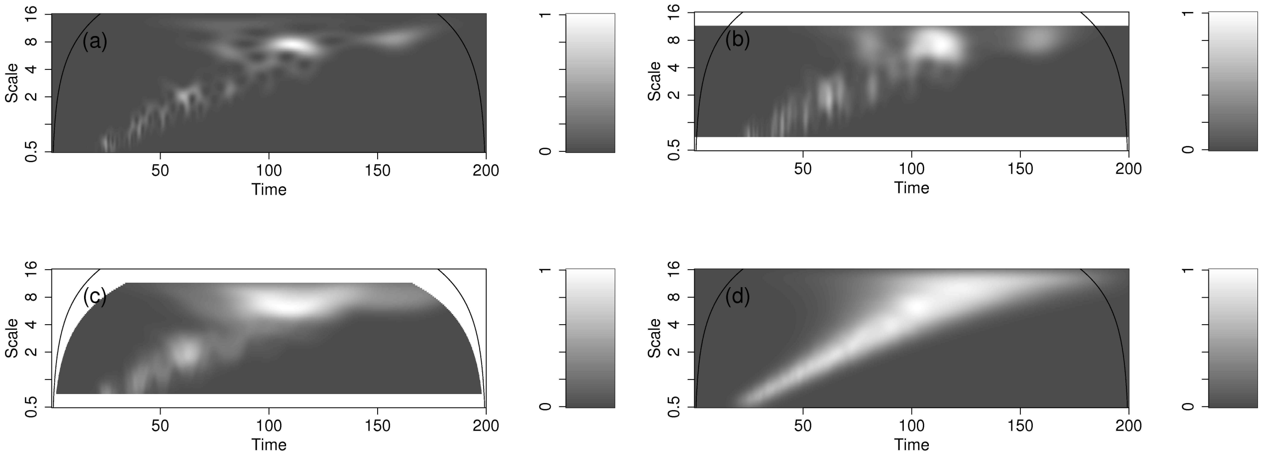

Figure 3.7: Estimation of the stochastic chirp based on the realization in Fig. 3.3(b). (a) the wavelet scalogram, i.e. the sample spectrum without averaging. (b) Averaged sample spectrum with $w_a/N_{voice} = 0.5$. (c) Averaged sample spectrum with $w_a/N_{voice} = 0.5$ and $w_b = 3$. (d) the spectrum estimated as the mean of 1,000 realizations without averaging.

yields a rather insufficient estimation of the real spectrum. Averaging, shown in (b) and (c), reduces the variance but produces a bias. The estimation based on 1,000 realizations without averaging (d) yields a pretty accurate result of the underlying process only subject to the inherent bias that decreases for small scales.

3.6 Significance Testing

Given a wavelet sample spectrum, one is often interested, if the estimate is compatible with some trivial background noise, or if it significantly deviates and thus represents some interesting features. This desire leads to the concept of statistical significance testing. In this section, we present the state-of-the-art significance tests for wavelet spectra. Based upon this test, we develop an areawise significance test that accounts for multiple testing effects. Both tests are compared in terms of sensitivity and specificity. Finally, we discuss the special problems coming along with significance testing of cross wavelet analysis.

3.6.1 Pointwise Testing of the Wavelet Spectrum

To our knowledge, Torrence and Compo [121] were the first to establish significance tests for wavelet spectral measures. They assumed a reasonable background spectrum for the null hypothesis and tested for every point in the time/scale plane separately (i.e. *pointwise*) whether the power exceeded a certain critical value corresponding to the chosen significance level (see Sec. 3.3.3). Since the critical values of the background model are difficult to be accessed analytically [67], they need to be estimated based on a parametric bootstrap:

- Choose a significance level α.
- Choose a reasonable model (e.g. an AR[1] process in case of climate data following Hasselmann [37]) as null hypothesis H_0 and fit it to the data.

- Estimate the $(1-\alpha)$-quantile S_{crit} (i.e. the critical value) of the corresponding background spectrum by Monte Carlo simulations. Depending on the chosen background model and the chosen normalization of the spectral estimator, the critical value in general depends on scale.
- Check for every point in the wavelet domain, whether the estimated spectrum exceeds the corresponding critical value. The set of all pointwise significant wavelet spectral coefficients is given as

$$P_{pw} = \{\, (b,a) \,|\, \widehat{S}_g(b,a) > S_{crit} \,\} \tag{3.34}$$

3.6.2 Areawise Testing of the Wavelet Spectrum

Multiple Testing, Inherent Correlations and Spurious Significance

The concept of pointwise significance testing always leads to the problem of multiple testing: Given a significance level α, a repetition of the test for N wavelet spectral coefficients leads to, on an average, αN false positive results (see Sec. 1.4). For any time/scale resolved analysis, a second problem comes into play: According to the reproducing kernel Eq. (3.10), neighboring times and scales of a wavelet transformation are correlated. Correspondingly, false positive results always occur as contiguous patches. These spurious patches reflect local spurious oscillations, which are randomly stable[7] for a short time.

For the inverse problem of interpreting data from a process with unknown spectrum, these effects mark an important problem: Which of the patches detected in a pointwise manner remain significant when considering multiple testing effects and the inherent correlations of the wavelet transformation? Referring back to Fig. 3.1(b) from Sec. 3.3 illustrates that a mere visual judgment based on a sample spectrum will presumably be misleading: Even in the case of white noise, the test described in Sec. 3.6.1 yields a large number of - by construction spuriously - significant patches.

Measuring Areawise Significance

We develop an *areawise test*, which utilizes information about the size and geometry of a detected patch to decide whether it is significant or not. The main idea is simple: If the inherent correlations are given by the reproducing kernel (Sec. 3.2.5), then also the typical patch area for random fluctuations is given by the reproducing kernel. Following the dilation covariance Eq. (B.2) and as illustrated in Fig. B.3, the typical patch width in time and scale direction grows linearly with scale.

However, investigating the wavelet spectral matrices Fig. 3.1(b) reveals that many spurious patches are not formed "typical" but rather arbitrary and complex. Patches might exhibit a large extend in one direction, but be very localized in the other direction (Patch A in Fig. 3.1(b)). Other patches might consist of rather small patches connected by thin "bridges" (Patch B in Fig. 3.1(b)). These patches are spurious even though their area might be large compared to the corresponding reproducing kernel. Thus, not only the area but also the geometry has to be taken into account.

Given the set of all patches with pointwise significant values, P_{pw} (see Eq. (3.34)), we define *areawise significant* patches in the following way: For every (a,b), we choose a *critical*

---

[7] As physicists, we would rather say coherent. We prefer to use the term stable to avoid confusion with the term coherence, which here refers to the interrelation between two processes

area $P_{crit}(b,a)$. It is given as the subset of the time/scale domain, where the reproducing kernel, dilated and translated to (b,a), exceeds the threshold of a certain critical level K_{crit}:

$$P_{crit}(b,a) = \{\, (b',a') \,|\, (K(b,a;b',a') > K_{crit} \,\} \tag{3.35}$$

Then the subset of additionally *areawise significant* wavelet spectral coefficients is given as the union of all critical areas that completely lie inside the patches of pointwise significant values:

$$P_{aw} \;=\; \bigcup_{P_{crit}(b,a)\,\subset\, P_{pw}} P_{crit}(b,a). \tag{3.36}$$

In other words: Given a patch of pointwise significant values, a point inside this patch is areawise significant, if any reproducing kernel (dilated according to the investigated scale) containing this point totally fits into the patch. Consequently, small as well as long but thin patches or bridges are sorted out as being insignificant.

Areawise Significance Level

The larger the critical area, the larger a patch needs to be to be detected by the test, i.e. the critical area is related to the significance level α_{aw} of the areawise test. We defined the latter one as follows: The characteristic functions of the pointwise and areawise patches, P_{pw} and P_{aw} read

$$\begin{aligned} \chi_{P_{pw}}(a,b) &= \begin{cases} 1 & \text{if} \quad (a,b) \in P_{pw}, \\ 0 & \text{otherwise} \end{cases} \\ \chi_{P_{aw}}(a,b) &= \begin{cases} 1 & \text{if} \quad (a,b) \in P_{aw}, \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{3.37}$$

Then the corresponding areas A_{pw} and A_{aw} result as

$$\begin{aligned} A_{pw} &= \int_{(a,b)} \frac{db\,da}{a^2} \chi_{P_{pw}}(a,b) \\ A_{aw} &= \int_{(a,b)} \frac{db\,da}{a^2} \chi_{P_{aw}}(a,b) \end{aligned} \tag{3.38}$$

Note, that on every scale a, the area is related to the corresponding measure a^2. We now define the *significance level of the areawise test* as

$$\alpha_{aw} = \langle \frac{A_{aw}}{A_{pw}} \rangle, \tag{3.39}$$

i.e. one minus the average ratio between the areas of areawise significant patches and pointwise significant patches.

The relation between the desired areawise significance level α_{aw} and the critical area P_{crit} of the reproducing kernel is rather non-trivial. As a matter of fact, we had to estimate the corresponding critical area P_{crit} as a function of a desired significance level α_{aw} by a root finding algorithm individually for every triplet (ω_0, w_a, w_b). The idea of this algorithm is outlined in App. B.4.1.

AR[1] parameter	0	0.1	0.2	0.5	0.9
P_{crit}	6.17 ± 0.36	7.01 ± 0.06	7.21 ± 0.36	6.94 ± 0.31	7.00 ± 0.14

Table 3.1: Critical area P_{crit} at scale $a = 1$ for different AR[1] processes for the wavelet scalogram with $\alpha_{pw} = 0.05$ and $\alpha_{aw} = 0.1$.

We investigated the dependency of the critical area P_{crit} on the process under investigation. Therefore, we applied the estimation procedure outlined above to different AR[1]-models representing different correlation times. The result for the wavelet scalogram with $\alpha_{pw} = 0.05$ and $\alpha_{aw} = 0.1$ is shown in Tab. 3.1. Due to its inherent stochasticity, the algorithm converges slowly to the global minimum and the estimates have got a large variance. However, no monotonous relation emerges, suggesting a negligible influence of the correlation length on the critical area.

Testing for Significant Areas

The actual areawise test is performed as follows:

- Perform the pointwise test according to Sec. 3.6.1 on the α level.
- Stretch the reproducing kernel for every scale according to Eq. (B.2), choose a significance level α_{aw} for the areawise test and the corresponding critical area $P_{crit}(b, a)$ of the reproducing kernel.
- Slide the critical area $P_{crit}(b, a)$ (for every scale the corresponding dilated version) over the wavelet matrix. Every point inside a patch is defined as areawise significant, if the critical area containing this point totally lays within the patch.

Figure 3.8 illustrates the areawise test based on the result of the pointwise test for the NINO3 index and the Gaussian white noise realization initially shown in Fig. 3.1. With $\alpha_{aw} = 0.1$, the areawise test is capable of sorting out about 90% of the spuriously significant area from the pointwise test. In case of the NINO3 data, remarkable structures remain, representing the major El Niño events. Some tempting patterns turn out to be indistinguishable from background noise. Finally, all the small patches have been sorted out. In case of the white noise realization, only a few of the numerous spurious patches are not rejected by the areawise test. Taking these two panels together, the areawise test obviously is capable of sorting out most false positive results and simultaneously detecting physical structures.

3.6.3 Sensitivity and Specificity of the Areawise Test

Real world processes, in particular of geophysical or physiological nature, often exhibit power on a wide range of scales, where only a narrow band of time localized oscillations might be of interest. The question arises, how strong the localization in time and scale might be in relation to the background noise to be in principle identifiable. This question addresses the sensitivity of the test. On the other hand, it is relevant to know, how many false positive results appear, when no significant structure is present. This question addresses the specificity of the test.

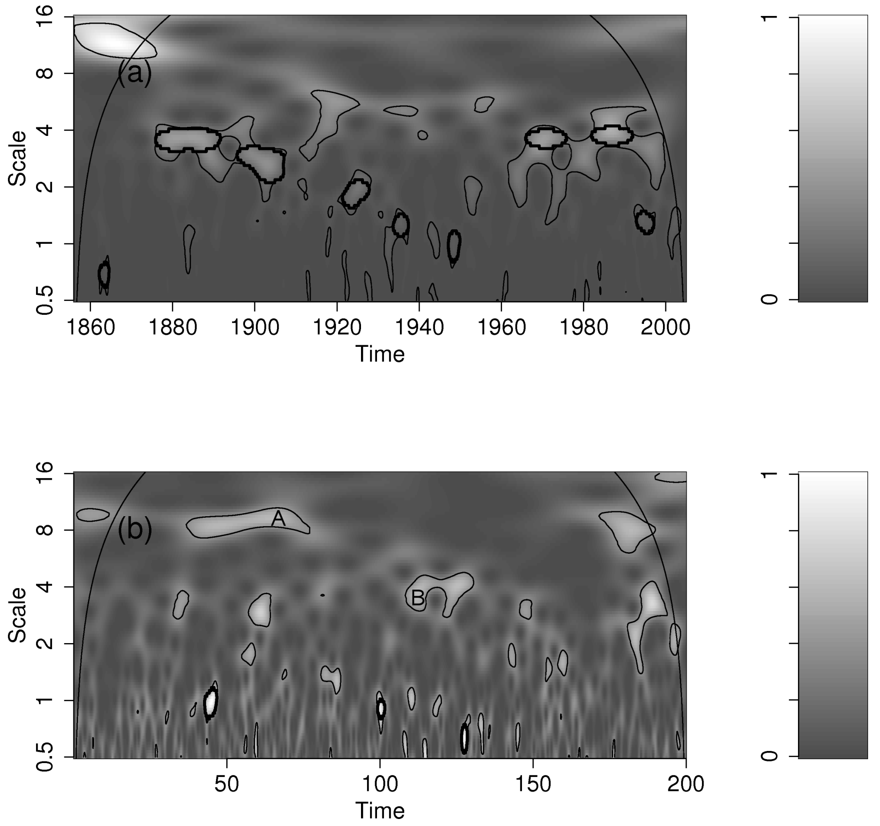

Figure 3.8: Areawise significance test performed on the examples from Fig. 3.1. (a) NINO3 index, (b) a realization of Gaussian white noise. Thin contour: pointwise significant patches, thick contour: areawise significant patches. Most of the by construction spurious patches are sorted out.

Definitions

To investigate these questions, we defined Gaussian bumps

$$\begin{aligned} m(b,a) &= \exp\left(-\frac{(b-b_0)^2}{2\sigma_b^2}\right) \\ &\cdot \exp\left(-\frac{(c\log(a)-c\log(a_0))^2}{2\sigma_a^2}\right), \end{aligned} \tag{3.40}$$

where b_0 and a_0 denote the center in time and scale respectively, whereas σ_b and σ_a define the width in time and scale direction. The logarithm of the scale provides a Gaussian bump in the typical logarithmic scale axis wavelet matrix. Realizations were calculated according to Eq. (3.14) with driving Gaussian white noise $\eta(t)$. The resulting time series was superimposed by independent background noise. As a simple model we chose Gaussian white noise $\xi(t) \sim \mathcal{N}(0,\sigma^2)$ with zero mean and variance σ_ξ^2. Figure 3.9 displays an example. The amplitude of the driving noise was chosen as $\sigma_\eta = 1$. However, the variance of the resulting bump is much lower (at the peak around $0.2\sigma_\eta$), as the bump is confined to a small spectral band. Thus,

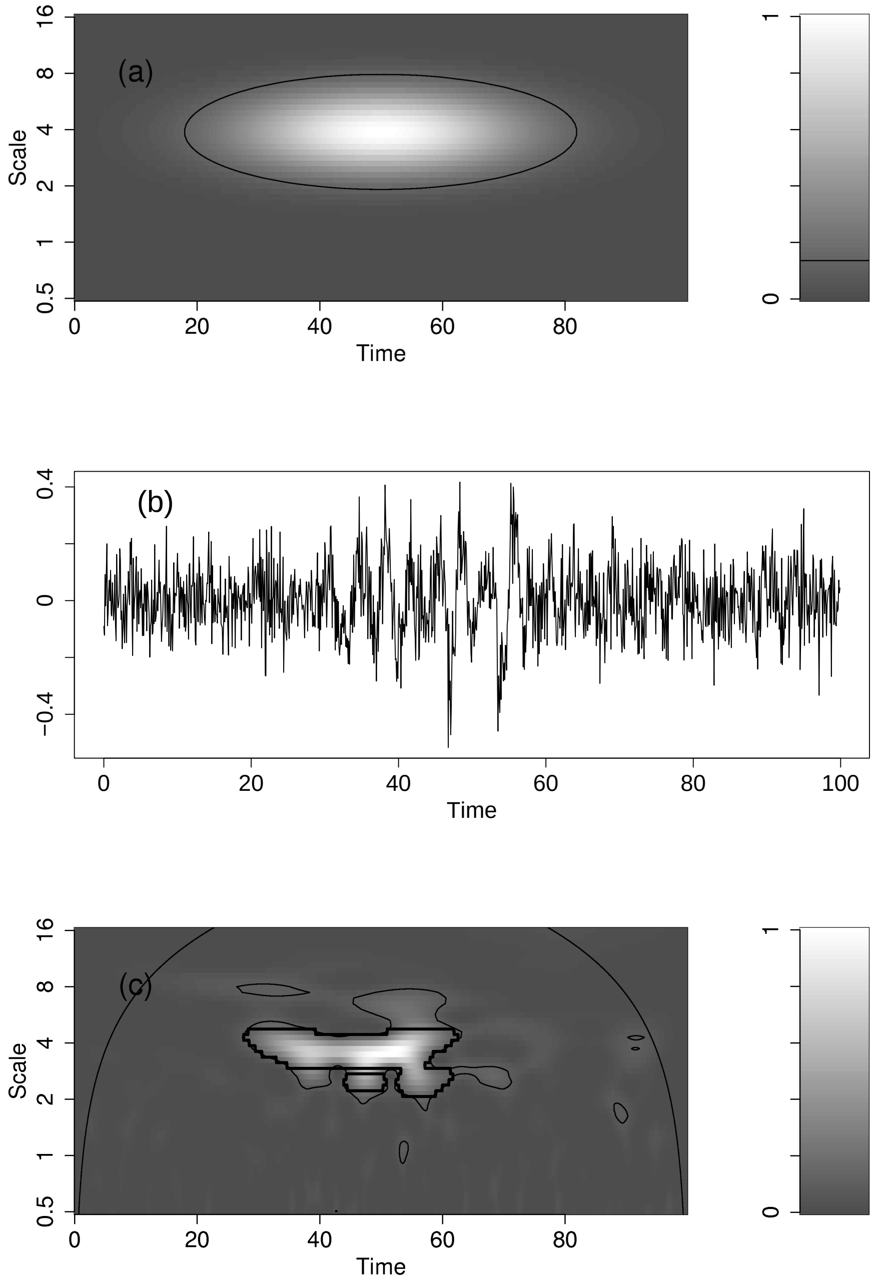

Figure 3.9: Gaussian bump with $b_0 = 50$, $a_0 = 4$, $\sigma_b = 16$ and $\sigma_a = 0.5$ superimposed by white noise. The variance of the driving noise was $\sigma_\eta = 1$, that of the background noise $\sigma_\xi = 0.1$. For details see text. (a) m(a,b), the contour-line marks $1/e^2$. (b) A realization in time domain using a Morlet wavelet with $\omega_0 = 6$. (c) The corresponding wavelet sample spectrum calculated using the same wavelet. Thin and thick lines surround pointwise and areawise significant patches, respectively.

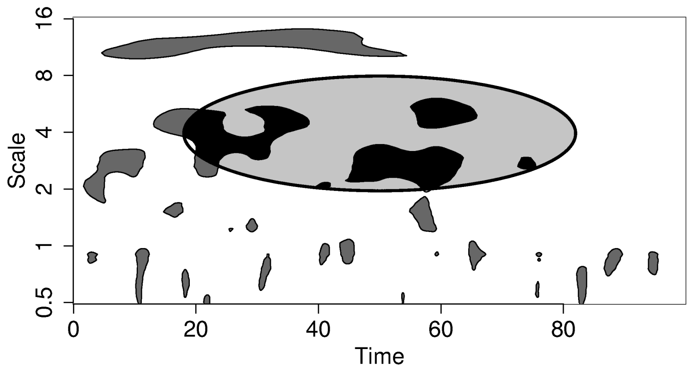

Figure 3.10: Pointwise test of a Gaussian bump with $b_0 = 50$, $a_0 = 4$, $\sigma_b = 16$ and $\sigma_a = 0.5$ superimposed by white noise. The variance of the driving noise was $\sigma_\eta = 1$, that of the background noise $\sigma_\xi = 0.2$. black: true positive (P_{TP}), white: true negative (P_{TN}), dark grey: false positive (P_{FP}), light grey: false negative (P_{FN}) patches.

the superimposed noise with $\sigma_\xi = 0.1$ represents a 50% noise-level in relation to the bump itself. Therefore, we define the signal to noise ratio at the peak as $SNR_{peak} = 0.2\sigma_\eta/\sigma_\xi$. For $\sigma_\xi = 0.2$, $SNR_{peak} = 1$. We performed the following study:

- We simulated Gaussian bumps of different widths σ_b and fixed $\sigma_a = 0.5$, superimposed by background noise with different variances σ_ξ. For each set of values (σ_b, σ_ξ), we simulated $N = 10,000$ realizations.
- To every realization, we applied the pointwise ($\alpha_{pw} = 0.05$) and subsequently the areawise test ($\alpha_{aw} = 0.1$). As null hypothesis we chose red noise fitted to the individual realization.

Based on this experiment, we compared the sensitivity and specificity of the areawise significance test to those of the pointwise test. We define the area of the bump P_B (i.e. the set of points where we assume H_0 as being wrong) and its complement P_{NB} as

$$\begin{aligned} P_B &= \{(a,b)\,|\,m(a,b) > 1/e^2\}, \\ P_{NB} &= \{(a,b)\,|\,m(a,b) \leq 1/e^2\}. \end{aligned} \tag{3.41}$$

The outcomes of the test we denote in the following way:

- True positive patches: $P_{TP} = P \cap P_B$ (black in Fig. 3.10),
- true negative patches: $P_{TN} = \bar{P} \cap P_{NB}$ (white),
- false positive patches: $P_{FP} = P \cap P_{NB}$ (dark grey),
- false negative patches: $P_{FN} = \bar{P} \cap P_B$ (light grey),

where P stands for either P_{pw} or P_{aw} and $\bar{P}$ denotes the complement. Using characteristic functions as in Eq. (3.37), we calculate the corresponding areas A_B, A_{NB}, A_{TP}, A_{TN}, A_{FP}

and A_{FN} as in Eq. (3.38). Then we can define the estimators

$$\begin{aligned} \widehat{SE} &= \frac{A_{TP}}{A_B} \quad \text{and} \\ \widehat{SP} &= \frac{A_{TN}}{A_{NB}}. \end{aligned} \tag{3.42}$$

Sensitivity

The sensitivity of the pointwise test is higher than that of the areawise test (see Tab. 3.2(a)), as the latter one sorts out small patches in the area of the bump. The sensitivity depends strongly on the signal to noise ratio.

For low background noise, $\sigma_\xi \ll \sigma_\eta$, both tests perform very well (Tab. 3.2(a)), although the part of the bump area not detected by the areawise test is around twice as large than that not detected by the pointwise test (because the areawise test sorts out small patches, Tab. 3.2(b))). As the noise-level increases to the order of the bump's driving noise, $\sigma_\xi \sim \sigma_\eta$, the sensitivity decreases rapidly. For a zero signal to noise ratio, $\sigma_\xi \gg \sigma_\eta$ (not shown), the sensitivity of both tests converges to the rate of purely randomly positive results, i.e. the sensitivity of the pointwise test converges to $\alpha_{pw} = 0.05$, that of the areawise test to $\alpha_{pw} \cdot \alpha_{aw} = 0.005$. However, the ratio between the parts of the area not detected by the two tests (Tab. 3.2(b)) converges to $(1 - \alpha_{pw})/(1 - \alpha_{pw} \cdot \alpha_{aw}) \approx 1 - \alpha_{pw} = 0.95$. In other words, for a very low signal to noise ratio, the performance of the pointwise test is not really better. It just detects patches that occur spuriously because of the dominant noise.

Trivially, small bumps nearly free from background noise are detected almost totally. This occurs because the small bumps are shorter than the reproducing kernel and thus get enlarged by the estimation. For large bumps, the sensitivity is in general lower. However, the decrease of the sensitivity with noise is much larger for small bumps than for large bumps. That means that small patches get rather invisible as they get superimposed by strong noise.

Specificity

The specificity of the areawise test is higher than that of the pointwise test (see Tab. 3.2(c)), as the latter one detects many more false positive patches outside the area of the bump. Whereas the specificity of the areawise test appears to be - almost independently of the signal to noise ratio - close to one, that of the pointwise test decreases for high background noise, as more and more spurious patches appear. For a zero signal to noise ratio, $\sigma_\xi \gg \sigma_\eta$ (not shown), the specificity is determined by purely randomly positive results only, i.e. for the pointwise test it converges to $1 - \alpha_{pw} = 0.95$, for the areawise test to $1 - \alpha_{pw} \cdot \alpha_{aw} = 0.995$.

At first sight, the difference between the two tests seems to be rather marginal, but taking into account the numbers of false positive results, an obvious difference arises: The ratio $A_{FP}(pw)/A_{FP}(aw)$ between the two tests for a high signal to noise ratio is ~ 1, as almost no false positive results appear in both cases. However, for a vanishing signal to noise ratio the ratio $A_{FP}(pw)/A_{FP}(aw)$ converges to $1/\alpha_{aw} = 10$. (Because of the small denominator, the estimated values are corrupted by a high uncertainty; the order of the values rather than the values itself is interesting). The specificity is - trivially - almost independent of the bump-width as it considers the area off the bump.

(a)		SNR_{peak}									
		∞		20		10		2		1	
		pw	aw	pw	aw	pw	aw	pw	aw	pw	aw
	2	0.95	0.89	0.83	0.66	0.92	0.84	0.58	0.30	0.26	0.06
	4	0.95	0.91	0.69	0.54	0.73	0.59	0.64	0.47	0.40	0.19
σ_b	8	0.76	0.66	0.65	0.53	0.61	0.47	0.59	0.44	0.35	0.19
	12	0.79	0.71	0.56	0.47	0.53	0.41	0.52	0.38	0.43	0.27
	16	0.71	0.62	0.58	0.46	0.31	0.20	0.45	0.31	0.39	0.23

(b)		SNR_{peak}				
		∞	20	10	2	1
	2	0.5	0.5	0.5	0.6	0.8
	4	0.5	0.7	0.6	0.7	0.7
σ_b	8	0.7	0.7	0.7	0.7	0.8
	12	0.7	0.8	0.8	0.8	0.8
	16	0.8	0.8	0.9	0.8	0.8

(c)		SNR_{peak}									
		∞		20		10		2		1	
		pw	aw	pw	aw	pw	aw	pw	aw	pw	aw
	2	0.98	0.98	0.99	0.99	0.98	0.99	0.93	0.99	0.94	0.99
	4	0.97	0.98	1.00	1.00	1.00	1.00	0.96	0.99	0.94	1.00
σ_b	8	0.99	0.99	1.00	1.00	1.00	1.00	0.97	1.00	0.95	1.00
	12	0.98	0.99	1.00	1.00	1.00	1.00	0.99	1.00	0.94	0.99
	16	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	0.94	0.99

(d)		SNR_{peak}				
		∞	20	10	2	1
	2	1.2	1.4	1.4	8.3	10.5
	4	1.2	1.6	1.5	7.9	11.2
σ_b	8	1.6	2.0	2.5	13.9	13.2
	12	1.5	2.3	3.6	16.6	9.8
	16	1.7	2.0	11.0	21.7	9.8

Table 3.2: (a) Sensitivity of the pointwise (pw) and the areawise (aw) test, (b) ratio of false negative results from pointwise test to areawise test $A_{FN}(pw)/A_{FN}(aw)$, (c) Specificity of the pointwise (pw) and the areawise (aw) test, (d) ratio of false positive results from pointwise test to areawise test $A_{FP}(pw)/A_{FP}(aw)$. The signal to noise ratio SNR_{peak} is given as the ratio between the signal level in the peak $0.2\sigma_\eta$ and the noise level σ_ξ (see also Fig. 3.9). All values are estimated based on 1,000 realizations of the corresponding bump.

Consequences

For data sets exhibiting a broad spectrum (i.e. a low signal to noise ratio), this study brings along important consequences: Many of the patches resulting from the pointwise patches are likely to be spurious. Applying the areawise test drastically increases the reliability of the interpretation.

Accounting for Distinct Localized Peaks

The areawise patch does not take into account the spectral value at a point (b, a); only information of the critical value contour-line is utilized to define the patch. Thus, a strongly localized patch formed by a high peak might be sorted out. This disadvantage is the cost for practically accounting for areal information. However, in principle there is a way to handle this problem by repeating the test for different pointwise significance levels α. The higher the level, the more localized patches might be identified.

3.6.4 Testing of Covarying Power

The Wavelet Cross Spectrum

Compared to testing the single wavelet spectrum, the inference of covarying power is rather non-trivial. Such as for the stationary Fourier cross spectrum and the covariance (its time domain counterpart), no significance test for the wavelet cross spectrum exists. Assume two processes exhibiting independent power at overlapping time and scale intervals. This power does not covary, i.e. information about one of the processes is not capable of predicting the other one. Hence, the real wavelet cross spectrum is zero. By contrast, the estimated wavelet cross spectrum always differs from zero. As it is not a normalized measure, it is impossible to decide whether a cross spectral coefficient is large because the one or the other process exhibits strong power or if actually covarying power does exist. Maraun and Kurths [67] illustrated this discussion and analyzed a prominent example. To overcome this problem, one normalizes the cross spectrum and tests against zero coherence.

Pointwise Testing of Wavelet Coherence

The structure of the test is similar to that developed for the wavelet spectrum, although two main differences exist. As the coherence is normalized to the single wavelet spectra, the critical value becomes independent of the scale as long as the smoothing is done properly according to Sec. 3.5.2, i.e. when the geometry of the reproducing kernel is accounted for.

In the case of Fourier analysis, the coherence critical value is independent of the processes to be compared, if they sufficiently well follow a linear description [9, 90]. This independency, however, holds exactly only in the limit of long time series. As wavelet analysis is a localized measure, this condition is not fulfilled. Using the Monte Carlo simulation based procedure as described in App. B.4.1, we investigated the process dependency of the critical values. Table 3.3 shows the estimated critical values for examples of different AR[1]-processes. The dependency on the smoothing parameters w_a and w_b can be seen comparing (a) and (b). The dependency on the process parameter a, however, is rather marginal; the difference of the critical values between the white noise case ($a = 0$) and the strongly correlated process ($a = 0.9$) is always less than 2%.

(a)	a	0	0.1	0.5	0.9
	10%-level	0.861	0.862	0.868	0.880
	5%-level	0.900	0.902	0.906	0.916
	1%-level	0.949	0.951	0.952	0.959
(b)	a	0	0.1	0.5	0.9
	10%-level	0.775	0.780	0.787	0.808
	5%-level	0.827	0.832	0.838	0.856
	1%-level	0.898	0.901	0.907	0.919

Table 3.3: Critical values of the squared wavelet coherence COH^2_{crit}, between two AR[1]-processes with identical parameters a, $x_i = ax_{i-1} + \eta_i$, for different significance levels and different a. Estimated for a Morlet wavelet, $\omega_0 = 6$ with (a) $w_a = 0.5$ octaves, $w_b = 0$ and (b) $w_a = 0.5$ octaves, $w_b = 1$.

Areawise Testing of Wavelet Coherence

Also here, an areawise test can be performed to sort out false positive patches being artefacts from time/frequency resolved analysis. The procedure is exactly the same as for the wavelet spectrum, only the critical patch-size $P_{crit}(b, a)$ has to be re-estimated. Areawise significant patches denote significant common oscillations of two processes. Here, common means that two processes exhibit a rather stable phase relation on a certain scale for a certain time interval.

However, this does not necessarily imply coherence in a strict sense. Processes oscillating on similar frequencies trivially exhibit patches indicating an intermittently similar phase evolution. The lengths of the patches are given by the decorrelation times of the single processes and the similarity of the concerned frequencies. Figure 3.11 illustrates this discussion. we simulated realizations of two AR[2] processes with slightly different parameters: $x_i = a_1 x_{i-1} + a_2 x_{i-2} + \eta_i$, with $a_1 = 1.95012$, $a_2 = -0.967216$ for the first and $a_1 = 1.95303$, $a_2 = -0.967216$ for the second process. With a sampling time $dt = 1/12$, this gives a common relaxation time of $\tau = 5$ and mean periods of $t_1 = 4$ and $t_2 = 4.4$, respectively. Even though the driving noise is independent, randomly common oscillations with a length related to the relaxation time and the difference in period occur.

If one is not only interested in deriving significant common oscillations, but also significant coherence in the sense of coupling between the processes, the areawise test has to be succeeded by another step: It has to be tested, if the time interval of the common oscillations is significantly long compared to typical randomly common oscillations of independent processes. In this study, we would like to restrain the discussion to outlining the basic idea: One has to construct a bootstrap ensemble representing the length distribution of randomly common oscillations of the two processes under the null hypothesis (i.e. independence). On the one hand, this can be realized by a parametric bootstrap, i.e. by fitting two sufficiently complex models to the data and then performing Monte Carlo simulations. Alternatively, one can apply a non-parametric bootstrap by constructing surrogate data of the two time series. A patch with a length exceeding a certain quantile of the length distribution then signifies coherence in a strict sense. For an overview about surrogate time series, see [112], for bootstrapping in general, see [17].

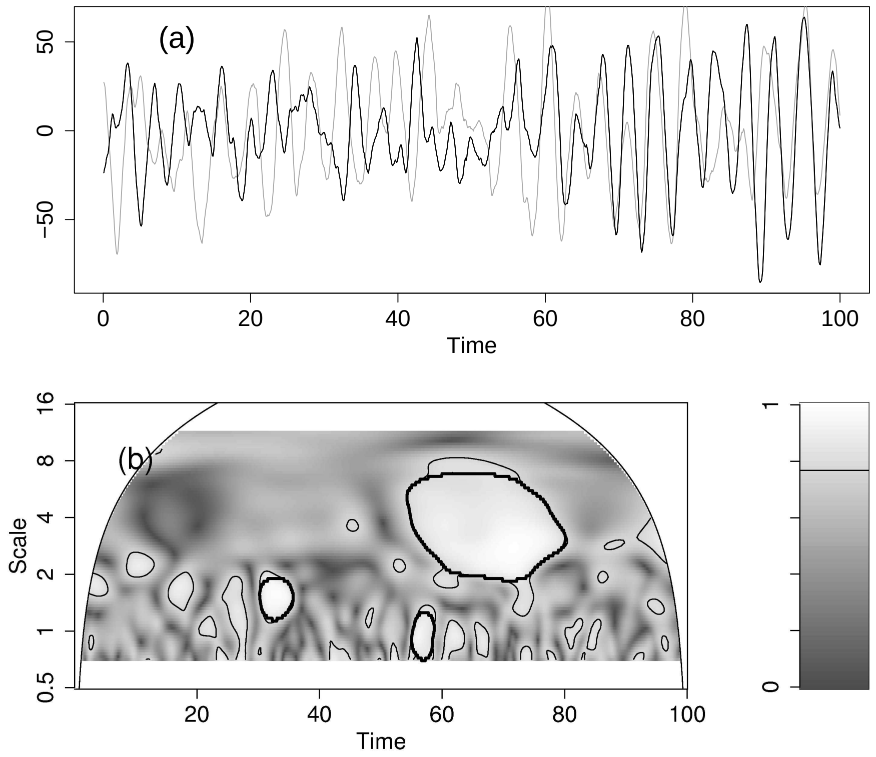

Figure 3.11: Areawise significance test for the coherence of two independent AR[2]-processes. (a) Time series, (b) wavelet coherence. Thin lines: pointwise test. Thick lines: Areawise test. On the main frequency of 1/4, randomly common oscillations produce a large false positive patch.

3.7 Applications

3.7.1 Wavelet Analysis of ENSO and NAO

In Sec. 3.3.3, we presented the wavelet cross spectral analysis between the NINO3 and the NAO index as performed by Huang et al. [45] and raised doubt about the reliability of the results. The related logical impossibility of performing a significance test against a zero non-normalized cross spectrum was discussed in detail in Sec. 3.6.4. In fact, nothing of the spurious structure visible in the cross spectral analysis (Fig. 3.2) is left after calculating the wavelet coherence instead and applying the areawise test, see Fig. 3.12. Apparently, the findings of Huang et al. [45] are a mere artefact of the non-normalized wavelet cross spectrum. Hence, an instantaneous influence from ENSO on NAO during strong El Niño events cannot be unambiguously concluded from continuous wavelet analysis.

However, even though wavelet coherency reveals no areawise significant result, we would like to comment on one of the pointwise significant patches. Between the mid-1930s and the mid-1940s a peak on the 4 years scale emerges. Because this patch is not areawise significant, it is not distinguishable from background noise. However, during this period other studies investigating spatio-temporal data find a significant coherent oscillation of ENSO and North

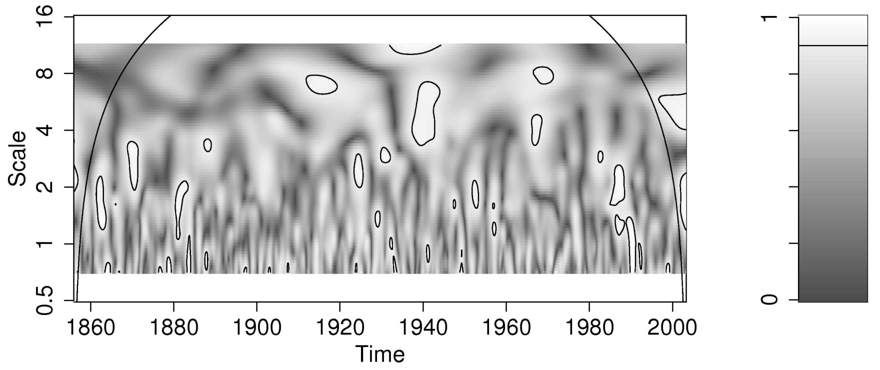

Figure 3.12: Squared wavelet sample coherency between the NINO3 index and the NAO index, calculated with a Morlet wavelet, $\omega_0 = 6$, $w_b = 0$, $w_a = 0.5/\text{octave}$, including the pointwise significance test (thin contour line). No patches appear to be areawise significant.

Atlantic climate [100]. This result exemplifies the limits of the areawise test: As confirmed by other researchers, a significant coupling between ENSO and NAO exists. The pointwise significance test is sensitive to detect the influence, but the areawise test shows that this patch is too small to be distinguished from noise. This decreased sensitivity is the cost for the capability of a reliable corroboration of a hypothesis by means of wavelet spectral analysis.

3.7.2 Wavelet Analysis in Hydrology

In hydrology, continuous wavelet transform became popular in different types of applications: In river regime characterization CWT is used to detect how discharge is related to climate variability indices [e.g. 62, 63] or to qualitatively analyze how certain features of the meteorological input time series are transferred to the hydrological system output [e.g. 27, 64]. This approach is also applied for the analysis of the hydrological functioning of karstic systems [e.g. 61]. Smith et al. [114] applied CWT to answer the question to which degree a catchment smoothes out the spatial variability of rainfall. Si and Zeleke [113] used wavelet coherency analysis to detect spatial scale and location dependent relationships between physical and hydraulic soil properties. The following applications exemplify the benefits gained by wavelet spectral analysis in hydrological studies [69].

Relation between Grand St. Bernard Temperature and Drance de Bagnes River Run-Off

Figure 3.13 displays the result for the relation between the monthly mean temperature at Grand St. Bernard and the monthly mean run-off of the closeby Drance de Bagnes river. Both time series are deseasonalized (for details, see App. A). After applying the areawise significance test, some interesting features remain: First, coherent deviations from the mean annual cycle are reflected in areawise significant patches at the one year scale, with a prominent occurrence in the mid-eighties. Furthermore a large significant peak shows up at the two year scale during the 1960's.

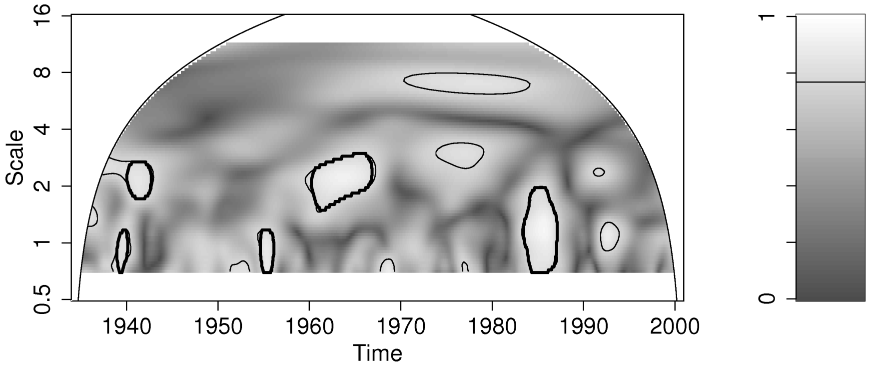

Figure 3.13: Squared wavelet sample coherency of the temperature at Grand St. Bernard vs. runoff at Drance de Bagnes river, calculated with a Morlet wavelet, $\omega_0 = 6$, $w_b = 1.5$, $w_a = 0.5$/octave, including the pointwise significance test (thin contour line) and the areawise significance test (thick contour line).

During the mid-eighties, snow accumulation in the Alps was above average during several winters [4], resulting in important snowpacks at the end of the accumulation season even at relatively low altitudes (some Alpine glaciers experienced visible surface increase during this period, [e.g. 36]). During the melt periods in the following summers, these seasonal snowpacks led to above average discharge, visible in the coherence peak in the annual cycle. In parallel, the snowmelt resulted in exceptional catchment saturation that in turn led to regular high flow events during important summer/autumn rainfall events. Accordingly, the same analysis applied to the maximum instead of mean monthly discharge (not shown here) shows the same significant peak at the same scale and location.

The peak at the two years scale in the 1960's results from a different mechanism: Several years during and preceding the peak were exceptionally cold, causing an above average snow accumulation [e.g. 4]. This can result in an increase of the elapsed time between the moment when water enters the hydrological system in form of snow and the moment when it leaves it as discharge: The seasonal snow storage property of the catchment extends on a longer temporal scale and a bi-annual cycle becomes temporarily superimposed on the normally strong annual discharge cycle.

Influence of NAO on Bourg St. Pierre Precipitation

The application of the areawise test to the wavelet sample spectrum of the precipitation at Bourg St. Pierre sorts out many false positive results (see Fig. 3.14); the result is barely distinguishable from the red noise background spectrum. However, a broad peak around the year 1940 at a scale of 4 years is identified as being significant. Additionally, some deviation from the annual cycle is discovered.

The significant peak in the monthly maximum precipitation at Bourg St. Pierre at the 4 year scale in the beginning of the 1940's also emerges in the coherence analysis between the same data and the NAO index (see Fig. 3.15). In general, high NAO values effect higher winter temperatures in Europe. Westerly storm tracks are deflected to the north causing higher than normal precipitation in Northern Europe and lower than normal precipitation in Southern

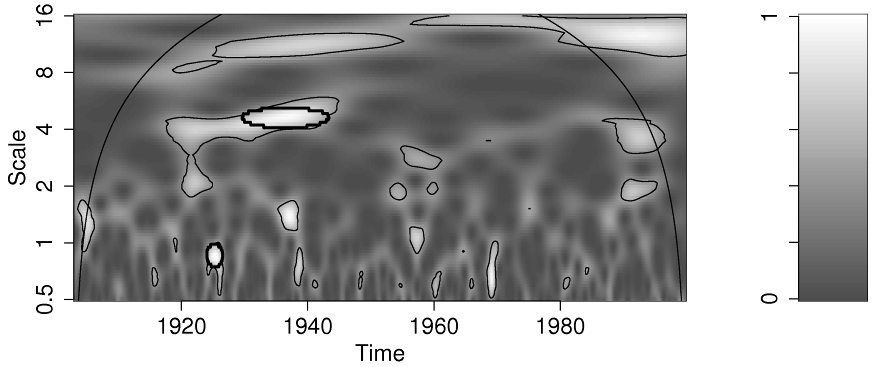

Figure 3.14: Wavelet sample spectra of the precipitation at the station Bourg St. Pierre calculated with a Morlet wavelet, $\omega_0 = 6$, without averaging, including the pointwise significance test (thin contour line) and the areawise significance test (thick contour line).

Europe [e.g. 47, 72, 127, 129]. In the years 1938/1939, the NAO exhibited relatively high values, leading to warm winters in Europe. The precipitation in Bourg St. Pierre during these years was relatively low. During the following two years, 1941 and 1942, NAO decreased to a distinct minimum, accompanied by severely cold European winters and high precipitation in Bourg St. Pierre. Afterwards, the phenomenon reversed again. This interannual oscillation exhibits a period of 4 years and might explain the significant peak in the wavelet sample spectrum of the precipitation at Bourg St. Pierre as well as that in the coherence analysis between the precipitation data and the NAO index.

3.8 Summary and Conclusions

Continuous wavelet transformation is a powerful and well studied mathematical instrument. However, the formulation of continuous wavelet spectral analysis so far had a rather tentative character. Open questions concerned the definition of nonstationary spectra in the wavelet domain by means of process realizations. Based on the previous measures, it was impossible to generate realizations of a particular nonstationary Gaussian process. The bias and variance of the wavelet sample spectrum had not been studied. Further problems were the effect of multiple testing in the conventional pointwise significance tests and the interpretation of cross wavelet analysis.

To meet these shortcomings, we presented a framework to define nonstationary Gaussian processes in wavelet domain: A driving Gaussian noise is transformed to the wavelet domain, multiplied with wavelet multipliers defining the desired nonstationary spectrum and finally transformed back to the time domain with a suitable reconstruction wavelet. These processes are completely defined by its wavelet multipliers and the wavelets used for the construction. We defined the process behavior in such a way that the dependency on the wavelets vanishes for small scales. This concept elegantly allows to define wavelet spectra and wavelet cross spectra. For the stationary case, these spectra are closely related to Fourier spectra.

Based on this framework, we investigated the variance and bias of continuous wavelet spectral estimation and developed an areawise significance test to overcome multiple testing

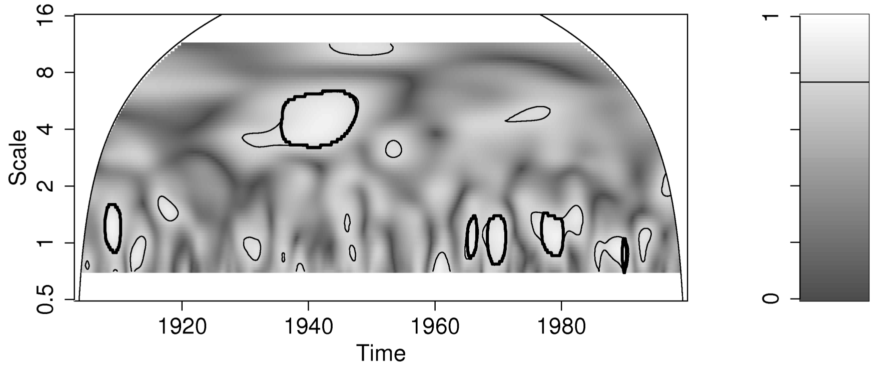

Figure 3.15: Squared wavelet sample coherency of the precipitation at Bourg St. Pierre vs. NAO-index calculated with a Morlet wavelet, $\omega_0 = 6$, $w_b = 1.5$, $w_a = 0.5$/octave, including the pointwise significance test (thin contour line) and the areawise significance test (thick contour line).

effects occurring in pointwise testing. Here, the existence of an extended reproducing kernel turned out to be an important property of time and frequency resolved analysis in general and wavelet spectral analysis in particular: It represents inherent correlations of any wavelet transformation and defines a minimum size of time and scale resolved structures. Its area scales quadratically with the investigated scale.

For kernel averaging to reduce the variance of a wavelet sample spectrum, the extension of the averaging kernel has to be chosen corresponding to the reproducing kernel on every scale. Otherwise, the variance and bias will change with scale. The reproducing kernel also gives a reasonable minimum length of the averaging kernel for effectively reducing the variance. In this context, we showed that the global wavelet spectrum is in general not an unbiased estimator of the Fourier spectrum.

Due to multiple testing effects and the reproducing kernel, significance testing has to be performed in two steps: A pointwise test investigates for every wavelet coefficient separately, if it exceeds a certain critical value. However, pointwise significant patches might be an artefact resulting from a combination of multiple testing and the inherent correlations given by the reproducing kernel; even wavelet spectra of white noise show typical spurious patches. Thus, a subsequent areawise significance test has to assess whether a patch exceeds a critical size given by the reproducing kernel. Smaller patches are in principle indistinguishable from noise.

We studied the areawise significance test in comparison to the conventional pointwise significance test in terms of sensitivity and specificity. As it sorts out patches small in relation to the reproducing kernel, the areawise test is less sensitive but more specific. Given observations with a broad spectrum, e.g. from geophysics or physiology, the conventional test mimics a misleading structure that is successfully uncovered by the areawise test. This two level significance test provides a reliable measure to corroborate a proposed hypothesis.

For the testing of coherency, even a third step needs to be developed: Patches "surviving" the areawise test signify a common oscillation on a certain scale for a certain time interval. However, common does not mean coherent in the sense of coupling. Processes exhibiting oscillations on similar frequencies trivially show patches of a certain length given by the decorrelation times of the single processes. Thus, to infer coherence in a strict sense, it needs to be tested, whether the patch is long in relation to typical random in-phase oscillations. Significance testing of the non-normalized wavelet cross spectrum turned out to be not feasible.

Contributions to climatological discussions illustrate the benefits and limits of wavelet spectral analysis: We repeated the analysis of Huang et al. [45] for the relation between ENSO and NAO. We could show that the proposed instantaneous influence from ENSO on NAO is a mere artefact of the non-normalized cross spectrum. However, this analysis also illustrates the limits of wavelet spectral analysis: A known correlation during a certain period is visible in the wavelet coherency, but not distinguishable from background noise. The analyses of hydrological time series from the Swiss Alps are typical for applications of wavelet spectral and cross spectral analysis in hydrology and related areas. Current hydrological wavelet studies (e.g. [63, 113]), however, still result in rather qualitative conclusions. The application of the areawise significance test to the real-world time series showed the benefit that hydrological time series analysis can gain from such methods. Quantitative spectral analysis could become a valuable instrument for data selection in model calibration or for the intercomparison of discharge regimes. Another interesting field of application is the assessment of the quality of hydrological model outputs: The areawise significance test enables to test whether the simulated spectrum is significantly different from the observed.

The presented framework opens a wide field of possible extensions and applications. The areawise test might be extended by a voicewise test: In some cases, where stationary variance exists on a certain scale (e.g. the annual cycle), the resulting wavelet sample spectrum exhibits a chain of short patches. Some of these patches will likely be too short to be areawise significant. Here, a test might investigate if the whole voice[8] on average exhibits significant power.

The definition of nonstationary Gaussian processes by means of wavelet multipliers is prototypically suitable for nonparametric bootstrapping in wavelet domain. Given a particular realization of a process, one can estimate its underlying wavelet spectrum that in turn can be used as wavelet multipliers to generate nonstationary surrogate data with almost identical spectral properties. This approach, however, needs further investigation. As shown in the example in Sec. 3.5.3, a single wavelet periodogram exhibits a huge variability and provides a very poor estimate of the underlying spectrum. Consequently, one has to average the result. Here, the trade-off between bias and variance has to be studied in detail.

Aside from the construction of nonstationary surrogate data, this approach allows to perform significance testing with a more complex nonstationary background spectrum. Often, the variance of the underlying process is superimposed by a trend. Testing against a stationary background spectrum would cause many false positive results in regions of high variance and many false negative results in regions with low variance. This problem can be solved by constructing a nonstationary background spectrum by means of wavelet multipliers.

Since wavelet coherence provides time and scale resolved information about the phase difference between two processes, one can formulate measures for phase coherence in the wavelet domain. In principle, it is also possible to investigate for $n : m$ phase synchronization (see the concepts section of the next chapter) by comparing different scales.

Many real world processes exhibit distributions far away from being Gaussian. Applying

---

[8] i.e. a stripe of constant scale

the significance test with a Gaussian background spectrum might not be suitable in this case and produce a lot of false positive results. Thus, an extension of the presented framework to non-Gaussian processes is required.

Chapter 4

Phase Coherence between ENSO and the Indian Summer Monsoon

4.1 Introduction

El Niño/Southern Oscillation (ENSO) and the Indian summer monsoon (ISM, below also referred to as monsoon) are the predominant climate phenomena in the Asian/Pacific region exhibiting oscillations on inter-annual scale with a large social and economical impact. While ENSO exhibits self-sustained oscillations of the tropical Pacific coupled ocean-atmosphere system, the monsoon performs oscillations driven by the annual cycle of the land vs. sea surface temperature gradient. For detailed reviews of ENSO and ISM refer to [86] or [10] and [132] or [25], respectively.

ENSO and the ISM have been known to be correlated on inter-annual time scales since the pioneering work of Walker at the beginning of the last century [e.g. 129]. He found a relationship of weak (strong) monsoons following low (high) values of the Southern Oscillation index, i.e. monsoon failure coinciding with El Niños and strong monsoons with La Niñas, respectively. This coupling faded during the second decade of the 19th century, only strengthening again in the 1960s. Since then, intensive studies have investigated the time dependency of the coupling between the two processes [e.g. 131, 123]. Recent work finds a weakening relation since the 1980s [60, 108]. Much attention has been concentrated on understanding the coupling mechanisms such as Pacific SST and Walker circulation anomalies as well as Eurasian snow-cover [e.g. 57, 102, 135].

The previous results are all based on linear time series analysis, either sliding correlations or wavelet coherence as discussed in the previous chapter. In the context of phase synchronization [104], the method of Hilbert phase analysis has been widely used to analyze coupling between self sustained oscillators. This method decomposes oscillation dynamics into time dependent amplitude and phase, making it possible to investigate the time resolved phase difference of the oscillators, irrespective of their amplitudes. Because the approach provides no frequency resolved information[1], the time resolution is optimized.

This chapter deals with the application of the Hilbert phase analysis to the ENSO monsoon coupling. By investigating for phase coherence, we aim to get a deeper insight than possible with linear methods and hence to contribute to current climatological discussions, for instance to the question, if the recent weakening of the coupling is without precedent.

---

[1] By calculating the phase derivative, only the instantaneous frequency is given

Furthermore, we intend to improve the accuracy of earlier findings and to provide ideas for the coupling mechanism. To account for typical interdecadal variability of the tropical ocean atmosphere system, we combine this method with a recently proposed geometrically motivated filter [79]. Before applying the method to the real data, we evaluate it with a simple toy model and study its behavior in terms of bias and variance. The results reproduce the findings of Torrence and Webster [123]. However, it turns out that Hilbert phase analysis provides a better time resolution and detects additional intervals of coupling invisible to linear methods. To corroborate the existence of two epochs of phase coherence, we develop a significance test of high specificity that considers the length distribution of intervals of random phase coherence. Finally, we discuss the possibility of a volcanic influence on the coupling.

In Sec. 4.2, we briefly review the physics of ENSO and monsoon and their coupling. In the subsequent section, we raise a set of questions we aim to discuss in the following study. Section 4.4 introduces the data. In Sec. 4.5, we present and evaluate the methodology we use to analyze the data. The analysis itself is presented in Sec. 4.6, the results are given in Sec. 4.7. The climatological relevance of our findings is discussed in Sec. 4.8, followed by a summary and conclusions.

4.2 Sketching the Physics of ENSO and Monsoon

In this study, the coupling between ENSO and monsoon takes center stage and the physics of the single processes are rather relegated to black boxes. However, to get an idea of possible interactions, it is nevertheless necessary to sketch the basic processes underlying ENSO and monsoon.

4.2.1 ENSO

ENSO is a self-sustained oscillator comprising the coupled tropical ocean atmosphere system. Bjerknes [8] described the basic mechanism almost 40 years ago: A sea surface temperature (SST) gradient between the western Pacific warm pool east of Indonesia and the eastern Pacific cold tongue off the Peruvian coast induces the easterly winds of the Walker circulation. These winds in turn stabilize the temperature gradient by driving surface water masses westwards and upwelling of cold deep water close to South America. Perturbations, induced by the annual cycle and noise, excite so-called Kelvin and Rossby waves. They propagate westward and eastward, respectively, and produce variability on interannual timescales. Every two to seven years, the Walker circulation exceptionally relaxes and the temperature gradient between east and west Pacific almost vanishes. This phenomenon sets on in early summer and usually peaks during winter time, giving it the name El Niño (Spanish: Christ child). Such events are accompanied by devastating rainfall in Peru and Ecuador, having a severe socio-economic impact. El Niño events are usually followed by a stronger than normal Walker circulation and temperature gradient, called La Niña.

Zebiak and Cane [134] developed the first idealized coupled ocean atmosphere model that was at least in principle capable of reproducing the tropical Pacific interannual variability. This model has been reduced to several conceptual models, following either the delay-oscillator [116] or the recharge oscillator paradigm [48] (for an overview please refer to the book of Dijkstra [19]). The equations of a noise driven delay oscillator we have studied in

Rainfall	Total	El Niño	La Niña
Below average	53	24	2
Above average	71	4	19
Deficient	22	11	2
Heavy	18	0	7

Table 4.1: Relation between ISM and ENSO. Given are the total numbers of monsoon years of a particular strength, and those falling into an El Niño or La Niña year (taken from Webster et al. [132])

another context [110] are listed in App. C.1. Philander [85] gives a comprehensive description of the ENSO phenomenon.

4.2.2 Indian Summer Monsoon

The annual monsoon winds are the large scale equivalent of the daily on- and off shore winds. The phenomenon arises from the different heat capacities of land and ocean. During winter, the land surface cools faster than the ocean surface, causing a pressure high on the Indian subcontinent and a pressure low on the Indian ocean. This gradient causes oceanward winds. During summer, however, the situation reverses: The land heatens up faster than the ocean, inducing a pressure low over India and a pressure high over the sea. Now the winds blow towards the Indian subcontinent, taking up moisture from the water surface that condensates over land and causes the important ISM rains. This system is a simple relaxator, forced by the annual cycle and characterized by a set of boundary conditions defining the amplitude of the monsoon. The most important are the position of the inter-tropical convergence zone (ITCZ), i.e. Indian ocean SST, the land surface temperature, which is influenced by the Tibetan snow cover, and the strength of the trade winds. The Indian agriculture heavily relies on the ISM rainfall. For a detailed review, please refer to Webster et al. [132] or Gadgil [25].

4.2.3 Coupling between ENSO and Indian Summer Monsoon

Since the pioneering works of Walker and Bliss [129], it is known that the Walker circulation and the Indian summer monsoon are correlated. Almost every El Niño event was preceded by a severe drought in India during summer[2], whereas many La Niña events were accompanied by strong Indian summer rainfall [132]. However, Tab. 4.1 shows that this teleconnection is not a one to one coupling, but rather a statistical relation influenced by other effects: On the one hand no heavy ISM rainfall and only four above average ISM rainfall events occurred during El Niño years. On the other hand, not all weak and deficient ISM coincide with El Niño. A similar relation holds for La Niña years. Hence the amplitude relation between the two processes is rather unstable.

As the strength and onset of the Indian summer monsoon is of great importance for Indian agriculture, several works investigate the predictability of the monsoon. Most of the monsoon predictors are manifestations of ENSO [59]. For a review of the ENSO monsoon teleconnection, please refer to [59, 131, 132]

---

[2] The lead of monsoon does not indicate a coupling direction from the monsoon towards ENSO, as the development of an El Niño event already starts in the preceeding summer and only culminates in winter.

Variability of the coupling

The link between ENSO and monsoon seems to be subject to variability on long time scales. This effect was first noticed by Walker and Bliss [130], who's forecast scheme happened to fail since the 1920s. Rasmusson and Carpenter [98] found an interdecadal variability of the coupling. The wavelet coherency analysis by Torrence and Webster [123] revealed a weakening of the relationship in the 1920s, then a strengthening in the 1960s, followed by a newly decrease in the 1980s. The significance of these and similar results based on correlation analysis is still under debate. Gershunov et al. [28] argues that a stable (true) correlation between two processes produces realizations that exhibit a time varying and partly insignificant sample correlation. Recent studies employing correlation analysis show evidence that the connection between ENSO and monsoon is weakening since the 1980s, arguably because of global warming [58, 60, 108].

Mechanisms

The mechanism that links ENSO to the Indian summer monsoon is still not fully understood. Several studies investigated the possible influence of ENSO on the ISM boundary conditions. One possible link is an ENSO induced shift of the ITCZ in the tropical Indian ocean, influencing directly the surface pressure gradient between the ocean and the Indian subcontinent [132]. Further on, Goswami [30] suggests a mechanism by influencing the tropical wind circulation cells: positive ENSO and thus Walker circulation anomalies close to the equator effect the monsoon Hadley circulation and lead to an increased ISM. Another mechanism links ENSO to the monsoon via the Tibetan snow cover: ENSO positively influences the snow accumulation on the Tibetan plateau. This in turn exhibits a direct influence onto the ISM, since a higher snow cover increases the albedo and thus decreases the absorption of solar radiation. Additionally, more heat is used to melt the snow during summer. These two effects lead to a smaller surface temperature rise, i.e. they weaken the land sea thermal contrast and thus the strength of the monsoon circulation [e.g. 133, 60, 135]. There is furthermore evidence, that the monsoon Hadley circulation (i.e. the north/south components of the winds) and the Walker circulation are subject to interdecadal variability, changing the influence of ENSO on the monsoon [56, 57].

4.3 Questions

From a climatological point of view, the following questions arise from the previous discussion:

- Is the interdecadal variability of the ENSO/monsoon coupling significant?
- What is the timing of this variability?
- What is the mechanism causing this variability?
- Is the relation weakening?
- If so, is the weakening without precedent or a rather typical behavior?

On the other hand, the problem of ENSO monsoon coupling is also of particular interest from a data analysis point of view. The question is, if there exists a method suitable to analyze the rather complicated coupling better than a simple year to year comparison or correlation

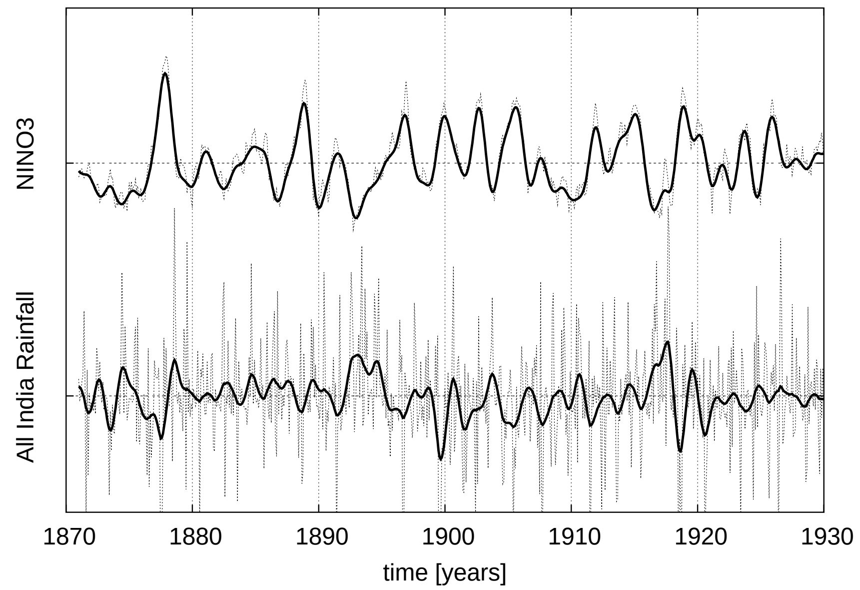

Figure 4.1: Section of the NINO3 (upper graph) and AIR anomalies (lower graph) time series. The dotted lines depict the raw data, the solid lines show the low-pass filtered data used for the further analysis.

analysis, possibly with a better time resolution than the wavelet coherency analysis. In Sec. 4.2.3, we showed that the amplitude relation between ENSO and monsoon is rather unstable: High rainfall might occur during non-La Niña events, deficient rainfall during non-El Niño years. This effect leads to a decrease in correlation or wavelet coherency, even though a clear influence exists. Thus, the question arises, if there is a method accounting for this problem.

4.4 Data

We used the NINO3 index (see App. A) as a measure for ENSO variability. The monsoon was represented by the monthly anomalies of the All India Rainfall (AIR) index (see App. A). We analyze the data in the period from Jan 1st 1871 to Dec 31st 2003. We want to emphasize that for a phase analysis, it is irrelevant whether the amplitudes of the physical processes are well represented by these simple indices, provided the phases of the dynamics are sufficiently well reproduced. Since our work focuses on the inference of phase relations of inter-annual oscillations, we low-pass filtered the data in the spectral domain by multiplying the Fourier transformation of the data with a hyperbolic tangent, i.e. high frequency variability with frequencies higher than 0.7 cycles per year is damped. Fig. 4.1 shows a section of the time series of the ENSO and monsoon data, clearly emphasizing the inter-annual oscillations of ENSO and the biennial oscillation of the monsoon [99].

4.5 Concepts and Methods

In the context of phase synchronization, the method of Hilbert phase analysis has been widely used to analyze coupling between self sustained oscillators. We aim to utilize this method for the investigation of the coupling between ENSO and monsoon. In this section, we sketch the phenomenon of phase synchronization and discuss Hilbert phase analysis. We extend the method by a curvature defined phase to account for long scale variability superimposed to the interesting dynamics. The resulting phase estimator is investigated in terms of bias and variance.

4.5.1 Phase Synchronization

Consider two (or more) self-sustained oscillators, i.e. two autonomous dissipative systems with energy supply oscillating on a limit cycle, torus or chaotic attractor in phase space. In many cases, the trajectory can be projected onto a two dimensional plane and then be described by a time varying amplitude $A(t)$ and phase $\Phi(t)$. If the considered oscillators exhibit slightly different eigen-frequencies ω_0, their phases will evolve independently and the phase difference will diverge. However, if the systems are brought into contact, one observes in general the following phenomena: If the weak coupling exceeds a certain threshold depending on the frequency mismatch $\Delta\omega$, the frequencies start to adjust, such that the phases get *locked*:

$$|n\Phi_1(t) - m\Phi_2(t)| < \epsilon, \tag{4.1}$$

with $n, m \in \mathbb{N}$. This phenomenon is called $(n : m)$ *phase synchronization* [104, 87]. For stronger coupling, complete synchronization might arise, i.e. the amplitudes might also adjust. For a comprehensive introduction and overview over synchronization phenomena in science, please refer to Pikovsky et al. [87].

4.5.2 Measures of Synchronization

For noisy systems, the condition Eq. (4.1) is in general not fulfilled: noise broadens the distribution of $\phi_{n,m}(t) = n\Phi_1(t) - m\Phi_2(t)$ and may even induce phase slips [118]. These effects and additional measurement noise require a statistical interpretation of synchronization. In fact, inferring synchronization from observed data is an estimation problem. Among others, the following statistical measures for the strength of synchronization have been suggested. Given a histogram p_k, $k = 1...N$ of measured phase differences $\phi_{n,m}(t)$, one can define a Shannon entropy based index [118]:

$$\rho = \frac{S_{max} - S}{S_{max}}, \tag{4.2}$$

where $S = -\sum_{k=1}^{N} p_k \ln p_k$ is the entropy of the histogram and $S_{max} = \ln N$. A second index might be defined based on directional statistics. Given a time series of phase differences $\phi_{n,m}(t_k) \in [0, 2\pi]$, $k = 1...N$, then

$$R = |\frac{1}{N}\sum_{k=1}^{N} e^{i\phi_{n,m}(t_k)}| \tag{4.3}$$

yields a value between 0 and 1, signifying vanishing and perfect synchronization, respectively [2]. For both measures, different significance tests have been suggested [2, 118].

4.5.3 Inference of Coupling Direction

Phase synchronization is caused by a mutual influence of the phases onto each other. This influence might be asymmetric. E.g. in the case of radio controlled clocks, the coupling is unidirectional. If the synchronization is perfect, the phases become indistinguishable and the coupling direction cannot be detected. However, in the presence of a perturbed phase relation, one can in principle study the strength and direction of the coupling. One ansatz is to estimate the mutual predictability [111]. Rosenblum and Pikovsky [103] suggested to study whether the phase dynamics of one oscillator are influenced by the phase of the other. This approach has successfully been applied to physiology [105] and neurology [14].

4.5.4 Hilbert Phase Estimation

Estimating the Phase from Univariate Data

To investigate for coherent phase relations, one has to derive the oscillation phases of the involved systems. The reconstruction of the phase $\Phi(t)$ of the oscillating process from a measured time series constitutes an inverse problem and one has to construct a suitable estimator $\hat{\Phi}(t)$ for the underlying phase (see sections 1.1 and 1.5). If one observes only a one dimensional time series $x(t)$ of each system, initially one has to find a suitable two dimensional embedding. For a comparison of different techniques, please refer to [93]. One common approach is to extend the time series into the complex plane, i.e. to construct a two dimensional analytical signal

$$\zeta(t) = x(t) + iy(t) \tag{4.4}$$

by use of the *Hilbert transformation* [24]

$$y(t) = H(x(t)) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau. \tag{4.5}$$

P.V. denotes the Cauchy principal value (for details refer to [104]). An estimator for the phase $\Phi(t)$ can then be defined as

$$\hat{\Phi}(t) = \arctan \frac{y(t)}{x(t)}. \tag{4.6}$$

Observed data from e.g. biological or geophysical systems are in general corrupted by noise. To estimate the underlying phase, the data hence have to be preprocessed. Rossberg et al. [106] developed a nice data driven filter to obtain an optimal phase estimation. We restrained ourselves to a simple static filter in the Fourier domain (see App. 4.4). We call the algorithm of estimating a phase based on Hilbert transformation *Hilbert phase analysis* (HPA).

A Curvature Defined Phase to Eliminate Low Frequency Variability

The approach described above is meaningful only if the embedded signal rotates around a fixed center. For geophysical signals exhibiting variations on a wide range of frequencies, especially on long time scales, this requirement is rarely fulfilled. Thus, suitable filtering is required to eliminate the low frequency variations of the center.

In a related setting, Osipov et al. [79] suggested the following approach: They aimed to analyze the phase of the funnel Roessler attractor, which exhibits a trajectory not always cycling around the unstable fixpoint $(0, 0)$ in the projection onto the (x, y)-plane (for the equations, see App. C.2). They defined a phase not based on the coordinates itself, but rather on the curvature of the trajectory. If the curvature is always positive, one can define a monotonically growing phase by considering the time derivative of the signal, $\Phi(t) = \arctan \frac{\dot{y}(t)}{\dot{x}(t)}$. As long as the curvature is positive, the transformed trajectory $(\dot{x}(t), \dot{y}(t))$ always circles around a fixed origin.

This approach using a *curvature defined phase* (CDP) can be interpreted as a sophisticated filter[3] that eliminates slow variations. Hence, it might be suitable for the definition of a phase estimator

$$\hat{\Phi}(t) = \arctan \frac{\dot{y}(t)}{\dot{x}(t)} \tag{4.7}$$

for the analysis of data superimposed by variability on long time scales.

Evaluating the Method

To study the performance of the estimator Eq. (4.7) for our individual application in terms of bias and variance, we constructed a simple toy model with a well known phase, resembling a typical geophysical oscillator. We designed the toy-model as follows:

- The oscillator was represented by the well known Roessler system (for the equations, see App. C.2):
- The weather noise was modeled by Gaussian white noise $\eta(t) \sim \mathcal{N}(0, \sigma_H^2)$ with $\langle \eta(t_1)\eta(t_2) \rangle = \delta(t_1 - t_2)$.
- Low frequency variability has been represented by a simple AR[2] process (see Sec. 1.2.2) with an overall variance σ_L^2 and a mean period T_L.

We simulated different time series of length $N = 24,000$ (with sampling time $dt = 1/12$, i.e. $T = 2,000$) with different parameters $\sigma_H \in [0, 1]$, $\sigma_L \in [0, 2]$ and $T_L \in [5, 50]$. Figure 4.2 shows an example with noise amplitude $\sigma_H = 0.2$ and low frequency variability (amplitude $\sigma_L = 1$, mean period $T_L = 20$). This parameter setting well resembles the setting of reconstructing the ENSO phase from the NINO3-data. To the set of simulated time series we applied the following steps:

1. Low-pass filter the data in the spectral domain. A smooth function (arcus tangens) damping frequencies > 0.7 is chosen.
2. In case of CDP, estimate derivatives by second order difference scheme and running mean with window width $2l + 1 = 13$ data points.
3. Embed by Hilbert transformation with phase defined according to Eq. (4.7).
4. Unwrap the phases (add 2π after each oscillation).
5. To estimate the phase velocity, another derivative is required.

We compared various filtering parameters and different approaches to estimate the derivatives and found the results being robust.

---

[3] in the Fourier domain, the derivative corresponds to a multiplication with $i\omega$

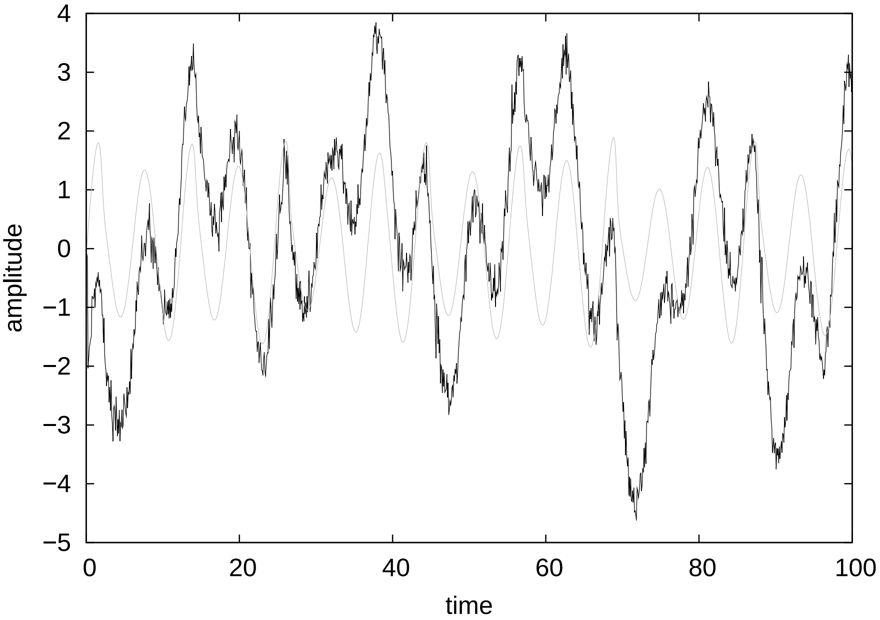

Figure 4.2: Black: First component of the Roessler oscillator (mean frequency 0.163248, normalized to unit variance), superimposed by Gaussian white noise (relative amplitude $\sigma_H = 0.2$) and low frequency variability (relative amplitude $\sigma_L = 1$, mean period $T_L = 20$). Grey: The reference first component of the Roessler oscillator (Section of the whole time series only).

Figure 4.3(a) displays the phase, estimated from the time series shown in Fig. 4.2 minus the underlying phase of the undisturbed Roessler oscillator. The black line depicts the estimation with the CDP Eq. (4.7); the standard estimation Eq. (4.6) is given by the grey line. Whereas the latter one over long intervals propagates with a too low phase velocity (i.e. frequency), because it only “sees” the strong oscillations of the superimposed low frequency variability, the former one stays within the same period as the true phase. These characteristics result in a strongly biased and almost unbiased frequency estimate, respectively, as shown in Fig. 4.3(b): The true frequency distribution of the underlying Roessler oscillator is plotted as a dotted line. The grey line displays the frequency distribution estimated with the standard phase definition. Because of the noise and the stochastic low frequency variability, the distribution is very much broadened in comparison to the true one. Additionally, the low frequency variability strongly shifts the estimation towards lower frequencies. The frequency distribution estimated with the CDP (black line) is also broadened in comparison to the underlying distribution. However, the bias towards lower frequencies is much smaller here.

Figure 4.4 presents a systematic study of the bias (left column) and variance (right column) of the two estimators for different parameter settings of the toy model. The first row shows the dependency on the white noise level in absence of low frequency variability. In this case, the conventional definition (grey line) appears to be the superior estimator: The bias is indistinguishable from zero and the variance is systematically smaller than that of the CDP (black line). For the latter definition, a bias appears for high noise levels exceeding a threshold of around 60%.

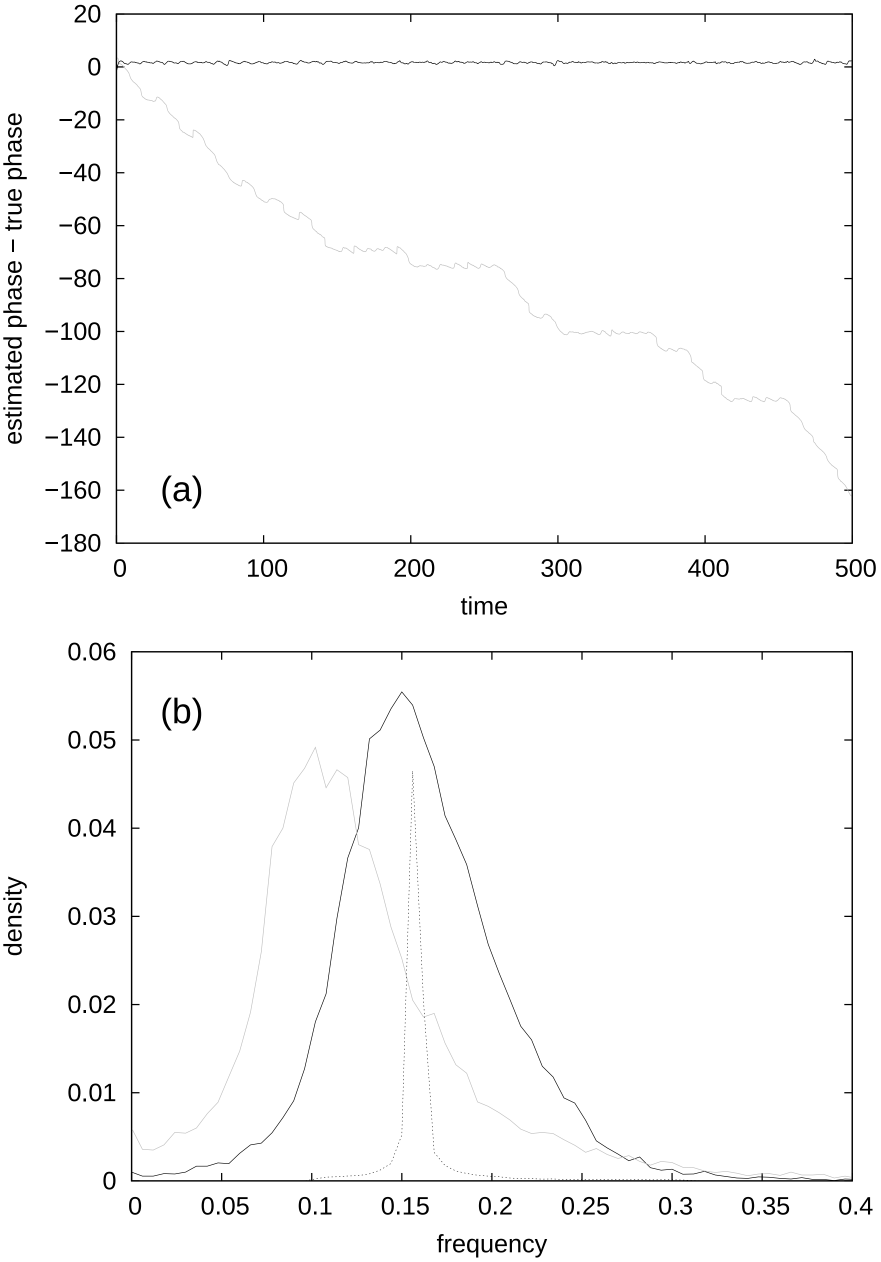

Figure 4.3: (a) Phase propagation of the example shown in Fig. 4.2 minus true phase propagation of reference Roessler oscillator. (b) Corresponding estimated frequency distribution. Grey: standard phase definition. Black: phase definition according to Osipov et al. [79]. Dotted: reference Roessler oscillator (only bottom panel, density divided by 10).

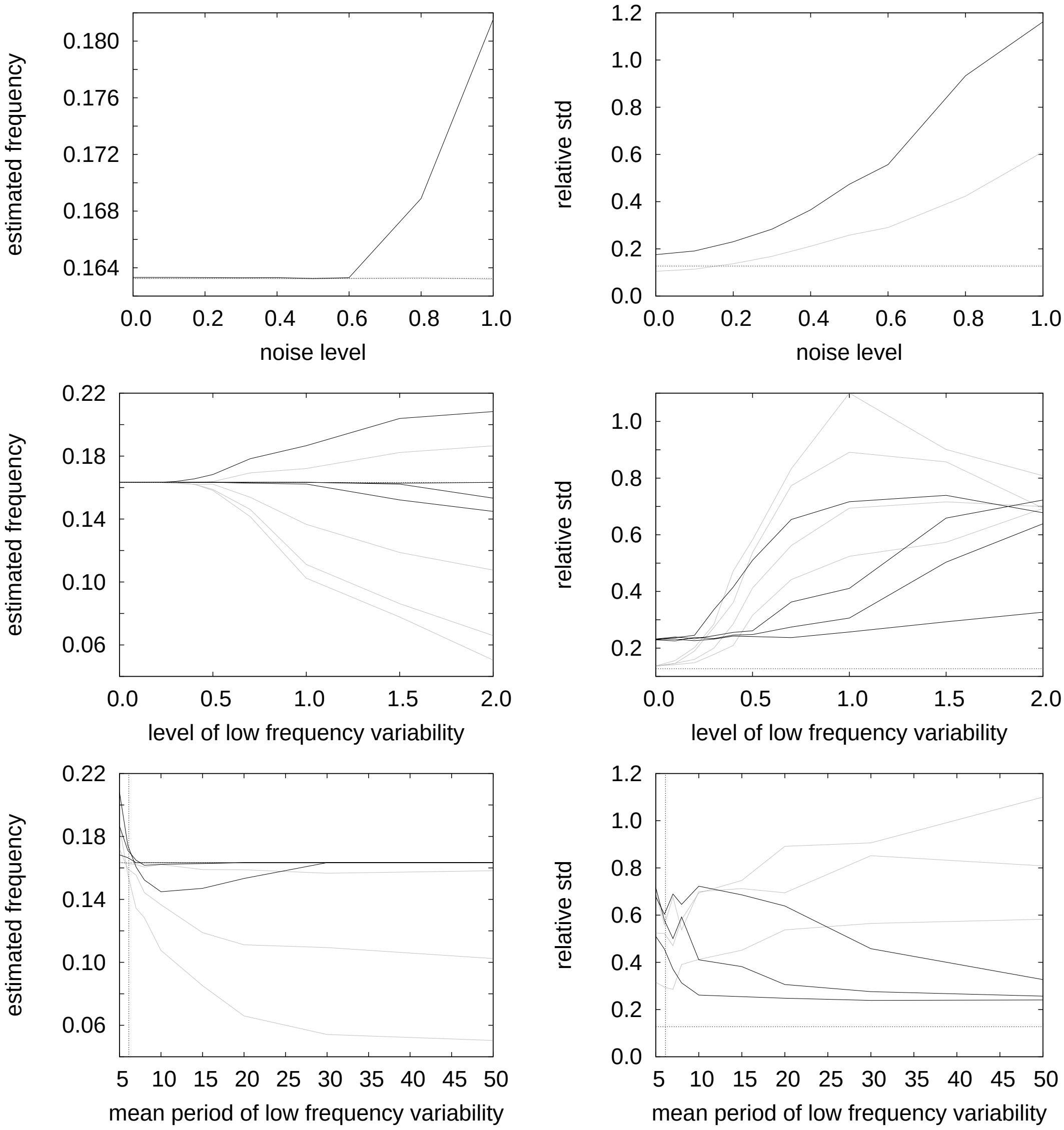

Figure 4.4: Bias and variance of frequency estimation for standard phase definition (grey) and curvature defined phase (CDP) according to Osipov et al. [79] (black). Left column: estimated frequency, the dotted line marks the true frequency of the Roessler oscillator. Right column: standard deviation of estimation, normalized to the true frequency. Top row: as function of noise level, no low frequency variability. Middle row: as function of low frequency variability level; each line represents a different mean period of low frequency variability. Bottom row: as function of low frequency mean period; each line represents a different level of intensity. Dotted line: true mean period of Roessler oscillator.

The second row shows the dependency on the level of the low frequency variability for a fixed white noise level $\sigma_H = 0.2$ and different mean periods T_L of the low frequency components. For $T_L = 5$, the low frequency variability exhibits almost the same mean period as the Roessler ($T_R \approx 6.1$). Thus, for large σ_L, a separation of both time scales is almost impossible for both methods and the estimate is biased towards higher frequencies (the two upper lines), although the bias for the curvature defined phase is lower than that of the standard phase estimation. In this case, the variance of the estimator is relatively high for both methods. For longer periods T_L (the lower lines), one observes a much stronger bias towards lower frequencies for the conventional phase estimator as for the curvature defined one. The variance of the latter estimator is much lower than that of the former one.

The third row shows the dependency on the period of the low frequency variability for a fixed white noise level $\sigma_H = 0.2$ and different levels of the low frequency components. The bias of the standard phase definition is always higher than that of the CDP and increases with the period depending also on the strength of the low frequency variability. The curvature defined estimator performs better: For low periods T_L, the scale separation is rather difficult, hence the bias towards the frequency $1/T_L$ is quite strong for high values of σ_L. For large T_L, however, the separation operates almost perfectly, resulting in a vanishing bias independent of σ_L. A similar picture arises for the variance: For the standard estimator, it simply increases with T_L and σ_L. For the curvature defined estimator instead, the variance exhibits a maximum for low periods T_L and decreases for higher T_L to a common value determined by the high frequency noise level σ_H.

Summarizing these observations, the standard phase estimator is the method of choice for noisy data not superimposed by low frequency variability. For typical geophysical data exhibiting large fluctuations on long time scales, however, the CDP estimator performs much better in terms of bias and variance. Not surprisingly, the quality of the estimation gets better, the longer the period of the corrupting components, as this results in local stationarity and vanishing derivatives of the low frequency trends.

4.6 Phase of ENSO and Indian Monsoon

We apply the HPA with a CDP to analyze the coupling between ENSO and ISM. As the monsoon does not perform self-sustained oscillations (see Sec. 4.2.2), no phase synchronization can arise. In this special case, the coupling reflects a different physical principle: As described in Sec. 4.2.3, ENSO influences the boundary conditions of the monsoon system. Thus (disregarding the possible influence of the monsoon on ENSO), we are concerned with modulation rather than with synchronization in a strict sense. Nevertheless, the method presented above might reveal additional valuable information about the coupling: A simple linear analysis does not decompose amplitude and phase. Even methods as wavelet analysis, which can investigate a time dependent phase, are only capable of deriving a 1:1 phase relation. In the case of the ENSO/ISM coupling, this property might mask a possible interrelation: As shown in Tab. 4.1 in Sec. 4.2.3, not all ISM deficiencies occur during El Niño events, and not all heavy ISM rainfall years coincide with La Niña years. As discussed in Sec. 4.3, this unstable amplitude relation leads to a decrease in correlation. The phase analysis, by contrast, completely disregards the amplitude. The method only investigates, whether there is an influence of ENSO on the phasing of monsoon, regardless of the particular strength.

To estimate the phases of ENSO and monsoon, we applied the steps listed in Sec. 4.5.4.

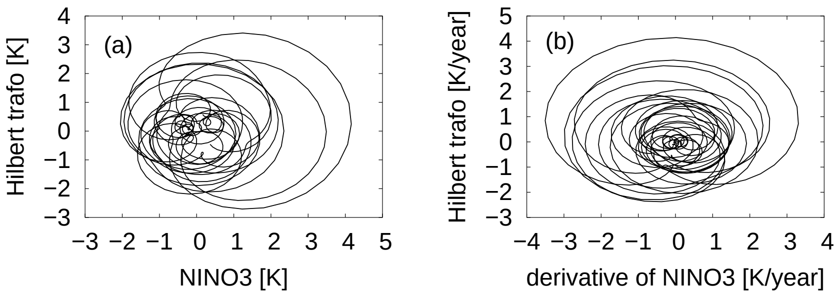

Figure 4.5: (a) Embedding of low-pass filtered NINO3 time series by Hilbert transformation. Many oscillations are not centered around a common center. (b) The same, but for the time derivative of the NINO3 time series. All pronounced oscillations circle around the origin.

Fig. 4.5(a) shows the embedding $(x(t), y(t))$ for the (low-pass filtered) NINO3 time series. Many oscillations do not cycle around a common center, such that the phase propagation will be underestimated. This observation justifies the application of the curvature motivated estimator Eq. (4.7) discussed above. Fig. 4.5(b) shows that the latter approach eliminates most of the slow variations, such that at least all large and pronounced oscillations are centered around a common origin and a meaningful phase can be defined. The derivative $\dot{x}(t)$ and its Hilbert transform $\dot{y}(t)$ are plotted in Fig. 4.6. The dotted line represents the envelope $A(t) = (\dot{x}(t)^2 + \dot{y}(t)^2)^{1/2}$, i.e. the amplitude of the oscillation. Finally, we calculated the phases according to Eq. (4.7) and unwrapped them by adding 2π after each oscillation. Additionally, we choose the envelope $A(t)$ as a measure to select regions of well defined phases: For time derivatives $\dot{x}(t)$ normalized to unit variance, $A(t) > 0.8$ appeared to be a reasonable choice to exclude spurious oscillations.

The results for the unwrapped phases are presented in Fig. 4.7. During the same time, the AIR performs more cycles than the NINO3 time series, resulting from the quite stable biennial oscillation. However, during some epochs, the phases seem to evolve similarly, suggesting to investigate for phase coherence.

4.7 Results

4.7.1 Phase Relation

To study the occurrence of phase coherence in more detail, we calculated the difference of the phases (Fig. 4.8(a)). Distinct epochs become visible: The plateaus from 1886 to 1908 and from 1964 to 1980 indicate phase coherence during these intervals (marked I). In the years 1908-1921, 1935-1943 and 1981-1991, the monsoon oscillates faster than ENSO, failing when ENSO peaks on inter-annual scales as during the phase coherent intervals, but with an additional peak in between (marked II). Here two times the phase of ENSO minus the monsoon phase yields a plateau[4] (see Fig. 4.8(b)). During the epochs discussed above, the phases of ENSO and monsoon are predominantly well defined (grey shading) and both systems exhibit distinct oscillations, which are also visible in the NINO3 time series itself (Fig. 4.8) or its wavelet

---

[4] Even though this might be no dynamic feature, we call this 2:1 coherence, as it produces a phase difference similar to that of 2:1 phase synchronization

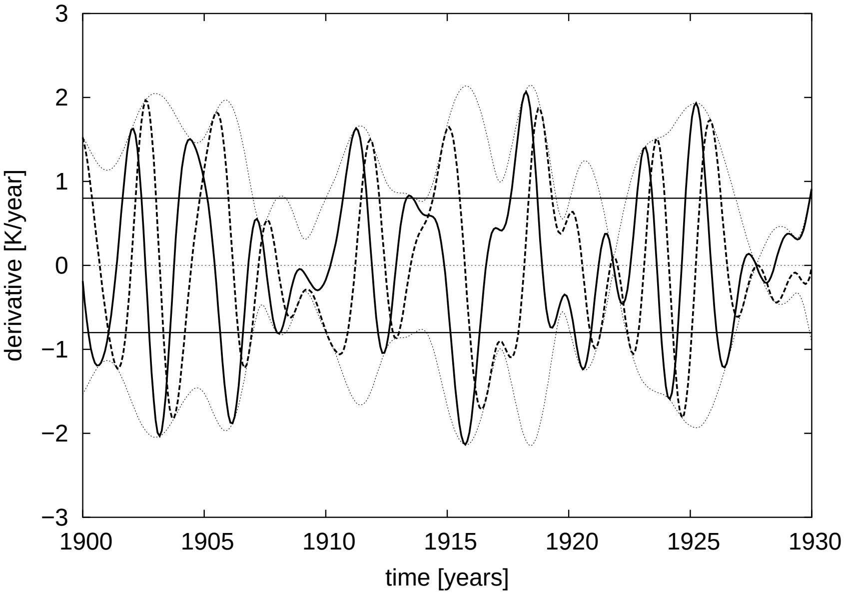

Figure 4.6: Smoothed Derivative of the normalized and low-pass filtered NINO3 time series (solid line, section), the corresponding Hilbert transformation (dashed line). The dotted line represents the amplitude A (for details see text). For $A > 0.8$ (solid horizontal lines), we consider the phase to be well defined.

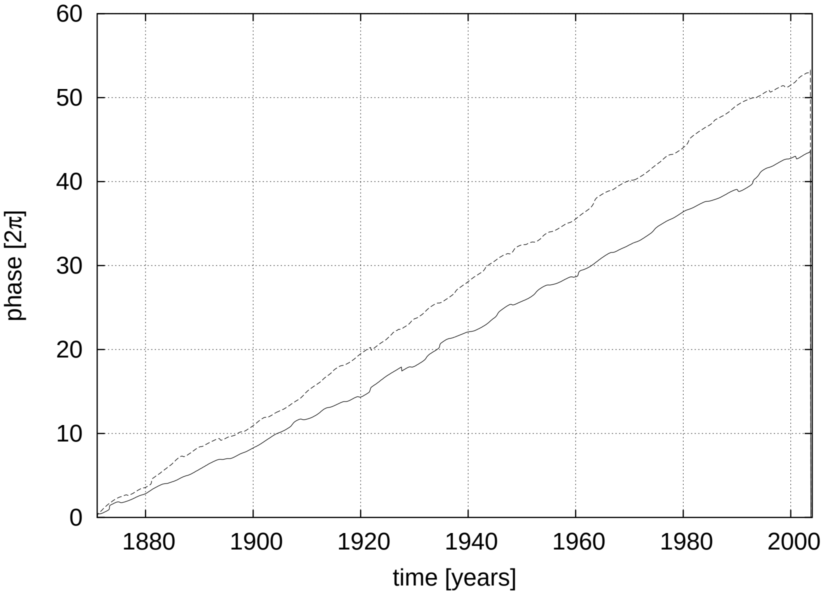

Figure 4.7: Phase propagation of NINO3 (solid) and monsoon (dashed). On average, the latter one oscillates faster. However, during some periods the phases propagate similarly, suggesting to investigate for phase synchronization.

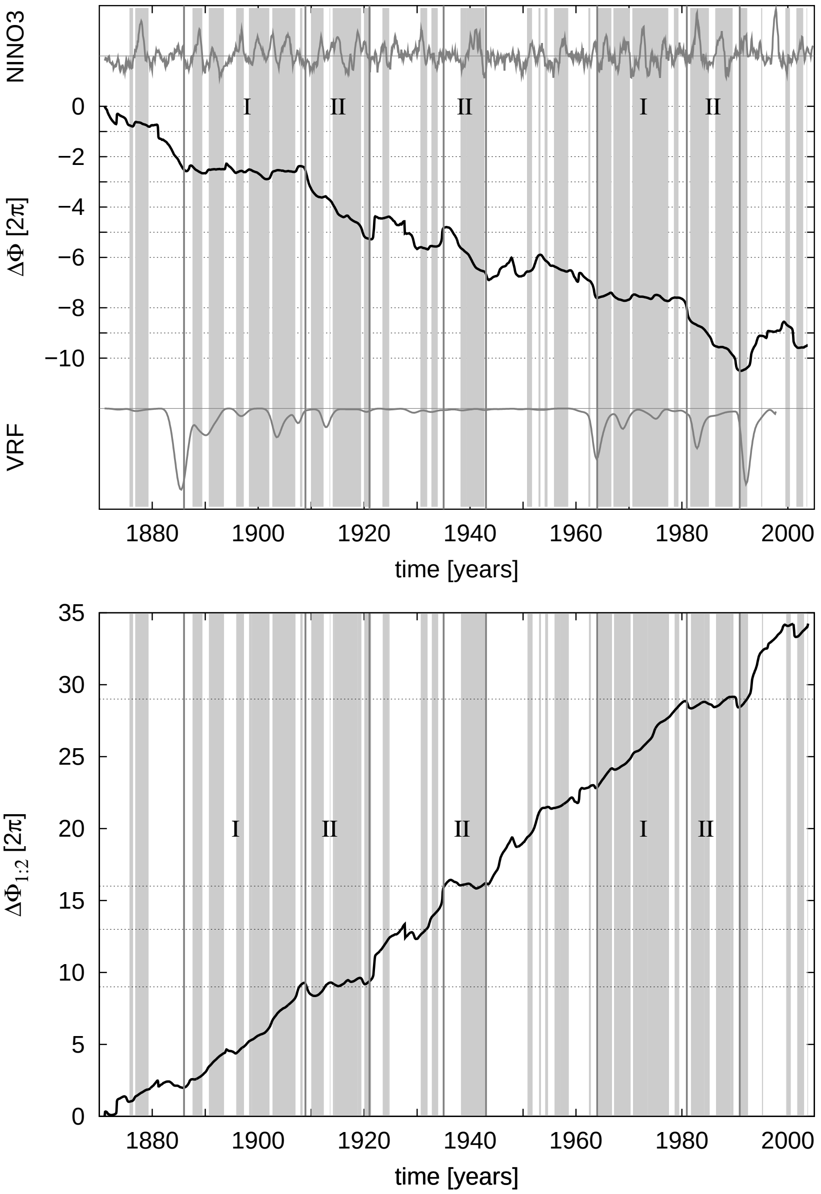

Figure 4.8: (a)Phase difference of ENSO and monsoon (black). Grey shading marks intervals of jointly well defined phases. 1886-1908 and 1964-1980 (I): plateaus indicate phase coherence. 1908-1921, 1935-1943 and 1981-1991 (II): monsoon oscillates with twice the phase velocity of ENSO (see also panel (b)). During these intervals, both systems exhibit distinct oscillations (NINO3 time series, upper graph). 1921-1935 and 1943-1963: phases are badly defined, both processes exhibit irregular oscillations of low variance (upper graph). Lower graph shows volcanic radiative forcing index (VRF). (b) Difference of two times the ENSO phase and monsoon phase.

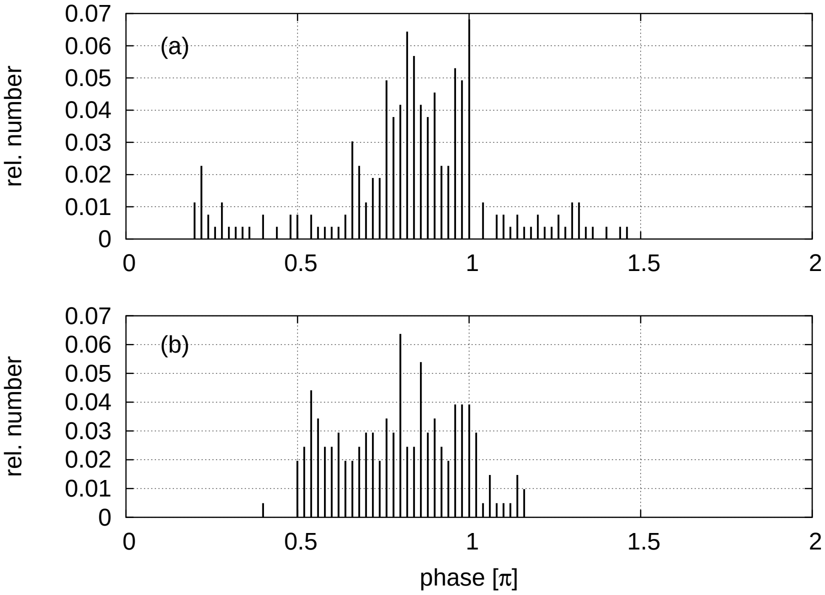

Figure 4.9: Histogram of phase differences for the two phase coherent intervals (a) 1886-1908, (b) 1964-1980. Both diagrams show peaks between $\pi/2$ and π, reflecting that ENSO and monsoon are anti-correlated.

spectrum [34]. During other times, especially 1921-1935 and 1943-1963, the phases are rather badly defined and both processes exhibit irregular oscillations of low variance. Fig. 4.9 shows the histogram of phase differences of the two intervals of phase coherence. Both diagrams show distributions far away from being uniform, with peaks between $\pi/2$ and π, reflecting that ENSO and monsoon are anti-correlated. The indeces defined in Sec. 4.5.2 for the two plateaus result as $\rho \approx 0.23$, $R \approx 0.76$ for the interval 1886-1908 and $\rho \approx 0.26$, $R \approx 0.86$ for the interval 1964-1980. Fig. 4.10 shows the histogram of differences between two times the ENSO phase and the monsoon phase for the three corresponding intervals. The diagrams show a distribution different from being uniform. In case of the epochs 1908-1921 and 1935-1943, also peaks are clearly visible; however, for the 1981-1991 interval, the distribution is rather flat. These findings are reflected in rather low indeces: $\rho \approx 0.19$, $R \approx 0.32$ for the interval 1908-1921, $\rho \approx 0.21$, $R \approx 0.22$ for the interval 1935-1943 and $\rho \approx 0.19$, $R \approx 0.35$ for the interval 1981-1991. As these histogram represent rather short intervals and keeping in mind the low indeces, one should interpret these results carefully.

4.7.2 Significance Testing

To evaluate the significance of the results presented in the previous section, one could test, if the indeces ρ and R significantly deviate from zero. Such a test, however, is rather unspecific. The problem one encounters is the same as discussed in Sec. 3.6.4 of chapter 3: Processes oscillating with similar frequencies trivially exhibit intervals of bounded phase differences, i.e. realizations will show plateaus of random phase "locking" with a characteristic length distribution depending on the frequency mismatch and the decorrelation times of the processes. Thus we developed a significance test that tests the lengths of the plateaus against the length distribution under the null hypothesis of mutually independent processes with

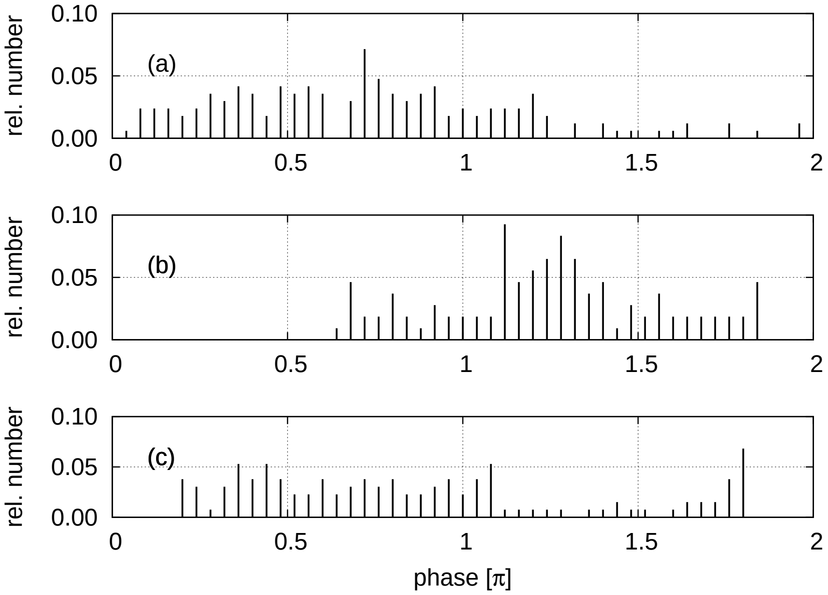

Figure 4.10: Histogram of differences between two times the ENSO phase and the monsoon phase, for the intervals (a) 1921, (b) 1935-1943 and (c) 1981-1991.

Period	2 years	3 years	4 years	5 years	6 years	7 years
ENSO	15/42	15/42	8/42	1/42	0	3/42
ISM	26/50	18/50	4/50	2/50	0	0

Table 4.2: Period distribution of ENSO and ISM estimated from the NINO3 index and the AIR index, respectively.

corresponding spectral properties.

We estimated the length distribution for two independent processes based on nonparametric bootstrapping. Actually, we have taken advantage of the fact that both processes are phase locked to the annual cycle, such that we could reduce the problem of constructing surrogate data to simulating annually resolved random time series with oscillations lasting for an integer number of years. These oscillations reflect the time between two El Niño events or between two monsoon failures, respectively. From the used data, we estimated the spectrum as follows: The distribution of times between two ENSO peaks and two monsoon failures is shown in Tab. 4.2. We simulated 1,000,000 pairs of annually resolved 150 years long symbolic time series, each containing integer values from 2 to 7 according to these probability distributions (e.g. 4 consecutive values of “4” mark one cycle of period 4 years and occur with a probability of 8/42 in case of ENSO). Fig. 4.11 shows the estimated probabilities for plateaus of a certain number (not years) of common oscillations to occur. The lengths of the observed plateaus correspond to 8 oscillations in the interval 1886-1908 and 6 oscillations in the interval 1964-1980. The probability that 8 successive oscillations appear in both time series at the same time with pairwise identical periods resulted as $p_8 \approx 0.0017$, the probability for a plateau of 8 randomly common oscillations resulted as $p_6 \approx 0.014$. Hence, the first plateau is significant on the 1%-level, the second on the 5% level. However, the probability that both plateaus occur randomly within 150 years is $p_1 p_2 \approx 2.3 \cdot 10^{-5}$, i.e. the result is significant at

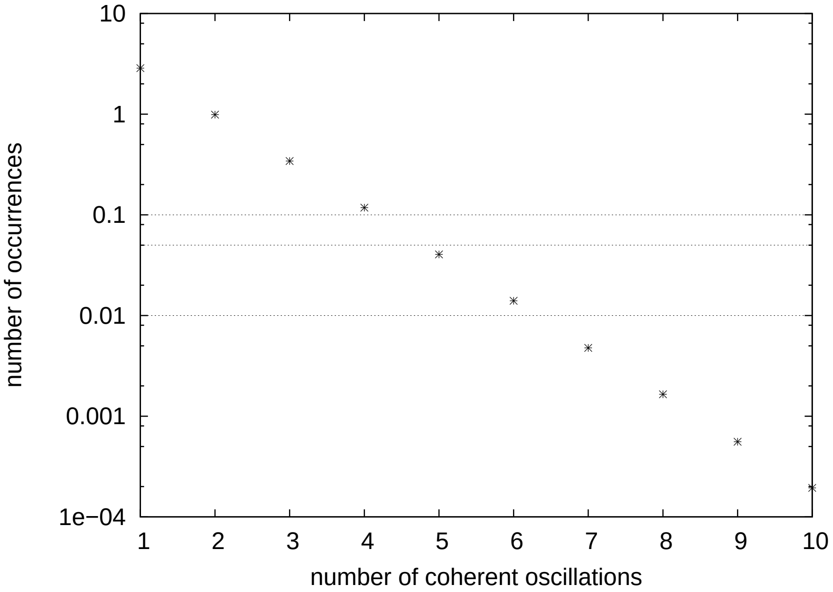

Figure 4.11: Probability of occurrences of phase coherent oscillations for a 150 year time series as a function of the number of coherent oscillations. The black dotted lines depict the 0.1, 0.05 and 0.01 levels, respectively.

least at the 1% level. As the intervals of 2:1 phase relation cover rather short intervals and furthermore only give low values for the synchronization indeces, we forbear from applying a significance test on these epochs.

4.7.3 Coupling Direction

Theoretically, our approach allows to investigate for the direction of coupling (see Sec. 4.5.3), i.e. whether the modulating influence of ENSO on monsoon is dominating or if a possible reverse influence has to be taken into account. However, as the coupling appears to be time dependent with stable periods of only around 200 correlated data points, we expect that the dynamical noise is too high to obtain any reliable results.

4.8 Climatological Discussion

4.8.1 Interdecadal Variability

We could identify distinct epochs, especially two intervals of phase coherence from 1886 to 1908 and from 1964 to 1980. Interestingly, these epochs show high variability and distinct oscillations of ENSO as well as monsoon, unlike the intervals 1921-1935 and 1943-1963, where both processes show low variability and no coherence. These results corroborate earlier findings of Torrence and Webster [123]: The intervals of phase coherence coincide with regions of high wavelet coherence. In contrast to previous correlation and coherence analysis, we could identify the following additional epochs of coupling by use of HPA: During the years 1908-1921, 1935-1943 and 1981-1991, the monsoon oscillates twice as fast as ENSO. During ENSO years, the monsoon rainfall also weakens. However, it additionally fails half an ENSO cycle later. Also unlike sliding correlation and wavelet analyses, which intrinsically have to average (here roughly two decades), we estimate an almost instantaneous phase of only the dominant oscillation and hence are able to estimate the onset of phase coherence with an accuracy of around one ENSO cycle.

The two main epochs of coupling appeared to be highly significant. However, the criticism of Gershunov et al. [28] that *changes* in correlation strength between ENSO and monsoon could be spurious due to stochastic fluctuations, also holds for these results: Also our method is not able to decide whether the temporary weakening of the coupling, i.e. the interdecadal variability, is significant. It might be that the periods without a stable phase relation, i.e. the intervals before 1886, from 1921 to 1935, from 1943 to 1963 and after 1991 in fact represent epochs of coupling that are only masked by noise. This question can only be addressed by employing a model of both ENSO and monsoon, representing the coupling and typical stochastic variability.

4.8.2 Temporarily Weakening Relationship with Precedents

As mentioned in Sec. 4.2.3, Kumar et al. [60] found that 21 year sliding correlations between ENSO and monsoon were high from 1856 to the early 1980’s, but decreased to insignificant values afterwards. They suggest that global warming may cause the decoupling. Since our approach decomposes signals into phase and amplitude, we gain a more detailed insight: After 1981, the monsoon was still coupled to the ENSO cycle until 1990. During this time the monsoon still fails after strong El Niños but peaks additionally shortly after (see Sec. 4.8.1). This behavior is invisible to correlation analysis, leading to the insignificant correlations observed in [60]. Only after 1990, also our method does not find any stable relation between the two processes. However, this behavior appears to be typical rather than a feature without precedent: A similar 2:1 coherence occurred in the periods 1908-1921 and 1935-1943, and a behavior resembling that after 1991 appeared also from 1921 to 1935 and before 1886. Thus, it is unlikely that the observed decrease in correlation really represents a distinct trend of a weakening relationship between ENSO and monsoon.

4.8.3 Possible Volcanic Forcing

The high time resolution allows us to precisely determine the onset of phase coherent intervals and thus to suggest a mechanism that might cause the coupling. The lower graph in Fig. 4.8 displays the volcanic radiative forcing index [109]. Interestingly both intervals of phase coherence coincide with periods of strong volcanic radiative forcing and start with two major eruptions, of Krakatau (1883, Sunda strait) and Mount Agung (1963, Bali), both located in southern Indonesia and exhibiting large climatic forcing [101].

We now introduce an idea, which might help to understand the recent findings of Adams et al. [1]: Volcanic forcing might not cause single ENSO events, but rather either increase the coupling between ENSO and monsoon, causing more regular oscillations of the total system, or cause more regular oscillations of one of the systems (probably ENSO), increasing the coupling between them. The climatic impact of volcanoes is based on radiative forcing mainly by sulphate aerosols. Blocking of short-wave radiation causes summer cooling (and winter warming) and overall global cooling lasting for 1-3 years. Multiple eruptions have caused cooling even on decadal scales, e.g. in the period from 1883 to 1912 and from 1963 on [e.g. 7]. This cooling effect could reduce the land/sea temperature gradient and thus make the monsoon more sensitive to ENSO influence [e.g. 60]. For a detailed review of possible mechanisms, refer to [101].

However, this idea needs further evaluation: A spatiotemporal analysis is required to corroborate the influence of volcanic cooling onto the boundary conditions of the ENSO monsoon system, especially on the Tibetan snow cover. This could be achieved by producing composites of temperature fields and possibly temperature gradient fields for the different intervals. Furthermore, numerical simulations could help to identify the proposed link in a model experiment.

4.9 Summary and Conclusions

To analyze the coupling between ENSO and the Indian summer monsoon, we applied Hilbert phase analysis, a method used in nonlinear time series analysis to study the relationship of two oscillating systems with respect to their phases but independently of their amplitudes. To account for interdecadal variability of the tropical ocean atmosphere system, we combined this method with a recently proposed curvature defined phase that represents a high pass filter. We tested the performance of this method in terms of bias and variance by applying it to a simple toymodel resembling the ENSO variability.

We identified distinct epochs of different phase relations. Two intervals of phase coherence from 1886 to 1908 and from 1964 to 1980 confirm earlier findings of Torrence and Webster [123]. These findings are corroborated by a significance test of high specificity accounting for typical random phase coherent intervals. In contrast to sliding correlation and wavelet analyses, we estimate an almost instantaneous phase of only the dominant oscillation and hence obtain a much higher time resolution. In fact, we are able to estimate the onset of phase coherence with an accuracy of around one ENSO cycle.

Kumar et al. [60] found a weakening relation between ENSO and monsoon since the early 1980's. Using the Hilbert phase analysis method that decomposes amplitude and phase, we get a more detailed insight: After 1981, the Monsoon still fails after strong El Niños but peaks additionally shortly after. This additional peak reduces the linear correlation, even though coupling is present. Only after 1991, no distinct relation between the phases of ENSO and monsoon exists. Both epochs, however, are typical for the system as similar intervals have occurred before. Thus, our findings indicate that the weakening of the relationship between ENSO and monsoon is not without precedent and might be only temporal.

The high time resolution allows us to precisely determine the onset of phase coherent intervals and thus to suggest a mechanism of volcanic forcing that might amplify the coupling. Volcanic radiative cooling might increase the Tibetan snow cover that in turn reduces the land sea temperature contrast and thus makes the monsoon more sensitive to ENSO influences. This idea needs further evaluation by means of detailed time series analysis and model simulations.

Chapter 5

Discussion, Conclusions and Outlook

5.1 Discussion and Conclusions

Understanding the climate system requires an interplay of theoretical studies, model experiments and time series analysis. The thesis at hand mainly aimed to contribute to methodological questions of the latter field. A second aspect was to apply the investigated methods to recent climatological problems. A main focus of my work was to study the reliable inference of particular process properties from an inverse problem point of view. This required methods of high specificity.

Together with Prof. Dr. Jens Timmer and Henning Rust, I studied the inference of long-range correlations by means of Detrended Fluctuation Analysis (DFA). Identifying this task as an inverse problem, we argued that power-law scaling of the fluctuation function and thus long-memory may not be assumed a priori but have to be established. This requires to investigate the local slopes of the fluctuation function. In fact, it is not sufficient to conclude scaling from a straight line fit to the fluctuation function in a log-log representation, as done in many publications. We accounted for the variability characteristic for stochastic processes by calculating empirical confidence regions. The comparison of a long-memory with a short-memory model showed that the inference of long-range correlations from a finite amount of data by means of DFA is not specific. Instead, we rather suggest to attempt to falsify long memory and hence to infer short range correlations. In this respect, we showed that a local slope larger than $\alpha = 0.5$ for large scales does not necessarily imply long-memory. We also demonstrated, that a finite scaling of the autocorrelation function is shifted to larger scales in the fluctuation function. Thus, it is not valid to directly transfer a typical time scale from a DFA result to an equivalent time scale in the autocorrelation function, as done in other publications. Based on these studies, we reviewed the DFA results for the Prague temperature data set and showed that long-range correlations cannot be concluded unambiguously.

In the second part of the thesis, I presented the results from a collaboration with Prof. Dr. Holschneider. This work was mainly motivated by the realization that, although the mathematical properties of CWT were well studied, the formulation of continuous wavelet spectral analysis had a rather tentative character. Open questions concerned the definition of non-stationary spectra in the wavelet domain and the bias and variance of the wavelet sample spectrum. Further problems were the effect of multiple testing in the conventional pointwise significance tests and the interpretation of cross wavelet measures. To meet these shortcomings, we proposed an equivalence class of nonstationary Gaussian stochastic processes defined in the wavelet domain. These processes are characterized by means of wavelet multipliers and exhibit well defined time dependent spectral properties; they allow one to generate realizations of any nonstationary Gaussian process. Based on this framework, we studied the dependency of the realizations of such processes on the wavelets used for the generation. We calculated bias and variance of the wavelet sample spectrum and investigated their behavior for certain asymptotic cases and in presence of averaging. Furthermore we showed that the global wavelet spectrum in general is not an unbiased estimator for the Fourier spectrum. To overcome the difficulties of multiple testing, we developed an areawise significance test that utilizes basic properties of continuous wavelet transform to decide whether a pointwise significant result is larger than a certain critical area and thus distinguishable from typical stochastic fluctuations. Using the framework of nonstationary Gaussian processes in wavelet domain, we compared the areawise test and the conventional pointwise test in terms of sensitivity and specificity. For the rather hypothetical case of a high signal to noise ratio, the pointwise test showed a higher sensitivity and a comparable specificity. However, for typical geophysical data exhibiting a low signal to noise ratio the sensitivity was slightly reduced in favor of an considerably improved specificity such that the areawise test remarkably increases the reliability of wavelet spectral analysis. We estimated the wavelet coherence between ENSO and NAO and could show that earlier findings proposing an instantaneous influence of strong El Niños on North Atlantic climate were a mere artefact. Finally, we presented applications to Hydrological questions, work that has been done together with Dr. Bettina Schaefli.

In the last part I studied the coupling between El Niño/Southern Oscillation (ENSO) and the Indian Monsoon on inter-annual time scales. I estimated the oscillating phases of both systems by means of Hilbert transformation, an approach that originates from phase synchronization analysis. This method allows one to investigate the relation of two oscillating systems with respect to their phases, independently of their amplitudes. To account for typical interdecadal variability, I combined the method with a recently introduced curvature defined phase. To evaluate the performance of the resulting technique, I constructed a simple toy model resembling the ENSO variability and studied the phase estimator in terms of bias and variance for different signal to noise ratios. I identified distinct epochs, especially two intervals of phase coherence, 1886-1908 and 1964-1980, confirming earlier findings from a new point of view. A significance test of high specificity corroborated these results. I also detected so far unknown periods of coupling which are invisible to linear methods. These findings contradict earlier propositions of an unprecedented weakening relationship between ENSO and the monsoon. Instead, they suggest that the decreasing correlation during the last decades has happened before and might be partly inherent to the ENSO/Monsoon system. Finally, I outlined a possible interpretation of how volcanic radiative forcing could cause the coupling.

These three studies investigate different aspects of time series analysis with a focus on climatological questions. However, a guiding idea was the study of the sensitivity and especially specificity of the methods applied. In the case of DFA, earlier research has almost completely disregarded this question. Instead of trying to corroborate or falsify the findings of long memory, the unspecific methodology forced every correlation structure into the procrustean bed of power laws. This thesis showed that a reliable inference of long memory for a typical amount of data is impossible by means of DFA. Also in case of wavelet spectral analysis, conventional tests were highly unspecific and produced many spurious results. The presented areawise test overcomes this disadvantage. Finally, the specific test I developed for the significance testing of phase coherence rejects typical random coherent oscillations and thus provides robust results.

In this context, it is important to point out the limits of time series analysis, especially in geosciences: Climate is a vast system consisting of a variety of interacting subsystems. This complexity is faced by limited observations. However, a sufficient amount of data is crucial for a reliable time series analysis. In case of DFA, it is not possible to discriminate between long memory and short memory based on the available time series. In wavelet analysis, the amount of data is also important for a high signal to noise ratio. However, since the analysis is local in time, the relevant limits are of different nature: To infer the frequency of a local oscillation, one has to observe this oscillation for a certain time. If a high time resolution is required, the frequency resolution is decreased and vice versa. Thus, also for a densely sampled time series, oscillations shorter than a critical length are in principle not distinguishable from noise. For the inference of coherence between two processes, one additionally has to exclude typical random coherent oscillations. Hence, even though wavelet analysis allows a time resolved investigation, the identification of localized properties is confined by fundamental limits. These arguments also hold for the inference of phase coherence by means of Hilbert phase analysis: To infer a constant phase relation, one has to establish constancy over a sufficiently long interval that exceeds typical randomly coherent intervals.

The previous discussion illustrates the benefits and limits of time series analysis. For a reliable inference, a researcher has to be aware of the sensitivity and specificity of the methodology he applies. Investigating complex and stochastic systems, a specific test takes the place of corroboration by the attempt of falsification. However, the results one obtains with either of the methods presented in this thesis are limited by the amount of data and fundamental methodological constraints and, eventually, they are of a merely statistical nature. As for any statistical result, it is up to the researcher to provide a reasonable interpretation. A statistical analysis in principle can barely be an end in itself but rather a part of the interplay between theoretical reasoning, model experiments and time series analysis to gain a deeper scientific understanding.

5.2 Outlook

The studies in this thesis raised a lot of outreaching questions. Some of them concern the limits of the particular method discussed above and require a reorientation. Others emerge from ideas and concepts proposed in this work and initiate a new field of research.

DFA previously had been shown to be a reliable estimator of the Hurst exponent of fractional Brownian motion. However, the *inference* of long memory by means of DFA appeared to be not feasible. Instead, the nonparametric DFA-procedure should be superseded by a parametric modeling approach. In the sense of Occam's razor, one fits different FARIMA models to the data and utilizes a particular information criterion (e.g. Bayes, Akaike) to select the optimal model. This model might contain short as well as long memory components. Such an approach is currently under development by Rust [107].

The definition of nonstationary Gaussian processes by means of wavelet multipliers is prototypically suitable for the generation of nonstationary surrogate data of an estimated wavelet spectrum. However, because averaging is needed to obtain an estimate of the underlying spectrum with a reasonably low variance, this approach needs to be further investigated in terms of variance and bias. A possible application is the significance testing with a non-stationary background spectrum. This idea might overcome the problems one faces when the variance of the observed process exhibits a trend. Cross wavelet analysis might be used to define measures for scale and time resolved phase coherence. Finally, the Gaussian processes in wavelet domain are easily extended to non-Gaussian noise. This extension might be useful for the analysis of processes showing non-Gaussian marginal distributions.

The phase coherence estimation by means of Hilbert transformation addresses a similar problem as the estimation of wavelet coherence. The latter one provides a time and scale resolved phase relation at the cost of an inherent uncertainty relation: A higher time resolution decreases the scale resolution and vice versa. Hilbert phase analysis abandons any scale resolved information in favor of an optimal time resolution. In this sense, it marks the opposite of stationary Fourier analysis. However, this advantage requires a signal with a pronounced main oscillation. A further study should systematically compare both approaches and identify which one is preferable in which setting.

Based on the phase coherence analysis of ENSO and the Indian summer monsoon, we suggested a possible influence of volcanic radiative forcing on the coupling between the two processes. However, this idea needs further evaluation: A spatiotemporal analysis is required to corroborate the influence of volcanic cooling onto the Tibetan snow cover, which is one possible link for the ENSO influence on the monsoon. Also, model simulations could help to identify the proposed link.

Beyond these tasks closely connected to my thesis, general climatological challenges await to be faced by future research. Climate models need to be further improved. Already included processes can be represented even more realistically and further important processes have to be captured. On the one hand, this trivially depends on computational power. On the other hand, processes on scales too small to be resolved by the models need to be better parameterized. This in turn necessitates reliable theoretical simplifications, but also the collection of sufficient data describing the individual processes. Eventually, to corroborate and improve the models, a dense network of observations of the relevant climatological variables is required. If one is interested in the climate's behavior on scales up to decades, this need might be met by direct observations. To understand processes on longer scales, however, further proxy data reaching back into the earth's history have to be systematically collected.

Appendix A

Data Sets

A.1 Large Scale Climate Indeces

A.1.1 El Niño/Southern Oscillation

To represent the ENSO phenomenon, I choose the common extended NINO3-index defined by Kaplan et al. [54] as the monthly SST anomalies with respect to the climatological mean from 1961 to 1990, integrated over the eastern tropical Pacific in the range of 150W to 90W and from 5N to 5S respectively. The resulting time series ranges from Jan 1st 1856 to Dec 31st 2003, see Fig. A.1(a).

A.1.2 North Atlantic Oscillation

A widely used NAO index was defined by Jones et al. [49] as the surface level pressure difference between Iceland and Gibraltar. The measurements were available from Jan 1st 1825 to Dec 31st 2003, see Fig. A.1(b).

A.1.3 Indian Monsoon

The Indian monsoon is represented by the deseasonalized monthly anomalies of the All India Rainfall (AIR) index defined by Mooley and Parthasarathy [75]. The data set covers the period from Jan 1st 1871 to Dec 31st 2003, see Fig. A.1(c).

A.1.4 Volcanic Radiative Forcing Index

The volcanic radiative forcing index has been defined by Sato et al. [109] and represents the optical depth of the atmosphere at $\lambda = 0.55\mu m$, normalized to the eruption of Krakatau in 1883. It provides a measure of the aerosols emitted by volcanic eruptions into the upper troposphere. See Fig. A.1(d).

A.2 Regional Climate data from the Swiss Alps

A.2.1 Precipitation at Bourg St. Pierre

Deseasonalized monthly maximum precipitation measured at Bourg St. Pierre, located in the south-western part of the Swiss Alps close to the pass of Grand St. Bernard, see Fig. A.2(a).

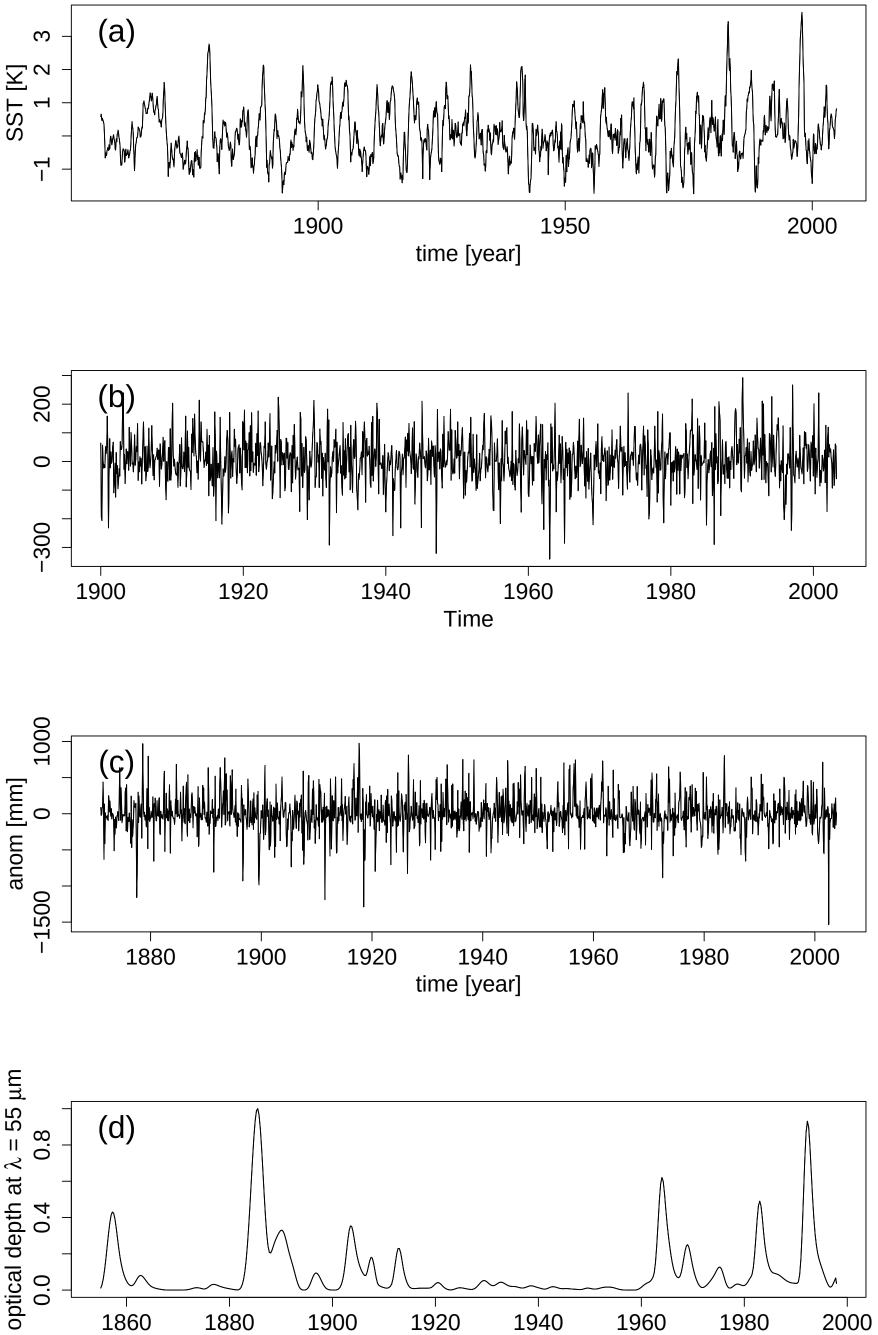

Figure A.1: (a) NINO3-index (SST anomaly in 150W to 90W and from 5N to 5S), (b) NAO-index (pressure gradient between Iceland and Gibraltar), (c) deseasonalized AIR-index (normalized all Indian rainfall anomalies), (d) Volcanic radiative Forcing index (optical depth at $\lambda = 55\mu m$ normalized to Krakatau eruption).

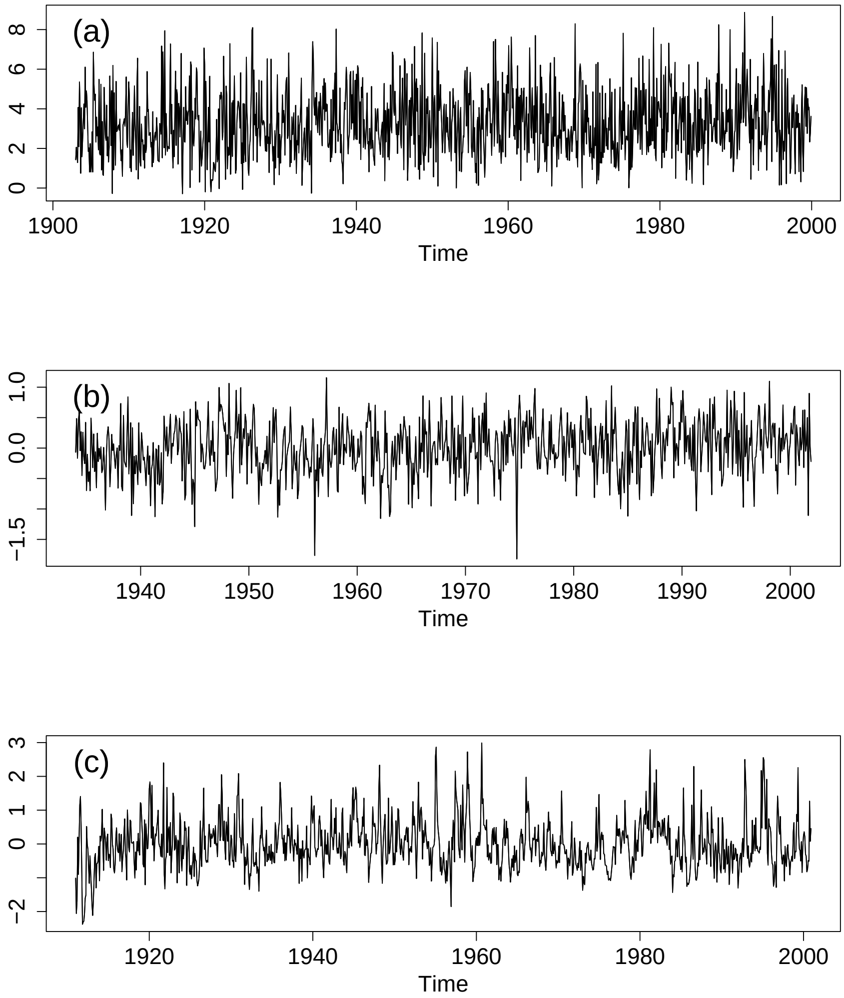

Figure A.2: (a) Monthly maximum deseasonalized precipitation at Bourg St. Pierre, (b) monthly mean deseasonalized temperature at Grand St. Bernard, (c) monthly mean deseasonalized run-off at Drance de Bagnes river

Deseasonalization has been carried out on the daily time series, i.e. every date has been normalized by subtracting the interannual mean value of this date and by dividing by the interannual variance. The data have been provided by the Swiss Meteorological Institute.

A.2.2 Temperature at Grand St. Bernard

Deseasonalized monthly mean temperature at Grand St. Bernard, see Fig. A.2(b). The deseasonalization has been performed as in the case of the precipitation data. These data have also been provided by the Swiss Meteorological Institute.

A.2.3 Run-Off of River Drance de Bagnes

Deseasonalized monthly mean run-off at the closeby Drance de Bagnes river that has a mean catchment altitude of 2630 m a.s.l. (catchment size 254 km^2) and a discharge regime strongly influenced by snow- and ice melt (see [42]), see Fig. A.2(c). For the deseasonalization, see the precipitation data. The construction of a dam in 1956 has been accounted for by calculating separate mean values for the intervals before and after the construction. The data have been provided by the Swiss Federal Office for the Environment (Hydrological Section).

Appendix B

Continuous Wavelet Spectral Analysis

B.1 Properties of the Transformation

B.1.1 Covariances

The covariances of the wavelet transformation $W_g s$ of a signal $s(t)$ with respect to the wavelet $g(t)$ corresponding to a translation $t - b'$ and dilation t/a' are given as follows:

$$\begin{aligned} W_g s[t-b'](b,a) &= W_g s[t](b-b',a) && \text{(B.1)} \\ W_g s[t/a'](b,a) &= W_g s[t](b/a',a/a') && \text{(B.2)} \end{aligned}$$

Here, the brackets [.] denote dependencies of a variable, whereas (.) denote dependencies of the resulting transformation. The invariant sets of the wavelet transformation are thus stripes of constant scale for translation and cones with edge in the origin for dilation. These properties constitute the affine group [40, 12].

B.1.2 Projection Property

On the one hand, applying the inverse transformation M_h to a wavelet transformation $W_g s(t)$ recovers the original signal, $M_h W_g s(t) = s(t)$, i.e. $M_h W_g = \mathbb{I}$. On the other hand, the CWT from 1-D to 2-D does not produce any new information, i.e. a continuous wavelet transform is neither orthogonal nor uncorrelated. Hence, not every function on the positive half-plane is a wavelet transformation. Taking an arbitrary function $f(b,a)$, the transformation $P_{g\to h} f(b,a) = W_h M_g f(b,a)$ to the time domain and back to the wavelet domain thus is a projector onto the space of all wavelet transformations[40]:

$$P^2_{g\to h} f(b,a) = P_{g\to h} f(b,a)$$

B.1.3 Reproducing Kernel

As a consequence of the projection property, App. B.1.2, $r(b,a)$ is a wavelet transformation, if and only if

$$r(b,a) = \int_0^\infty \frac{da'}{a'} \int_0^\infty db' \frac{1}{a'} K_{g,h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right) r(b',a'). \tag{B.3}$$

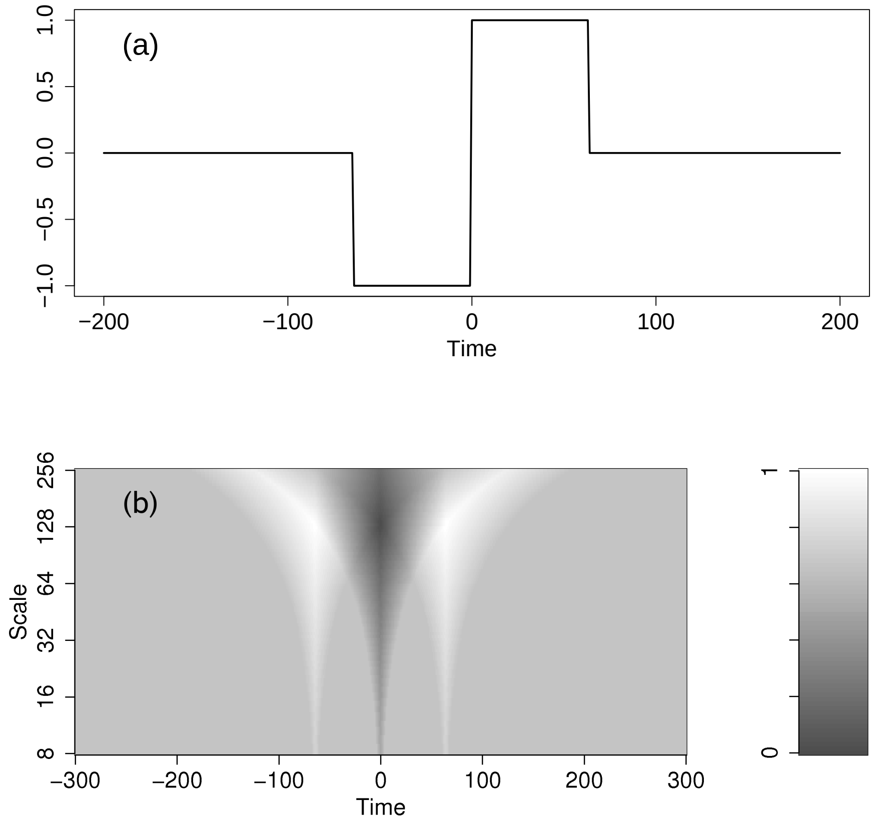

Figure B.1: Haar wavelet. (a) Representation in time domain for $a = 128$, (b) reproducing kernel for $a = 128$.

For details see [40]. The function $K_{g,h}((b-b')/a', a/a')$ is given as

$$K_{g,h}\big((b-b')/a', a/a'\big) = W_g h((b-b')/a') \tag{B.4}$$

and is called the reproducing kernel [40]. As an example, the reproducing kernel of the Morlet wavelet is given in App. B.1.5.

B.1.4 Example Wavelets

The Haar Wavelet

The Haar wavelet is a real wavelet defined as [40]:

$$g(t) = \begin{cases} -1 & \text{for} \quad -1/2 \le t < 0 \\ 1 & \text{for} \quad 0 \le t < 1/2 \\ 0 & \text{otherwise} \end{cases} \qquad \hat{g}(\omega) = 2i\frac{1-\cos\omega}{\omega} \tag{B.5}$$

Because it is very localized in the time domain (see Fig. B.1(a)), it exhibits a poor frequency localization $\hat{g}(\omega)$. This is also reflected in the reproducing kernel (see Fig. B.1(b)), that ranges over the whole scale axis. Thus, the Haar wavelet is suitable to detect very time localized structures like jumps inside a time series, but it is not suitable for a time scale analysis.

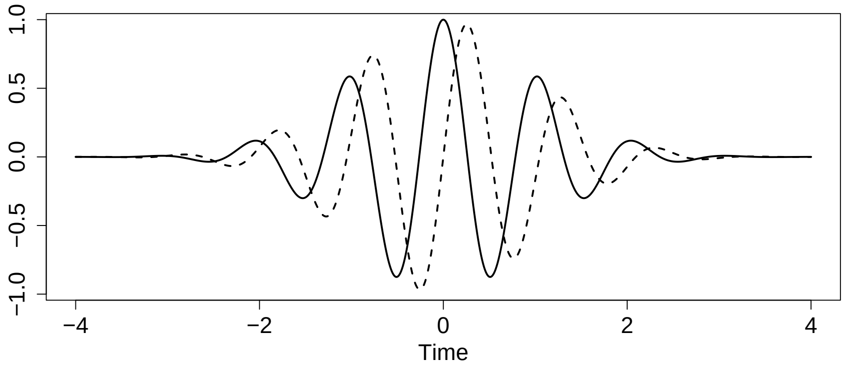

Figure B.2: Morlet wavelet in time domain for $\omega_0 = 6$ at scale $a = 1$. Solid line: real part, dashed line: imaginary part.

B.1.5 The Morlet Wavelet

The Morlet wavelet is defined as a complex exponential localized with a Gaussian envelope [40]:

$$g(t) = e^{i\omega_0 t} e^{-t^2/2}, \qquad \hat{g}(\omega) = \sqrt{2\pi} e^{-(\omega-\omega_0)^2/2}, \tag{B.6}$$

where ω_0 adjusts the time/scale resolution, see Fig. B.2. Strictly speaking, the Morlet wavelet is not progressive, as it has contributions on negative frequencies. Since its mean does not vanish, the Morlet wavelet is not even a wavelet. However, for sufficiently large ω_0, the negative frequency components and the mean almost vanish. For small ω_0 instead, one might observe artefacts. Using Eq. (B.4), the reproducing kernel of the Morlet wavelet results as

$$C(s_1, s_2, t_1, t_2) = \sqrt{\frac{2s_1 s_2}{s_1^2 + s_2^2}} \exp\left\{ i\omega_0 \frac{s_1 + s_2}{s_1^2 + s_2^2}(t_2 - t_1) \right\} \exp\left\{ -\frac{1}{2} \frac{(t_2 - t_1)^2 + \omega_0^2 (s_2 - s_1)^2}{s_1^2 + s_2^2} \right\} \tag{B.7}$$

(for the calculation, see Maraun and Kurths [67]) and is plotted in Fig. B.3.

B.1.6 Wavelet Transformation of Discrete Sampled Data

When estimating any continuous wavelet spectral measure from a discretely sampled time series, one is in principle able to calculate an infinite number of scales starting from the lowest one corresponding to the Nyquist frequency and ending at the length of the time series. However, the amount of independent scale information is limited by the reproducing kernel (see App. B.1.3). Thus, a reasonable choice of scales is the following:

$$a_j = a_0 2^{\frac{i-1}{N_{voice}}} \tag{B.8}$$

with $i = 1 \ldots N_{voice} \cdot N_{octave} + 1$, $a_{min} = a_0$ and $a_{max} = a_0 2^{N_{octave}}$. Here, a_0 corresponds to a frequency lower than or equal to the Nyquist frequency, N_{octave} denotes the number of octaves (i.e. powers of two), and N_{voice} the number of voices (i.e. calculated scales) per octave. Too high values of N_{voice} give no new information but only increase computational cost. We call the matrix given by all calculated wavelet spectral coefficients in the time/scale-plane *wavelet spectral matrix*.

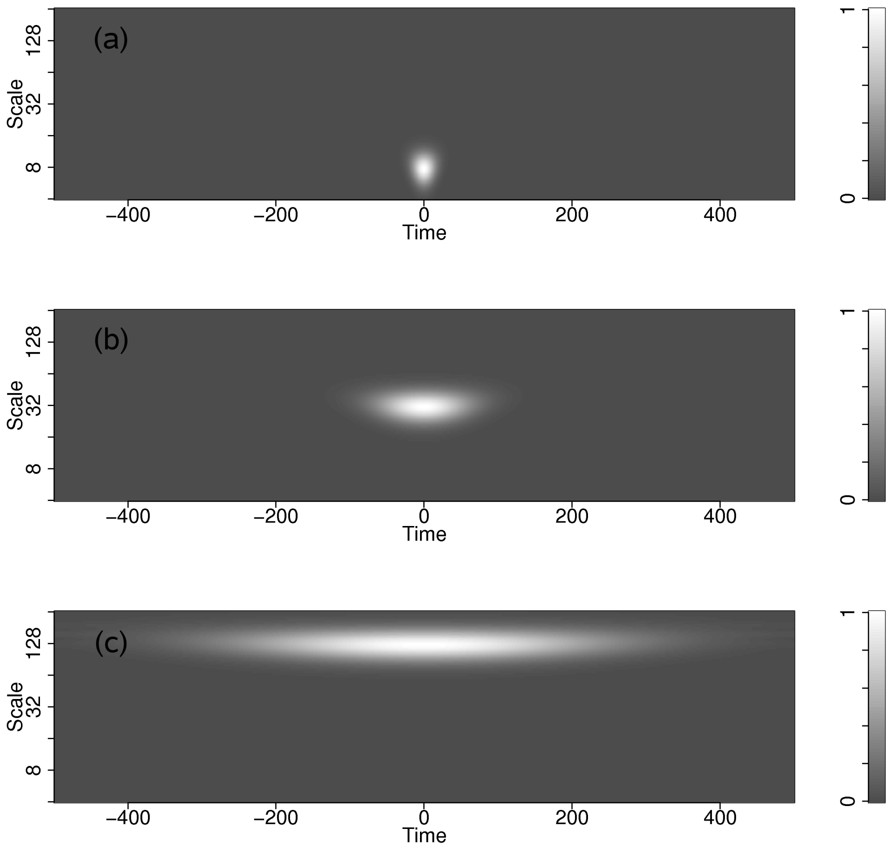

Figure B.3: Reproducing kernel of the Morlet wavelet for three different scales: (a) $a = 8$, (b) $a = 32$ and (c) $a = 128$. The width in time and in scale direction increases linearly with scale (i.e. in scale direction it appears constant on a logarithmic scale axis)

B.1.7 Cone of Influence

When calculating the wavelet transformation in the beginning or end of a finite data set, the analyzing wavelet is cut off[1]. Close to the boundaries this effect causes a strong bias of the estimation. To account for this boundary effect, one evaluates after which time a wavelet envelope has decayed to $1/e^2$ and uses this range to define the so-called cone of influence. Outside this cone with a distance from the boundaries proportional to scale, usually marked by a black line, wavelet coefficients have to be interpreted carefully.

---

[1] To avoid computational errors, one applies zero-padding, i.e. extends the time series with zeros to the left and to the right.

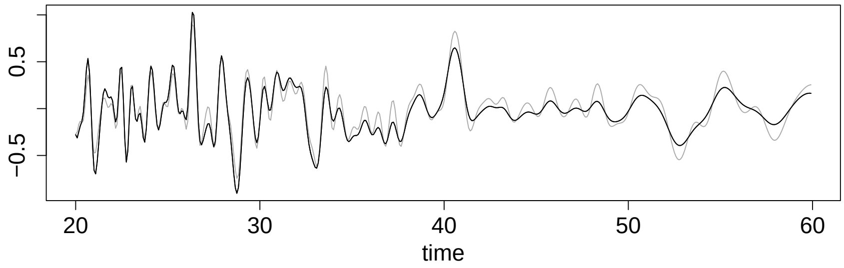

Figure B.4: Gaussian chirp, calculated from the same white noise realization but filtered with Morlet wavelets with different parameters $\omega_0 = 12$ (black) and $\omega_0 = 6$ (grey).

B.2 Locally Stationary Wavelet Processes

The idea presented in Chap. 3 is different from the approach of Nason and von Sachs [77], Nason et al. [78], who define stochastic processes by superimposing weighted wavelet "atoms",

$$X_t = \sum_{j=-J}^{-1} \sum_{k\in\mathbb{Z}} \omega_{j,k}\psi_{j,k}(t)\xi_{jk}, \quad \text{for} \quad t = 0, ..., T-1, \tag{B.9}$$

where ξ_{jk} denote random variables, the $\psi_{j,k}(t)$ are discrete non-decimated wavelets with amplitudes $\omega_{j,k}$. These wavelets are time continuous but confined to dyadic scales.

B.3 Properties of Gaussian Processes in Wavelet Domain

B.3.1 The Dependency on the Wavelet

Given a realization of white noise $\eta(t)$, the difference between the realizations $s_g(t)$ and $s_h(t)$ for a certain $m(b,a)$ but different wavelets g and h reads

$$\begin{aligned}\Delta_{g,h}(t) &= M_g m W_g \eta(t) - M_h m W_h \eta(t)\\ &= \underbrace{M_h W_h}_{\mathbb{I}} M_g m W_g \eta(t) - M_h m W_h \underbrace{M_g W_g}_{\mathbb{I}} \eta(t)\\ &= M_h \left[W_h M_g m \, - \, m W_h M_g\right] W_g \eta(t)\\ &= M_h \left[P_{g\to h} *, m\, \cdot\,\right] W_g \eta(t),\end{aligned}$$

with $P_{g\to h} = W_h M_g$. Figure B.4 shows an example of a Gaussian chirp, calculated from the same white noise realization but filtered with Morlet wavelets with different parameters $\omega_0 = 12$ (black) and $\omega_0 = 6$ (grey).

The commutator in the previous equation is given by the integral kernel

$$\left[m(b', a') \, - \, m(b,a)\right] P_{g\to h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right)$$

Developing $m(b', a')$ into a Taylor series around (b, a) gives

$$\begin{aligned}= \Big[\, & m(b,a) \,+\, (b-b')\partial_b m(b,a) + (a'-a)\partial_a m(b,a) + O((a-a')^2 + (b-b')^2) \\ & - m(b,a) \,\Big]\, P_{g\to h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right)\end{aligned}$$

For $P_{g\to h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right)$ sufficiently localized around (b, a), this reads

$$\approx \quad (b-b')\,\partial_b m(b,a)\, P_{g\to h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right) + (a'-a)\,\partial_a m(b,a) P_{g\to h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right)$$

With $P'_{g\to h}(b,a) = \frac{1}{a} b P_{g\to h}(b,a)$ and $P''_{g\to h}(b,a) = \left(\frac{1}{a} - 1\right) P_{g\to h}(b,a)$ we finally get

$$= \quad a\,\partial_b m(b,a) P'_{g\to h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right) \,+\, a\,\partial_a m(b,a) P''_{g\to h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right) \tag{B.10}$$

To ensure asymptotic independence of the chosen wavelet, it is necessary that $\Delta_{g,h}(t)$ vanishes for small scales. This is ensured in the following way:

- $P'_{g\to h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right)$ and $P''_{g\to h}\left(\frac{b-b'}{a'}, \frac{a}{a'}\right)$, given by the wavelets g and h, have to be sufficiently localized.
- $a\,\partial_b m(b,a)$ and $a\,\partial_a m(b,a)$ have to vanish for small scales. This is fulfilled for processes exhibiting the asymptotic behavior given by Eqn. 3.16 and 3.17.

Both asymptotic cases Eqn. 3.16 and 3.17 are equivalent in the sense that more and more reproducing kernels fit into local structures of the process; if the number of reproducing kernels in a local structure in time b or scale a direction is given as

$$N_{b,a} \sim \frac{1}{a\partial_{b,a} m(b,a)},$$

then the following equations for the behavior defined in Eqn. 3.16 and 3.17 hold:

$$\begin{aligned} &\text{I.} \quad N_{b,a} \sim \frac{1}{a^{\epsilon}} \to \infty \quad \text{for} \quad a \to 0, \\ &\text{II.} \quad N_{b,a} \sim 2^{\lambda} \to \infty \quad \text{for} \quad \lambda \to \infty. \end{aligned} \tag{B.11}$$

B.3.2 The Relation to Fourier Spectra

For stationary processes, i.e. $m(b,a) \equiv m(a)$, Eq. (3.14) in the Fourier domain reads

$$\begin{aligned}\hat{s}(\omega) &= \int_0^\infty \frac{da}{a} \frac{1}{\sqrt{a}} \hat{h}(a\omega) m(a) \hat{g}(a\omega) \hat{\eta}(\omega) \\ &= \underbrace{\left[\int_0^\infty \frac{da}{a} \frac{1}{\sqrt{a}} G(a\omega) m(a)\right]}_{f(\omega)} \hat{\eta}(\omega)\end{aligned}$$

Here, we abbreviate $G(a\omega) = \hat{h}(a\omega)\hat{g}(a\omega)$. In this context, $(\hat{\cdot})$ refers to the Fourier transformation. The term $f(\omega)$ denotes Fourier multipliers representing the Fourier spectrum of the process $m(a)$.
Developing $m(a)$ into a Taylor series around $2\pi/\omega$, $m(a) = m(2\pi/\omega) + (a - 2\pi/\omega)m'(2\pi/\omega) + O((a - 2\pi/\omega)^2)$ leads to

$$f(\omega) \approx m\left(\frac{2\pi}{\omega}\right)\left[\int_0^\infty \frac{da}{a}\frac{1}{\sqrt{a}}G(a\omega)\right] + m'\left(\frac{2\pi}{\omega}\right)\left[\int_0^\infty \frac{da}{a}\frac{1}{\sqrt{a}}G(a\omega)\left(a - \frac{2\pi}{\omega}\right)\right]$$

We factor out $2\pi/\omega$ in the second integral. If $G(a\omega)$ is well localized, the integrals might be considered as being constant. Finally we obtain

$$f(\omega) \approx m\left(\frac{2\pi}{\omega}\right)C_1 + \frac{2\pi}{\omega}m'\left(\frac{2\pi}{\omega}\right)C_2$$

As expected, the Fourier spectrum is given by the wavelet spectrum plus a correction term. The latter depends on the localization of the used wavelets and on the slope of the wavelet spectrum. For high frequencies the difference vanishes, if $m'(2\pi/\omega) < O(\omega)$ and if the process behaves as defined by Eq. (3.16). In case of a process behaving as defined by Eq. (3.17), the difference vanishes for $\lambda \to \infty$.

B.3.3 Distribution of the Wavelet Sample Spectrum

The Fourier (cross) periodogram asymptotically exhibits a χ^2 distribution with two degrees of freedom [9]. For the localized wavelet transformation, this property holds in a good approximation. Because the Fourier periodogram is asymptotically uncorrelated, smoothing over neighboring frequencies preserves a χ^2 distribution with the number of degrees of freedom depending on the length of the smoothing window. As wavelet coefficients are correlated due to the reproducing kernel, this result does no longer trivially hold for the wavelet sample spectrum.

The theorem of Ogasawara & Takahashi [97] states, when a sum of squares of correlated Gaussian distributed random variables is χ^2 distributed. Given a vector $\mathbf{Y}$ of Gaussian random variables

$$\mathbf{Y} \sim \mathcal{N}_n(0, \mathbf{\Sigma}) \tag{B.12}$$

with a covariance matrix $\mathbf{\Sigma}$, the product $\mathbf{Y}^T\mathbf{A}\mathbf{Y}$ is χ^2 distributed, when

$$\mathbf{\Sigma A \Sigma A \Sigma} = \mathbf{\Sigma A \Sigma} \tag{B.13}$$

holds.
In the case of smoothing neighboring values of the wavelet periodogram, we are interested in the sum of squares (in this example smoothing over three neighboring scales is considered. Generalization follows straight forward)

$$\mathbf{Y}^T\mathbf{A}\mathbf{Y} = W_1(s_1)W_1(s_1)^* + W_1(s_2)W_1(s_2)^* + W_1(s_3)W_1(s_3)^*. \tag{B.14}$$

This relation is obtained by defining

$$\mathbf{Y} = \begin{pmatrix} W_1(s_1) \\ W_1(s_1)^* \\ W_1(s_2) \\ W_1(s_2)^* \\ W_1(s_3) \\ W_1(s_3)^* \end{pmatrix} \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{B.15}$$

The covariance matrix then reads

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma & 0 & \sigma_{12} & 0 & \sigma_{13} & 0 \\ 0 & \sigma & 0 & \sigma_{12} & 0 & \sigma_{13} \\ \sigma_{12} & 0 & \sigma & 0 & \sigma_{23} & 0 \\ 0 & \sigma_{12} & 0 & \sigma & 0 & \sigma_{23} \\ \sigma_{13} & 0 & \sigma_{23} & 0 & \sigma & 0 \\ 0 & \sigma_{13} & 0 & \sigma_{23} & 0 & \sigma \end{pmatrix} \tag{B.16}$$

where the single entries can be calculated according to Eq. (3.10).
For these conditions, Eq. (B.13) reads

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & s_{1213} & 0 & s_{121323} & 0 & s_{131223} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & s_{121323} & 0 & s_{1223} & 0 & s_{231213} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & s_{131223} & 0 & s_{231213} & 0 & s_{1323} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{B.17}$$

with $s_{ijkl} = \sigma^2 + \sigma_{ij}^2 + \sigma_{kl}^2$ and $s_{ijklmn} = 2\sigma\sigma_{ij} + \sigma_{kl}\sigma_{mn}$ and is never fulfilled. Thus the smoothed wavelet sample spectrum is not χ^2 distributed, but rather follows a generalized χ^2-distribution.

B.3.4 Bias of the Spectral Estimator

The bias of the wavelet scalogram reads

$$\begin{aligned} \mathrm{Bias}(\widehat{S}_g(b,a)) &= \langle\, |W_k M_h\, m(b,a)\, W_g \eta(t)\,|^2 \,\rangle - |m(b,a)|^2 \\ &= W_k M_h\, \bar{W}_k \bar{M}_h\, m(b_1,a_1)\bar{m}(b_2,a_2)\, W_g \bar{W}_g\, \langle \eta(t_1)\eta(t_2) \rangle \;-\; |m(b,a)|^2. \end{aligned}$$

With $\langle \eta(t_1)\eta(t_2)\rangle = \delta(t_1 - t_2)$, $\sigma_\eta = 1$ and $W_g g((t_2 - b_2)/a_2) = K((b_1 - b_2)/a_2, a_1/a_2)$ we get

$$= W_k M_h\, \bar{W}_k \bar{M}_h\, m(b_1,a_1)\bar{m}(b_2,a_2)\, K\left(\frac{b_1 - b_2}{a_2}, \frac{a_1}{a_2}\right) \;-\; |m(b,a)|^2.$$

Developing $m(b_{1,2}, a_{1,2})$ into a Taylor series around (a,b), i.e.
$m(b_{1,2}, a_{1,2}) \approx m(a,b) + (b_{1,2} - b)\,\partial_b m(b_{1,2}, a_{1,2}) + (a_{1,2} - a)\,\partial_a m(b_{1,2}, a_{1,2})$, and writing
$(b_{1,2} - b)\,\partial_b m(b_{1,2}, a_{1,2}) + (a_{1,2} - a)\,\partial_a m(b_{1,2}, a_{1,2}) = f_{1,2}$,
$W_k M_h \bar{W}_k \bar{M}_h K((b_1 - b_2)/a_2, a_1/a_2) = C$, this leads to

$$\begin{aligned} \mathrm{Bias}(\widehat{S}_g(b,a)) \;\approx\; & (C-1)\,|m(b,a)|^2 \\ & -\; W_k M_h\, \bar{W}_k \bar{M}_h \big[m(b_1,a_1) f_2 + m(b_2,a_2) f_1 + f_1 f_2 \big] K\left(\frac{b_1 - b_2}{a_2}, \frac{a_1}{a_2}\right) \end{aligned}$$

If the wavelets are properly normalized, such that $C = 1$, the bias reduces to the second term. Following the same reasoning as in App. B.3.1, the bias vanishes for $a \to 0$ or $\lambda \to \infty$.

B.4 Estimating the Patch-size

B.4.1 Stochastic Root Finding

The significance level α_{aw} of the areawise test is a function of the critical area P_{crit}. Unfortunately, this function is not accessible analytically, such that it is impossible to choose a desired significance level α_{aw} and then straightforwardly calculate the corresponding critical area. In fact, one has to employ a root finding algorithm that solves the equation

$$f(P_{crit}) - \alpha_{aw} = 0$$

The estimation for $f(P_{crit})$ results from Monte Carlo simulations and thus is stochastic itself - conventional root finding algorithms fail to solve this problem. Thus we developed an ad hoc iterative procedure:

1. Choose three reasonable initial guesses for P_{crit} and estimate α_{aw} based on Monte Carlo simulations.

2. Assume a locally linear behavior of $f(.)$ around the root and fit a straight line to the three outcomes.

3. As next guess, choose the root of the straight line and estimate an improved α_{aw} based on Monte Carlo simulations.

4. Go back to 2., fit the straight line to all previous iterates (past iterates might be given a lower weight).

5. Choose a termination criterion, e.g. a desired accuracy or a maximum number of iterations.

Appendix C

Phase Coherence

C.1 A Stochastically Perturbed ENSO Delay Oscillator

Galanti and Tziperman [26] suggested a delayed oscillator exhibiting phase locking to the annual cycle as a conceptual model for ENSO. In another context, we modified this model slightly to account for stochastic wind perturbations [110]. The thermocline depth anomaly is modeled as

$$\begin{aligned}
h(t) &= e^{-\epsilon_m(\tau_1+\tau_2)} r_w r_e h_{eE}(t-\tau_1-\tau_2) \\
&- e^{-\epsilon_m(\tau_2+\frac{\tau_1}{2})} r_w \frac{1}{\beta\rho}\tau_W\tau_1 \cdot A^*\{\mu(t-\tau_2-\frac{\tau_1}{2})b_0T(t-\tau_2-\frac{\tau_1}{2})+A_{W_n}(t-\tau_2-\frac{\tau_1}{2})\} \\
&+ e^{-\epsilon_m\frac{\tau_2}{2}}\frac{\tau_W\tau_2}{\rho C_o}\{\mu(t-\frac{\tau_2}{2})b_0T(t-\frac{\tau_2}{2})+A_{W_e}(t-\frac{\tau_2}{2})\}. \qquad \text{(C.1)}
\end{aligned}$$

The parameters of the model are listed in Tab. C.1. Representing seasonal coupling, μ varies according to a sine function. The SST at the eastern boundary is given by the differential equation

$$\partial_t T = -\epsilon_T T - \gamma\frac{\bar{\omega}}{H_1}(T - T_{sub}(h)). \qquad \text{(C.2)}$$

Modeling the temperature of the upwelling water as a hyperbolic tangent introduces the main nonlinearity in the model. Equations C.1 and C.2 constitute a delay differential equation.

C.2 The Roessler Oscillator

Equations of the Roessler oscillator:

$$\begin{aligned}
\dot{x} &= -y - z \\
\dot{y} &= x + ay \\
\dot{z} &= b + (x - c)z
\end{aligned}$$

For the toy model, we choose the chaotic regime with parameters $a = 0.1$, $b = 0.1$ and $c = 14$. For $b = 1$ and $c = 8.5$, the Roessler attractor is funnel for a below a critical threshold $a_c \approx 0.19$.

Parameter	Description
a, b	scaling parameters of additional noise drive
A_{W_e}	equatorial additional noise drive
A_{W_n}	northern additional noise drive
A^*	relates non-equatorial windstress to equatorial SST
b_0	mean ocean atmosphere coupling strength
C_o	Kelvin wave propagation speed
dt	wind affected fraction of basin crossing time
ε	strength of seasonal variation
ϵ_m	damping coefficient of the ocean
ϵ_T	thermal damping coefficient
h	thermocline depth anomaly
H_1	upwelling layer depth
μ	seasonal varying coupling
ρ	density of ocean water
$r_{e/w}$	coastal boundary wave reflection coefficients
$\tau_{1,2}$	basin crossing time ($_1$ Rossby, $_2$ Kelvin)
T	temperature anomaly at eastern boundary
T_{sub}	temperature anomaly at depth H_1
$\bar{\omega}$	mean upwelling

Table C.1: Parameters of the conceptual ENSO delay oscillator.

Appendix D

Software

D.1 Wavelet Analysis

For the wavelet spectral analysis we developed the package bivar, written for the open source statistical environment R [94]. The software is based on routines written by Carmona et al. [12], included in their package Rwave. The package consists of object oriented functions to perform areawise significance tests for wavelet spectra and wavelet coherence and may be downloaded from the webpage

www.agnld.uni-potsdam.de/∼maraun/wavelets

It is planned to publish this package in the Comprehensive R Archive Network (CRAN).

D.2 DFA and Phase Analysis

The Routines for estimating the detrended fluctuation and the phase differences are written in C using the numerical recipes [89] and the implementation of the Hilbert transformation by Michael Rosenblum. We have collected these algorithms together with other data analysis software in the program nola (available on inquiry).

Bibliography

[1] J.B. Adams, M.E. Mann, and C.M. Ammann. Proxy evidence for an El Niño like response to volcanic forcing. *Nature*, 426:274–278, 2003.

[2] C. Allefeld and J. Kurths. Testing for phase synchronization. *Int. J. Bif. Chaos*, 14(2): 405–416, 2004.

[3] D. Avnir, O. Birham, D. Lindar, and O. Malcai. Is the geometry of nature fractal? *Science*, 279:39, 1998.

[4] M. Beniston. Mountain Climates and Climatic Change: An Overview of Processes Focusing on the European Alps. *Pure and Appl. Geophys.*, 162(8-9):1587–1606, 2005.

[5] J. Beran. *Statistics for Long-Memory Processes*. Monographs on Statistics and Applied Probability. Chapman & Hall, 1994.

[6] J. Beran, R.J. Bhansali, and D. Ocker. On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes. *Biometrika*, 85(4): 921–934, 1998.

[7] C. Bertrand, J.-P. van Ypersele, and A. Berger. Volcanic and solar impacts on climate since 1700. *Climate Dynamics*, 15(5):355–367, 1999.

[8] J. Bjerknes. Atmospheric teleconnections from the equatorial Pacific. *Mon. Wea. Rev*, 97:163–172, 1969.

[9] P. J. Brockwell and R. A. Davis. *Time Series: Theory and Models*. Heidelberg: Springer, 1991.

[10] M.A. Cane. The evolution of El Niño, past and future. *Earth Planet. Sci. Lett.*, 230: 227–240, 2005.

[11] J. G. Caputo, B. Malraison, and P. Atten. Determination of attractor dimension and entropy for various flows: An experimentalist's viewpoint. In G. Mayer-Kress, editor, *Dimensions and Entropies in Chaotic Systems*, volume 32 of *Springer Series in Synergetics*. Springer Verlag, Berlin, 1986.

[12] R. Carmona, W.-L. Hwang, and B Torresani. *Practical Time-Frequency Analysis. Gabor and Wavelet Transforms with an Implementation in S.* Academic Press, 1998.

[13] Z. Chen, P.Ch. Ivanov, K. Hu, and H.E. Stanley. Effect of nonstationarities on detrended fluctuation analysis. *Phys. Rev. E*, 65:041107, 2002.

[14] L. Cimponeriu, M.G. Rosenblum, T. Fieseler, J. Dammers, M. Schiek, M. Majtanik, P. Morosan, A. Bezerianos, and P.A. Tass. Inferring asymmetric relations between interacting neuronal oscillators. *Progress of Theoretical Physics Suppl.*, 150:22–36, 2003.

[15] P.M. Crowley, D. Maraun, and D. Mayes. How hard is the euro area core? A wavelet analysis of growth cycles in Germany, France and Italy. *submitted to Econ. J.*, 2006.

[16] I. Daubechies. *Ten Lectures on Wavelets.* Society for Industrial and Applied Mathematics, 1992.

[17] A.C. Davison and D.V. Hinkley. *Bootstrap Methods and their Application.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.

[18] M. Denker and G. Keller. Rigorous statistical procedures for data from dynamical systems. *J. Stat. Phys*, 44:67–93, 1986.

[19] H.A. Dijkstra. *Nonlinear Physical Oceanography*, volume 28 of *Atmospheric and Oceanographic Sciences Library.* Springer, 2005.

[20] P.J. Easterbrook, J.A. Berlin, R. Gopalan, and D.R. Matthews. Publication Bias in Clinical Research. *Lancet*, 337(8746):867–872, 1991.

[21] P. Elek and L. Márkus. A long range dependent model with nonlinear innovations for simulating daily river flows. *Natural Hazards and Earth System Sciences*, 4:227–283, 2004.

[22] K. Fraedrich and Richard Blender. Scaling of atmosphere and ocean temperature correlations in observations and climate models. *Phys. Rev. Lett.*, 90:108501, 2003.

[23] U. Frisch. *Turbulence : The Legacy of A. N. Kolmogorov.* Cambridge University Press, 1995.

[24] D. Gabor. Theory of communication. *J. IEEE*, 93:429–457, 1946.

[25] S. Gadgil. The Indian Monsoon and its Variability. *Annu. Rev. Earth Planet. Sci.*, 31: 429–467, 2003.

[26] E. Galanti and E. Tziperman. ENSO's phase locking to the Seasonal Cycle in the Fast-SST, Fast-Wave and Mixed-Mode Regimes. *J. Atmos. Sci.*, 57:2936–2950, 2000.

[27] C. Gaucherel. Use of wavelet transform for temporal characterisation of remote watersheds. *J. Hydrol.*, 269(3-4):101–121, 2002.

[28] A Gershunov, N. Schneider, and T. Barnett. Low-frequency modulation of the ENSO-Indian monsoon rainfall relationship: Signal or noise? *J. Clim.*, 14(11):437, 2001.

[29] L. Giratis, P. Kokoszka, and R. Leipus. Testing for long memory in the presence of a general trend. *J. Appl. Prob.*, 38:1033–1054, 2001.

[30] B.N. Goswami. Interannual variations of Indian summer monsoon in a GCM: External conditions versus internal feedbacks. *J. Clim.*, 11(4):501–522, 1998.

[31] R. B. Govindan, D. Vyushin, A. Bunde, S. Brenner, S. Havlin, and H.-J. Schellnhuber. Global climate models violate scaling of the observed atmospheric variability. *Phys. Rev. Lett.*, 89(2):028501, 2002.

[32] C.W.J Granger. Long memory relationships and the aggregation of dynamic models. *J. of Econometrics*, 14:227 – 238, 1980.

[33] B.T. Grenfell, O.N. Bjœrnstad, and J. Kappey. Travelling waves and spatial hierarchies in measles epidemics. *Nature*, 414:716–723, 2001.

[34] D. Gu and S.G.H. Philander. Secular changes of annual and interannual variability in the tropics during the past century. *J. Climate*, 8:864–876, 1995.

[35] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13:49–52, 1902.

[36] W. Haeberli, R. Frauenfelder, and M. Hoelzle, editors. *Glacier mass balance bulletin no. 7 (2000-2001).* IAHS - UNEP - UNESCO - WMO, Zürich, Switzerland, 2003.

[37] K. Hasselmann. Stochastic climate models: Part I. Theory. *Tellus*, 28(6):473–485, 1976.

[38] G.H. Haug, D. Günther, L.C. Peterson, D.M. Sigman, K.A. Hughen, and B. Aeschlimann. Climate and the Collapse of the Maya Civilization. *Science*, 299:1731–1735, 2003.

[39] J.M. Hausdorff and C.-K. Peng. Multiscaled randomness: A possible source of 1/f noise in biology. *Phys. Rev. E*, 54(2):2154, 1996.

[40] M. Holschneider. *Wavelets: an analysis tool.* Oxford University Press, 1998.

[41] J. Honerkamp. *Statistical Physics.* Springer, Berlin, 1998.

[42] P. Horton, B. Schaefli, B. Hingray, A. Mezghani, and A. Musy. Assessment of climate change impacts on Alpine discharge regimes with climate model uncertainty. *Hydrol. Proc.*, in press, 2006.

[43] J.R.M. Hoskins. Fractional differencing. *Biometrica*, 68:165–176, 1981.

[44] K. Hu, P. Ch. Ivanov, Z. Chen, P. Carpena, and H. E. Stanley. Effect of trends on Detrended Fluctuation Analysis. *Phys. Rev. E*, 64(011114), 2001.

[45] J. Huang, K. Higuchi, and A. Shabbar. The relationship between the North Atlantic Oscillations and El Niño-Southern Oscillation. *Geophys. Res. Lett.*, 25(14):2707–2710, 1998.

[46] L. Hudgins, C.A. Friehe, and M.E. Mayer. Wavelet Transform and Atmospheric Turbulence. *Phys. Rev. Lett.*, 71(20):3279–3282, 1993.

[47] J.W. Hurrel and H. van Loon. Decadal variations in climate associated with the north Atlantic oscillation. *Clim. Change*, 36:301–326, 1997.

[48] F.-F. Jin. An equatorial ocean recharge paradigm for ENSO. 1. Conceptual model. *J. Atmos. Sci.*, 54(7):811–829, 1997.

[49] P.D. Jones, T Jónsson, and D. Wheeler. Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and South-West Iceland. *Int. J. Climatol.*, 17:1433–1450, 1997.

[50] G. Kaiser. *A friendly Guide to Wavelets*. Birkhäuser, 1994.

[51] M. Kallache, H.W. Rust, and J. Kropp. Trend assessment: Applications for hydrology and climate research. *Nonlinear Proc. Geoph.*, 12:201–210, 2005.

[52] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde. Detecting long-range correlations with detrended fluctuation analysis. *Physica A*, 295: 441, 2001.

[53] H. Kantz and T. Schreiber. Dimension estimates and physiological data. *Chaos*, 5: 143–153, 1995.

[54] A. Kaplan, M.A. Cane, Y. Kushnir, A.C. Clement, M.B. Blumenthal, and B. Rajagopalan. Analyses of global sea surface temperature 1856-1991. *J. Geophys. Res.*, 103:18567–18589, 1998.

[55] E. Koscielny-Bunde, A. Bunde, S. Havlin, H. E. Roman, Y. Goldreich, and H.-J. Schellnhuber. Indication of a universal persistence law governing atmospheric variability. *Phys. Rev. Lett.*, 81(3):729, 1998.

[56] V. Krishnamurthy and B.N. Goswami. Indian Monsoon-ENSO Relationship on Interdecadal Timescale. *J. Clim.*, 13(3):579–595, 2000.

[57] R. Krishnan and M. Sugi. Pacific decadal oscillation and variability of the Indian summer monsoon rainfall. *Clim. Dyn.*, 21:233–242, 2003.

[58] K.K. Kumar and R. Kleeman. Epochal Changes in Indian monsoon-ENSO precursors. *Geophys. Res. Lett.*, 26(1):75–78, 1999.

[59] K.K. Kumar, M.K. Sonan, and K.R. Kumar. Seasonal forecasting of Indian summer monsoon rainfall. *Weather*, 50:449–466, 1995.

[60] K.K. Kumar, B. Rajagopalan, and M.A. Cane. On the Weakening Relationship Between the Indian Monsoon and ENSO. *Science*, 284:2156–2159, 1999.

[61] D. Labat, R. Ababou, and A. Mangin. Rainfall-runoff relations for karstic springs. Part II: continuous wavelet and discrete orthogonal multiresolutions analyses. *J. Hydrol.*, 238:149–178, 2000.

[62] D. Labat, J. Ronchail, J. Callede, J.L. Guyot, E. de Oliveira, and W. Guimaraes. Wavelet analysis of Amazon hydrological regime variability. *Geophys. Res. Lett.*, 31(2): L02501, 2004.

[63] D. Labat, J. Ronchail, and J.L. Guyot. Recent advances in wavelet analyses: Part 2–Amazon, Parana, Orinoco and Congo discharges time scale variability. *J. Hydrol.*, 314(1-4):289–311, 2005.

[64] M. Lafrenière and M. Sharp. Wavelet analysis of inter-annual variability in the runoff regimes of glacial and nival stream catchments, Bow Lake, Alberta. *Hydrol. Proc.*, 17: 1093–1118, 2003.

[65] Ta-Hsin Li. Wavelet Spectrum and Its Characterization Property for Random Processes. *IEEE Transactions on Information Theory*, 48(11):2922–2937, 2002.

[66] B.B. Mandelbrot and J. W. van Ness. Fractional brownian motions, fractional noises and applications. *SIAM Rev.*, 10:422–437, 1968.

[67] D. Maraun and J. Kurths. Cross wavelet analysis. Significance testing and pitfalls. *Nonlin. Proc. Geoph.*, 11(4):505–514, 2004.

[68] D. Maraun and J. Kurths. Epochs of phase coherence between El Niño/ Southern Oscillation and Indian monsoon. *Geophys. Res. Lett.*, 32:L15709, doi:10.1029/2005GL023225, 2005.

[69] D. Maraun and B. Schaefli. Wavelet Spectral and Cross Spectral Analysis in Hydrology. Springer, 2006.

[70] D. Maraun, H.W. Rust, and J. Timmer. Tempting long-memory. On the interpretation of DFA results. *Nonlin. Proc. Geoph.*, 11(4):495–503, 2004.

[71] D. Maraun, J. Kurths, and M. Holschneider. Nonstationary Gaussian processes in wavelet domain: Definitions, estimation and significance testing. *submitted to Phys. Rev. E*, 2006.

[72] J. Marshall, Y. Kushnir, D. Battisti, P. Chang, A. Czaja, R. Dickson, J. Hurrel, M. McCartney, R. Saravanan, and M. Visbeck. North Atlantic Climate Variability: Phenomena, Impacts and Mechanisms. *Int. J. Climatol.*, 21:1863–1898, 2001.

[73] A. I. McLeod and K. W. Hipel. Preservation of the Rescaled Adjusted Range. 1. A Reassessment of the Hurst Phenomenon. *Wat. Resour. Res.*, 14(3):491–508, 1978.

[74] A. Montanari, R. Rosso, and M. S. Taqqu. A seasonal fractional ARIMA model applied to the Nile River monthly flows. *Wat. Resour. Res.*, 36(5):1249–1259, 2000.

[75] D.A. Mooley and B. Parthasarathy. Fluctuations in all-India summer monsoon rainfall during 1871-1978. *Climate Change*, 6:287–301, 1984.

[76] H. Moritz. General Considerations Regarding Inverse and Related Problems. In *Inverse Problems: Principles and Applications in Geophysics, Technology and Medicine*. Akademie Verlag, 1993.

[77] G.P. Nason and R. von Sachs. Wavelets in time series analysis. *Phil. Trans. R. Soc. Lond. A*, 357(1760):2511–2526, 1999.

[78] G.P. Nason, R. von Sachs, and G. Kroisandt. Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *J. Roy. Stat. Soc. B*, 62:271–292, 2000.

[79] G.V. Osipov, B. Hu, C.S. Zhou, M.V. Ivanchenko, and J. Kurths. Three types of transitions to phase synchronization in chaotic oscillators. *Phys. Rev. Lett.*, 91:024101, 2003.

[80] A. Papoulis. *Probability, Random Variables and Stochastic Processes.* New York: McGraw-Hill, 3rd edition, 1991.

[81] J.D. Pelletier and D. Turcotte. Self-Affine Time Series: II. Applications and Models. *Advances in Geophysics*, 40:91–166, 1999.

[82] C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, M. Simons, and H. E. Stanley. Finite-size effects on long-range correlations: Implications for analyzing DNA sequences. *Phys. Rev. E*, 47(5):3730–3733, 1993.

[83] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. Mosaic organization of DNA nucleotides. *Phys. Rev. E*, 49(2):1685, February 1994.

[84] D.B. Percival and A.T. Walden. *Wavelet Methods for Time Series Analysis.* Cambridge Univ. Press, 2000.

[85] S.G. Philander. *El Niño, La Niña, and the Southern Oscillation.* Academic Press, 1990.

[86] S.G. Philander. A review of tropical ocean-atmosphere interactions. *Tellus*, 51A-B: 71–90, 1999.

[87] A.S. Pikovsky, M.G. Rosenblum, and J. Kurths. *Synchronization: A universal concept in nonlinear sciences.* Cambridge Univ. Press, 2001.

[88] K.R. Popper. *The Logic of Scientific Discovery.* Hutchinson & Co, 1959.

[89] W. H. Press, S. A. Teukolsky, W. T. Vetter, and B. P. Flannery. *Numerical Recipes in C.* Cambridge University Press, Cambridge, 2nd edition, 1992.

[90] M. B. Priestley. *Spectral Analysis and Time Series.* Academic Press, 1992.

[91] M. B. Priestley. *Non-linear and Non-Stationary Time Series Analysis.* Academic Press, 1998.

[92] L. Qiu and M.-H. Er. Wavelet spectrogram of noisy signals. *Int. J. Electronics*, 79(5): 665–677, 1995.

[93] R.Q. Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger. Performance of different synchronization measures in real data: A case study on electroencephalographic signals. *Phys. Rev. E*, 65:041903, 2002.

[94] R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2006. URL `http://www.R-project.org`. ISBN 3-900051-07-0.

[95] S. Rahmstorf, D. Archer, D.S. Ebel, O. Eugster, J. Jouzel, D. Maraun, U. Neu, G.A. Schmidt, J. Severinghaus, A.J. Weaver, and J. Zachos. Cosmic rays, carbon dioxide, and climate. *Eos, Trans. AGU*, 85(4):38, 2004.

[96] S. Rahmstorf, D. Archer, D.S. Ebel, O. Eugster, J. Jouzel, D. Maraun, U. Neu, G.A. Schmidt, J. Severinghaus, A.J. Weaver, and J. Zachos. Reply to comment on "cosmic rays, carbon dioxide, and climate". *Eos, Trans. AGU*, 85(48):511, 2004.

[97] C. Rao. *Linear Statistical Inference and its Applications.* Wiley, 1965.

[98] E.M. Rasmusson and T.H. Carpenter. The Relationship Between Eastern Equatorial Pacific Sea Surface Temperatures and Rainfall over India and Sri Lanka. *Mon. Wea. Rev.*, 111(3):517–528, 1983.

[99] E.M. Rasmusson, X. Wang, and C.F. Ropelewski. The biennial component of ENSO variability. *J. Mar. Syst.*, 1:71–96, 1990.

[100] N. Rimbu, G. Lohmann, T. Felis, and J. Pätzold. Shift in ENSO Teleconnections Recorded by a Northern Red Sea Coral. *J. Clim.*, 16:1414–1422, 2002.

[101] A. Robock. Volcanic eruptions and climate. *Rev. Geophys.*, 38(2):191–219, 2000.

[102] A. Robock, M. Mu, K. Vinnikov, and D. Robinson. Land Surface Conditions over Eurasia and Indian Summer Monsoon Rainfall. *J. Geophys. Res.-Atmos.*, 108(D4):663, 2002.

[103] M. G. Rosenblum and A. S. Pikovsky. Detecting direction of coupling in interacting oscillators. *Phys. Rev. E*, 64:045202(R), 2001.

[104] M.G. Rosenblum, A.S. Pikovsky, and J. Kurths. Phase synchronization of chaotic oscillators. *Phys. Rev. Lett*, 76(11):1804–1807, 1996.

[105] M.G. Rosenblum, L. Cimponeriu, A. Bezerianos, A. Patzak, and R. Mrowka. Identification of coupling direction: Application to cardiorespiratory interaction. *Phys Rev. E*, 65(4):041909, 2002.

[106] A.G. Rossberg, K. Bartholomé, and J. Timmer. Data-driven optimal filtering for phase and frequency of noisy oscillations: Application to vortex flow metering. *Phys. Rev. E*, 69:016216, 2004.

[107] H.W. Rust. *In preparation.* PhD thesis, University of Potsdam, 2006.

[108] S. Sarkar, R.P. Singh, and M. Kafatos. Further evidences for the weakening relationship of Indian rainfall and ENSO over India. *Geophys. Res. Lett.*, 31:L13209, 2004.

[109] M. Sato, J.E. Hansen, M.P. McCormick, and J.B. Pollack. Stratospheric aerosol optical depths, 1850-1990. *J. Geophys. Res.*, 98(22):987–994, 1993.

[110] J. Saynisch, J. Kurths, and D. Maraun. A Conceptual ENSO Modell under Realistic Noise Forcing. *accepted for publication in Nonlin. Proc. Geophys.*, 2006.

[111] T. Schreiber. Measuring information transfer. *Phys. Rev. Lett.*, 85(2):461–464, 2000.

[112] T. Schreiber and A. Schmitz. Surrogate time series. *Physica D*, 142(3-4):346–382, 2000.

[113] B.C. Si and T.B. Zeleke. Wavelet coherency analysis to relate saturated hydraulic properties to soil physical properties. *Wat. Res. Res.*, 41(11):W11424, 2005.

[114] M.B. Smith, V.I. Koren, Z. Zhang, S.M. Reed, J.-J. Pan, and F. Moreda. Runoff response to spatial variability in precipitation: an analysis of observed data. *J. Hydrol.*, 298(1-4):267–286, 2004.

[115] J.-D. Stanley, M.D. Krom, R.A. Cliff, and J.C. Woodward. Nile Flow Failure at the End of the Old Kingdom, Egypt: Strontium Isotopic and Petrologic Evidence. *Geoarchaeology*, 18(3):395–402, 2003.

[116] M.J. Suarez and P.S. Schopf. A delayed action oscillator for ENSO. *J. Atmos. Sci.*, 45: 3283–3287, 1988.

[117] M. S. Taqqu, V. Teverovsky, and W. Willinger. Estimators for long-range dependence: An empirical study. *Fractals*, 3(4):785–798, 1995.

[118] P. Tass, M.G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H.-J. Freund. Detection of n:m Phase Locking from Noisy Data: Application to Magnetoencephalography. *Phys. Rev. Lett.*, 81(15):3291–3294, 1998.

[119] J. Timmer and M. König. On generating power law noise. *A&A*, 300(3):707–710, 1995.

[120] J. Timmer, S. Haeussler, M. Lauk, and C. H. Luecking. Pathological tremors: Deterministic chaos or nonlinear stochastic oscillators? *Chaos*, 10:278–288, 2000.

[121] C. Torrence and G. Compo. A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.*, 79:61–78, 1998.

[122] C. Torrence and P.J. Webster. The annual cycle of persistence in the El Niño Southern Oscillation. *Quart. J. Roy. Met. Soc.*, 124:1985–2004, 1998.

[123] C. Torrence and P.J. Webster. Interdecadal Changes in the ENSO-Monsoon System. *J. Clim.*, 12:2679–2690, 1999.

[124] K.E. Trenberth. Persistence of Daily Geopotential Heights over the Southern Hemisphere. *Month.Wea.Rev.*, 113:38–53, 1985.

[125] A.A. Tsonis and J.B. Elsner. Testing for scaling in natural forms and observables. *J. Stat. Phys.*, 81(5/6):869–880, 1995.

[126] A.A. Tsonis, P.J. Roeber, and J.B. Elsner. On the existence of spatially uniform scaling laws in the climate system. In M.M. Novak, editor, *Paradigms of complexity*, pages 25–28. World Scientific, Singapore, 2000.

[127] H. van Loon and J.C. Rogers. Seesaw in winter temperatures between Greenland and Northern Europe. 1. General description. *Mon. Weath. Rev.*, 106:296–310, 1978.

[128] H. von Storch and F.W. Zwiers. *Statistical Analysis in Climate Research*, chapter Specific Statistical Concepts in Climate Research. Cambridge University Press, Cambridge, 1999.

[129] G.T. Walker and E.W. Bliss. World Weather V. *Mem. Roy. Meteor. Soc.*, 4:53–84, 1932.

[130] G.T. Walker and E.W. Bliss. World Weather VI. *Mem. Roy. Meteor. Soc.*, 4:119–139, 1937.

[131] P.J. Webster and S. Yang. Monsoon and ENSO: Selectively interactive systems. *Q.J.R. Meteorol. Soc.*, 118:887–926, 1992.

[132] P.J. Webster, V.O. Magaña, T.N. Palmer, J. Shukla, R.A. Tomas, M. Yanai, and T. Yasunari. Monsoons: Processes, predictability, and the prospects for prediction. *J. Geophys. Res.*, 103(C7):14451–14510, 1998.

[133] S. Yang. ENSO-snow-monsoon associations and seasonal-interannual predictions. *Int. J. Climatol.*, 16(2):125–134, 1996.

[134] S.E. Zebiak and M.A. Cane. A model El Niño Southern Oscillation. *Mon. Wea. Rev.*, 115:2262–2278, 1987.

[135] H. Zhao and G.W.K. Moore. On the relationship between Tibetan snow cover, the Tibetan plateau monsoon and the Indian summer monsoon. *Geophys. Res. Lett.*, 31: L14204, 2004.

Acknowledgments

I would like to thank my supervisor Prof. Dr. Kurths for giving me the possibility to work in his group and for his support and guidance through the last three years.

Further, I thank Prof. Dr. Matthias Holschneider for the good collaboration and his guidance in the field of wavelet analysis. Without his help, the chapter about wavelet analysis would have been much shorter and shirtsleeve.

I am grateful for the time series education I experienced in the lectures of Prof. Dr. Jens Timmer. I also would like to thank him for the good cooperation and his advice during this thesis.

Special thanks go to Henning Rust, for the intensive cooperation throughout the last years. I really appreciate his advice and the inspiring discussions either on my balcony, in various cafes in Berlin or Potsdam or in the statcrew-seminars.

I wish to express my gratitude to Dr. Udo Schwarz for his tremendous patience during endless discussions, in which I could benefit a lot from his scientific knowledge.

I very much appreciate the advice and the motivation I received from Prof. Dr. Stefan Rahmstorf, also his confidence during our cooperation.

I am grateful for discussions with Prof. Dr. Gerrit Lohmann, especially during the time he invited me to Bremerhaven.

I would like to acknowledge the collaboration with Dr. Bettina Schaefli and Prof. Dr. Patrick Crowley.

I especially thank Dr. Sven Titz for the guidance through the first year of my work in Potsdam and for all his advice also after he left the group.

For many helpful discussions and advice, I also would like to thank Prof. Dr. Rainer Dahlhaus, Prof. Dr. Arkady Pikovsky, Dr. Michael Rosenblum, Malaak Kallache, Boris Orlowski, Dr. Suhasini Subba Rao and all the others I cannot remember in this moment.

I very much appreciate the help of James Ong, who read all my recent papers and corrected my English before submission, as well as the constant and reliable help of Birgit Voigt and Jörg-Uwe Tessmer.

I thank Dr. Ernest Montbrió, Dr. Gert Zöller, Gorka Zamora, María Herrojo, Antonio Gámez, Yong Zhou, Natalia Tukhlina, Komalapriya Chandrasekaran, Dr. Murilo Baptista, Lucia Zemanova, Dr. Ralf Steuer and all my former and present colleagues in the group and especially in my office for the nice time we had.

My very personal thanks and wishes are dedicated to Antje Manthey.