# Novel Feature for Quantifying Temporal Variability of Poincaré Plot: A Case Study

CK Karmakar, AH Khandoker, J Gubbi, M Palaniswami

The University of Melbourne, Melbourne, VIC, Australia

## Abstract

*The Poincaré plot of RR intervals is one of the most popular techniques used in heart rate variability (HRV) analysis. The standard descriptors $SD1$ and $SD2$ of Poincaré plot represents the distribution of signal by quantifying spatial (shape) information. The present study proposes a novel descriptor, Complex Correlation Measure (CCM), to quantify changes in temporal structure of points of Poincaré plots. To compare performance of $CCM$ with standard Poincaré descriptor $SD1$ and $SD2$, we have calculated ROC area for each descriptor between Normal Sinus Rhythm (NSR) and Congestive Heart Failure (CHF) subjects. The RR intervals of 54 NSR subjects and 29 CHF subjects from Physionet NSR and CHF database are used. The $p$ value obtained from chi-square analysis between two groups was found significant only for $CCM$ ($p$=9.07E-14). The largest ROC area between two groups was for $CCM$ (0.92) which indicate that $CCM$ can be used as a significant feature for detecting pathology.*

## 1. Introduction

In general, Poincaré plot is a two dimensional plot constructed by plotting consecutive points of RR time-series on phase space or cartesian plane [1]. It has been shown to reveal patterns of heart rate dynamics resulting from nonlinear processes [2, 3]. It is extensively used for qualitative visualization of physiological signal. It is commonly applied to asses the dynamics of heart rate variability (HRV) [2, 4–7]. Tulppo *et. al.* [2] fitted an ellipse to the shape of the Poincaré plot and defined two standard descriptors of the plot $SD1$ and $SD2$ for quantification of the Poincaré plot geometry. These standard descriptor represent the minor axis and the major axis of the ellipse respectively as shown in figure 1. The description of $SD1$ and $SD2$ in terms of linear statistics, given by Brennan *et. al.* [3], shows that the standard descriptors guide the visual inspection of the distribution. In case of HRV, it reveals a useful visual pattern of the RR interval data by representing both short and long term variations of the signal [2, 3]. In [8], the authors examined the theoretical demand with different lags and showed that there is a curvilinear relationship between lag Poincaré plot indices for normal subjects, which is lost in Congestive Heart Failure (CHF) patients. Therefore, measurement from a series of lagged Poincaré plots (multiple lag correlation) can potentially provide more information about the behavior of Poincaré plot than the conventional *lag-1* plot measurements [9].

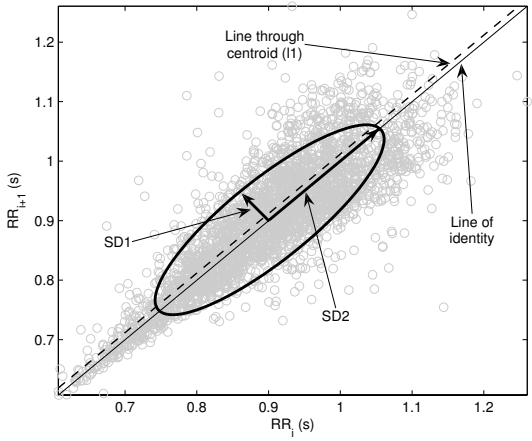


Figure 1. A standard Poincaré plot of RR intervals of a healthy person (N=2000).

Two basic descriptors of the plot are $SD1$ and $SD2$ and their mathematical derivation can be found in [3]. The line of identity is the $45^0$ imaginary diagonal line on the Poincaré plot and the points falling on the imaginary line has the property $RR_n = RR_{n+1}$. $SD1$ measures the dispersion of points perpendicular to the line of identity, whereas $SD2$ measures the dispersion along the line of identity. Fundamentally, $SD1$ and $SD2$ of Poincaré plot is directly related to the basic statistical measures, standard deviation of RR interval (SDRR), and standard deviation of the successive difference of RR interval (SDSD), which is given by the relation shown in equation 1 and equation 2.

$$SD1^2 = \frac{1}{2}SDSD^2 \quad (1)$$

$$SD2^2 = 2SDRR^2 - \frac{1}{2}SDSD^2 \quad (2)$$

The above equation sets are derived for unit time delay Poincaré plot.Researchers have shown interest in plots with different time delays to get a better insight in the time-series signal. Usually the time delay is multiple of the cycle length or the sampling time of the signal [10].

The lack of temporal information is the primary limitation of the standard descriptors of the Poincaré plot. $SD1$ and $SD2$ represents the distribution of signal in 2D space and carries only spatial (shape) information. It should be noted that many possible RR interval series result in identical plot with exactly similar $SD1$ and $SD2$ values in spite of different temporal structure. Therefore, to reflect temporal variation, we developed a descriptor to incorporate multiple lag correlation information, which we call as *Complex Correlation Measure* (CCM). The proposed descriptor is not only related to the standard descriptors, but it also embeds temporal information, which can be used in quantification of the temporal dynamics of the system. In this paper, we aim to evaluate all three descriptors ($SD1$, $SD2$ and $CCM$) of the Poincaré plot of RR intervals, and compare their performance in differentiating CHF from normal subjects.

## 2. Methods

### 2.1. Subjects

In order to validate the proposed measure - $CCM$, two case studies were conducted on RR interval data. The data from MIT-BIH Physionet database are [11] used in the experiments. Fewer attempts are made by medical fraternity to utilize Poincaré plot to evaluate CHF. In this study, we have analyzed the performance of $CCM$ and compared it with that of $SD1$ and $SD2$ for recognizing congestive heart failure using HRV signal.

In this study, we have used $54$ long-term ECG recordings of subjects in normal sinus rhythm (30 men, aged $28.5$ to $76$, and $24$ women, aged $58$ to $73$) from Physionet Normal Sinus Rhythm database [11]. Furthermore, we have also used $29$ long-term ECG recordings of subjects (aged 34 to 79) with CHF (NYHA classes I, II and III) from Physionet Congestive Heart Failure database.Same ECG acquisition with beat annotations were used as discussed in previous case study. Similar to previous case study, *lag-1* Poincaré plots were constructed for both normal and CHF subjects and the new descriptor $CCM$ was computed as per traditional descriptors.

### 2.2. Complex correlation measure ($CCM$)

$CCM$ measures the point-to-point variation of the signal rather than gross description of the Poincaré plot. It is computed in a windowed manner which embeds the temporal information of the signal. A moving window of three consecutive points from the Poincaré plot are considered and the temporal variation of the points are measured. If three points are aligned on a line then the value of the variation is zero, which represents the linear alignment of the points. Moreover, since the individual measure involves three points of the two dimensional Poincaré plot, it is comprised of at least four different points of the time series for lag $m = 1$ and at most six points in case of lag $m \geq 3$. Hence the measure conveys information about four different lag correlation of the signal. If Poincaré plot is composed of $N$ points then the temporal variation of the plot, termed as $CCM$, is composed of all overlapping three point windows and can be calculated as:

$$CCM(m) = \frac{1}{C_n(N-2)} \sum_{i=1}^{N-2} \|A(i)\| \quad (3)$$

where $m$ represents lag of Poincaré plot, $A(i)$ represents area of the *i-th* triangle and $C_n$ is the normalizing constant which is defined as, $C_n = \pi * SD1 * SD2$, represents the area of the fitted ellipse over Poincaré plot. The length of major and minor axis of the ellipse are $2SD1$, $2SD2$, where $SD1$, $SD2$ are the dispersion perpendicular to the line of identity (minor axis) and along the line of identity (major axis) respectively.

Since RR intervals are discrete signal, the autocorrelation at lag $m = j$ can be calculated as:

$$\gamma_{RR}(j) = \sum_{n=1}^{N} RR_n RR_{n+j} \quad (4)$$

Finally, $CCM(m)$ can now be expressed as a function of autocorrelation at different lags. Hence,

$$CCM(m) = F[\gamma_{RR}(0), \gamma_{RR}(m-2), , \gamma_{RR}(m-1), \gamma_{RR}(m+1), \gamma_{RR}(m+2)]$$

where, $\gamma_{RR}(m)$ represents the *lag-m* autocorrelation of the RR interval time series. In the above equation $CCM(m)$ represents the point-to-point variation of the Poincaré plot with lag $m$ as a function of autocorrelation of lags $0$, $m-2$, $m-1$, $m+1$ and $m+2$. This supports that $CCM$ is measured using multiple lag correlation of the signal rather than single lag. For the conventional *lag-1* Poincaré plot $CCM(1)$ can be represented as:

$$CCM(1) = F\left[\gamma_{RR}(-1), \gamma_{RR}(0), \gamma_{RR}(2), \gamma_{RR}(3)\right]$$

### 2.3. ROC area analysis

In order to provide the discriminative performance of all measures, receiver-operating curve (ROC) analysis was

used [12], with the areas under the curves for each feature represented by the ROC area. A ROC area value of 0.5 means that, the distributions of the features are similar in two groups with no discriminatory power. Conversely, a ROC area value of 1.0 would mean that the distributions of the features of the two groups do not overlap at all. ROC plots are used to gauge the predictive ability of a classifier over a wide range of threshold values. A threshold value was applied such that a value below the threshold was assigned into one category whereas a value equal to or above the threshold was assigned into another category. ROC curves were plotted using results to examine qualitatively the effect of threshold variation on the classification performance. The area under ROC curve was approximated numerically using the trapezoidal rules [12] where the larger the ROC area the better the discriminatory performance.

## 2.4. Statistical analysis

In this study we have used Chi-square test to determine whether the descriptors values are independent from each other for NSR and CHF group. It suits our case studies as the sample size is small.

## 3. Results

The box-whiskers plot of all descriptors for normal and CHF subjects are shown in Figure 2. Figure 2*A*, represents BW plot for $log(SD1)$ and it is apparent that boxes (interquartile range) of normal and CHF subjects are overlapping. The BW of normal subjects is completely overlapped with the box and whisker (lower quartile) of the CHF subjects. In figure 2*B*, the box-whiskers plot of $log(SD2)$ is shown and boxes are apparently non-overlapped. But the BW plot of normal subjects mostly overlaps with the whisker (upper quartile) of the CHF subjects. In figure 2*C*, the BW plot of $log(CCM)$ is shown to be non-overlapping and only the upper quartile (box) and whisker of normal subjects are overlapped with the whisker (lower quartile) of the CHF subjects.

The values of the mean and the standard deviation for both types of subjects are shown in table 1. Last row represents the $p$ value obtained from Chi-square analysis between two groups for $SD1$, $SD2$ and $CCM$. Though $SD2$ and $CCM$ shows similar difference among the mean of two subject groups, the *standard deviation* of $CCM$ is lower which concentrates the distribution of $CCM$ values around mean comparing with that of $SD2$. As a result we obtained significant $p$ value from chi-square analysis between Normal and CHF subject groups only for $CCM$ as shown in table 1. Maximum ROC area, 0.92. for differentiating CHF from NSR groups was found for $CCM$. Though ROC area for $SD2$ was 0.90, the $p$ values calculated using Chi-square test was insignificant. The lowest ROC area obtained for $SD1$ was 0.71.

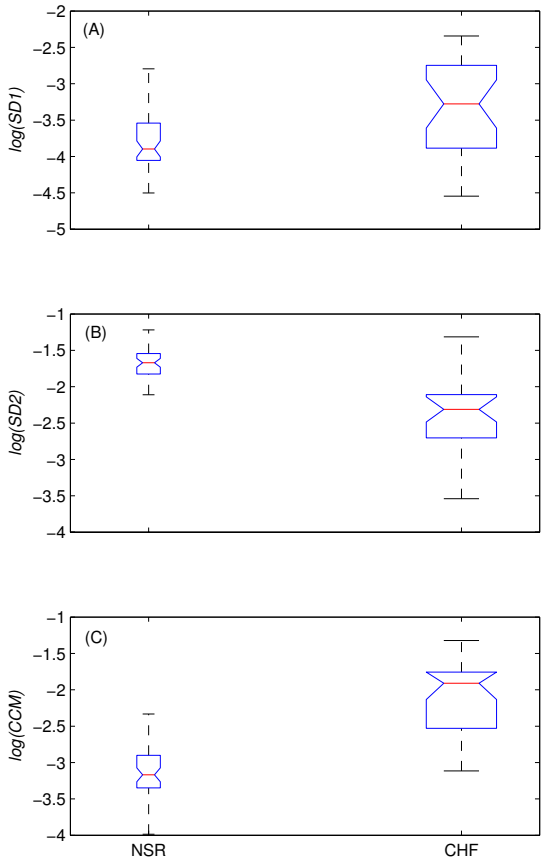


Figure 2. The distribution of descriptors are shown using Box-whiskers (BW) plot (without outliers) of (A) log(SD1), (B) log(SD2) and (C) log(CCM) for Normal Sinus Rhythm (NSR, n = 54) and Congestive Heart Failure (CHF, n = 29) subjects.

Table 1. *Mean* $\pm$ *Standard deviation* of $SD1$, $SD2$ and $CCM$ for normal sinus rhythm (NSR) and congestive heart failure (CHF) subjects. ROC (receiver-operating characteristic) area and $p$ value of Chi-square analysis are given in the last two rows.

| | $SD1$ | $SD2$ | $CCM$ |
|---|---|---|---|
| NSR | $0.03 \pm 0.02$ | $0.19 \pm 0.04$ | $0.05 \pm 0.03$ |
| CHF | $0.04 \pm 0.02$ | $0.11 \pm 0.06$ | $0.14 \pm 0.06$ |
| ROC area | 0.71 | 0.90 | 0.92 |
| $p$ value | 3.57E-02 | 4.62E-02 | 6.22E-08 |

## 4. Discussion and conclusions

The main motivation for using Poincaré plot is to visualize the variability of any time series signal. In addition to this qualitative approach, we propose a novel quantitative measure, $CCM$, to extract underlying temporal dynamics in a Poincaré plot. Both $SD1$ and $SD2$ are second order statistical measures [3], which are used to quantify the dispersion of the signal perpendicular and along the line of identity respectively. Moreover, $SD1$ and $SD2$ are functions of $lag - m$ correlation of the signal for any $m$ lag Poincaré plot. In contrast, $CCM$ is a function of multiple lag ($m-2$, $m-1$, $m$, $m+1$, $m+2$) correlations.

In the presented case study, we have shown as to how Poincaré plot can be used to characterize CHF subjects from normal subjects using RR interval time series. Compared to $SD2$, $SD1$ and $CCM$ values were found to be higher in CHF subjects. This findings might indicate that the short term variation in HRV is higher in CHF subjects, however, the long term variation is reduced. Since $CCM$ captures the signal dynamics at short level (i.e, 3 points of the plot), it appears to be affected by short term variation of the signal in CHF subjects. In the case of recognition of CHF subjects, although $SD2$ showed good result $CCM$ was found to be more significant as shown in table 1.

Above discussion indicates that $CCM$ is an additional descriptor of Poincaré plot with $SD1$ and $SD2$. This also implies that $CCM$ is a more consistent descriptor compared to $SD1$ and $SD2$. Our primary motivation for detecting pathology with a novel descriptor like $CCM$ rather than by observing visual pattern is achieved as shown by the case studies described. In this study, we have not looked at the physiological interpretation of $CCM$ which remain to be studied in future.

The proposed Complex Correlation Measure is based on the limitation of standard descriptors $SD1$ and $SD2$. The analysis carried out confirms the hypothesis that $CCM$ measures the temporal variation of the Poincaré plot. In contrast to the standard descriptors, $CCM$ evaluates point-to-point variation of the signal rather than gross variability of the signal. Besides the mathematical definition of $CCM$ and analyzing properties of the measure, we have also evaluated the performance of $CCM$ using real world case studies. In future, $CCM$ may be used as an efficient feature for pathology detection.

## References


[1] Liebovitch LS, Scheurle D. Two lessons from fractals and chaos. Complexity 2000;5:34–43.

[2] Tulppo MP, Makikallio TH, Takala TES, Seppanen T, V. HH. Quantitative beat-to-beat analysis of heart rate dynamics during exercise. Am J Physiol 1996;271:H244–H252.

[3] Brennan M, Palaniswami M, Kamen P. Do existing measures of poincare plot geometry reflect nonlinear features of heart rate variability. IEEE Trans on Biomed Engg 2001; 48:1342–1347.

[4] Acharya UR, Joseph KP, Kannathal N, Lim CM, Suri JS. Heart rate variability: a review. Medical and Biological Engineering and Computing 2006;44(12):1031–1051.

[5] Tulppo MP, Makikallio TH, Seppanen T, Airaksinen JKE, V. HH. Heart rate dynamics during accentuated sympathovagal interaction. Am J Physiol 1998;247:H810–H816.

[6] Toichi M, Sugiura T, Murai T, Sengoku A. A new method of assessing cardiac autonomic function and its comparison with spectral analysis and coefficient of variation of r-r interval. J Auton Nerv Syst 1997;62:79–84.

[7] Hayano J, Takahashi H, Toriyama T, Mukai S, Okada A, Sakata S, Yamada A, Ohte N, Kawahara H. Prognostic value of heart rate variability during long-term follow-up in chronic haemodialysis patiens with end-stage renal disease. Nephrol Dial Transplant 1999;14:1480–1488.

[8] Thakre TP, Smith ML. Loss of lag-response curvilinearity of indices of heart rate variability in congestive heart failure. BMC Cardiovascular Disorders 2006;6(27).

[9] Lerma C, Infante O, Perez-Grovas H, Jose MV. Poincare plot indexes of heart rate variability capture dynamic adaptations after haemodialysis in chronic renal failure patients. Clin Physiol Funct Imaging 2003;23(2):72–80.

[10] Brennan M, Palaniswami M, Kamen P. Poicare plot interpretation using a physiological model of hrv based on a network of oscillators. Am J Physiol Heart Circ Physiol 2002; 283:1873–1886.

[11] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. Circulation 2000 (June 13);101(23):e215–e220. Circulation Electronic Pages: http://circ.ahajournals.org/cgi/content/full/101/23/e215.

[12] Hanley JA, McNeil BJ. The meaning and use of the area under receiver operating characteristic (roc) curve. Radiology 1982;143:29–36.



Address for correspondence:

Chandan Karmakar
Dept. of EEE, The University of Melbourne
Parkville, VIC 3010, Australia
c.karmakar@ee.unimelb.edu.au