

The Effect of Aging on Nonlinearity and Stochastic Nature of Heart Rate Variability Signal Computed using Delay Vector Variance Method

Srinivas Kuntamalla
Department of Physics
National Institute of Technology
Warangal, INDIA - 506004

L. Ram Gopal Reddy
Department of Physics
National Institute of Technology
Warangal, INDIA - 506004

ABSTRACT

Heart rate variability analysis is fast gaining acceptance as a potential non-invasive means of autonomic nervous system assessment in research as well as clinical domains. In this study, a new nonlinear analysis method is used to detect age related changes in the degree of nonlinearity and stochastic nature of heart rate variability signals. The data obtained from an online and widely used public database (i.e., MIT/BIH physionet database), of young and elderly subjects is used in this study. The method used is the delay vector variance (DVV) method, which is a unified method for detecting the presence of determinism and nonlinearity in a time series and is based upon the examination of local predictability of a signal. From the results it is clear that there is no significant change in the minimum target variance values for young and elderly subjects and also the values are very small, which indicates that there is a strong deterministic component over the stochastic one in both the groups. There is a significant decrease in the degree of nonlinearity from younger to elder subjects (p- value, 0.0002). This indicates that there is no change in the stochastic or deterministic nature of the signals but there is a considerable change in the degree of nonlinearity with aging.

General Terms

Nonlinear analysis, Heart rate variability.

Keywords

Nonlinearity, stochastic nature, heart rate variability, delay vector variance

1. INTRODUCTION

Heart rate variability (HRV) analysis is gaining acceptance as a potential non-invasive means of autonomic nervous system assessment in research as well as clinical domains. Usually HRV data is obtained from R peaks of electrocardiogram (ECG). The heart beat instants are taken at these points and consequently the beat to beat intervals are determined as the time interval from one R peak to the next one. Therefore, these intervals are called R-R intervals and are plotted against their beat number, which is called a tachogram (Figure 1). In 1996, the Taskforce of the ESC/NASPE published standards in HRV analysis proposing several time and frequency parameters based on short-term (5-min) and long-term (24-h) HRV data [1]. Although HRV has been the subject of many clinical studies investigating a wide spectrum of cardiological and non-cardiological diseases and clinical conditions, a general consensus of the practical use of HRV in medicine has been reached only in two clinical scenarios:

depressed HRV can be used as a predictor of risk after acute myocardial infarction and as an early warning sign of diabetic neuropathy. The HRV can be analyzed using several methods which are broadly classified as time domain and frequency domain methods. Time domain measures are simple statistical operations on R-R intervals, such as standard deviation of normal R-R intervals (SDNN), root mean square of successive R-R interval differences (RMSSD) and the percentage change of normal R-R intervals that differ by > 50 ms (PNN50) etc. Frequency domain analysis includes FFT or AR based power spectral density measures which provide information on how variance distributes as a function of frequency. Three main spectral components are distinguished in a spectrum calculated from short-term recordings in both absolute and normalized units: very low frequency (VLF) (≤ 0.04 Hz), Low frequency (LF) (0.04 - 0.15 Hz), and high frequency (HF) (0.15 - 0.4 Hz) components. Heart rate dynamics are nonlinear in nature and it is proved that nonlinear analysis of it provides more appropriate information for the physiological interpretation of heart rate variability [2]. However, variety of contradictory reports in this domain indicates that there is a need for a more rigorous investigation of methods. The nonlinear analysis of HRV is a valuable tool in both clinical practice and physiological research reflecting the ability of the cardiovascular system [3]. Poincare plot, Approximate Entropy, Sample Entropy, Detrended Fluctuation Analysis, Correlation Dimension and Sequential trend analysis are some of the nonlinear analysis methods of HRV.

In recent years, as the presence of nonlinearity and determinism in a biomedical signal is used as an index for risk stratification in many diseases [4], detecting the nature of physiological time series received large attention. Physiological time series are generated by complex systems for which it is not possible to solve or even set up the equations governing the dynamics, and generally assumed that such time series significantly display nonlinearity. There are two types of methods available for testing the nonlinearity in a time series [5]

- (i) fitting a linear or nonlinear model to the time series and their accuracies are evaluated
- (ii) comparing nonlinearity measures computed for the signal with those computed for linearised versions of the signal (surrogates)

The delay vector variance (DVV) is a unified method for detecting the presence of determinism and nonlinearity in a time series and is based upon the examination of local predictability of a signal. Further, it spans the complete range of local linear

models due to the standardization to the distribution of pair wise distances between delay vectors. This approach is easy to interpret, because it produces diagrams (DVV-plot and DVV Scatter plot) that give information about the prevalence of stochastic or deterministic components and linear or nonlinear nature of signals. The DVV-plot characterizes a time series in a standardized manner and DVV Scatter plot employs the concept of surrogate data.

This paper describes the parameters from delay vector variance method which characterize the nonlinearity and stochastic nature of the HRV signal. First the DVV method is introduced and the parameters that describe the nonlinearity and determinism are presented. Then the parameters of DVV are calculated from HRV signals pertaining to healthy young individuals and healthy old persons. Finally these parameters for the two groups are statistically compared

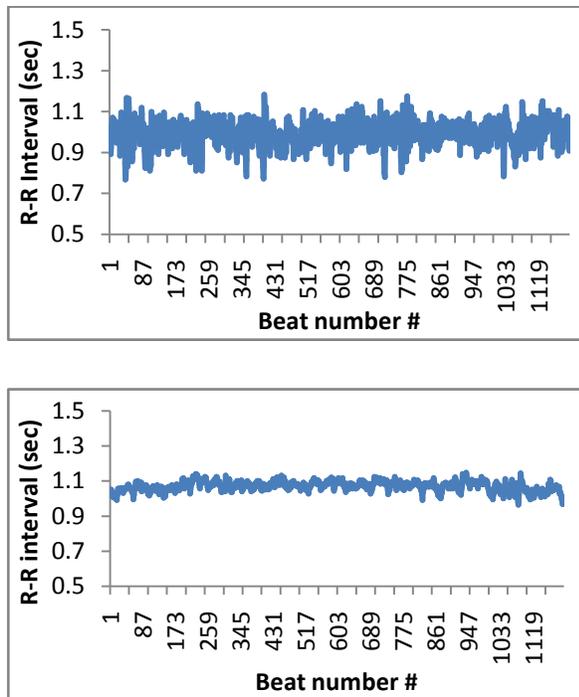


Figure 1: R-R interval tachograms of a healthy young subject (top) and a healthy elder subject (bottom)

2. METHODS

2.1 Data Group

The data base used in this study is publicly available FANTASIA data base from Physionet website [6]. This data base contains twenty young (21 - 34 years old) and twenty elderly (68 - 85 years old) rigorously-screened healthy subjects electrocardiographic (ECG), and respiration signals of 120 minutes duration. In half of each group, the recordings also include an uncalibrated continuous non-invasive blood pressure signal. Each subgroup of subjects includes equal numbers of men and women.

To maintain wakefulness all the subjects are made to watch the movie fantasia (Disney, 1940) in a resting. All the signals were digitized at 250 Hz. The database also contains heart beat instances annotated using an automated arrhythmia detection algorithm, and each beat annotation was verified by visual inspection.

From this database 19 records from young group and 19 records from elderly group are considered for study. The two records f2o08 and f2y09 are not used because of lot of ectopic beats.

2.2 Surrogate Data Generation

A surrogate signal is a realization of null hypothesis, which in this paper is that the original signal is linear. There are two main methods of generating a surrogate series

- (i) Phase randomization method
- (ii) Iterative Amplitude Adjusted Fourier Transform

Any stationary linear process can be fully specified by its auto correlation function. Thus computing the Fourier transform of the original series and replacing the phases with random number, uniformly distributed between 0 and 2π and taking inverse transform produces a surrogate series which is having same autocorrelation function. This method of generating surrogate series by adding random phases to the Fourier Transform is called Phase randomization. Nevertheless, surrogate data generated using phase randomization has a Gaussian distribution of values, whereas HRV series do not have a normal distribution [7].

Therefore, the Iterative Amplitude Adjusted Fourier Transform approach described in [8] is used for this study because it yields superior results and overcomes the normal distribution problem. This approach retains the amplitude spectrum and distribution of the original signal. The algorithm consists of a simple iteration scheme. A sorted list of the values $\{s_n\}$ and the squared amplitudes of the Fourier transform of $\{s_n\}$, $\{S_k^2\}$ are calculated initially. Beginning with a random shuffle (without replacement) $\{s_n^{(0)}\}$ of the data, $\{s_n^{(i)}\}$ is brought to the desired sample power spectrum. This is achieved by taking the Fourier transform of $\{s_n^{(i)}\}$, replacing the squared amplitudes $\{S_k^{2(i)}\}$ by $\{S_k^2\}$, and then transforming back. The phases of the complex Fourier components are kept unchanged. Thus the first step enforces the correct spectrum but usually the distribution will be modified. Therefore, as the second step, rank order the resulting series in order to assume exactly the values taken by $\{s_n\}$. As the spectrum of the resulting $\{s_n^{(i+1)}\}$ will be modified again the above two steps have to be repeated several times.

2.3 Delay Vector Variance Method

A time series can be represented in phase space conveniently using time delay embedding. When time delay is embedded into a time series it can be represented by a set of delay vectors (DVs) of a given dimension. If m is the dimension of the delay vectors then it can be expressed as $X(k) = [x_{(k-m\tau)} \dots x_{(k-\tau)}]$, where τ is the time lag. Now for every DV $X(k)$, there is a corresponding target, namely the next sample x_k . A set $\beta_k(m, d)$ is generated by grouping those DVs that are within a certain Euclidean distance (d) to DV $X(k)$. This Euclidean distance will be varied in a manner standardized with respect to the distribution of pair wise distances between DVs. Now for a given embedding dimension m , a

measure of unpredictability σ^{*2} (target variance) is computed over all sets of β_k . The variation of the standardized distance enables the complete range of pair wise distances to be examined. The procedure for Delay Vector Variance method can be summarized as below [9][10].

The mean μ_d and the standard deviation σ_d are computed over all pair wise Euclidean distances between DVs given by $\|x(i) - x(j)\|$ ($i \neq j$)

The sets $\beta_k(m, d)$ are generated such that $\beta_k = \{x(i) | \|x(k) - x(i)\| \leq d\}$ i.e., sets which consist of all DVs that lie closer to $X(k)$ than a certain distance d , taken from the interval $[\mu_d - n_d \sigma_d ; \mu_d + n_d \sigma_d]$ where n_d is a parameter controlling the span over which to perform DVV analysis.

For every set $\beta_k(m, d)$ the variance of the corresponding targets $\sigma_k^2(m, d)$ is computed. The average over all sets $\beta_k(m, d)$ is divided by the variance of the time series signal, σ_k gives the inverse measure of predictability, namely target variance $\sigma^{*2}(m, d)$. The variance is computed only if there are at least 30 DVs in a set $\beta_k(m, d)$.

The plot of target variance $\sigma^{*2}(m, d)$ as a function of spans d for a given dimension m is called DVV plot and as the distance is standardized it is easy to interpret the plot. An example of DVV plot is shown in figure 2.

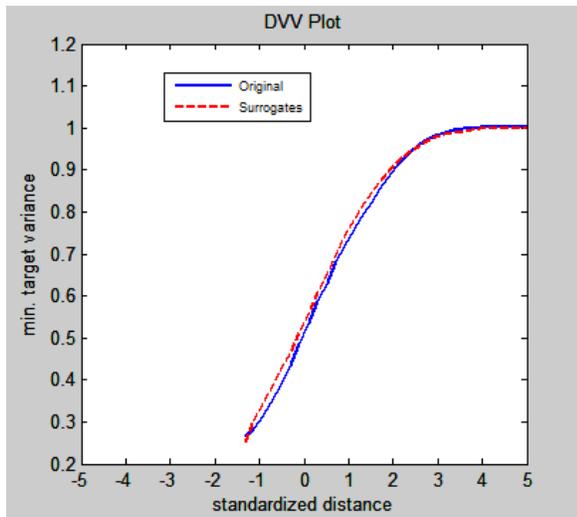


Figure 2: DVV plot for an elderly subject (f2o03) (solid curve) and average DVV plots computed over 25 surrogates (dashed curve)

The minimum target variance, $\sigma_{min}^{*2}(m)$ is a measure of noise present in the time series. The amount of noise is the prevalence of the stochastic component. The presence of strong deterministic component will lead to small target variance for small spans (on the left hand side of the DVV plot). At the extreme right, the DVV plots smoothly converges to unity, since for maximum spans, all DVs belong to the same universal set, and the variance of the targets is equal to the variance of the time series.

With the help of the original and the surrogate time series DVV scatter diagram can be produced, where the horizontal axis corresponds to the DVV plot of the original time series and the

vertical to that of the surrogate time series. If the surrogate time series signal yields DVV plots similar to that of the original time series, the ‘DVV scatter diagram’ (figure 3) coincides with the bisector line and original signal is linear. If the surrogate time series yield DVV plots not similar to that of the original time series, the curve will deviate from the bisector line and original time series is non-linear. Thus the deviation from the bisector line is an indication of nonlinearity, and can be quantified as root mean square error (RMSE) between the σ^{*2} s of the original time series and the σ^{*2} s averaged over the DVV plots of the surrogate data.

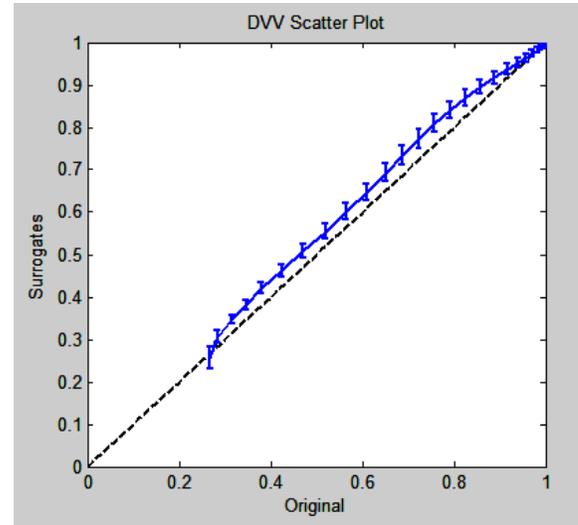


Figure 3: DVV scatter plot for an elderly subject (f2o03)

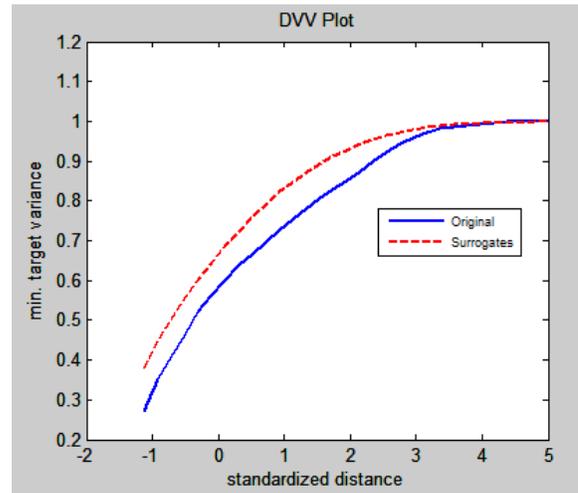


Figure 4: DVV plot for a young subject (f1y06) (solid curve) and average DVV plots computed over 25 surrogates (dashed curve)

2.4 Feature Extraction

In this method the parameters of interest with respect to DVV plots are minimum target variance of original signal, minimum target variance of average DVV plots computed over surrogate signals and their difference (MTVSOD, minimum target variance

difference between surrogates and original). The parameter with respect to DVV scatter plot is rms error value i.e., root mean square error (RMSE) between the σ^{*2} s of the original time series and the σ^{*2} s averaged over the DVV plots of the surrogate data. In this paper the value of n_d is fixed at 5 to make the DVV plot smoothly converge to unity at extreme right, since all DVs belong to the same set for maximum span. The value of τ is taken as 1 and the value of m is taken as 5. Here no procedure is adopted for calculating optimum time delay embedding dimension as we are comparing the parameters of young and elderly subjects, there must be a common dimension for all the subjects. Therefore, all the parameters are computed for given dimension ($m=5$) and given distance parameter ($n_d=5$) and $\tau=1$. The number of surrogates generated are 25 using IAAFT method. As reported by Temujin Gautama et al. [9], the length of the time series must be greater than 1000, otherwise there is a profound effect on minimum target variance. The length of time series considered in this paper is 1200 samples. The feature extraction and implementation of method is done in Matlab 7.3.

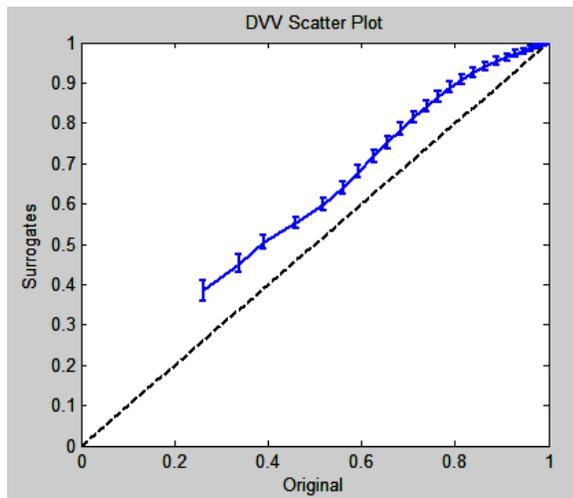


Figure 5: DVV scatter plot for a young subject (f1y06)

3. RESULTS

Figure 2 shows a DVV plot for an elderly subject (f2o03) (solid curve) and average DVV plots computed over 25 surrogates (dashed curve). The two curves almost coincide and has a σ_{min}^{*2} value of 0.253 for surrogates and 0.261 for original signal. There is no much difference between these two values.

Figure 3 shows a DVV scatter plot for an elderly subject (f2o03), whose rms error value is 0.0182. Figure 4 shows a DVV plot for a young subject (f1y06) (solid curve) and average DVV plots computed over 25 surrogates (dashed curve). There is a clear difference seen between the two curves and the σ_{min}^{*2} value for surrogates is 0.386 and for original signal is 0.279. There is a lot of difference between these two values.

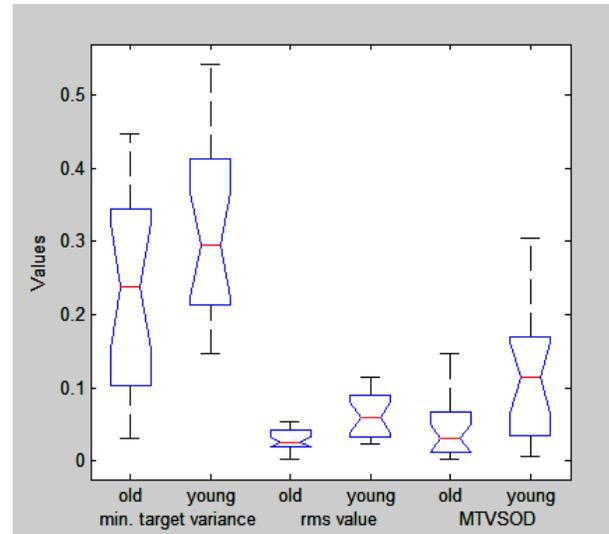


Figure 6: Box – Whiskers plot for the parameters determined from DVV plots and DVV scatter plots for young and old groups.

Figure 5 shows a DVV scatter plot for a young subject (f1y06). The rms error value for this subject is 0.1342, which is a large value compared to the rms value in figure 3.

Box – Whiskers plot for the parameters determined from DVV plots and DVV scatter plots for young and old groups is drawn to compare the values. The parameters extracted from these plots are tested for null hypothesis using a significance test (T-test). T-test is the most commonly used method to evaluate the differences in means between the two groups. The significance level for rejection of null hypothesis is set to 0.01 in this study. The P value < 0.01 is considered to be statistically significant. The p-values are given in Table 1 along with the mean and standard deviation values.

4. CONCLUSION

From the plots and the results it is clear that there is no significant change in the minimum target variance values for young and elderly subjects and the values are very small, which indicates that there is a strong deterministic component over the stochastic one. The results show that there is a significant decrease in the degree of nonlinearity from younger to elder subjects (p-value, 0.0002). Significant difference in the differences of minimum target variance of original and surrogates is also seen. This indicates that there is no change in the stochastic or deterministic nature of the signals but there is a considerable change in the degree of nonlinearity with aging.

Table 1 Mean(± std. dev.) values of the DVV parameters

	min. target variance	MTVSOD	rms error value
Elderly	0.0283(±0.014)	0.0475(±0.045)	0.237(±0.132)
Young	0.0620(±0.031)	0.1128(±0.081)	0.325(±0.118)
p-value (T-test)	0.063	0.0042	0.0002

5. ACKNOWLEDGMENTS

The authors are thankful to Danilo P. Mandec et al. for providing the DVV toolbox for Matlab under GNU general public license.

6. REFERENCES

- [1] Task Force of the European Society of Cardiology and North American Society of Pacing and Electrophysiology, . ‘Heart rate variability Standards of measurements, physiological interpretation, and clinical use’,. Eur Heart J., 17, pp 354-381, 1996.
- [2] K. Hnatkova, *et al.*, “Numeric processing of Lorenz plots of RR intervals from long-term ECGs : Comparison with time-domain measures of heart rate variability for risk stratification after myocardial infarction,” Journal of Electrocardiology, vol. 28, pp. 74-80, 1995
- [3] C. Karmakar, *et al.*, "Defining asymmetry in heart rate variability signals using a Poincaré plot," Physiological measurement, vol. 30, 2009
- [4] Phyllis K Stein and Anand Reddy, “Nonlinear heart rate variability and risk stratification in cardiovascular disease”, Indian Pacing. and Electrophysiology Journal, 2005.
- [5] T. Schreiber and A. Schmitz, “Surrogate time series”, Physica D, vol. 142, no.3-4, pp346-382, 2000.
- [6] N. Iyengar, C.-K. Peng, R. Morin, A.L. Goldberger, L.A. Lipsitz, “Age-related alterations in the fractal scaling of cardiac inter beat interval dynamics”, Am. J. Physiol., 271, p 1078, 1996.
- [7] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian and J.D. Farmer, “Testing for nonlinearity in time series: The method of surrogate data”, Physica D, 58, pp 77-94, 1992.
- [8] T. Schreiber and A. Schmitz, “Improved surrogate data for nonlinearity tests”, Phys.Rev. Lett., 77, pp 635-638,1996.
- [9] Temujin Gautama, Danilo P. Mandic , Marc M. Van Hulle, “The delay vector variance method for detecting determinism and nonlinearity in time series”, Physica D, 190, pp 167–176, 2004
- [10] Temujin Gautama, Danilo P. Mandic , Marc M. Van Hulle, “A Novel method for determining the nature of time series”, IEEE trans. Biomed. Engg., vol. 51, no. 5, pp 728-736, 2004.