Cross-frequency interactions in air temperature records

Milan Paluš

Institute of Computer Science
Academy of Sciences of the Czech Republic
Prague, Czech Republic
E-Mail: `mp@cs.cas.cz`

DAMES 2012, Potsdam, October 8, 2012

Complex dynamics in the atmosphere

- 1980's low dimensional chaos (Nicolis & Nicolis, Tsonis & Elsner, Fraedrich)
  - criticism by Grassberger, Lorenz
  - Paluš & Novotná 1994: linearity of $x(t) – x(t + \tau)$ interactions in temperature data
  - $x(t) – y(t)$ interactions – next talk by J. Hlinka
- Tsonis 2012 subsystems of low dimensionality
- long-range dependence, fractality (Koscielny-Bunde et al., 1998, Bunde & Havlin 2002, Eichner et al., 2002 ...)
  - criticism by Maraun et al., 2004
- multifractality (Schmitt et al. 1995, Ashkenazy et al., 2003, Zhou et al., 2010)

Cyclic phenomena hidden in colored noise

- Monte Carlo SSA (Ghil et al., Allen & Smith)
- Enhanced MC SSA (Paluš & Novotná) cycles of higher regularity than filtered noise
- detection and extraction of cycles $\longrightarrow$ interactions
  - Feliks et al.: phase synchronization in 7–8 yr cycles between areas – teleconnection
  - Paluš & Novotná: phase synchronization in 7–8 yr cycles between temperature and solar/geomagnetic activity – solar-terrestrial relations
  - Stein et al. 2011: phase synchronization of El-Niño and annual cycle

Emergent phenomena in complex systems

COMPLEX DYNAMICS
Not explained by a sum of properties of system components

INTERACTIONS OF SYSTEM COMPONENTS
EMERGENT PHENOMENA

STUDY OF INTERACTIONS

- clues to understanding complex behaviour
- facts for model building
- characterization – diagnostics

Measures from Information Theory

variable X, probability distribution $p(x)$

$$H(X) = -\sum_{x \in \Xi} p(x) \log p(x) \tag{1}$$

$$H(X, Y) = -\sum_{x \in \Xi} \sum_{y \in \Upsilon} p(x, y) \log p(x, y) \tag{2}$$

$$H(Y|X) = -\sum_{x \in \Xi} \sum_{y \in \Upsilon} p(x, y) \log p(y|x) \tag{3}$$

Mutual Information, Conditional MI

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{4}$$

$$I(X;Y|Z) = H(X|Z) + H(Y|Z) - H(X,Y|Z) \tag{5}$$

$$I(X;Y|Z) = I(X;Y;Z) - I(X;Z) - I(Y;Z) \tag{6}$$

Estimating H(X), I(X,Y) from time series

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Physics Reports 441 (2007) 1–46

PHYSICS REPORTS

www.elsevier.com/locate/physrep

Causality detection based on information-theoretic approaches in time series analysis

Katerina Hlaváčková-Schindler[a,*], Milan Paluš[b], Martin Vejmelka[b], Joydeep Bhattacharya[a,c]

[a] *Commission for Scientific Visualization, Austrian Academy of Sciences, Donau-City Str. 1, A-1220 Vienna, Austria*
[b] *Institute of Computer Science, Academy of Sciences of the Czech Republic Pod Vodárenskou věží 2, 18207 Praha 8, Czech Republic*
[c] *Department of Psychology, Goldsmiths College, University of London, New Cross SE14 6NW, London, UK*

Accepted 24 December 2006
Available online 6 February 2007
editor: I. Procaccia

Abstract

Synchronization, a basic nonlinear phenomenon, is widely observed in diverse complex systems studied in physical, biological and other natural sciences, as well as in social sciences, economy and finance. While studying such complex systems, it is important not only to detect synchronized states, but also to identify causal relationships (i.e. who drives whom) between concerned (sub) systems. The knowledge of information-theoretic measures (i.e. mutual information, conditional entropy) is essential for the analysis of information flow between two systems or between constituent subsystems of a complex system. However, the estimation of these measures from a set of finite samples is not trivial. The current extensive literatures on entropy and mutual information estimation provides a wide variety of approaches, from approximation-statistical, studying rate of convergence or consistency of an estimator for a general distribution, over learning algorithms operating on partitioned data space to heuristical approaches. The aim of this paper is to provide a detailed overview of information theoretic approaches for measuring causal influence in multivariate time series and to focus on diverse approaches to the entropy and mutual information estimation.

CONDITIONAL MUTUAL INFORMATION

$$I(\vec{Y}(t); \vec{X}(t+\tau)|\vec{X}(t)) =$$

$$I\Big(\big(y(t), y(t-\rho), \ldots, y(t-(m-1)\rho)\big); x(t+\tau)|$$
$$\big(x(t), x(t-\eta), \ldots, x(t-(n-1)\eta)\big)\Big)$$

$$= I\Big(y(t); x(t+\tau)|\big(x(t), x(t-\eta), \ldots, x(t-(n-1)\eta)\big)\Big)$$

SCALE-SPECIFIC INTERACTIONS

OSCILLATORY PROCESS – specific frequency

BROAD-BAND SIGNALS

- DIGITAL FILTERING
- WAVELET DECOMPOSITION
- EMPIRICAL MODE DECOMPOSITION
- SINGULAR SPECTRUM ANALYSIS

- SCALE-SPECIFIC SYNCHRONIZATION
- SCALE-SPECIFIC GRANGER CAUSALITY
- CROSS-SCALE INTERACTIONS
- CROSS-FREQUENCY COUPLING

Phase dynamics approach

ANALYTIC SIGNAL

$$\psi(t) = s(t) + j\hat{s}(t) = A(t)e^{j\phi(t)} \tag{7}$$

INSTANTANEOUS PHASE

$$\phi(t) = \arctan \frac{\hat{s}(t)}{s(t)} \tag{8}$$

INSTANTANEOUS AMPLITUDE

$$A(t) = \sqrt{\hat{s}(t)^2 + s(t)^2} \tag{9}$$

FILTERING $\longrightarrow$ HILBERT TRANSFORM

COMPLEX CONTINUOUS WAVELET TRANSFORM

Cross-frequency interactions

- phase–phase
- amplitude–amplitude
- phase–amplitude
  - neurophysiology: phase of slow oscillations (δ, θ) modulates the amplitude of fast oscillations (γ)

CAUSAL PHASE –> AMPLITUDE INTERACTIONS
in about a century long records of daily near-surface air temperature records from European stations

- phase ϕ_1 of slow oscillations (around 10 year period)
- amplitude A_2 of higher-frequency variability (periods 5 years and less)
- $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), \ldots, A_2(t-m\eta))$
- testing using surrogate data approach
  - Fourier transform (FT) surrogate data (Theiler et al.)
  - multifractal (MF) surrogate data (Paluš)

TESTING INTERACTIONS WITH & WITHIN MULTISCALE PROCESSES

PRL **101,** 134101 (2008)   PHYSICAL REVIEW LETTERS   week ending 26 SEPTEMBER 2008

Bootstrapping Multifractals: Surrogate Data from Random Cascades on Wavelet Dyadic Trees

Milan Paluš*

Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic

(Received 30 March 2007; revised manuscript received 21 June 2008; published 25 September 2008)

A method for random resampling of time series from multiscale processes is proposed. Bootstrapped series—realizations of surrogate data obtained from random cascades on wavelet dyadic trees—preserve the multifractal properties of input data, namely, interactions among scales and nonlinear dependence structures. The proposed approach opens the possibility for rigorous Monte Carlo testing of nonlinear dependence within, with, between, or among time series from multifractal processes.

DOI: 10.1103/PhysRevLett.101.134101   PACS numbers: 05.45.Tp, 05.45.Df, 89.75.Da

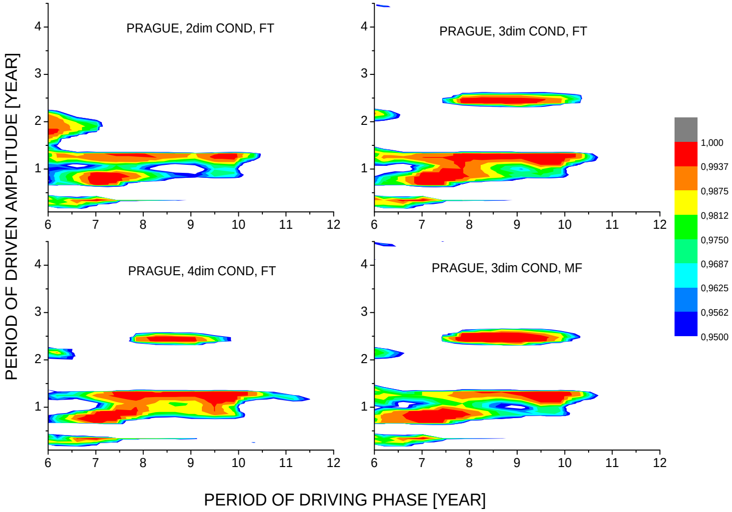
The estimation of any quantity from experimental data, with the aim to characterize an underlying process or its change, is incomplete without assessing the confidence of the obtained values or significance of their difference from natural variability. With the increasing performance and availability of powerful computers, Efron [1] proposed to replace (not always possible) analytical derivations based on (not always realistic) narrow assumptions by computational estimation of empirical distributions of quantities under interest using so-called Monte Carlo randomization procedures. In statistics, the term "bootstrap" [2] is coined for random resampling of experimental data, usually with

data in combinations with some constraints. Possible nonlinear dependence between a signal $s(t)$ and its history $s(t-\eta)$ is destroyed, as well as interactions among various scales in a potentially hierarchical, multiscale process. Multiscale processes that exhibit hierarchical information flow or energy transfer from large to small scales, successfully described by using the multifractal concepts (see [7] and references therein) have been observed in diverse fields from turbulence to finance [8], through cardiovascular physiology [9] or hydrology, meteorology, and climatology [10]. Angelini *et al.* [11] express the need for resampling techniques in evaluating data from atmospheric turbulence

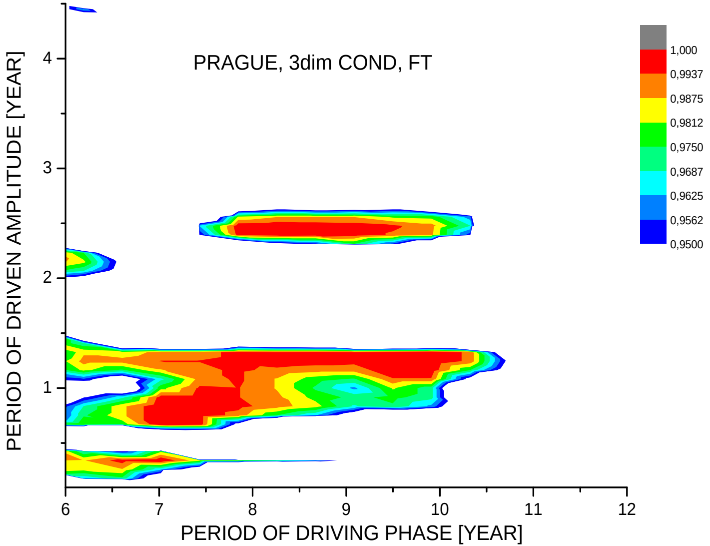
CAUSAL PHASE –> AMPLITUDE INTERACTIONS

- $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), \ldots, A_2(t-m\eta))$
- series length 32768
- forward lags $\tau = 1 - 750$ days
- backward condition lags η = 1/4 of the slow period
- Gaussian process estimator
- conditioning dimension: stable results from 3
- raw data include annual cycle
- seasonal mean and variance removed before surrogate randomization
- seasonal mean and variance added back to surrogate realizations

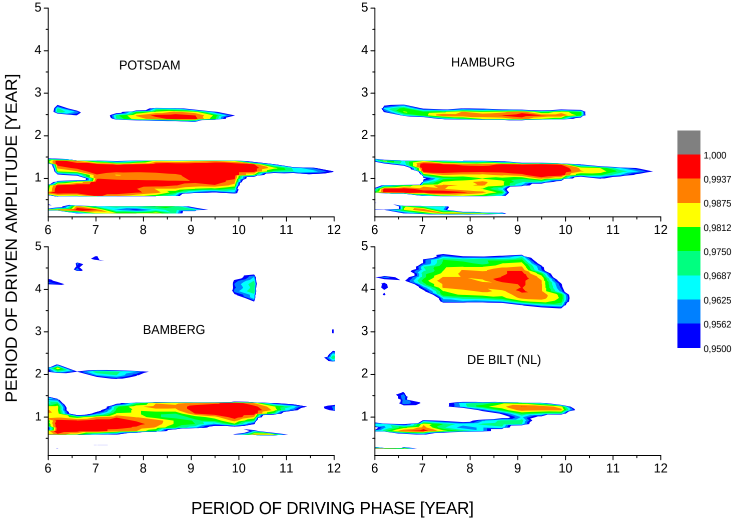
CAUSAL PHASE –> AMPLITUDE INTERACTIONS



CAUSAL PHASE –> AMPLITUDE INTERACTIONS



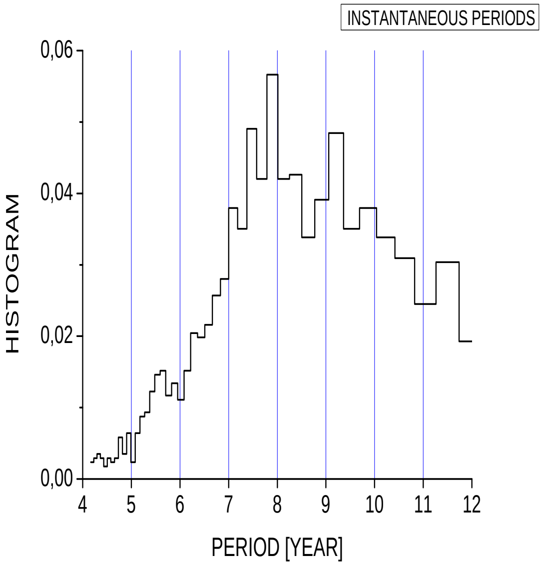
CAUSAL PHASE –> AMPLITUDE INTERACTIONS



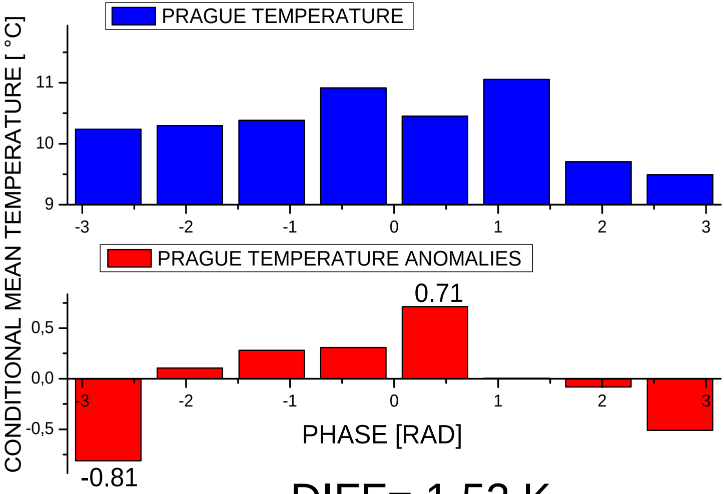
EFFECT PHASE –> AMPLITUDE COUPLING

- HOW TO QUANTIFY THE EFFECT OF PHASE –> AMPLITUDE COUPLING ?
- EXTRACT THE CYCLE WITH PERIOD AROUND 8 YEARS
- EXTRACT ITS PHASE
- DIVIDE THE PHASE INTO 8 BINS
- COMPUTE CONDITIONAL TEMPERATURE MEANS $< T|\phi \in (\phi_1, \phi_2) >$

SSA-extracted "7–8 yr cycle"



EFFECT PHASE –> AMPLITUDE COUPLING



DIFF= 1.52 K

Thank you for your attention

Support: Czech Science Foundation,
Project No. P103/11/J068.

NONLINEAR DYNAMICS WORKGROUP
http://ndw.cs.cas.cz
SW for interaction analysis
SW for network analysis
Preprints