Biomed Tech 2013; 58 (Suppl. 1) © 2013 by Walter de Gruyter · Berlin · Boston. DOI 10.1515/bmt-2013-4172

EXPLORING EFFECTIVE CONNECTIVITY BY A GRANGER CAUSALITY APPROACH WITH EMBEDDED DIMENSION REDUCTION

Britta Pester[1], Lutz Leistritz[1], Herbert Witte[1] and Axel Wismueller[2]

[1]Bernstein Group for Computational Neuroscience Jena, Institute of Medical Statistics, Computer Sciences and Documentation, Jena University Hospital, Friedrich Schiller University Jena, Germany

[2]Department of Imaging Sciences, Department of Biomedical Engineering, University of Rochester Medical Center, Rochester, New York, USA

Britta.Pester@mti.uni-jena.de

Abstract: *We propose applying the linear Granger Causality concept to very high-dimensional time series. The approach is based on integrating dimensionality reduction into a multivariate time series model. If residuals of dimensionality reduced models can be transformed back into the original space, prediction errors in the high–dimensional space may be computed, and a Granger Causality Index (GCI) is properly defined. We provide a proof–of–principle, and compare the results with the classical GCI.*

Keywords: *Granger Causality Index, Dimension reduction*

Introduction

A basic problem in quantifying directed information transfer is the consideration of effective connectivity in very high–dimensional (HD) systems. Currently, HD systems are transformed into a lower dimensional system, e.g. by Principal or Independent Component Analysis (PCA, ICA), and the connectivity structure of derived components is studied. Here the drawback is that a revealed interaction cannot be readily transferred back into the original HD space. Thus, directed interactions between the original network nodes are not revealed, which limits the interpretation of identified interaction patterns. Granger Causality (GC) is a suitable concept for assessing connectivity structures between time series. One popular approach uses principles of prediction [1], whereby application of a straightforward generalization to general time series models is enabled, providing an appropriate definition of prediction errors. Instead of analyzing interactions between derived components, a large scale GC (lsGC) approach preserves the interpretability of the original network nodes. The idea of that approach is the integration of a dimension reduction into a multivariate time series model, which allows computation of prediction errors in the original HD space.

Methods

A D–dimensional, p–th order MVAR process is given by $\mathbf{Y}(n) = \sum_{r=1}^{p} \mathbf{A}^r \mathbf{Y}(n-r) + \mathbf{E}(n),\ n = 1, \ldots, N$, with AR–parameters $\mathbf{A}^r \in \mathbb{R}^{D\times D}$ and a zero mean, uncorrelated noise process $\mathbf{E}$. In the case of HD data, a simple AR estimation is not possible as computational capacity rapidly meets its limits. Thus, in a first stage PCA serves as a preprocessing step for dimension reduction: $\mathbf{X} = \mathbf{W}\mathbf{Y}$, with $\mathbf{Y} = \big(\mathbf{Y}(1), \ldots, \mathbf{Y}(N)\big)$, the principal component (PC) matrix $\mathbf{X} \in \mathbb{R}^{D\times N}$, and the mixing matrix $\mathbf{W} \in \mathbb{R}^{D\times D}$. Let $\mathbf{X}^C$ and $\mathbf{W}^C$ be the reduced PC and mixing matrices consisting of the first C rows of $\mathbf{X}$ and $\mathbf{W}$, respectively. $\mathbf{X}^C(n)$ is now MVAR–modeled, and the modeled time series $\hat{\mathbf{X}}^C(n)$ is afterwards transformed back into the original HD space via left multiplication of the pseudo inverse $\mathbf{W}^{C+}$ of $\mathbf{W}^C$. The residuals of the whole model are then gained by $\hat{\mathbf{E}} = \mathbf{W}^{C+}\hat{\mathbf{X}}^C - \mathbf{Y}$. For GCI computations, the processing of the reduced data $\mathbf{Y}^{d-}$, where the d-th row of $\mathbf{Y}$ is deleted, can be performed in two different ways:

(a) *Multi PCA* (mPCA): for every $\mathbf{Y}^{d-}$ a separate PCA is performed, i.e. $\mathbf{X}_m^{d-} = \mathbf{W}_m^{d-}\mathbf{Y}^{d-}$, where $\mathbf{X}_m^{d-}$ and $\mathbf{W}_m^{d-}$ are calculated anew by PCA for each d. After reducing $\mathbf{W}_m^{d-}$ to dimension C and estimating the corresponding AR model, the modeled series $\hat{\mathbf{X}}_m^{d-}(n)$ can be calculated.

(b) *Single PCA* (sPCA): only one PCA is applied before eliminating rows of $\mathbf{Y}$, and modifications of the mixing matrix $\mathbf{W}$ are used for the dimension reduction of $\mathbf{Y}^{d-}$, i.e. $\mathbf{W}$ is reduced to $\mathbf{W}_s^{d-} \in \mathbb{R}^{C\times D-1}$ by eliminating the last $D-C$ rows and the d-th column. Now $\mathbf{X}_s^{d-} = \mathbf{W}_s^{d-}\mathbf{Y}^{d-}$ serves for the AR parameter estimation resulting in the modeled series $\hat{\mathbf{X}}_s^{d-}(n)$.

The residuals amount to $\hat{\mathbf{E}}_{m/s}^{d-} = \mathbf{W}_{m/s}^{d-\ +} \cdot \hat{\mathbf{X}}^{d-} - \mathbf{Y}^{d-}$. The lsGCI from d_1 to d_2 is then defined by $\gamma_{d_2\leftarrow d_1} = \ln\left(\hat{\Sigma}_{d_2}^{d_1-}/\hat{\Sigma}_{d_2}\right)$, where $\hat{\Sigma}_{d_2}$ and $\hat{\Sigma}_{d_2}^{d_1-}$ are the d_2-th diagonal entries of the covariance matrices of $\hat{\mathbf{E}}$ and $\hat{\mathbf{E}}^{d_1-}$.

Data

To compare the lsGCI with the conventional GCI, we considered a time series dimensionality that functions with both approaches. We realized 50–dimensional stationary MVAR processes of order two and various N between 125 and 1000. Thereby, the entire network structure was given by five pairwise different internal networks $\boldsymbol{N}_1, \ldots \boldsymbol{N}_5$ with ten nodes each (Fig. 1). The corresponding AR-parameters were chosen according to the AR-model of Baccala et al. [2], Fig. 4, and were scaled by factor 0.5 to ensure the stationarity of the entire process. The internal networks $\boldsymbol{N}_k$

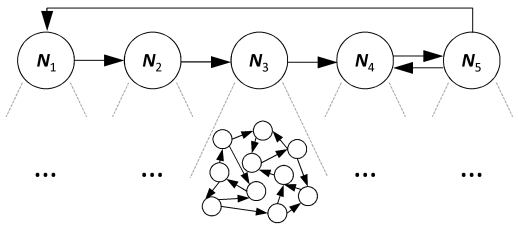

Figure 1: Network structure with five internal networks.

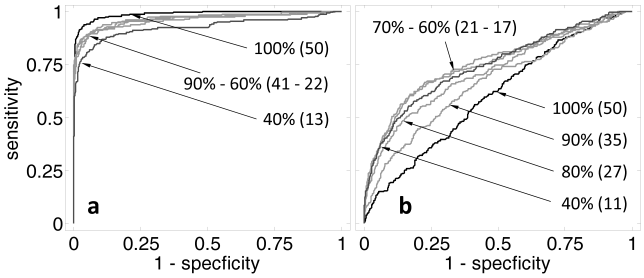

Figure 3: ROC curves for $N = 500$ (a) and $N = 125$ (b). The percentages specify the amount of variance explanation; numbers in parentheses specify C.

incorporate 20 directed edges by setting the associated first order AR-parameter to 0.2. The in- and out–degree of each node equals two. Finally, there are 20 randomly generated directed edges from nodes of $\boldsymbol{N}_1$ to nodes of $\boldsymbol{N}_2$, from nodes of $\boldsymbol{N}_2$ to nodes of $\boldsymbol{N}_3$, etc. (see Fig. 1). The added $E_d(n)$ were i.i.d. N(0,1) for all d and n.

Results

To evaluate the novel approach and to assess the effects of methodological differences by applying PCA we used the concept of ROC curves. Thereby (ls)GCIs serve as realizations of the test variable, and the status variable is defined by the presence (positive) or absence (negative) of an edge. First, the discriminative power of both PCA embeddings was investigated for different time series lengths, and differing amounts of variance explanation. As shown in Fig. 2 the sPCA approach exhibited primarily larger areas under the ROC curve (AUC). This finding was also confirmed for all other investigated time series lengths. Thus all subsequent analyses were performed with the sPCA approach. Secondly, ROC curves were used to analyze differ-

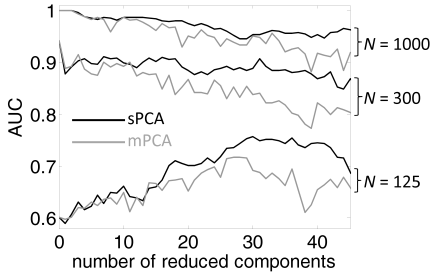

Figure 2: AUCs for different dimension reduction degrees.

ing amounts of variance explanation. Fig. 3 depicts examples for $N = 500$ and $N = 125$. As expected for large N, the dimension reduction resulted in an inferior performance depending on the number of components reduced. However, this performance is within acceptable limits for reasonable dimensionality reductions (Fig. 3a). For small N, the embedded dimension reduction yielded a performance increase (Fig. 3b). Finally, we considered the detection accuracy after significance testing (Tab. 1) ($\alpha = 5\%$, adjusted for multiple comparisons by False Discovery Rate).

Discussion

PCA is appropriate to extend linear GCI to HD time series. It reduces HD into lower-dimensional (LD) time series of

Table 1: Sensitivities and specificities after significance testing. The column '$\%$' specifies the variance explanation.

N	C	$\%$	Sens.	Spec.
	50	100	70.5	97.7
500	41	90.6	70.0	97.6
	27	70.3	65.4	97.6
	50	100	0.4	97.8
125	35	90.7	4.1	97.7
	21	70.8	6.8	97.5

PCs. LD time series are AR modeled, and the model residuals are transformed back into the original HD space. This transformation offers a better interpretability of results, enabling analysis of interactions between components of the original time series vs. between derived components (PCs). Alternative dimensionality reductions could also be considered if a back-transformation of the model residual from a temporary LD to the original HD space is allowed. An embedded dimension reduction appears to the quality of the network identification when enough time series samples are available, yet classical GCI still performs well. For shorter time series an embedded PCA seems to result in an improvement, most likely due to smaller AR parameter matrices and reduced estimator variances.

Acknowledgement

This study was supported by the grant 01GQ1202 of the Federal Ministry of Education and Research (Germany), as well as by the NIH grant 1R01DA034977-01 (USA).

Bibliography

[1] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969.

[2] L. A. Baccala and K. Sameshima, "Partial directed coherence: a new concept in neural structure determination," *Biological Cybernetics*, vol. 84, no. 6, pp. 463–474, 2001.

Biomed Tech 2013; 58 (Suppl. 1) © 2013 by Walter de Gruyter · Berlin · Boston. DOI 10.1515/bmt-2013-4173

COORDINATION AND TIMING OF HEART RATE COMPONENTS OF CHILDREN AND ADOLESCENTS WITH TEMPORAL LOBE EPILEPSY

Schiecke K[1], Wacker M[1], Benninger F[2], Feucht M[2], Witte H[1]

[1]Institute of Medical Statistics, Computer Sciences and Documentation, Jena University Hospital, Friedrich Schiller University Jena, Germany
[2]Epilepsy Monitoring Unit, Department of Child and Adolescent Neuropsychiatry, University Hospital Vienna, Austria

Karin.Schiecke@mti.uni-jena.de

Abstract: *The aim of our study was to reveal specific patterns of the heart rate variability (HRV) during pre-ictal, ictal, and post-ictal periods in epileptic children. The continuous Morlet-wavelet transform was adapted to explore the time-frequency characteristics of the HRV (scalogram, linear and non-linear phase locking, and band-power analyses). The empirical mode decomposition was used to separate HRV components (e.g. blood-pressure-related waves and respiratory sinus arrhythmia). Their time-variant non-linear predictability was analysed (point prediction error). Timing and coordination of HRV components occurs 100 s before seizure onset producing a higher degree of synchronization and a higher predictability of HRV. The combined use of advanced linear and non-linear methods is crucial for this result.*

Keywords: *children, epilepsy, heart rate variability, signal-adaptive decomposition, time-frequency analysis*

Introduction

HRV analysis in epilepsy has been carried out with two major clinical objectives. One is to reveal causes for sudden unexpected death in epileptic patients. Another focus is using *HRV* analysis as a tool for automated seizure prediction. The pre-ictal, ictal and post-ictal *HRV* courses have been investigated by using time- and frequency-domain features. Features of the time-frequency domain have received much less attention. Non-linear *HRV* analyses are frequently applied also in epileptic patients. However, the methods are usually time-invariant (for stationary signals) and not frequency-selective. The aim of this study is to demonstrate that combinations of time-variant, frequency selective, linear and non-linear analysis methods can be beneficially applied for *HRV* analysis in epileptic patients. Our working hypothesis is that phase properties of and between *HRV* components react sensitively before *EEG* seizure occurs. This hypothesis is based on our findings with regard to *EEG* burst activity which is accompanied by strong phase coupling reactions [1]. Two rhythms are of particular interest: The Traube-Hering-Mayer waves found in blood pressure, which occur also in the *HRV* (low-frequency range *LF*: 0.04 - 0.15 Hz), and the respiratory sinus arrhythmia (*RSA*) (high-frequency range *HF*: 0.15 - 0.4 Hz). The *HRV* data were previously analysed (visual analysis of heart rate) and the results were published by Mayer et al. [2].

Methods

Subjects and HRV computation: The *HRV* data of 18 patients were analysed (median age 9 years 4 months, range 6 years 10 months to 18 years 0 month). Presurgical evaluation was performed following a standard epilepsy surgery protocol. *EEG* and *ECG* data were recorded referentially against *Pz* (filter 1 to 70 Hz; sampling frequency 256 Hz). Seizure onset and termination in the *EEG* were determined independently by two reviewers. *EEG* and *ECG* samples including 10 minutes epochs (5 minutes before (pre-ictal state) and 5 min after the seizure onset (seizure and post-ictal state) were analysed. *QRS* detection was performed and used for the heart rate computation. The low-pass filtered event series (*LPFES*) was utilized applying the French-Holden algorithm. The final *HRV* representations were down sampled to 8 Hz.
Continuous Morlet transform (MWT) and derived time-variant parameters: The frequency-dependent complex analytic signal of the *HRV* is computed by using the *MWT*. Power and phase information of the complex analytic signal can be extracted. The scalogram *S* and the phase-locked scalogram S_{PL} were estimated. From *S* and S_{PL} the mean band power for each sampling point is computed for the frequency bands 0.04 - 0.15 Hz (*LFP*) and 0.15 - 0.4 Hz (*HFP*) according to the task force standards. Amplitude-independent phase-locking effect were analysed by the phase-locking index *PLI*. Quadratic phase coupling (*QPC*) between both frequency bands given above are computed by using time-variant mean bi-amplitude (*mBA*) and (normalised) mean bi-coherence (*mBC*) in the region of interest (*ROI)* [3].
Empirical mode decomposition (EMD) and derived time-variant parameters: The *EMD* decomposes the *HRV* into intrinsic mode functions (*IMF*s). *EMD* preserves non-linear properties of the separated components. Therefore, averaged time-courses of non-linear point prediction error (*PPE*) of the *IMF*s were calculated. The *PPE* computation procedure was described by Schwab et al. [4]. High *PPE* denotes low non-linear predictability and vice versa.
Statistics: The non-corrected Rayleigh test (α=10%) was used to create a trigger threshold of strong phase-locking in a *PLI* time-frequency map. In order to estimate confidence tubes of the mean time-courses of the extracted parameters a Bootstrap approach was used (1000 bootstraps, sample size 600 s, lower bound 5% / upper bound 95% percentile).

Results

HR courses: In Fig. 1A the *HR* courses of all seizures and the averaged *HR* are depicted. The averaged *HR* is characterised by a slow increase 60 s before seizure onset, and fast acceleration followed by deceleration during seizure.
MWT-based analysis: In non-phase locked *S* (Fig. 1 B) the occurrence of strong Mayer-wave related *LF* and *RSA*-related *HF* component became most apparent between 200 and 300 s, collapsed with the onset of seizure, recurred with strong power at 380 s (end of seizure) and got less pronounced up to the end of analysis interval.

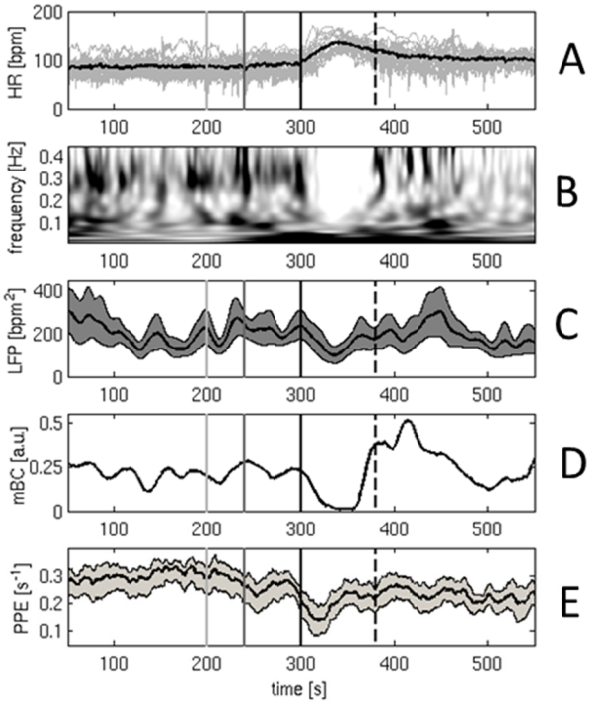

Figure 1: Original *HR* courses and derived parameters. "Bar code" of specific time points: black line - seizure onset, dashed black line - seizure termination, dark grey line - start of pre-onset acceleration, light grey line - start of timing and coordination of *HR* components. A) *HR* courses (grey: 18 seizures, bold black: averaged HR). B) *MWT*-based time-frequency related *S* analysis (black: high/white: low power). C) mean *LFP* band course and confidence tube. D) time-variant *mBC* in the *ROI*. E) mean *PPE* course and confidence tube of *EMD*-based *IMF2* component (*RSA*-related).

Power ridges in phase-locked scalogram S_{PL} could be observed at similar time-pattern as in *S*, and *LF* component related peaks in the *PLI* also occurred more often between 200 and 300 s (not shown). *LFP* (Fig. 1 C) as well as *HFP* band ranges (not shown) differ significantly during seizure in comparison to pre-ictal period (240 to 300 s) and parts of post-ictal period.
The *mBA* (not shown) and *mBC* (Fig. 1 D) revealed lower values during pre-ictal period, increased with a peak at the beginning of pre-onset acceleration of *HR* (240 s) and at seizure onset, and dropped down during seizure. The termination of seizure is accompanied by a rise of *mBC*.

EMD-based analysis: *EMD* separated the *HRV* into *IMF*s which are connected to specific components of *HR* (e.g. *IMF2*: *HF* component (≈0.3 Hz) / *IMF3*: *LF* component (≈0.1 Hz)). Averaged *PPE* of the original *HRV* (not shown) started to decrease (=higher predictability) at 260 s, showed significant increase during seizure and significant difference between pre- and post-ictal period (incomplete recuperation). The predictability of *IMF2* (Fig. 1 E) increased around 200 s (decrease of *PPE*), decreased shortly between 240 and 300s and significantly increased during seizure. Mean *PPE* is lower during post-ictal in comparison to pre-ictal period (no significant difference).

Discussion

An epileptic seizure is a time-dependent process, and time-variant *HRV* analysis may provide better information on the dynamics of coupling mechanisms between relevant cortical structures and the autonomic nervous system. The most striking clinical result of our methodological study is that timing and coordination of *HRV* components arises 100 s before *EEG*-seizure onset producing a higher degree of synchronization and a higher predictability of *HRV*. Certainly, this result is based on grand mean results and cannot be generalized to each patient/seizure. However our working hypothesis, that phase properties of the *HRV* components can be utilized for the description of these reactions/couplings, was confirmed. Our processing scheme is composed of time-variant analysis methods and the results complement each other. By means of the *MWT* the classical (time-invariant) power parameters can be made available as parameter courses. In addition, time-variant phase-locking and quadratic phase coupling can be examined. The results serve as a gold standard for signal-adaptive approaches like *EMD*. Although the *EMD* has method-intrinsic drawbacks, e.g. *EMD*'s filter characteristics are not always sufficient for biomedical applications and mode-mixing occurs, *EMD* based approaches provide advantages which should be used for *HRV* analysis.

Acknowledgement

This study was supported by the DFG (Wi 1166/10-2).

Bibliography

[1] Schwab, K., Skupin, H. et al.: Coordination of the EEG and the heart rate of preterm neonates during quiet sleep, *Neurosci. Lett.* , vol. 465, pp. 252-256, 2009
[2] Mayer, H., Benninger, F. et al.: EKG abnormalities in children and adolescents with symptomatic temporal lobe epilepsy, *Neurology*, vol. 63, pp. 324-328, 2004
[3] Witte, H., Putsche, P. et al.: Time-variant analysis of phase couplings and amplitude-frequency dependencies of and between frequency components of EEG burst patterns in full-term newborns, *Clin. Neurophysiol.*, vol. 122, pp. 253-266, 2011
[4] Schwab, K., Groh, T. et al.: Nonlinear analysis and modeling of cortical activation and deactivation patterns in the immature fetal electrocorticogram, *Chaos*, vol. 19, pp. 015111, 2009