Med Biol Eng Comput (2006) 44: 321–330
DOI 10.1007/s11517-006-0038-0

ORIGINAL ARTICLE

Niels Wessel · Hagen Malberg · Robert Bauernschmitt
Alexander Schirdewan · Jürgen Kurths

Nonlinear additive autoregressive model-based analysis of short-term heart rate variability

Received: 8 August 2005 / Accepted: 27 February 2006 / Published online: 29 March 2006

Abstract In this contribution we test the hypothesis that nonlinear additive autoregressive model-based data analysis improves the diagnostic ability based on short-term heart rate variability. For this purpose, a nonlinear regression approach, namely, the maximal correlation method is applied to the data of 37 patients with dilated cardiomyopathy as well as of 37 age- and sex-matched healthy subjects. We find that this approach is a powerful tool in discriminating both groups and promising for further model-based analyses.

Keywords Nonlinear regression · Heart rate variability · Modeling · Time series analysis

1 Introduction

Most cardiovascular diseases are characterized by a slow progression leading to possibly abrupt qualitative changes [24, 49]. This pathogenesis is a dynamic process and the associated cardiovascular parameters, such as heart rate variability (HRV), exhibit complex dynamics [3, 7, 25, 30]. Accordingly, the central requirement of nonlinear dynamics aims exactly at such physiological phenomena. First applications of nonlinear dynamics tools in medical diagnostics met with success in the last 10 years [11, 20]. So far, however, either only data were analyzed [8, 13, 19, 28, 31, 42, 47, 48] or models developed [12, 14, 32, 45, 46]. The purpose of this paper therefore is, to take a qualitatively new step for HRV data: the combination of data analysis and modeling. Model-based nonlinear-dynamic data analysis of non-invasively measured biosignals could lead to an improved diagnostic ability and to a better understanding of the cardiovascular regulation. Applications of the research results are manifold: Monitoring-, diagnosis-, course- and mortality prognoses as well as the early detection of heart diseases. As regards clinical applications, there is an extraordinary demand for new computer-controlled diagnostic methods to obtain a more exact and differentiated picture of the possibly damaged heart.

The analysis of HRV has become a powerful tool for the assessment of autonomic control. HRV measurements have proven to be independent predictors of sudden cardiac death after acute myocardial infarction, chronic heart failure or dilated cardiomyopathy [17, 21, 27, 33, 34, 38]. Moreover, it has been shown that short-term HRV analysis already has an independent prognostic value in risk stratification apart from that of clinical and functional variables [23]. However, the underlying regulatory mechanisms are still poorly understood. Short-term heart rate (HR) regulation is accomplished mainly by neural sympathetic- and parasympathetic-mediated cardiac baroreflexes and peripheral vessel resistance, whereas long-term regulation is achieved by hormonal pathways as well as other systems like the renin-angiotensin-system [4]. To gain more insight into cardiovascular regulation, in this paper a nonlinear additive autoregressive (NAAR) model is estimated nonparametrically. The nonlinear regression method we are using was introduced by Breiman and Friedman [6] in 1985 and already successfully applied to NAAR modeling of riverflow data in 1993 [9].

The paper is organized as follows. In Sect. 2 we shortly describe the data, HRV parameters as well as the nonlinear regression approach used. In Sect. 3 the

N. Wessel (✉) · J. Kurths
Institute of Physics, University of Potsdam,
Am Neuen Palais 10, 14415 Potsdam, Germany
E-mail: wessel@agnld.uni-potsdam.de

R. Bauernschmitt · N. Wessel
Clinic for Cardiovascular Surgery,
German Heart Center Munich, Lazarettstr. 36,
80636 Munich, Germany

A. Schirdewan · N. Wessel
Medical Faculty of the Charité, Franz Volhard Clinic, Helios Klinikum-Berlin, Wiltbergstr. 50, 13125 Berlin, Germany

H. Malberg
Institute for Applied Computer Science/Automation (AIA), University of Karlsruhe (TH), Kaiserstraße 12, 76128 Karlsruhe, Germany

results of standard HRV analysis, nonlinear modeling and simulating shall be presented. Finally, in Sect. 4 the results shall be discussed.

2 Methods

This study is aimed at investigating short-term HR recordings of 37 patients with dilated cardiomyopathy (DCM, age: 50.2 ± 9.1 years) in comparison to 37 healthy subjects (52.1 ± 6.7 years). The diagnosis of DCM was established in according to the recommendations for medical testing of patients with cardiomyopathy. The DCM patients had a left ventricular ejection fraction of 22.3 ± 5.7%, a New York Heart Association class of II–III and were getting the present standard therapy [ACE-inhibitor (31), beta-blocker (16), diuretics (27)]. All patients were in sinus rhythm and had no concomitant diseases. To minimize the probability for cardiac diseases in the healthy volunteers group we recruited them from an occupational health center. They had normal findings in echocardiography, bicycle ergometry, and Holter ECG for many years. No control subject had a history of cardiac diseases or symptoms. Patient characteristics are summarized in Table 1. All measurements were performed under comparable ambient conditions (before noon, after having a small breakfast, same place, temperature) with a standard ambulatory ECG (100 Hz) for 30 min under supine position. The consumption of alcohol and tobacco was forbidden 24 h before the measurement.

Ventricular premature beats as well as artifacts were filtered out using a special adaptive filtering procedure described in [39]. Figure 1a and b give examples of these data sets. To demonstrate later on where the nonlinear regression approach is adequate, a time series of a patient with atrial fibrillation from another study is given in Fig. 1c.

Table 1 Patient characteristics

Group	Control, $n = 37$	DCM, $n = 37$
Age (years)	50.2 ± 9.1	52.1 ± 6.7
Sex		
Male	26	27
Female	11	10
SBP (mmHg)	119.0 ± 18.6	110.6 ± 12.4*
DBP (mmHg)	73.2 ± 7.8	70.5 ± 9.9
HR (bpm)	69.5 ± 9.4	75.4 ± 7.6*
Echo		
LVEF (%)	–	22.3 ± 5.7
LVDD (mm)	–	67.8 ± 8.6
Pharmacotherapy		
ACEI	–	31
β-blocker	–	16
Diuretic	–	27

SBP systolic blood pressure, *DBP* diastolic blood pressure, *HR* heart rate, *LVEF* left ventricular ejection fraction, *LVDD* left ventricular diastolic diameter, *ACEI* angiotensin-converting enzyme inhibitor
*$P < 0.05$ versus control group

2.1 Heart rate variability analysis

Heart rate variability is calculated in the time and frequency domain regarding the Task Force HRV [17] and by methods of nonlinear dynamics. Analyzing HRV, the following standard parameters are calculated from the time series: MeanNN (mean value of normal beat-to-beat intervals): Inversely related to mean HR. sdNN (SD of intervals between two normal *R*-peaks): Gives an impression of the overall circulatory variability. sdaNN (the SD of successive 5 min' NN-interval mean values): quantifies long-range variabilities. rmssd (root mean square of successive RR-intervals) and pNN50 (percentage of RR-interval-differences greater than 50 ms): quantifying short-range variabilities. Shannon (the Shannon entropy of the histogram): quantification of RR-interval distribution. Apart from the time-domain parameters mentioned above, the HRV analysis focused on high-frequency (HF), low-frequency (LF) and very-low-frequency (VLF) components expressed by the normalized values LF/*P*, HF/*P*, VLP/*P*, and the ratio LF/HF (where *P* is the total power). The number of ventricular ectopic beats is automatically counted and denoted by noNNtime (cumulative time of not normal beat-to-beat intervals). Finally, HRV is analyzed by methods of nonlinear dynamics, especially symbolic dynamics [22, 36, 40, 41]: In this analysis the time series are transformed into symbol sequences with symbols from a given alphabet. Some details are lost in this process; however, the advantage is that the coarse dynamic behavior can be analyzed. The used parameter 'Polvar20' (Probability of low variability, 20 ms difference) characterizes short phases of low variability from successive symbols of a simple alphabet, consisting of only the symbols 0 and 1, where 0 stands for a small difference of less than 20 ms between two successive beat-to-beat intervals, and where 1 represent cases when the difference between two successive beat-to-beat intervals exceeds this limit, specifically given the time series $x_1,x_2,...,x_N$ one obtains the symbol series s_n

$$s_n = \begin{cases} 1: & |x_n - x_{n-1}| \geq 20\,\text{ms} \\ 0: & |x_n - x_{n-1}| < 20\,\text{ms} \end{cases} \quad (1)$$

Words consisting of the unique symbols all 0 or all 1 are counted. To obtain a statistically robust estimate of the word distribution, we restrict ourselves to the 64 words defined by six consecutive symbols. 'Polvar20' represents the probability of the word 000000 occurring and thus detects even intermittently decreased HRV.

2.2 Model-based data analysis

We generally assume that there are HR time series X_i, $i = 1,...,n$ (series of beat-to-beat-intervals as in Fig. 1). In this paper the following class of NAAR-models shall be considered

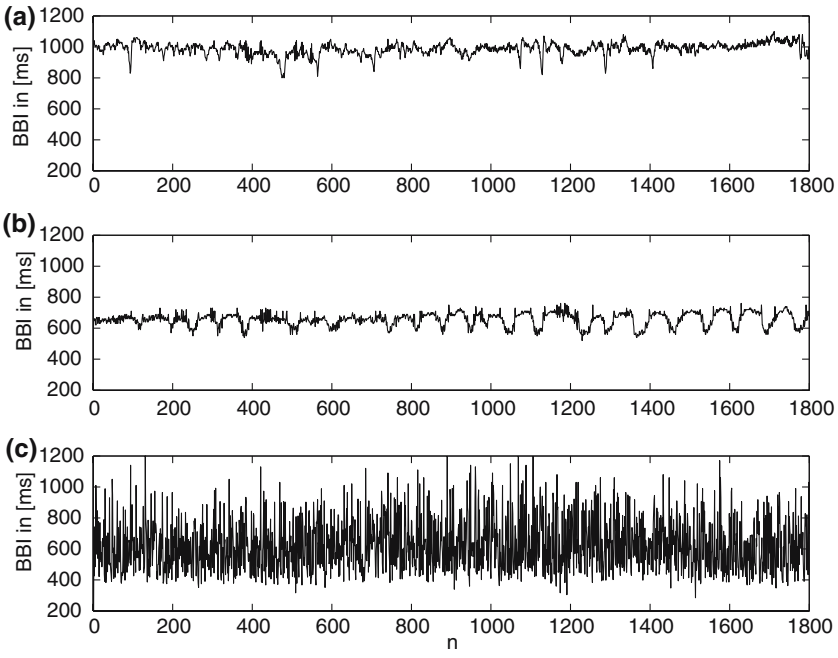

Fig. 1 Thirty minutes tachograms from a healthy volunteer (**a**), a patient with dilated cardiomyopathy (**b**) and a patient with atrial fibrillation (**c**) [beat-to-beat-interval (BBI)]

$$\theta(X_j) = \sum_{i=1}^{k} \phi_i(X_{j-i}), \quad j = k+1, \ldots, n. \tag{2}$$

where k denotes the history length. The alternating conditional expectations (ACE) algorithm [6] described below is applied as a nonparametric approach to estimate the possibly nonlinear transformations ϕ_i and θ in Eq. 2. The easiest model of this class is the linear autoregressive model

$$X_j = \sum_{i=1}^{k} a_i \cdot X_{j-i}, \quad j = k+1, \ldots, n. \tag{3}$$

The concept of maximal correlation is a very powerful criterion to measure the dependence in particular of nonlinear related variables [29]. The main idea of this approach is to measure the maximized correlation of properly transformed variables.

Given a real variable X_j and a n_i-dimensional vector $X = (X_1, \ldots, X_{n_i})$ in the additive model 2. Then, the maximal correlation is defined by

$$\begin{aligned} \Psi(X_j, X) &:= |\rho(\theta^*(X_j), \phi^*(X))| \\ &= \max_{\theta,\phi} |\rho(\theta(X_j), \phi(X))| \end{aligned} \tag{4}$$

where ρ denotes the correlation coefficient. The functions θ^* and ϕ^*, which fulfil the maximal condition (4), are called optimal transformation and represent an estimation of the model 2. To estimate them nonparametrically, we use the ACE-algorithm [6]. Afterwards, to enable a classification of these transformations, a polynomial fitting was performed to parametrize the model functions. For more details to the ACE-algorithm see the Appendix.

The maximal correlation and optimal transformation approach were applied recently to nonlinear dynamic systems to identify delay in lasers [35] and partial differential equations in fluid dynamics [37]. The ACE algorithm turned out to be a very efficient tool for nonlinear data analysis [16, 35, 39, 43, 44].

3 Results

3.1 Heart rate variability

The DCM group was characterized by a increased HR (consequently, by a lower meanNN) and a decreased HRV as reflected by the parameters sdNN and sdaNN5 (see Table 2). Interestingly, both groups hardly differed in the respiration-induced very short variability range (rmssd). Both, the increased HR and the decreased HRV lead to significantly different Shannon entropies of the RR-interval distribution of the groups. In frequency analysis, an increased VLF band is observed for the DCM group, whereas the LF band is decreased. Again, no significant differences result in the respiration-induced and vagally mediated HF band. The significant differences of LF/HF between both groups is due to the lower LF band in the DCM group exclusively. Surprisingly, parameters from symbolic dynamics did not show any significant difference between both groups (only as a trend 'Polvar20': $p = 0.057$). Finally, the DCM patients show a higher level of ventricular ectopy (NoNNtime).

3.2 Model-based data analysis

For simplicity reasons, we started with a linear autoregressive model given in Eq. 3. For different model orders up to 10, linear autoregressive modeling was performed and the model coefficients themselves were considered to

Table 2 Results of heart rate variability (HRV) analysis

	Control	DCM	p
meanNN	903.48 ± 129.90	810.07 ± 106.16	0.0013*
sdNN	50.70 ± 19.86	41.04 ± 20.76	0.0099
sdaNN5	25.10 ± 12.60	17.34 ± 11.53	0.0008*
rmssd	29.84 ± 15.17	24.25 ± 8.95	NS
pNN50	0.085 ± 0.109	0.039 ± 0.039	NS
Shannon	2.21 ± 0.38	1.99 ± 0.40	0.0123
LF/HF	2.89 ± 2.15	1.97 ± 1.44	0.0270
LF/P	0.24 ± 0.10	0.18 ± 0.09	0.0038*
HF/P	0.12 ± 0.07	0.13 ± 0.08	NS
VLF/P	0.40 ± 0.12	0.50 ± 0.19	0.0416
NoNNtime	17.0 ± 23.3	38.3 ± 38.2	0.0064
Polvar20	0.07 ± 0.15	0.11 ± 0.14	NS

Mean value ± SD, two-sided Mann–Whitney U test
*Significant after Holm correction [18])

be parameters for discrimination between DCM and controls. The model coefficients however, were significant only up to the order 2 (cf. Table 3). The residual variance η was significantly lower in the DCM group, certainly caused by the lower HRV in this group. The absolute values of the coefficients a_1 and a_2 were higher for the controls, indicating a higher dependence from the two predecessors for healthy subjects.

Based on the findings from linear autoregressive modeling, a NAAR-model of order 2 given in Eq. 5 was investigated.

$$X_j = \phi_1(X_{j-1}) + \phi_2(X_{j-2}), \quad j = 3, \ldots, n. \tag{5}$$

The left hand side of Eq. 5 is the identity and we are interested in estimating the functional dependence of X_j on its two predecessors. Figure 2 gives this functional relationship that was estimated from the time series given in Fig. 1. We see a nonlinear functional relation for both the healthy and the DCM subject in Fig. 2a and b. For the subject with atrial fibrillation in Fig. 2c, however, no real functional relationship could be estimated. In this case, the random influence is too strong, the RR-intervals are nearly uncorrelated. This example shows that there has to be a minimal relation between the data at least for applying this nonlinear regression approach.

The concept of maximal correlation and optimal transformations is a nonparametric approach which allows to estimate in particular nonlinear dependencies between two variables. However, in this paper the aim was to quantify these relations. Therefore, polynomial fitting of third order was performed. The fitting coefficients themselves were considered to be parameters in cardiac diagnostics.

$$X_j = a_1 \cdot X_{j-1}^3 + a_2 \cdot X_{j-1}^2 + a_3 \cdot X_{j-1} + a_4 + b_1 \cdot X_{j-2}^3 + b_2 \cdot X_{j-2}^2 + b_3 \cdot X_{j-2} + b_4 \quad j = 3, \ldots, n. \tag{6}$$

Using higher orders for polynomial fitting turned out not to be useful in this study. Figure 2 gives the fitted polynomials in addition to the optimal transformation. These fitting parameters did not differ statistically between DCM and control for ϕ_1. Table 4 gives the fitting parameters for ϕ_1 and ϕ_2 and we see that parameter b_3 has the smallest p-value of all considered parameters in this study.

To compare these fitting parameters directly with standard HRV parameters we performed a tenfold bootstrap approach. We repeatedly analyzed subsamples of the data based on discriminant function analyses. Each subsample was a random sample with 70% replacement from the full sample (resampling). We used ten different permutation of the data as training sets and the best three parameters to separate the groups were chosen. The parameter b_3 was automatically selected in nine of ten validation runs which demontrates the reliability of our results. The best combination of parameters to separate controls from DCM were the fitting parameter b_3, the normalized very low frequency 'VLP/P' as well as the number of ventricular premature beats, quantified by the preprocessing parameter 'noNNtime'. The averaged classification rates of this tenfold bootstrap validation were 82.1% for the training set (optimistic estimation) and 73.2% for the 30% test set (realistic estimation).

Figure 3 visualizes the polynomial fitting of ϕ_1 and ϕ_2 from Eq. 5 based on group averages. While the optimal transformations ϕ_1 are similar for the DCM and the control group (except the small difference at larger beat-to-beat-intervals—leading a_4 close to be significant), the behavior of ϕ_2 is totally different in both groups. The coefficients b_1 and b_2, representing the quadratic and the cubic part of the fitting and therefore lower than b_3 and b_4, both are higher in absolute values for the DCM subjects and lead to a more oscillating behavior of this fitting for this group. The coefficient b_3, representing the linear part of the fitting, is $b_3 = 0.045$ for the control subjects and, hence, higher than $b_3 = -0.61$ for the DCM group. This means that we have a linear tendency for a slight increase of ϕ_2 with increasing beat-to-beat-interval, whereas, the DCM group exhibits a clear decrease. The dominant part for the discrimination of both groups however, is the quadratic term b_2 which is extremely higher in the DCM group. Thus, for the interpretation of these results all fitting coefficients have to be considered together.

To investigate the stability of our approach we performed Monte-Carlo simulations based on the group

Table 3 Results of linear autoregressive model analysis

	Control	DCM	p
a_1	0.96 ± 0.08	0.90 ± 0.08	0.0011*
a_2	−0.05 ± 0.12	0.02 ± 0.10	0.0068*
η	1,766 ± 1159	1,111 ± 539	0.0026*

Mean value ± SD, two-sided Mann–Whitney U test. η represents the residual variance
*Significant after Holm correction

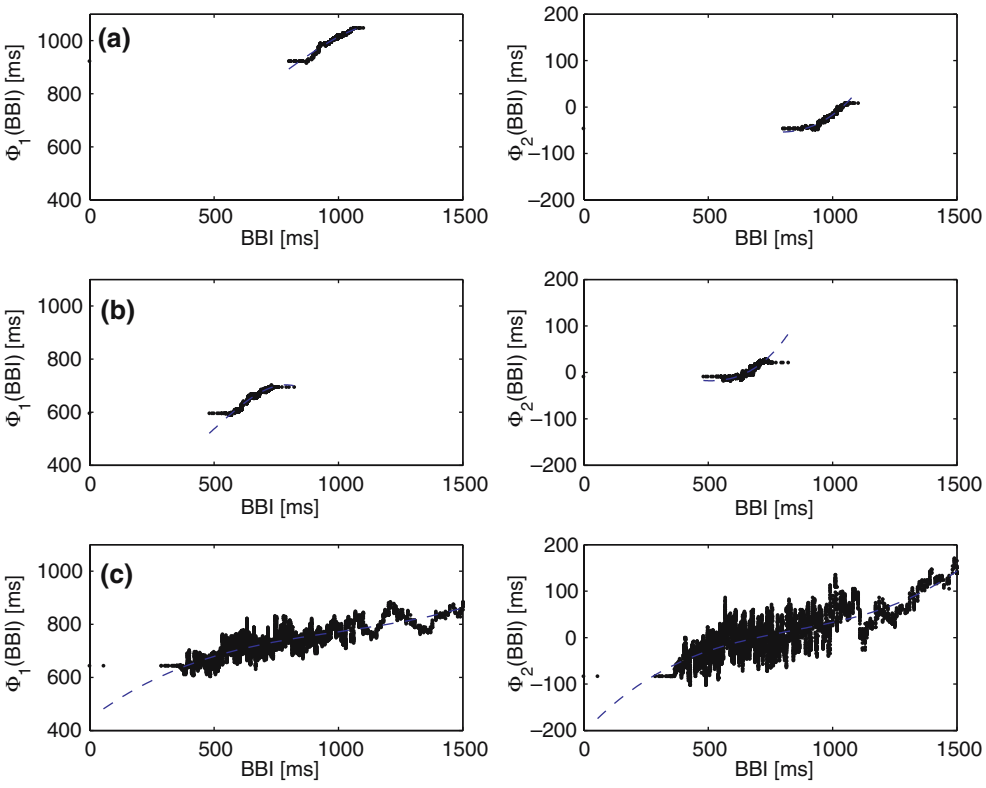

Fig. 2 Optimal transformations as well as estimated polynomials for the tachograms (**a**), (**b**) and (**c**) from Fig. 1 (BBI)

averaged polynomials given in Fig. 3. Two examples are given in Fig. 4. Initial values X_1 and X_2 for both simulations were 1,000 as one can expect from this figure. Then all simulated values were calculated as follows

$$X_j = a_1 \cdot X_{j-1}^3 + a_2 \cdot X_{j-1}^2 + a_3 \cdot X_{j-1} + a_4 + b_1 \cdot X_{j-2}^3 + b_2 \cdot X_{j-2}^2 + b_3 \cdot X_{j-2} + b_4 + \eta \quad j = 3, \ldots, n. \tag{7}$$

were η is a normally distributed random number with mean zero and SD 15 ms. Interestingly, the different mean HRs in both groups are achieved already after a few iterations. Due to the strong relationship each HR value to its two predecessors, we get a more homogeneous behavior in the control group, whereas the DCM simulation is characterized by some drops (e.g. around sample 1,100). From these simulations again the polynomial coefficients were estimated using the approach described. Figure 5 shows these functions which are very similar to the polynomials in Fig. 3. There is only a small horizontal shift of ϕ_1 in the DCM and of ϕ_2 in the control group, the qualitative behavior is remained. Thus, our method is able to consistently obtain important information from the time series, which indicates a possible use for surrogate data analysis.

Finally, we investigated whether our developed models are useful for the forecasting of HR data. For this reason, one-step predictions based on two predecessors X_{j-1} and X_{j-2} and the optimal transformations estimated from Eq. 5 were performed. Figure 6 gives the results of these simulations from the tachograms introduced in Fig. 1. We see adequate estimations for the control as well as for the DCM series, the time series are very similar to the ones given in Fig. 1. For the patient with atrial fibrillation, however, no comparable results were obtained. The one-step predictions fail due to the fact that the optimal transformations do not represent a real functional relationship.

Table 4 Results of NAAR-model-based HRV analysis

	Control	DCM	p
Ψ	0.844 ± 0.082	0.841 ± 0.084	NS
a_1	−1.3E−06 ± 7.4E−07	−1.3E−06 ± 1.2E−06	NS
a_2	3.1E−03 ± 1.4E−03	2.8E−03 ± 1.8E−03	NS
a_3	−1.621 ± 0.749	−1.356 ± 0.701	NS
a_4	800.7 ± 118.8	745.1 ± 98.2	0.061
b_1	−5.9E−08 ± 8.8E−07	−7.8E−07 ± 1.2E−06	0.0019*
b_2	1.9E−05 ± 1.6E−03	1.4E−03 ± 1.9E−03	0.0011*
b_3	0.045 ± 0.744	−0.606 ± 0.758	0.0006*
b_4	−4.92 ± 27.59	−13.72 ± 17.69	NS

Mean value ± SD, two-sided Mann–Whitney U test
*Significant after Holm correction

4 Discussion

The main purpose of this paper was to test whether model-based data analysis improves the results of diagnostic ability based on short-term HRV. Hence, firstly standard HRV analysis was performed: The DCM patients are mainly characterized by a increased HR, a

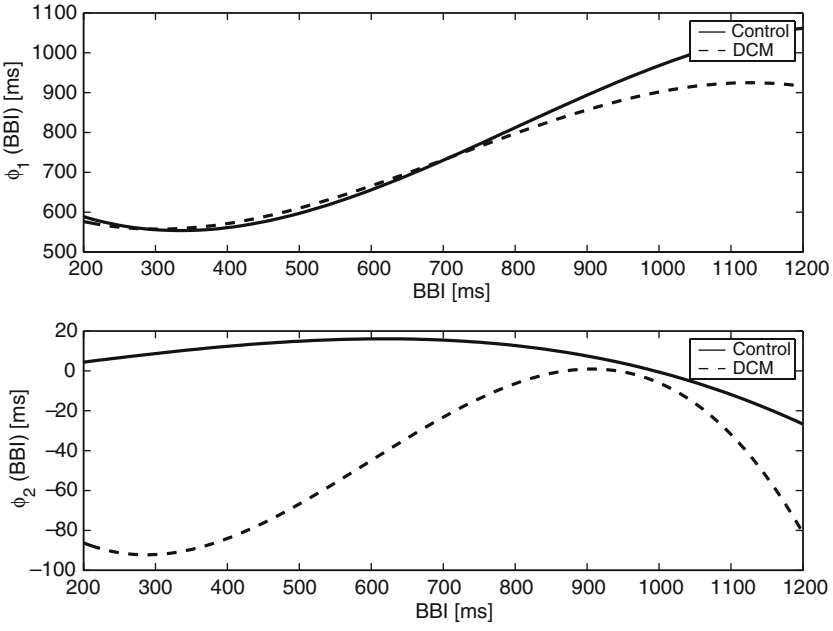

Fig. 3 Averaged polynomials of the optimal transformations ϕ_1 and ϕ_2 for the DCM (*dashed line*) as well as for the control group (*solid line*) (BBI)

decreased LF-component as well as a higher level of ventricular ectopy compared to controls. In a next step a linear autoregressive model was fitted to the data, a model order of 2 turned out to be most effective in discriminating both groups. HR variability reflects the complex interactions of many different control loops of the cardiovascular system. In relation to the complexity of the sinus node activity modulation system, a more predominantly nonlinear behavior has to be assumed. To this effect, a nonlinear regression approach, namely, the maximal correlation method was applied for non-parametric data-driven NAAR-modeling. Using this approach, no a priori assumptions about the dynamic system that regulates HR have to be made. To enable model-based data analysis, however, the estimated functional relationships have to be quantified. For this reason, polynomial fitting was performed and the fitting coefficients themselves were considered to be parameters in cardiac diagnostics. We found that this NAAR-model-approach is a powerful tool in discriminating cardiac patients from healthy controls. A tenfold bootstrap validation analysis revealed that one fitting parameter discriminates best in comparison to all parameters in this study. It must be mentioned here that these parameters, of course, are correlated with HR variability parameters. The highest correlation was found between a_4 and the mean HR with $r = 0.91$. The coefficient a_4 represents the offset of the polynomial

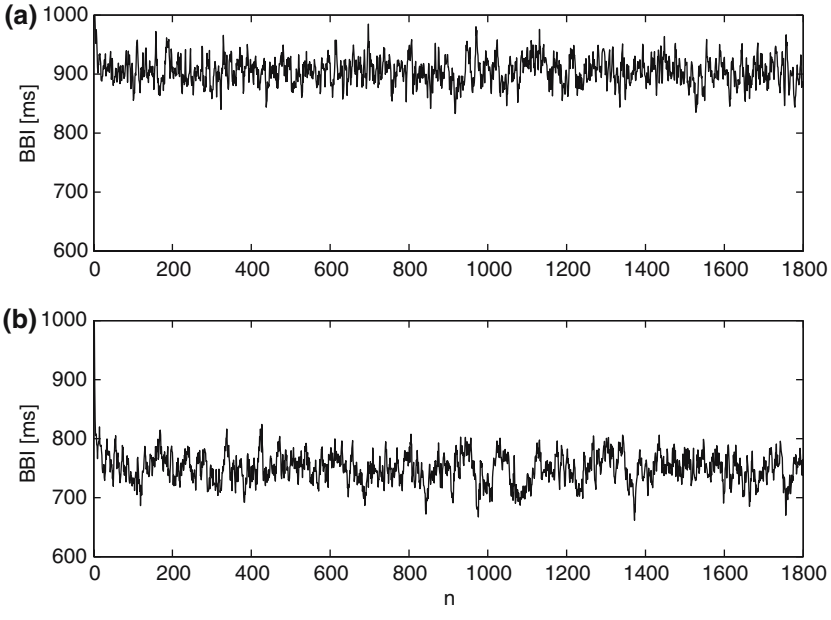

Fig. 4 Monte-Carlo simulations using the polynomials from Fig. 3 for the control (**a**) as well as for the DCM group (**b**) (BBI)

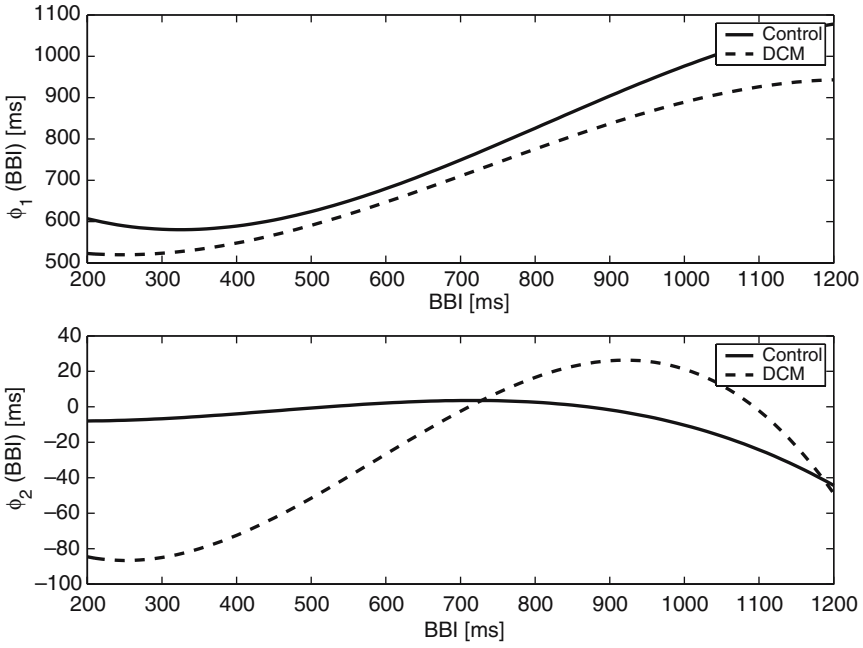

Fig. 5 Polynomials estimated from the Monte-Carlo simulations in Fig. 4 for the DCM (*dashed line*) as well as for the control group (*solid line*) (BBI)

fitting, and thus, the minimum beat-to-beat-interval, which obviously is correlated with the mean heart rate. The most significant coefficient b_3, however, did not show such high correlations with HRV parameters, the highest was found for LF/P with $r = 0.50$. The highest correlation for the coefficients b_1 and b_2 were $r = -0.46$ (to LF/P) and $r = 0.41$ (to meanNN), respectively. In this regard, the decrease of low frequency together with the apparently less effective baroreflex regulation in DCM patients may explain the differences in ϕ_2 between both groups. For the controls, we got a strong relationship between one beat-to-beat-interval and its pre-predecessor, whereas this clear relation is weakened for the DCM group. The Monto-Carlo simulations performed in this study confirmed these findings, the control subjects show a more homogeneous and stable behavior in HRV. Our results clearly indicate, that the NAAR-model parameters contain important additional information in comparison to HRV parameters. The simulation study showed that one-step predictions based on optimal transformations gave promising results. For subjects with atrial fibrillation or ventricular ectopy, however, no real functional relationship could be estimated. A precondition for applying this method are at least minimal

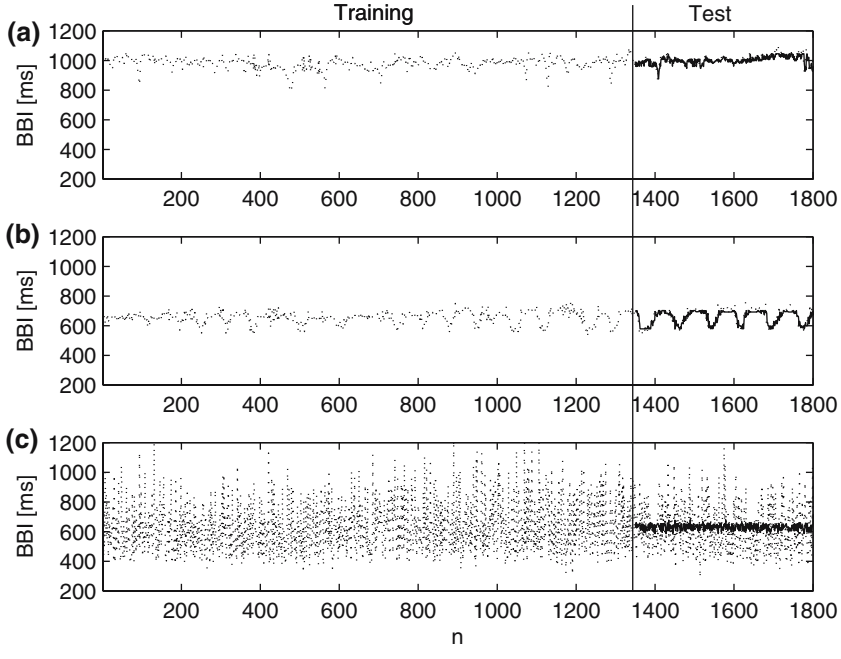

Fig. 6 One step predictions (*solid line*) of the tachograms (**a**), (**b**) and (**c**) from Fig. 1 using the optimal transformation estimated from the training set (first three quarters of the original data set—*dotted line*) and the last two values (BBI)

correlations in the data and the exlusion of artifacts. All results remain to be validated on a larger data base and it will have to be investigated whether more sophisticated models, such as polynomial fittings or NARMA-models [1, 2, 5, 10, 15, 26], improve model-based data analysis in risk stratification of cardiac patients. Summarizing the results of this paper, we find that the ACE algorithm is a powerful tool to estimate the transformations θ and ϕ in the analyzed model and, thus, enables the qualitatively new step: the combination of data analysis and modeling. Model-based nonlinear-dynamic data analysis of non-invasively measured biosignals lead to an improved clinical diagnostics in this study. La Rovere et al. [23] showed that short-term HRV analysis has an independent prognostic value in risk stratification apart from that of clinical and functional variables. Our introduced methodology is able to detect changes in cardiovascular variability due to cardiac dysfunction exemplary for DCM patients. It has to be mentioned here, that this study was an exclusively methodological investigation and has minor relevance for diagnosing DCM patients. However, improving these modeling techniques and understanding the underlying mechanisms may lead to the detection of cardiac dysfunction based on Holter monitoring in the future. Therefore, our approach seems to be very promising also for other biosignal processing, e.g. blood pressure variability or multivariate cardiovascular modeling with nonlinear interactions and feedback loops.

Acknowledgments We would like to thank Henning Voss for providing the Matlab-code and very helpful discussions as well as the Deutsche Forschungsgemeinshaft (DFG BA 1581/4-1, BR 1303/8-1, KU 837/20-1) for supporting the study.

5 Appendix: the ACE algorithm

This appendix provides a short description of the ACE algorithm of Breiman and Friedman [6], the computer programs used can be obtained from the authors (http://tocsy.agnld.uni-potsdam.de/). In the following section, the same notations as introduced in Sect. II are used.

Generally, the estimation of functions that are optimal for correlation is equivalent to the estimation of functions that are optimal for regression. Therefore, the problem

$$\Psi(S, T_1, \ldots, T_k) = \max_{\theta,\phi_i} \left| \rho(\theta(S), \sum_{i=1}^{k} \phi_i(T_i)) \right| \tag{8}$$

may also be expressed as the regression problem

$$E\left[\left(\theta(S) - \sum_{i=1}^{k} \phi_i(T_i) \right)^2 \right] != \min . \tag{9}$$

Here, the functions θ and ϕ_j ($j=1,\ldots,k$) are varied in the space of Borel measurable functions, and the constraints onto these functions are that they have vanishing expectation and finite variances to exclude trivial solutions.

For the one-dimensional case ($k=1$), the ACE algorithm works as follows: When denoting the conditional expectation of $\phi_1(T_1)$ with respect to S by $E[\phi_1(T_1)|S]$, then the function $\bar{\phi}_0(S) = E[\phi_1(T_1)|S]$ minimizes (9) with respect to $\theta(S)$ for a given $\phi_1(T_1)$. Similarly, $\bar{\phi}_1(T_1) = E[\theta(S)|T_1]/||E[\theta(S)|T_1]||$, where the norm is defined by $||Z|| = \sqrt{\mathrm{var}[Z]}$, minimizes (9) with respect to $\phi_1(T_1)$ for a given $\theta(S)$, keeping $E[\phi_1^2(T_1)]=1$. Now, the ACE algorithm consists of the following iterative procedure: Starting with the initial function

$$\phi_1^{(1)}(T_1) = E[S|T_1], \tag{10}$$

from $i=2$ it is calculated recursively

$$\theta^{(i)}(S) = E[\phi_1^{(i-1)}(T_1)|S] \tag{11}$$

and

$$\phi_1^{(i)}(T_1) = E[\theta^{(i)}(S)|T_1]/||E[\theta^{(i)}(S)|T_1]||, \tag{12}$$

until $E[(\phi_1^{(i)}(T_1) - \theta^{(i)}(S))^2]$ fails to decrease. The limit values are then estimates for the optimal transformations θ and ϕ_1. For the minimization of the right hand side of Eq. 9 a double-loop algorithm is used. In the additional inner loop, the functions

$$\phi_j^{(i)}(T_j) = E\left[\theta^{(i)}(S) - \sum_{p \neq j} \phi_p^{(i,i-1)}(T_p) \mid T_j \right]$$

are calculated. In the sum, the superscript "$.^{(i)}$" is used for $p < j$ and "$.^{(i\text{-}1)}$" for $p > j$. For $k > 1$, the ACE algorithm works similarly.

There are several possibilities of estimating conditional expectations from finite data-sets. In our examples, local smoothing of the data is used. This smoothing can be achieved with different kernel estimators. We use a simple boxcar window, i.e. the conditional expectation value $E[y|x]$ is estimated at each site i via

$$\hat{E}[y|x_i] = \frac{1}{2N+1} \sum_{j=-N}^{N} y_{i+j}$$

for a fixed half window size N. In all examples of this paper, $n=30$ is used to account for a reliable estimate of the mean value.

Furthermore, to allow for a better estimation in the case of inhomogeneous distributions, the data are transformed to have rank-ordered distributions prior to the application of the ACE algorithm [i.e. the data-set X is sorted in ascending order, resulting in the vector Y and all further calculations are performed with the corresponding index vector I, where $Y=X(I)$]. This allows for a more precise estimation of expectation values independently of the form of the data distribution and

simplifies the algorithm considerably. It is allowed, since the rank transformation is invertible and the maximal correlation is invariant under invertible transformations by definition. Proofs of convergence and consistency of the function estimates are given in Ref. [6].

References

1. Aguirre LA, Billings SA (1995) Improved structure selection based on term clustering. Int J Control 62:569–587
2. Aguirre LA, Correa MV, Cassini CCS (2002) Nonlinearities in narx polynomial models: representation and estimation. IEE Proc Part D Control Theory Appl 149:343–348
3. Baselli G, Cerutti S, Badilini F, Biancardi L, Porta A, Pagani M, Lombardi F, Rimoldi O, Furlan R, Malliani A (1994) Model for the assessment of heart period and arterial pressure variability interactions and of respiration influences. Med Biol Eng Comput 32:143–152
4. Berntson GG, Bigger JT, Ckberg DL, Grossman P, Kaufmann PG, Malik M, Nagaraja HN, Porges SW, S.J.P, P. H. Stone, an der Molen MW (1997) Heart rate variability: origins, methods, and interpretative caveats. Psychophysiology 34:623–648
5. Billings SA, Leontaritis IJ (1982) Parametric estimation techniques for nonlinear systems. In: Proceedings of 6th IFAC symposium ident., parameter estimation, Washington, pp 427–432
6. Breiman L, Friedman J (1985) Estimating optimal transformations for multiple regression and correlation. J Am Stat Assoc 80:580–619
7. Censi F, Calcagnini G, Lino S, Seydnejad SR, Kitney RI, Cerutti S (2000) Transient phase locking patterns among respiration, heart rate and blood pressure during cardiorespiratory synchronisation in humans. Med Biol Eng Comput 38:416–426
8. Cerutti S, Carrault G, Cluitmans PJ, Kinie A, Lipping T, Nikolaidis N, Pitas I, Signorini MG (1996) Non-linear algorithms for processing biological signals. Comput Methods Programs Biomed 51:51–73
9. Chen R, Tsay R (1993) Nonlinear additive arx models. J Am Stat Assoc 88:955–967
10. Chon KH, Kanters JK, Cohen RJ, Holstein-Rathlou NH (1997) Detection of chaotic determinism in time series from randomly forced maps. Phys D 99:471–486
11. Focus issue: dynamical disease: mathematical analysis of human illness. Chaos 5:1–344
12. Focus issue: mapping and control of complex cardiac arrhythmias (2002) Chaos 12:732–981
13. Goldberger AL, Rigney DR, Mietus J, Antman EM, Greenwald S (1988) Nonlinear dynamics in sudden cardiac death syndrome: heartrate oscillations and bifurcations. Experientia 44:983–987
14. Gray RA, Pertsov AM, Jalife J (1998) Spatial and temporal organization during cardiac fibrillation. Nature 392:75–78
15. Haber R, Keviczky L (1976) Identification of nonlinear dynamic systems. In: IFAC symp. ident. syst. parameter estimation, Tbilisi, Georgia, pp 62112
16. Hastie T, Tibshirani R (1990) Generalized additive models. Chapman and Hall, London and New York
17. Heart rate variability standards of measurement, physiological interpretation and clinical use (1996) Task force of the european society of cardiology and the north american society of pacing and electrophysiology. Circulation 93:1043–1065
18. Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70
19. Ivanov PC, Amaral LA, Goldberger AL, Havlin S, Rosenblum MG, Struzik ZR, Stanley HE (1999) Multifractality in human heartbeat dynamics. Nature 399:461–465
20. Kantz H, Kurths J, Mayer-Kress GE (1998) Nonlinear analysis of physiological data. Springer, Berlin Heidelberg New York
21. Kleiger RE, Miller JP, Bigger J, Moss A (1987) Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. Am J Cardiol 59:256–262
22. Kurths J, Voss A, Witt A, Saparin P, Kleiner H, Wessel N (1995) Quantitative analysis of heart rate variability. Chaos 5:88–94
23. La Rovere MT, Pinna GD, Maestri R, Mortara A, Capomolla S, Febo O, Ferrari R, Franchini M, Gnemmi M, Opasich C, Riccardi P, Traversi E, Corbelli F (2003) Short-term heart variability predicts sudden cardiac death in chronic heart failure patients. Circulation 107:565–570
24. Lakatta EG, Levy D (2003) Arterial and cardiac aging: major shareholders in cardiovascular disease enterprises: Part ii: the aging heart in health: links to heart disease. Circulation 107:346–354
25. Lombardi F (2002) Clinical implications of present physiological understanding of hrv components. Card Electrophysiol Rev 6:245–249
26. Lu S, Chon KH (2003) Nonlinear autoregressive and nonlinear autoregressive moving average model parameter estimation by minimizing hypersurface distance. IEEE Trans Sig Proc 51:3020–3026
27. Malik M, Padmanabhan V, Olson WH (1999) Automatic measurement of long-term heart rate variability by implanted single-chamber devices. Med Biol Eng Comput 37:585–594
28. Poon CS, Merrill CK (1997) Decrease of cardiac chaos in congestive heart failure. Nature 389:492–495
29. Renyi A (1970) Probability theory. Akadémiai Kiadó, Budapest
30. Rompelman O, Coenen AJ, Kitney RI (1977) Measurement of heart-rate variability: Part 1—comparative study of heart-rate variability analysis methods. Med Biol Eng Comput 15:233–239
31. Schmidt G, Malik M, Barthel P, Schneider R, Ulm K, Rolnitzky L, Camm AJ, Bigger JTJ, Schomig A (1999) Heart-rate turbulence after ventricular premature beats as a predictor of mortality after acute myocardial infarction. Lancet 353:1390–1396
32. Special issue virtual tissue engineering of the heart (2003) Int J Bifurcat Chaos 13:3531–3804
33. Szabo BM, van Veldhuisen DJ, van der Veer N, Brouwer J, De Graeff PA, Crijns HJ (1997) Prognostic value of heart rate variability in chronic congestive heart failure secondary to idiopathic or ischemic dilated cardiomyopathy. Am J Cardiol 79:978–980
34. Tsuji H, Larson MG, Venditti FJ, Manders ES, Evans JC, Feldman CL, Levy D (1996) Impact on reduced heart rate variability on risk for cardiac events. Circulation 94:2850–2855
35. Voss H, Kurths J (1997) Reconstruction of non-linear time delay models from data by the use of optimal transformations. Phys Lett A 234(5):336–344
36. Voss A, Kurths J, Kleiner HJ, Witt A, Wessel N, Saparin P, Osterziel KJ, Schurath R, Dietz R (1996) The application of methods of non linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death. Cardiovasc Res 31:419–433
37. Voss HU, Kolodner P, Abel M, Kurths J (1999) Amplitude equations from spatiotemporal binary-fluid convection data. Phys Rev Lett 83(17):3422–3425
38. Wessel N, Voss A, Kurths J, Schirdewan A, Hnatkova K, Malik M (2000) Evaluation of renormalised entropy for risk stratification using heart rate variability data. Med Biol Eng Comput 38:680–685
39. Wessel N, Voss A, Malberg H, Ziehmann C, Voss HU, Schirdewan A, Meyerfeldt U, Kurths J (2000) Nonlinear analysis of complex phenomena in cardiological data. Herzschr Elektrophys 11:159–173
40. Wessel N, Ziehmann C, Kurths J, Meyerfeldt U, Schirdewan A, Voss A (2000) Short-term forecasting of life-threatening cardiac arrhythmias based on symbolic dynamics and finite-time growth rates. Phys Rev E 61:733–739

41. Wessel N, Schirdewan A, Kurths J (2003) Intermittently decreased beat-to-beat variability in congestive heart failure. Phys Rev Lett 91:119801
42. Wessel N, Malberg H, Walther T (2004) Heart rate turbulence higher predictive value than other risk stratifiers? Circulation 109:e150–e151
43. Wessel N, Aßmus J, Weidermann F, Konvicka J, Nestmann S, Neugebauer R, Schwarz U, Kurths J (2004) Modeling thermal displacements in modular tool systems. Int J Bifurcat Chaos 14:2125–2132
44. Wessel N, Konvicka J, Weidermann F, Nestmann S, Neugebauer R, Schwarz U, Wessel A, Kurths J (2004) Predicting thermal displacements in modular tool systems. Chaos 14:23–29
45. Wikswo JPJ, Lin SF, Abbasm RA (1995) Virtual electrodes in cardiac tissue: A common mechanism for anodal and cathodal stimulation. Biophys J 69:2195–2210
46. Witkowski FX, Leon LJ, Penkoske PA, Giles WR, Spano ML, Ditto WL, Winfree AT (1998) Spatiotemporal evolution of ventricular fibrillation. Nature 392:78–82
47. Yamamoto Y, Hughson RL (1994) On the fractal nature of heart rate variability in humans: effects of data length and beta-adrenergic blockade. Am J Physiol 266:R40–R49
48. Yang AC, Hseu SS, Yien HW, Goldberger AL, Peng CK (2003) Linguistic analysis of the human heartbeat using frequency and rank order statistics. Phys Rev Lett 90:108103
49. Zipes DP, Wellens HJ (1998) Sudden cardiac death. Circulation 98:2334–2351