Evaluating complexity of short-term heart period variability through predictability techniques

Alberto Porta

Department of Biomedical Sciences for Health
Galeazzi Orthopedic Institute
University of Milan
Milan, Italy

Introduction

There is an increasing interest in evaluating short term complexity of heart period variability in humans mainly due to its relationship with cardiac neural regulation, pathology and aging

Traditional approaches quantify complexity in terms of information carried by the samples (i.e. entropy-based approaches)

However, complexity can be estimated in terms of predictability of future samples when a certain amount of previous values are given (the smaller predictability, the larger complexity)

Primary aims

To propose tools assessing complexity of heart period variability via predictability-based approaches

To demonstrate that this approach is strongly linked to the methods based on conditional entropy

Secondary aims

To show that complexity analysis of heart period variability is helpful to distinguish healthy subjects from pathological patients

To demonstrate that complexity analysis of heart period variability can be fully exploited under uncontrolled experimental conditions and during daily activities

Outline

1) Predictability approach based on conditional distribution and uniform quantization

2) Conditional entropy approach based on uniform quantization

3) Predictability approach based on conditional distribution and k nearest neighbors

4) Conditional entropy approach based on k nearest neighbors

5) Application to 24h Holter recordings of heart period variability obtained from healthy subjects and chronic heart failure population

Outline

1) Predictability approach based on conditional distribution and uniform quantization

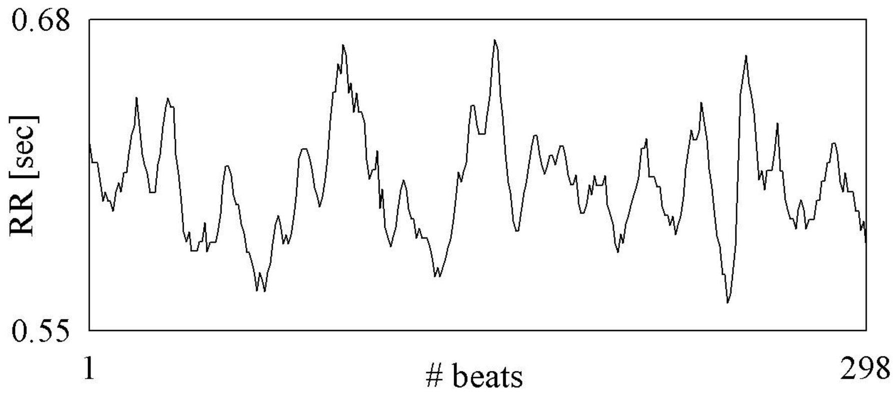
2) Conditional entropy approach based on uniform quantization

3) Predictability approach based on conditional distribution and k nearest neighbors

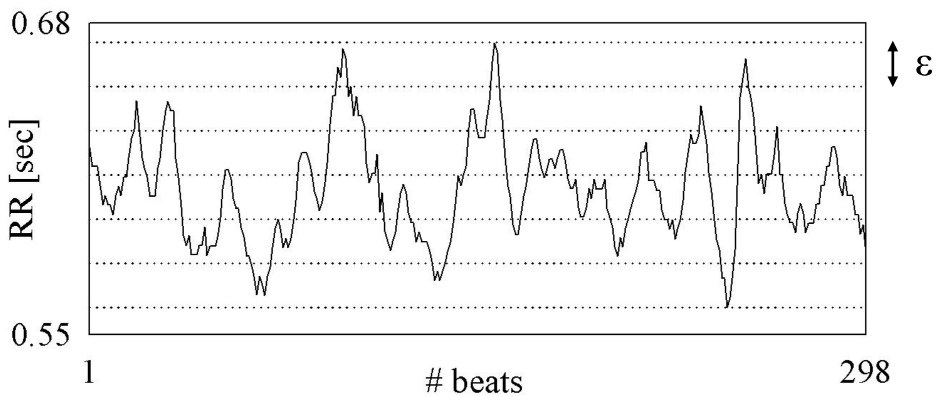
4) Conditional entropy approach based on k nearest neighbors

5) Application to 24h Holter recordings of heart period variability obtained from healthy subjects and chronic heart failure population

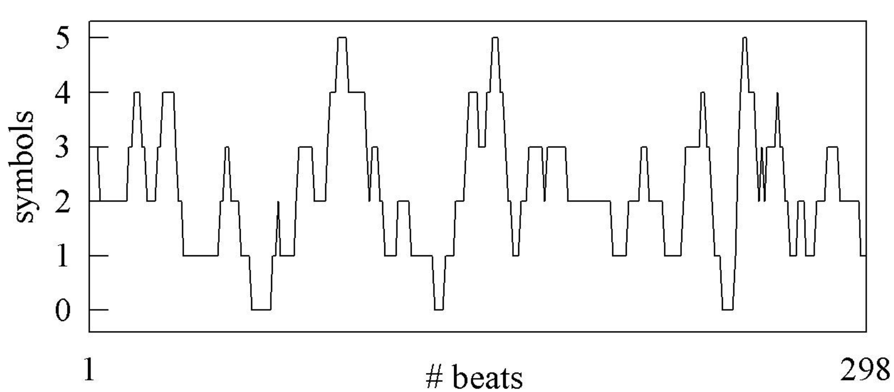
Uniform quantization



$\{RR(i), i=1,\ldots,N\}$ with $RR(i) \in R$

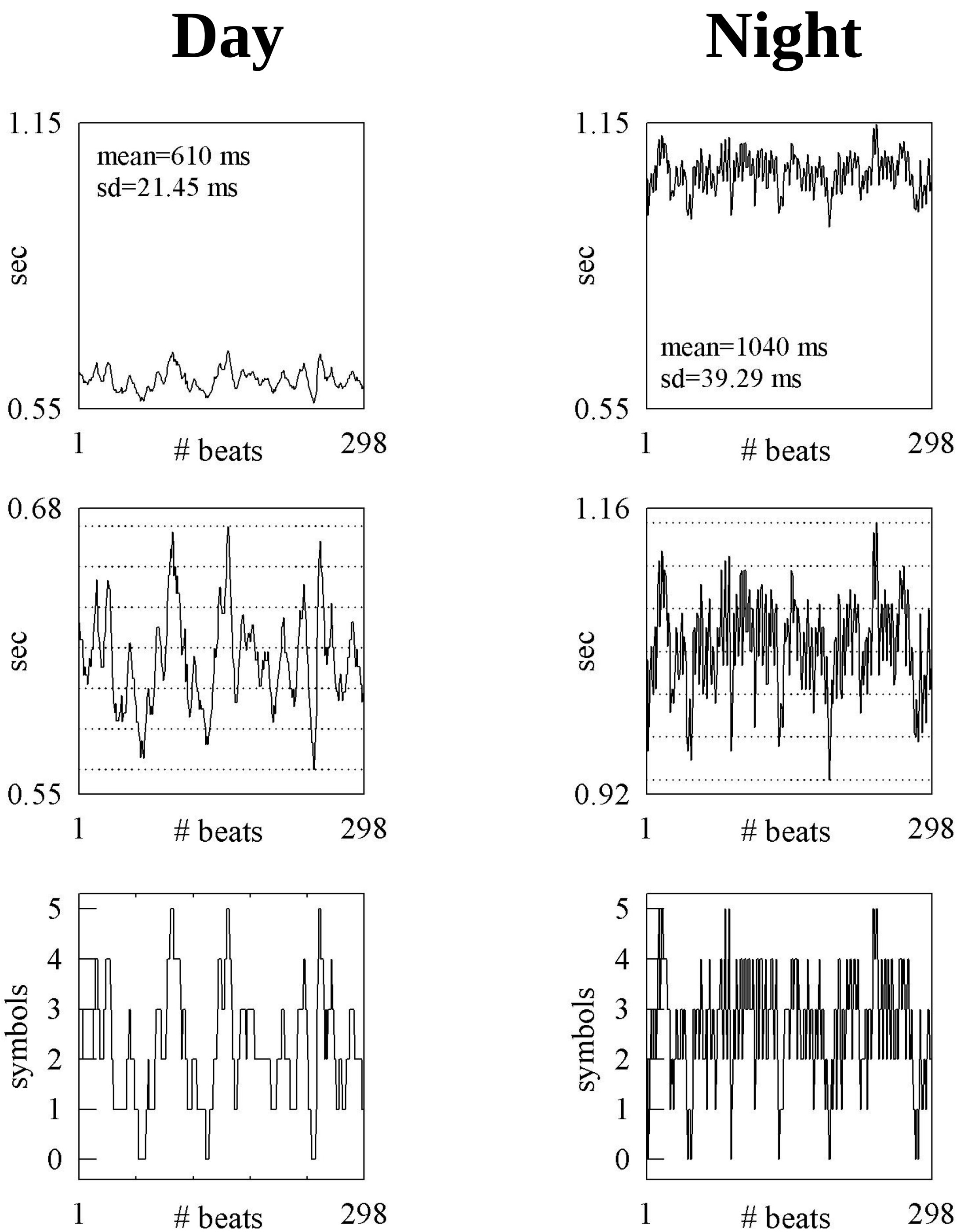


$$q=6 \text{ with } \varepsilon = \frac{\max(RR)-\min(RR)}{q}$$



$\{RR_q(i), i=1,\ldots,N\}$ with $RR_q(i) \in I$

$0 \le RR_q(i) \le q-1$


Day
Night
1.15
0.55
mean=610 ms
sd=21.45 ms
sec
1
beats
298
1.15
0.55
mean=1040 ms
sd=39.29 ms
sec
1
beats
298
0.68
0.55
sec
1
beats
298
1.16
0.92
sec
1
beats
298
5
4
3
2
1
0
symbols
1
beats
298
5
4
3
2
1
0
symbols
1
beats
298

Pattern construction

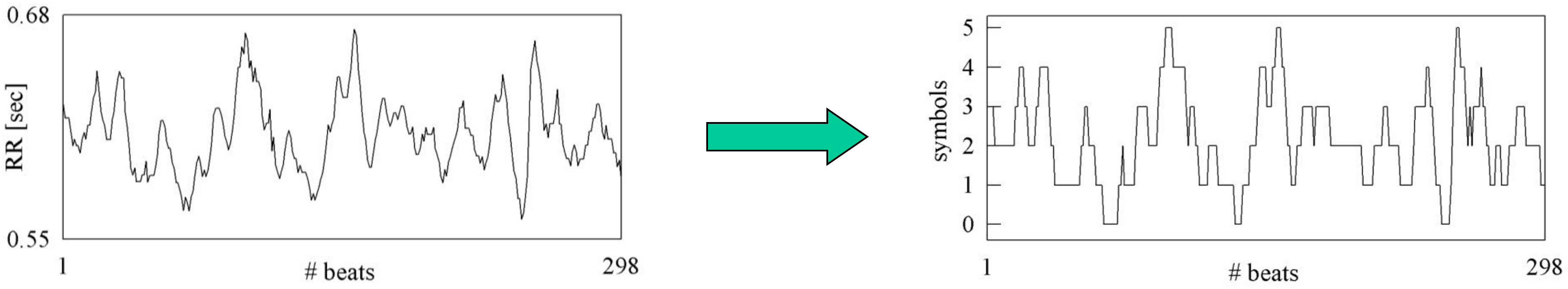
$$f: \ \{RR_q(i),\ i=1,\ldots,N\} \xrightarrow{f} \{RR_{q,L}(i),\ i=1,\ldots,N-L+1\}$$

with $RR_{q,L}(i) = (RR_q(i), RR_q(i-\tau), \ldots, RR_q(i-(L-1)\tau))$

$$0 \leq RR_q(i) \leq q-1$$

When $\tau=1$, $RR_{q,L}(i)$ is a feature extracted from the series

Example of pattern construction (L=3)



$$\{RR_q(i)\} = \{3, 3, 3, 3, 2, 2, 1, \ldots\}$$

$$(3,3,3)$$

$$(3,3,3)$$

$$(3,3,2)$$

$$(3,2,2)$$

$$(2,2,1)$$

$$\ldots$$

$$\{RR_{q,L}(i)\} = \{(3,3,3),(3,3,3),(3,3,2),(3,2,2),(2,2,2),\ldots\}$$

Transformation of a pattern into an integer

g: $RR_{q,L}(i)=(RR_q(i),RR_q(i-1),\ldots,RR_q(i-L+1)) \in I^L \xrightarrow{g} h_{q,L}(i) \in I$

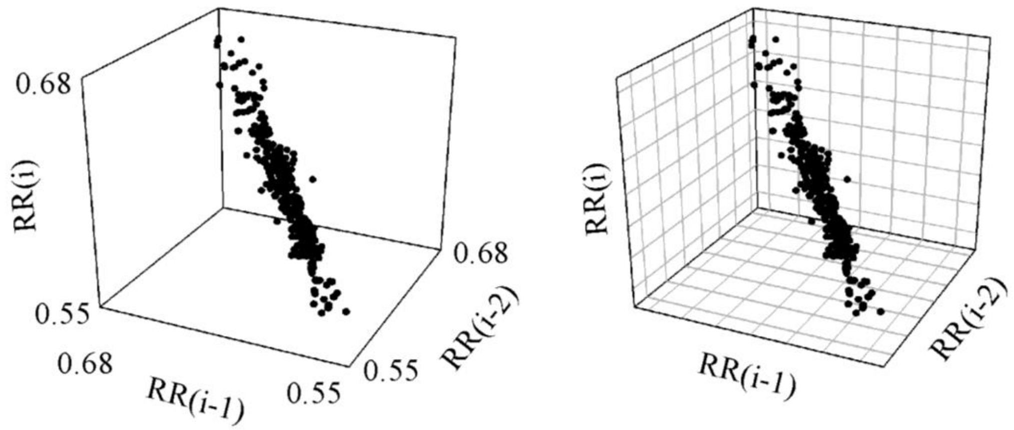
$$h_{q,L}(i) = RR_q(i)\cdot q^{L-1} + RR_q(i-1)\cdot q^{L-2} + \ldots + RR_q(i-L+1)\cdot q^0$$

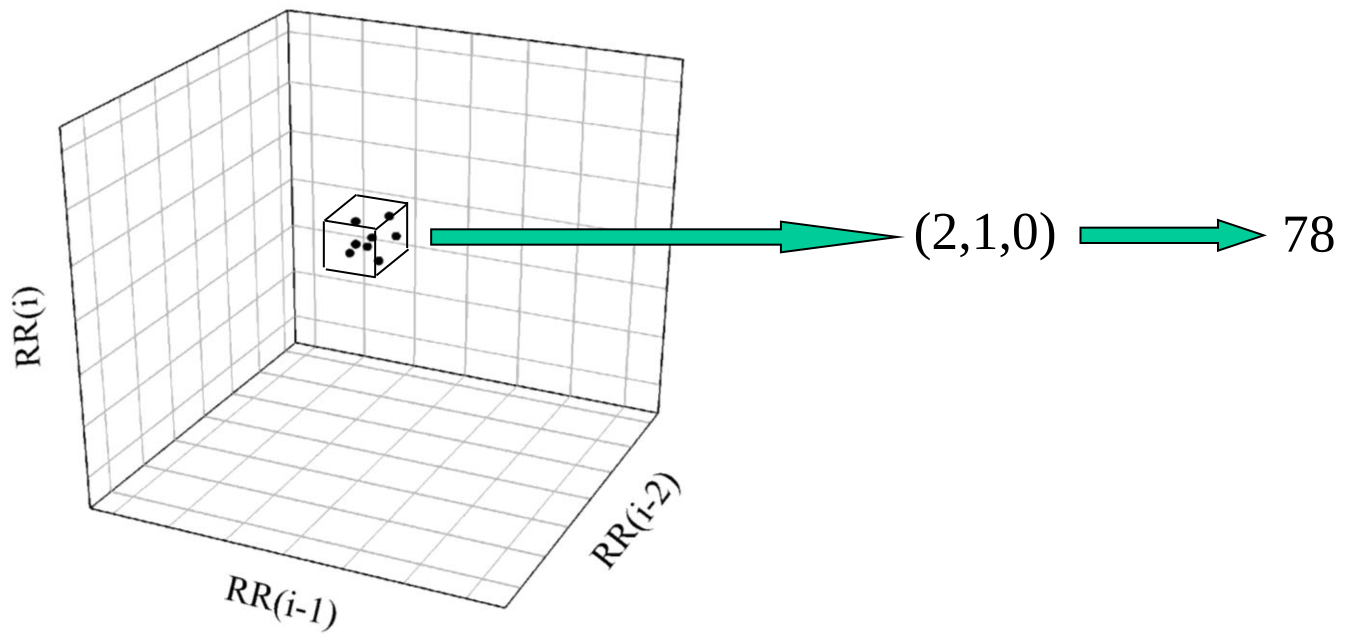
$$0 \leq h_{q,L}(i) \leq q^L-1$$

Example

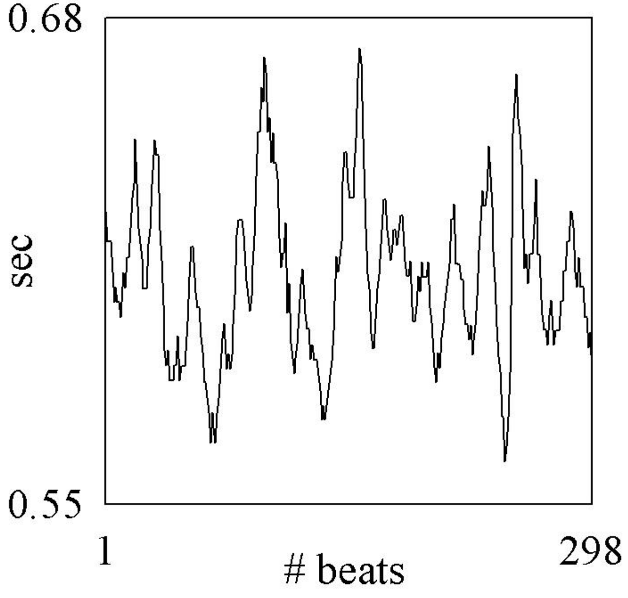
with L=3 and q=6 $\quad (2,0,5) \xrightarrow{g} 2\cdot 6^2+0\cdot 6^1+5\cdot 6^0 = 77$

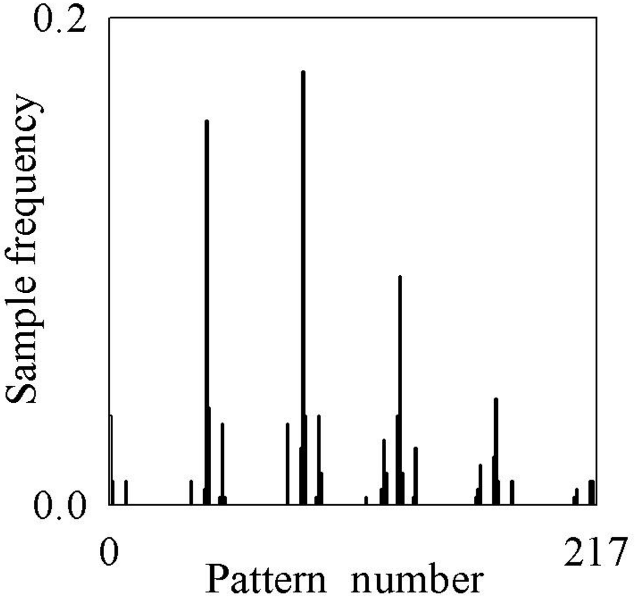
Uniform quantization in 3-dimensional embedding space





Example of pattern distribution in a 3-dimensional embedding space





Porta A et al, IEEE Trans Biomed Eng, 48:1282-1291, 2001

Toward the assessment of complexity based on prediction

Uniform quantization (in general any type of coarse graining) of the embedding space provides the basis for

1) entropy-based approaches

2) prediction techniques

Transforming any L-dimensional quantized pattern into a 2-dimensional one

$$RR_{q,L}(i) = (RR_q(i), RR_q(i-1), \ldots, RR_q(i-L+1)) = (RR_q(i), RR_{q,L-1}(i-1))$$

L-dimensional pattern $\quad\quad$ 2-dimensional pattern

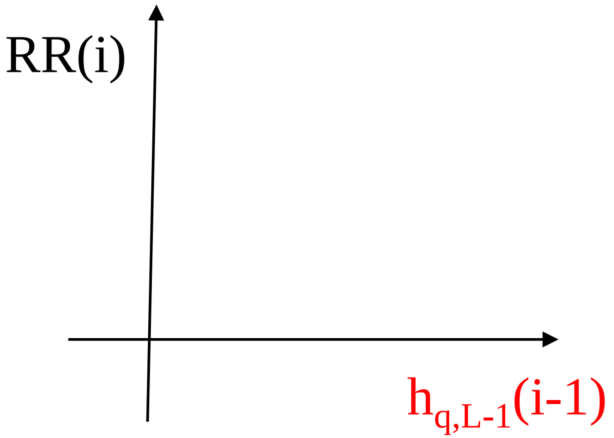
$$(RR_q(i), RR_{q,L-1}(i-1)) \longrightarrow (RR_q(i), h_{q,L-1}(i-1))$$

Conditional distribution of the current sample given L-1 previous values

Given the transformation

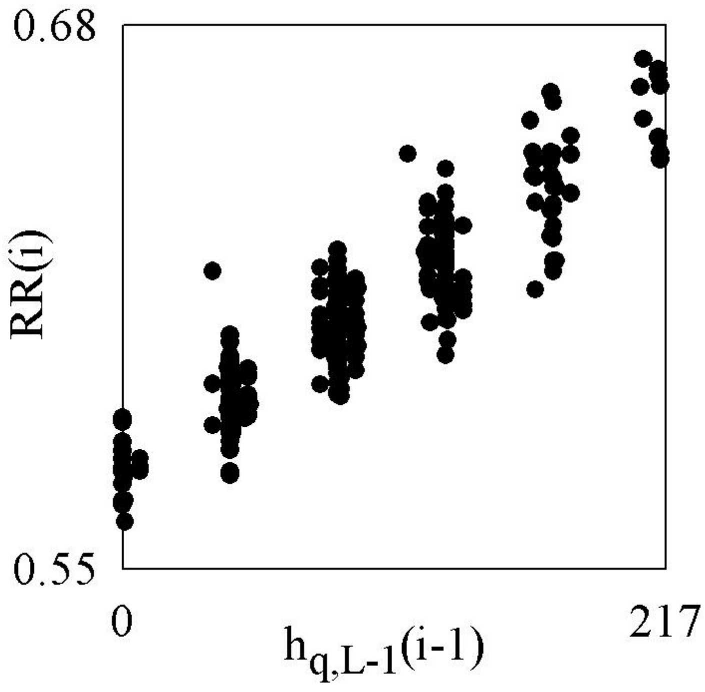
$$(RR_q(i), RR_{q,L-1}(i-1)) \longrightarrow (RR_q(i), h_{q,L-1}(i-1))$$

the conditional distribution of the current sample given L-1 previous values can be drawn in the plane

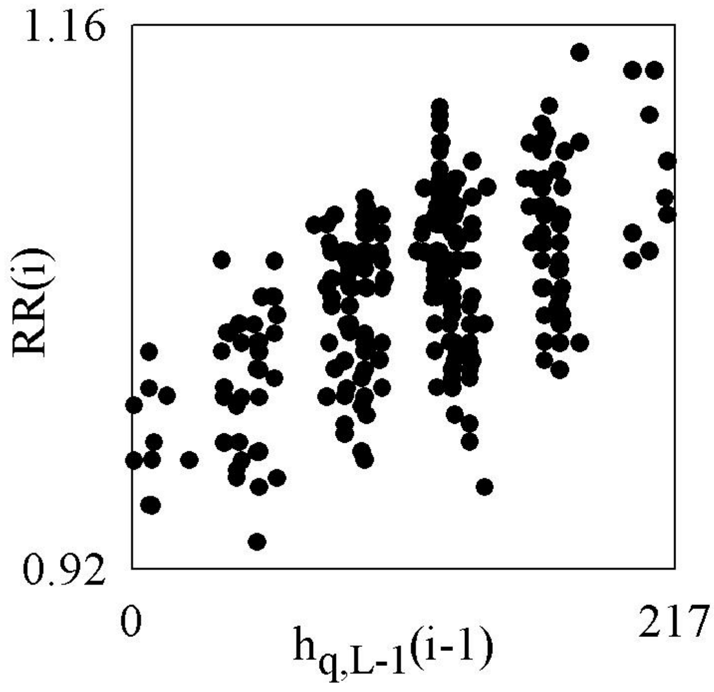


Examples of conditional distribution of the current heart period given three past RR intervals (L=3)

Day



Night



Porta A et al, Chaos, 17, 015117, 2007

Prediction based on conditional distribution: the uniform quantization (UQ) approach

Predictor

$$\hat{RR}(i/L-1) = \text{median}(RR(j)/RR_{q,L-1}(j-1) = RR_{q,L-1}(i-1))$$
$$= \text{median}(RR/h_{q,L-1}(i-1))$$

Defined the prediction error as

$$e(i) = RR(i) - \hat{RR}(i)$$

the mean square prediction error (MSPE) is

$$MSPE_{UQ}(L) = \frac{1}{N-L}\sum_{i=L}^{N} e^2(i) \qquad \text{with } 0 \leq MSPE_{UQ}(L) \leq MSD$$

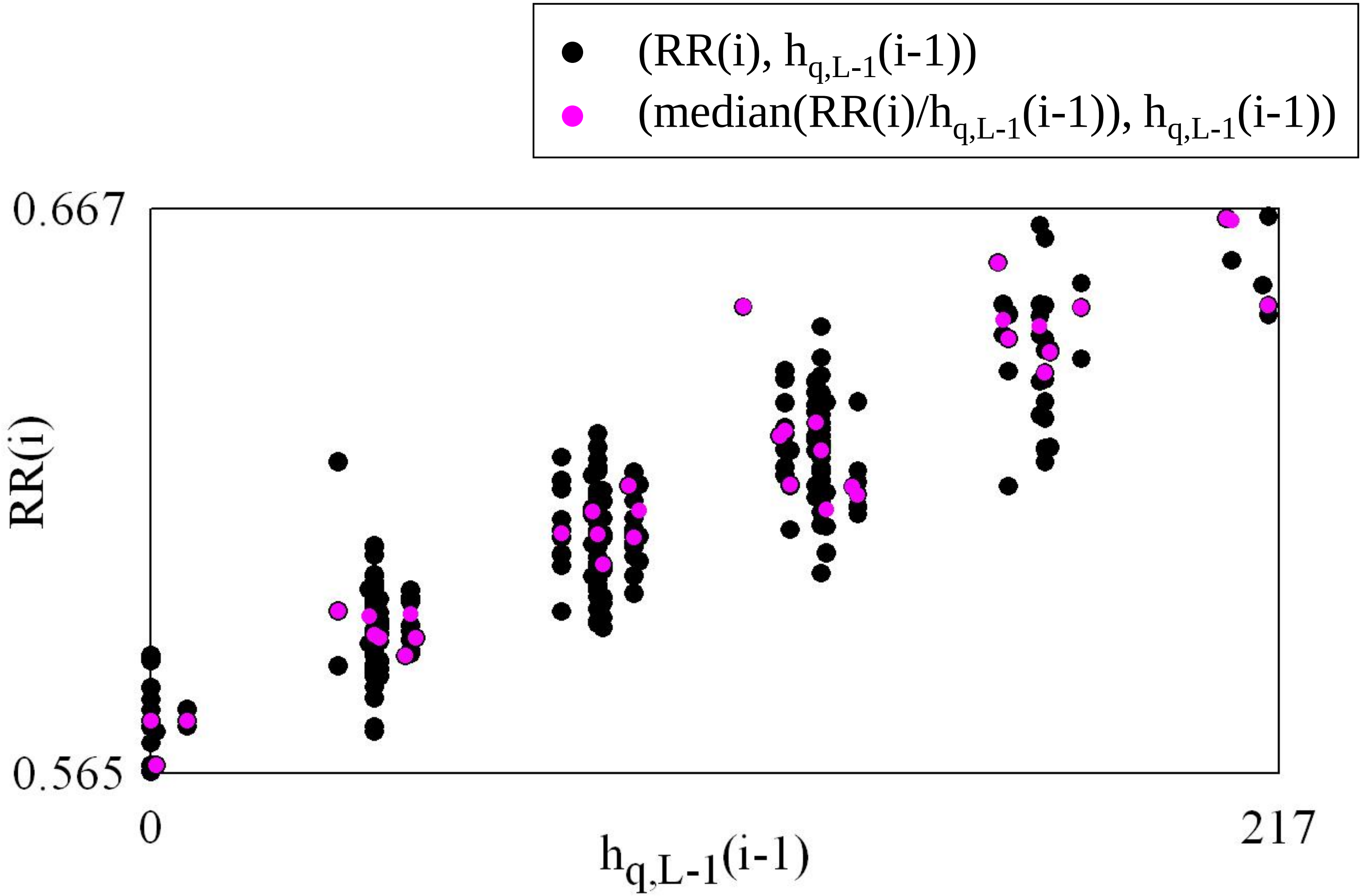
where $MSD_{UQ} = \frac{1}{N-1}\sum_{i=1}^{N}(RR(i)-RR_m)^2$ and RR_m= median(RR)

$MSPE_{UQ}(L) = 0 \Rightarrow$ perfect prediction

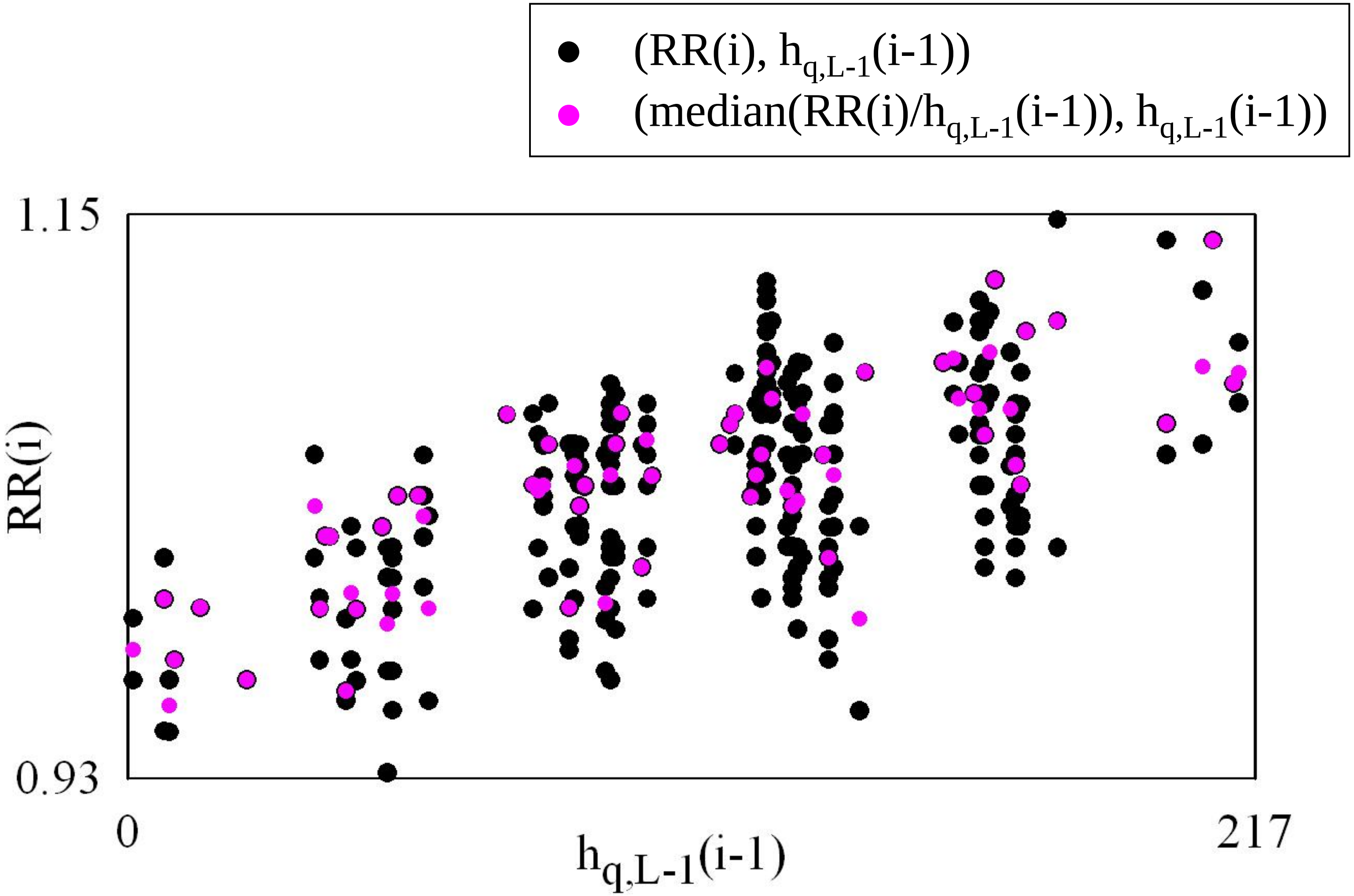
$MSPE_{UQ}(L) = MSD_{UQ} \Rightarrow$ null prediction

Porta A et al, IEEE Trans Biomed Eng, 47, 1555-1564, 2000

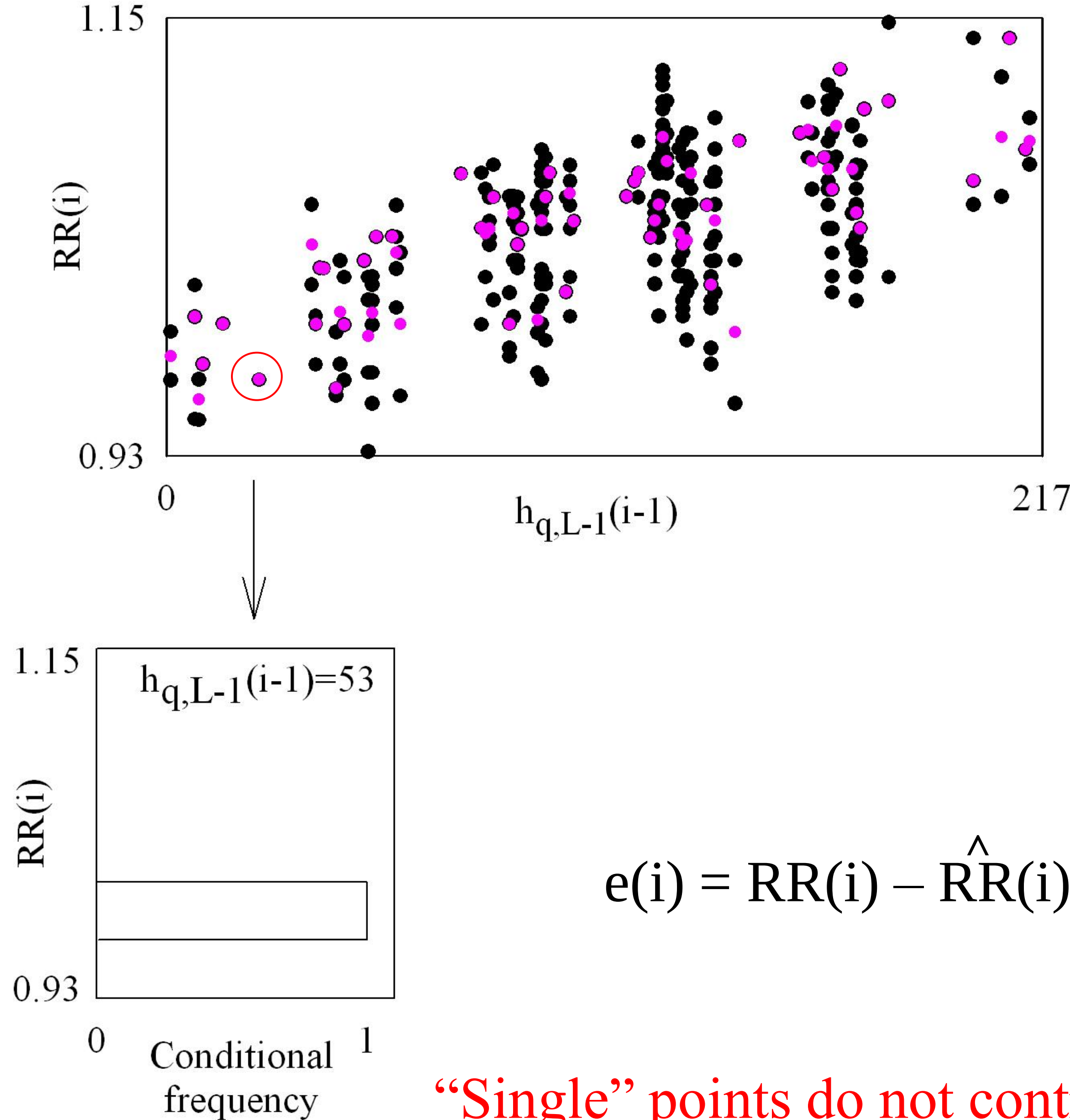
Examples of prediction based on conditional distribution with L=3 during daytime



Examples of prediction based on conditional distribution with L=3 during nighttime



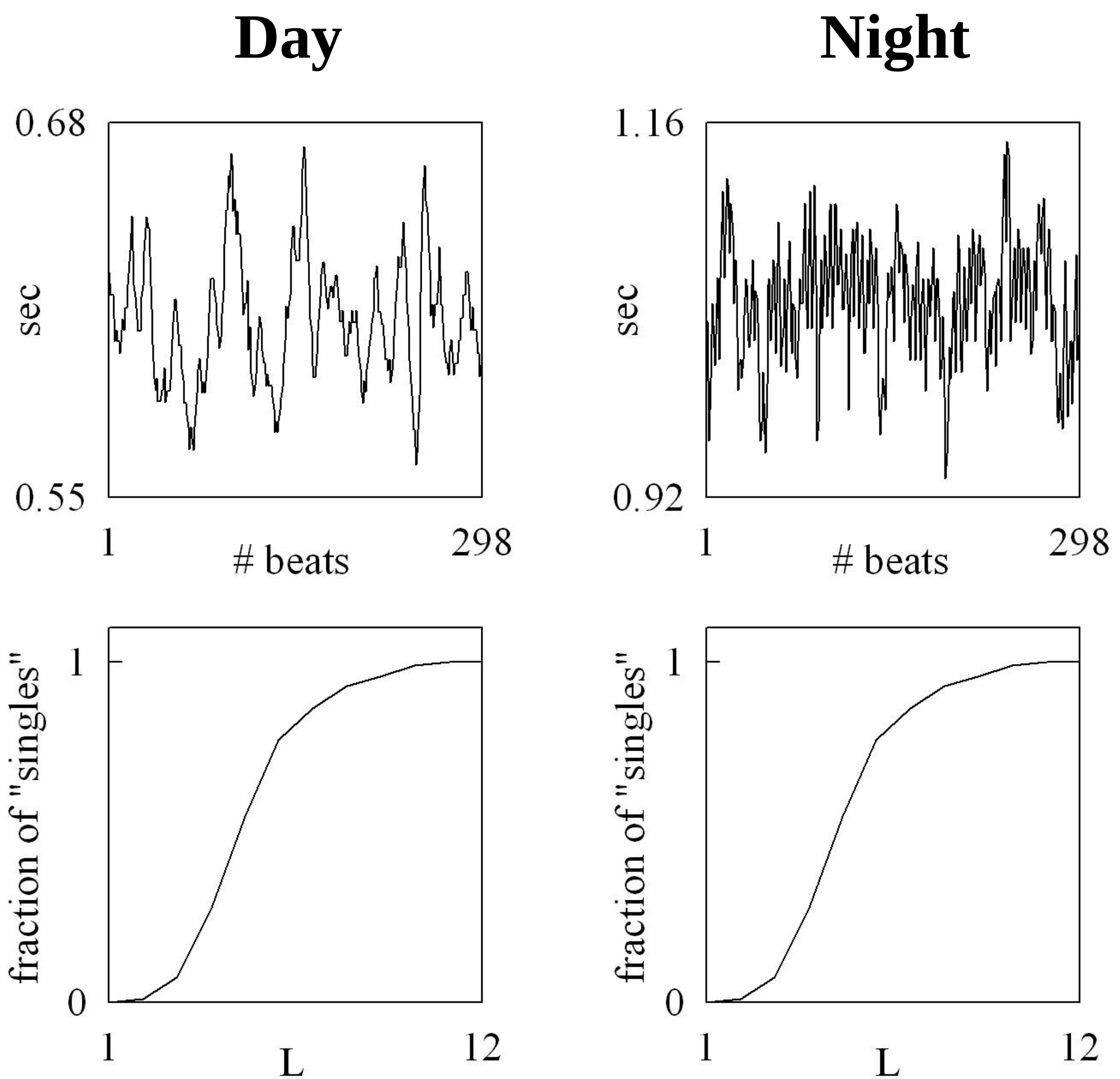
Overfitting



$$e(i) = RR(i) - \hat{R}R(i) = 0$$

"Single" points do not contribute to MSPE

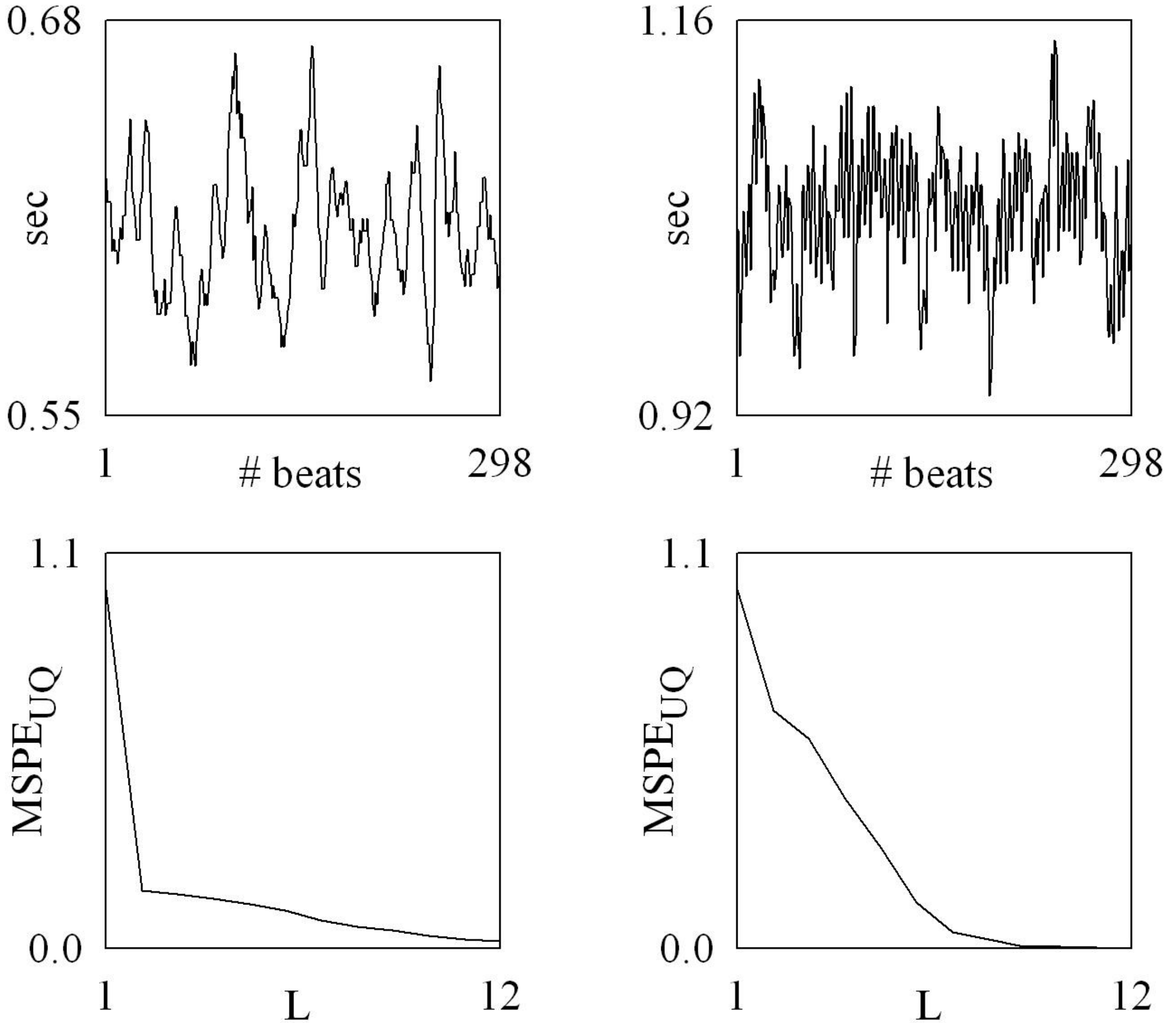
Course of single patterns with pattern length



Fraction of “singles” ⟶ 1 with L

Mean square prediction error

Day **Night**



$MSPE_{UQ}(L) \longrightarrow 0$ with L

Corrected mean square prediction error ($CMSPE_{UQ}$) and normalized $CMSPE_{UQ}$ ($NCMSPE_{UQ}$)

$$CMSPE_{UQ}(L) = MSPE_{UQ}(L) + MSD \cdot fraction(L)$$

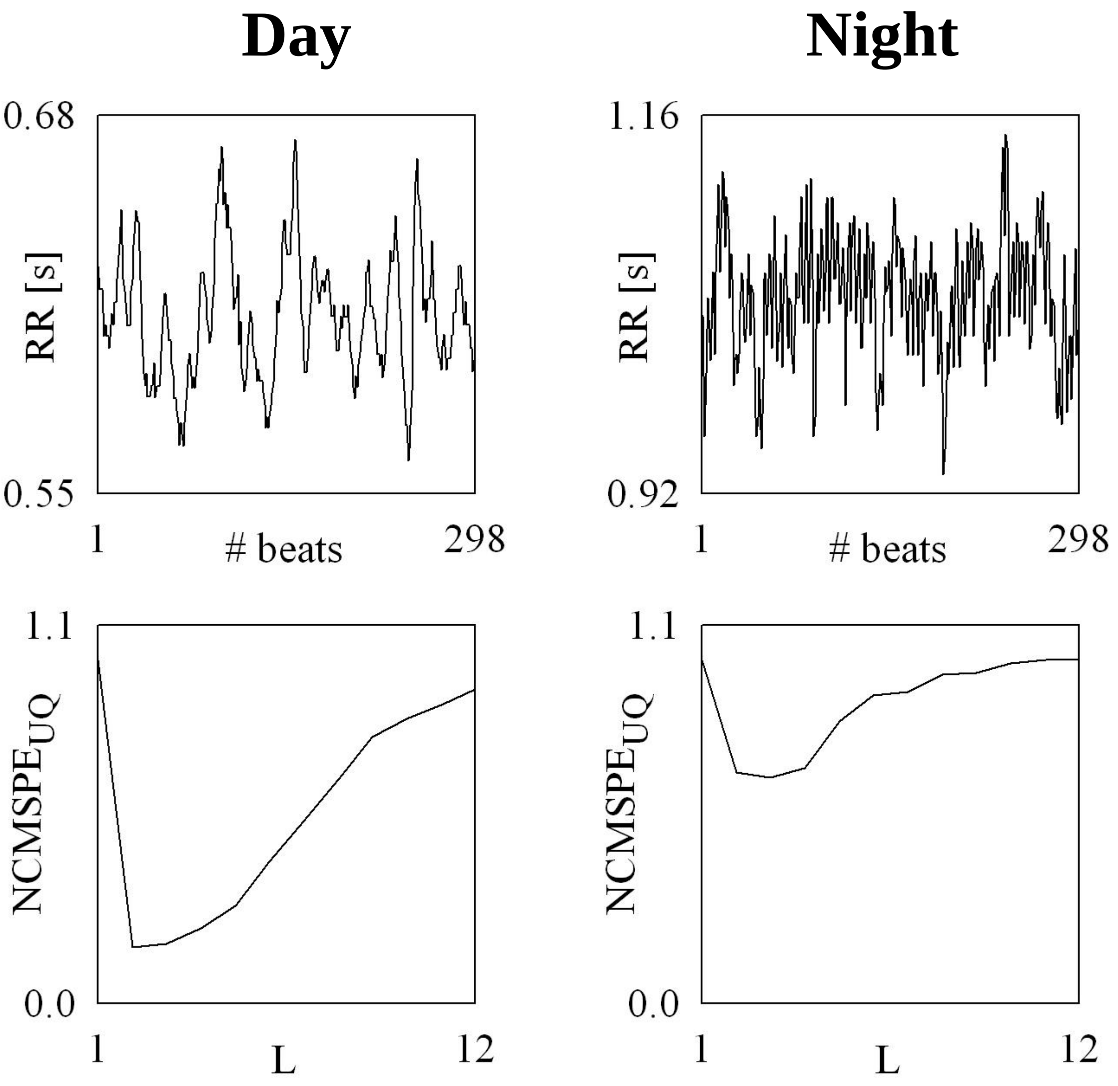
with $0 \le CMSPE_{UQ}(L) \le MSD$

$$NCMSPE_{UQ}(L) = \frac{CMSPE_{UQ}(L)}{MSD}$$

$0 \le NCMSPE_{UQ}(L) \le 1$

Porta A et al, IEEE Trans Biomed Eng, 47, 1555-1564, 2000

Normalized unpredictability index ($NUPI_{UQ}$)



$$NUPI_{UQ}=\min(NCMSPE_{UQ}(L))$$

$$0 \le NUPI_{UQ} \le 1$$

Outline

1) Predictability approach based on conditional distribution and uniform quantization

2) Conditional entropy approach based on uniform quantization

3) Predictability approach based on conditional distribution and k nearest neighbors

4) Conditional entropy approach based on k nearest neighbors

5) Application to 24h Holter recordings of heart period variability obtained from healthy subjects and chronic heart failure population

Conditional entropy

$$CE(L) = -\Sigma p(RR_{q,L-1}(i-1)) \cdot SE(RR_q/RR_{q,L-1}(i-1))$$

where

$$SE(RR/RR_{q,L-1}(i-1)) = \Sigma p(RR_q(i)/RR_{q,L-1}(i-1)) \cdot \log(RR_q(i)/RR_{q,L-1}(i-1)))$$

with $0 \leq CE(L) \leq SE(RR)$

and $SE(RR) = -\Sigma p(RR_q(i)) \cdot \log(p(RR_q(i)))$

Conditional entropy

Given the transformation g: $RR_{q,L-1}(i) \xrightarrow{g} h_{q,L-1}(i)$

$$CE(L) = -\Sigma p(h_{q,L-1}(i-1)) \cdot SE(RR_q/h_{q,L-1}(i-1))$$

where

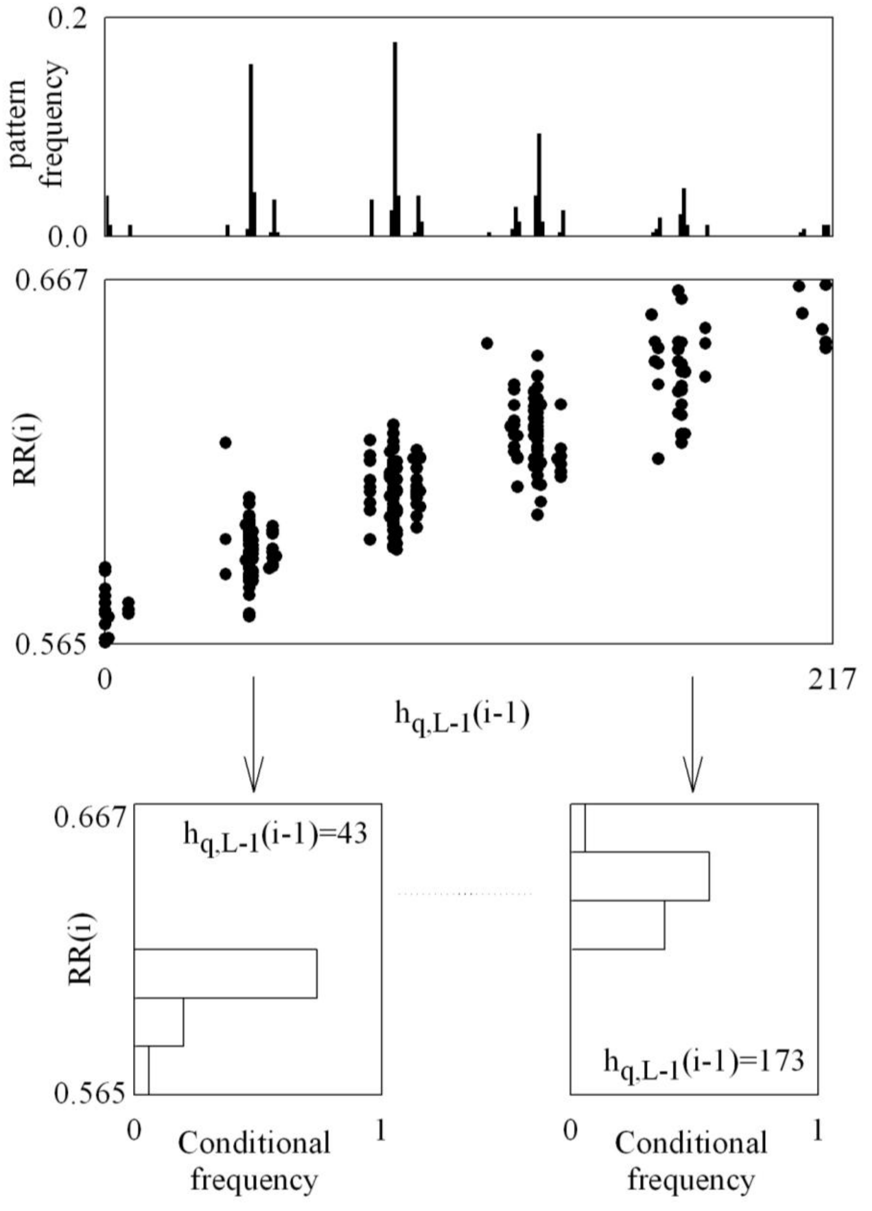
$$SE(RR/h_{q,L-1}(i-1)) = \Sigma p(RR_q(i)/h_{q,L-1}(i-1)) \cdot \log(RR_q(i)/h_{q,L-1}(i-1)))$$

with $0 \leq CE(L) \leq SE(RR)$

and $SE(RR) = -\Sigma p(RR_q(i)) \cdot \log(p(RR_q(i)))$

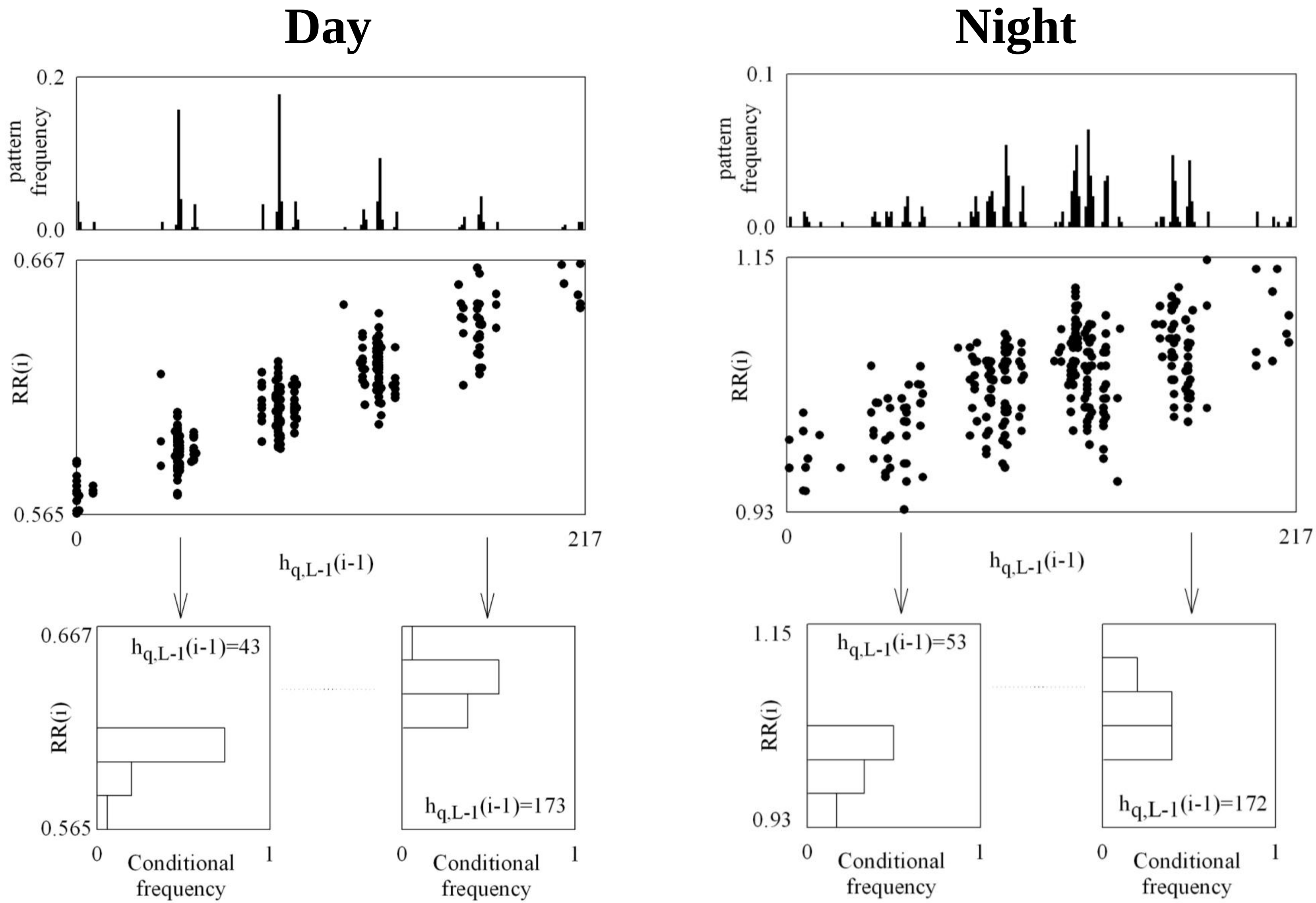
Porta A et al , Biol Cybern, 78:71-78, 1998
Porta A et al , Med Biol Eng Comput, 38, 180-188, 2000

Example of calculation of conditional entropy (L=4)



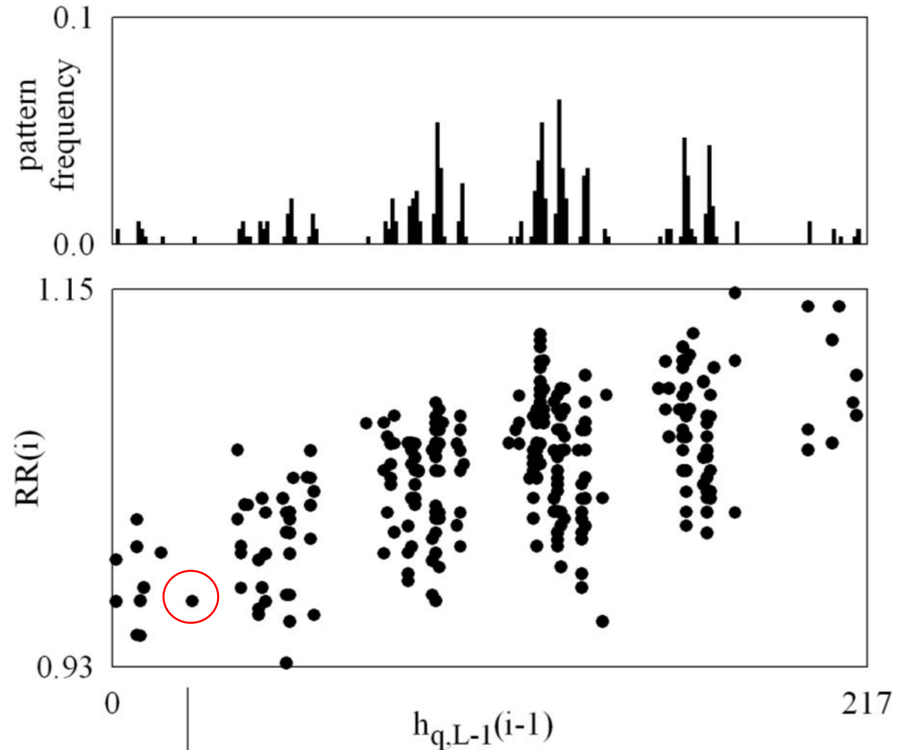
Porta A et al, Chaos, 17, 015117 2007

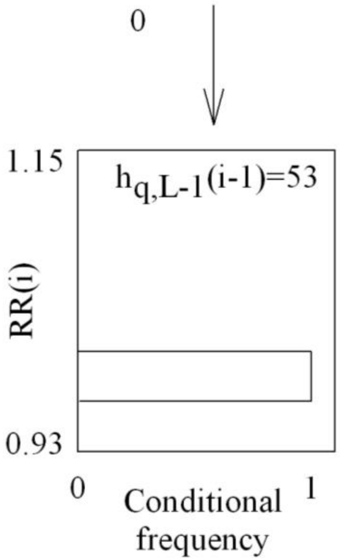
Example of calculation of conditional entropy (L=4) during daytime and nighttime



CE(L=4) during daytime < CE(L=4) during nighttime

Bias of conditional entropy (L=4)

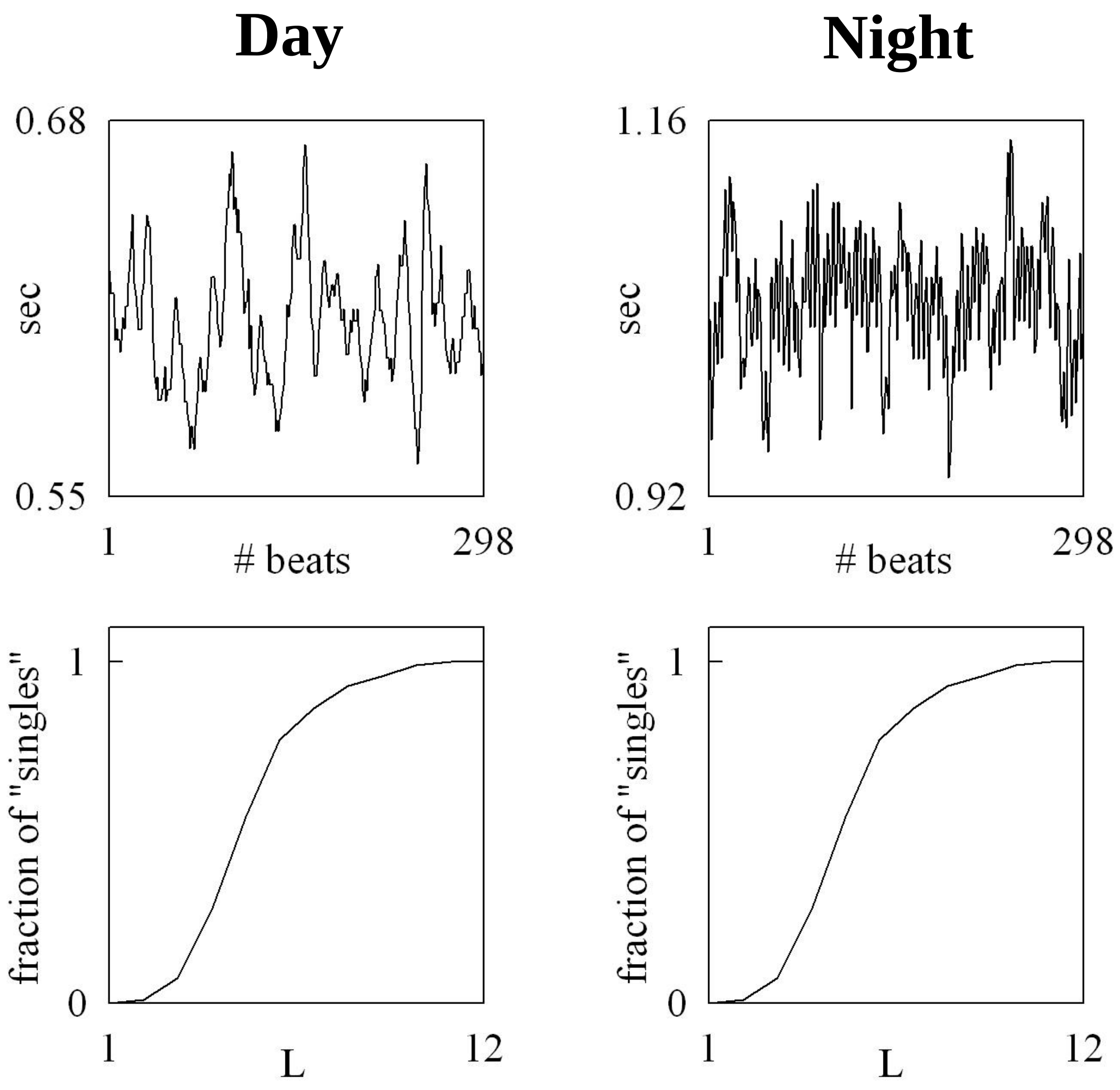




$SE(RR_q/h_{q,L-1}(i-1))=0$

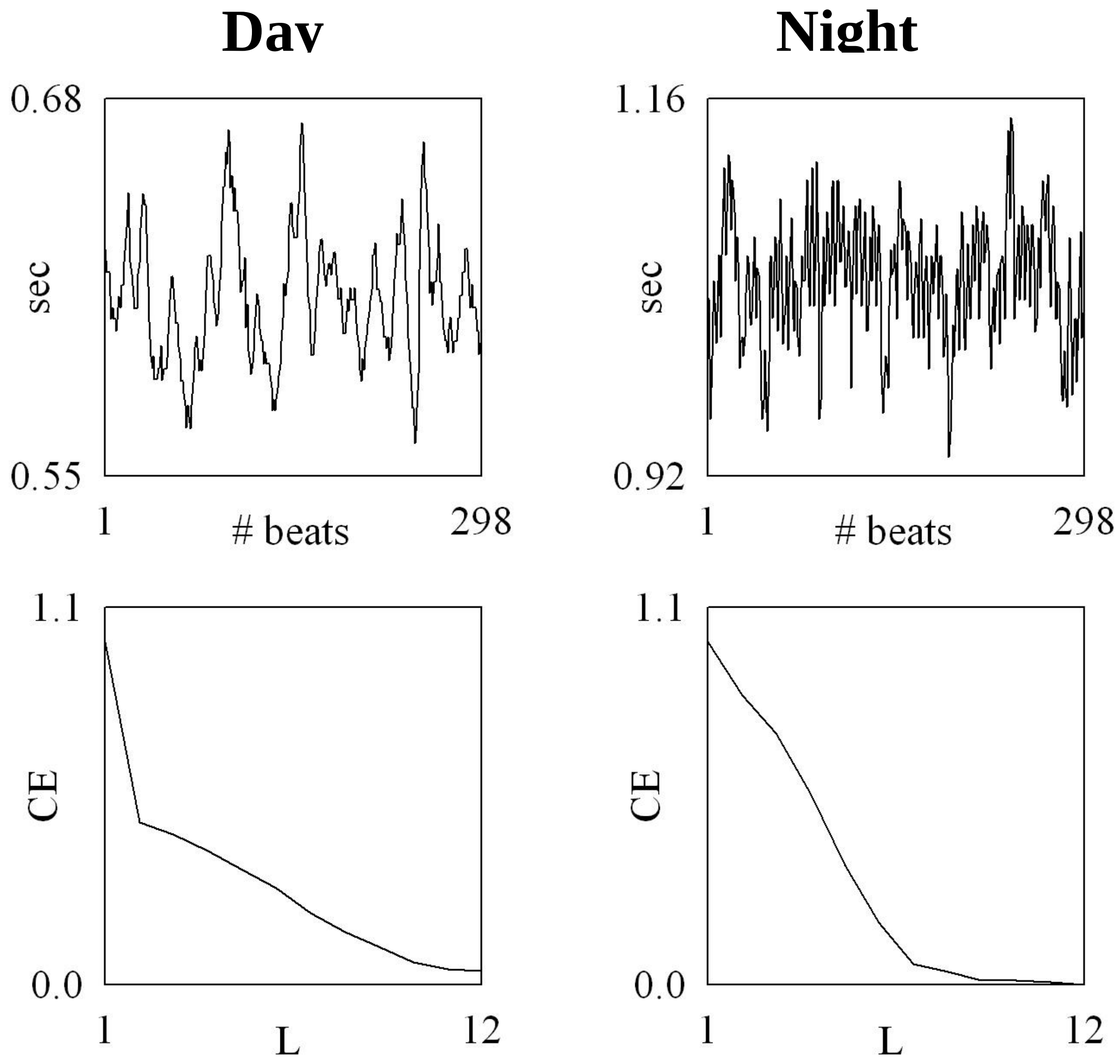
"Single" points do not contribute to CE

Course of single patterns with pattern length



Fraction of “singles” ⟶ 1 with L

Conditional entropy



CE(L) $\longrightarrow$ 0 with L

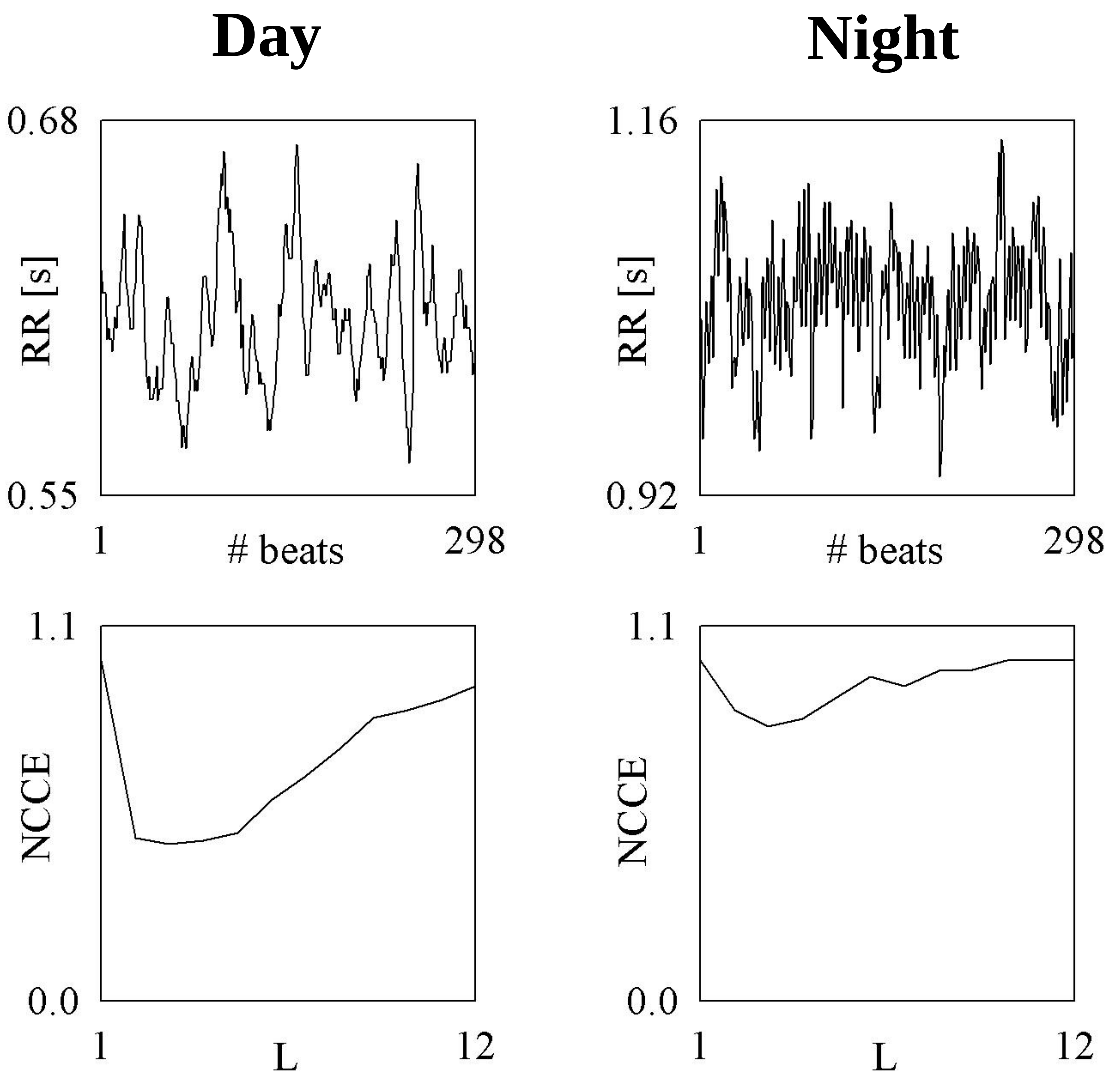
Corrected conditional entropy (CCE) and normalized CCE (NCCE)

$$\mathrm{CCE(L) = CE(L) + SE(L{=}1)\cdot fraction(L)} \qquad 0 \le \mathrm{CCE(L)} \le \mathrm{SE(L{=}1)}$$

$$\mathrm{NCCE(L)} = \frac{\mathrm{CCE(L)}}{\mathrm{SE(L{=}1)}} \qquad 0 \le \mathrm{NCCE(L)} \le 1$$

Porta A et al , Biol Cybern, 78:71-78, 1998
Porta A et al , Med Biol Eng Comput, 38, 180-188, 2000

Normalized complexity index (NCI_{UQ})



NCI_{UQ}=min(NCCE(L)) with $0 \leq NCI \leq 1$

Outline

1) Predictability approach based on conditional distribution and uniform quantization

2) Conditional entropy approach based on uniform quantization

3) Predictability approach based on conditional distribution and k nearest neighbors

4) Conditional entropy approach based on k nearest neighbors

5) Application to 24h Holter recordings of heart period variability obtained from healthy subjects and chronic heart failure population

Prediction based on conditional distribution: the k-nearest-neighbor (KNN) approach

Predictor

$\hat{RR}(i/L\text{-}1)$ = median($RR(j)/RR_{L-1}(j-1)$ belongs to the set of the k nearest neighbors of $RR_{L-1}(i-1)$)

Defined the prediction error as

$$e(i) = RR(i) - \hat{RR}(i)$$

the mean square prediction error ($MSPE_{KNN}$) is

$$MSPE_{KNN}(L) = \frac{1}{N-L}\sum_{i=L}^{N} e^2(i) \quad \text{with } 0 \le MSPE_{KNN}(L) \le MSD$$

where $MSD = \frac{1}{N-1}\sum_{i=1}^{N}(RR(i)-RR_m)^2$ and RR_m= median(RR)

$MSPE_{KNN}(L) = 0 \Rightarrow$ perfect prediction

$MSPE_{KNN}(L) = MSD \Rightarrow$ null prediction

A. Porta et al, IEEE Trans Biomed Eng, 54:94-106 , 2007

Normalized k-nearest-neighbor mean square prediction error (NKNNMSPE)

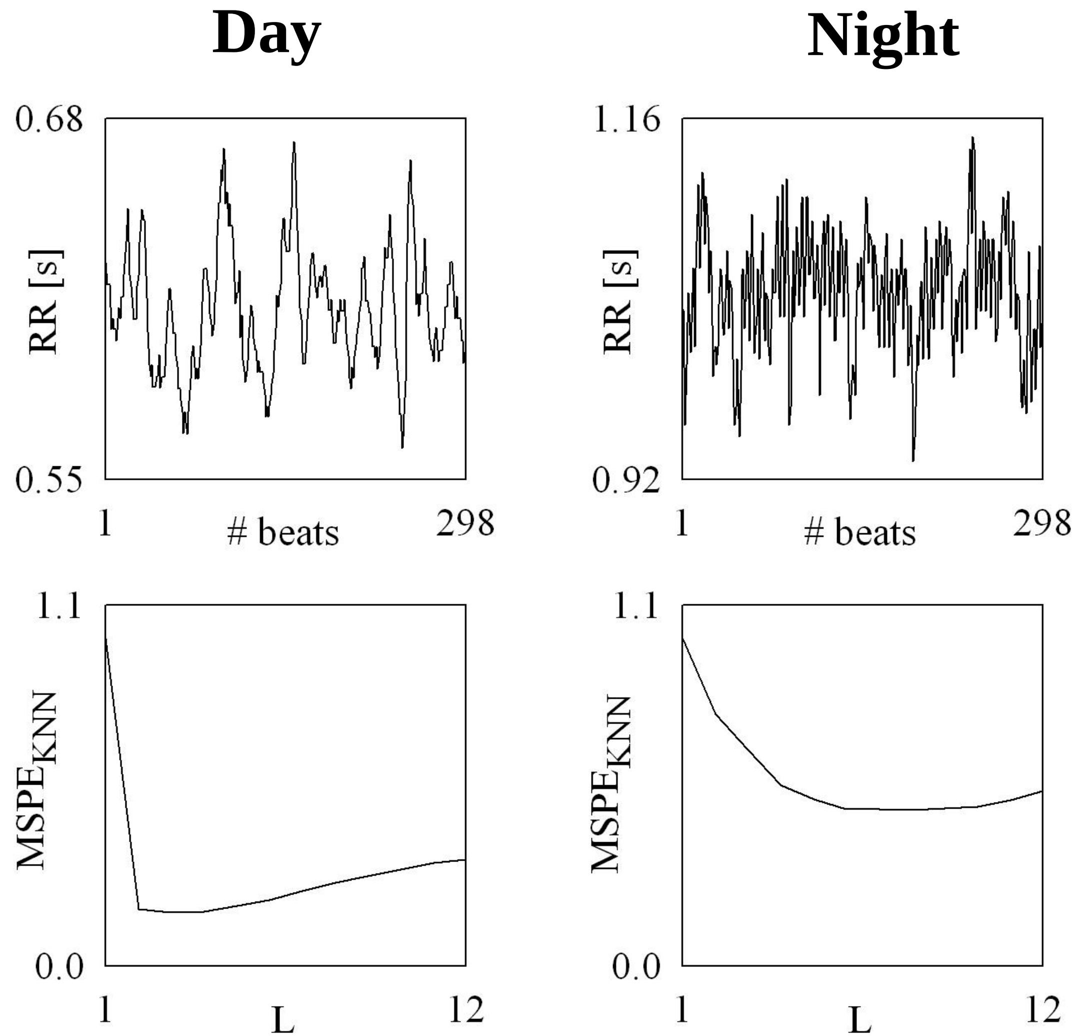
$$\mathrm{NMSPE}_{\mathrm{KNN}}(L) = \frac{\mathrm{MSPE}_{\mathrm{KNN}}(L)}{\mathrm{MSD}}$$

with $0 \leq \mathrm{NMSPE}_{\mathrm{KNN}}(L) \leq 1$

$\mathrm{NMSPE}_{\mathrm{KNN}}(L) = 0$ ⇨ perfect prediction

$\mathrm{NMSPE}_{\mathrm{KNN}}(L) = 1$ ⇨ null prediction

Normalized unpredictability index based on k-nearest-neighbor approach



$NUPI_{KNN}=\min(NMSPE_{KNN}(L))$ with $0 \leq NUPI_{KNN} \leq 1$

A. Porta et al, IEEE Trans Biomed Eng, 54:94-106 , 2007

Outline

1) Predictability approach based on conditional distribution and uniform quantization

2) Conditional entropy approach based on uniform quantization

3) Predictability approach based on conditional distribution and k nearest neighbors

4) Conditional entropy approach based on k nearest neighbors

5) Application to 24h Holter recordings of heart period variability obtained from healthy subjects and chronic heart failure population

K-nearest-neighbor conditional entropy (KNNCE)

$$\mathrm{KNNCE}(L) = \frac{1}{N-L+1}\sum_{i=L}^{N}\mathrm{SE}(RR/RR_{L-1}(i-1))$$

where

$\mathrm{SE}(RR/RR_{L-1}(i-1))$ is the Shannon entropy of conditional distribution of $RR(j)$ given that $RR_{L-1}(j-1)$ belongs to the set of k-nearest-neighbors of $RR_{L-1}(i-1)$

with $0 \leq \mathrm{KNNCE}(L) \leq \mathrm{SE}(RR)$

and $\mathrm{SE}(RR) = -\sum p(RR(i)) \cdot \log(p(RR(i)))$

Porta A et al, Physiol Meas, 34:17-33, 2013

Normalized k-nearest-neighbor conditional entropy (KNNCE)

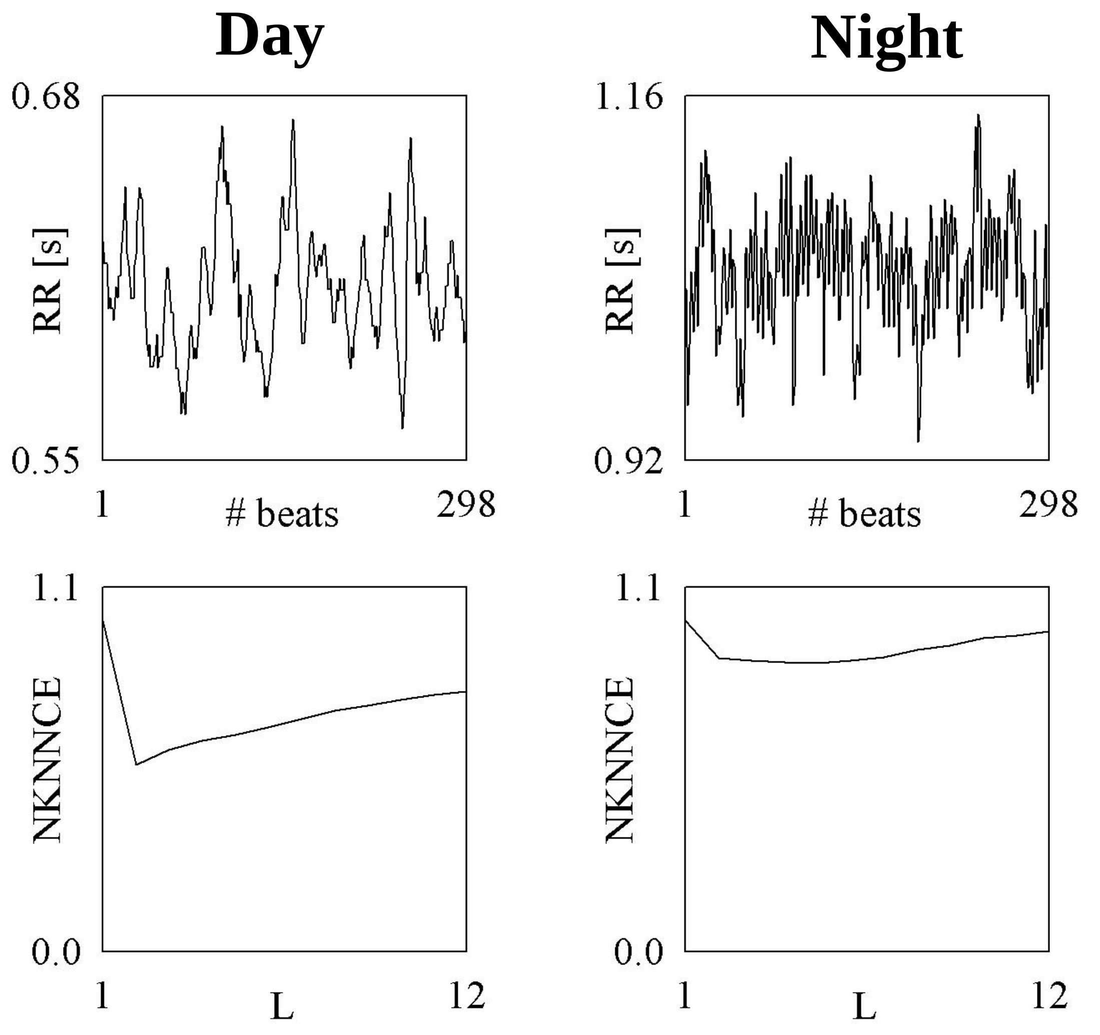
$$\text{NKNNCE(L)} = \frac{\text{KNNCE(L)}}{\text{SE(RR)}}$$

with $0 \leq \text{NKNNCE(L)} \leq 1$

NKNNCE(L) = 0 ⇨ null information, perfect prediction

NKNNCE(L) = 1 ⇨ maximum information, null prediction

Normalized complexity index based on k-nearest-neighbor approach



NCI_{KNN}=min(NKNNCE(L)) with $0 \leq NCI_{KNN} \leq 1$

Porta A et al, Physiol Meas, 34:17-33, 2013

Outline

1) Predictability approach based on conditional distribution and uniform quantization

2) Conditional entropy approach based on uniform quantization

3) Predictability approach based on conditional distribution and k nearest neighbors

4) Conditional entropy approach based on k nearest neighbors

5) Application to 24h Holter recordings of heart period variability obtained from healthy subjects and chronic heart failure population

Experimental protocol

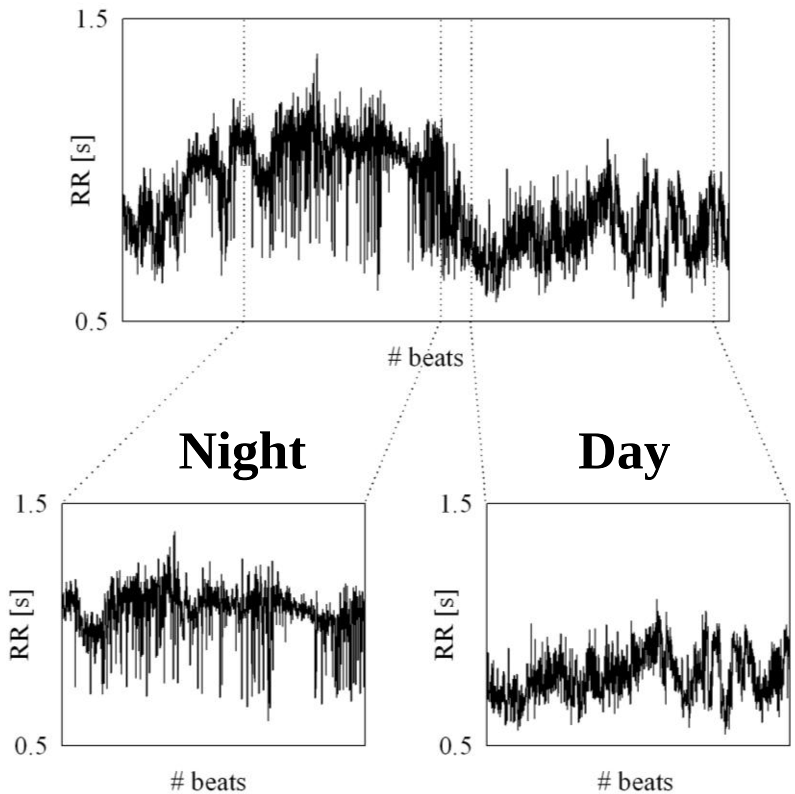
12 normal (N) subjects (aged 34 to 55)
13 chronic heart failure (CHF) patients (aged 33 to 56)

CHF patients are 2 in NYHA class I, 2 in NYHA class II, 9 in NYHA class III). Ejection fraction ranges from 13% to 30%, median=25%

ECGs were recorded for 24h with a standard analogue Holter recorder.

ECGs were sampled at 250 Hz and QRS detection was automatically performed by the software of the device

Analysis of 24h Holter recordings of heart period variability



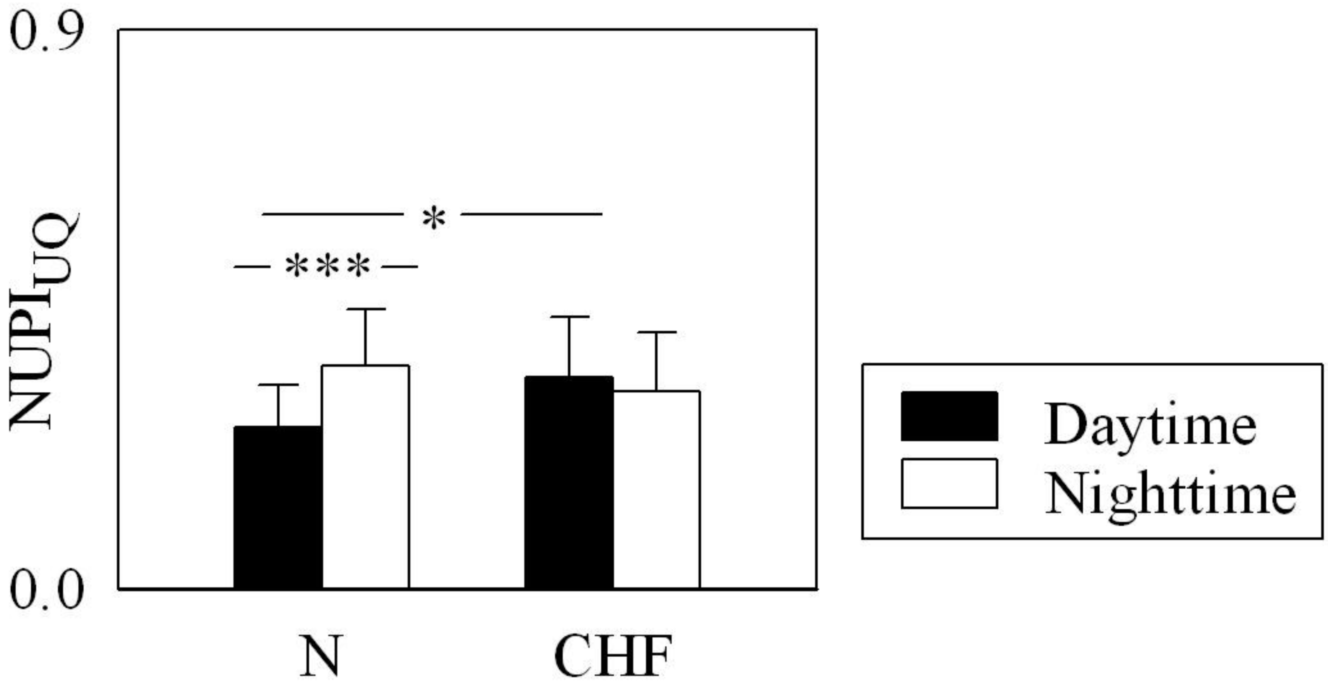
Day: from 09:00 AM to 07:00 PM

Night: from 00:00 AM to 05:00 AM

$NUPI_{UQ}$, $NUPI_{KNN}$, NCI_{UQ}, NCI_{KNN} were calculated iteratively over sequences of 300 samples with 50% overlap

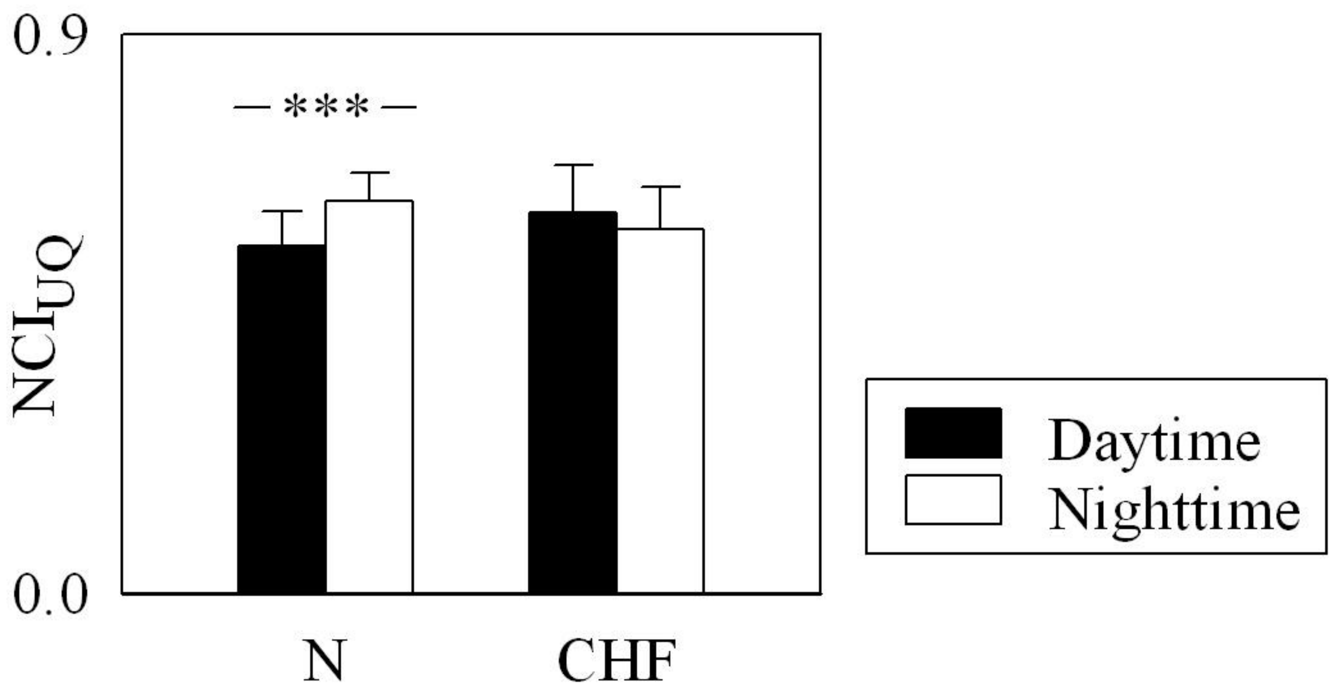
The median of the distribution of $NUPI_{UQ}$, $NUPI_{KNN}$, NCI_{UQ}, NCI_{KNN} during daytime and nighttime was assessed

$NUPI_{UQ}$ in N subjects and CHF patients



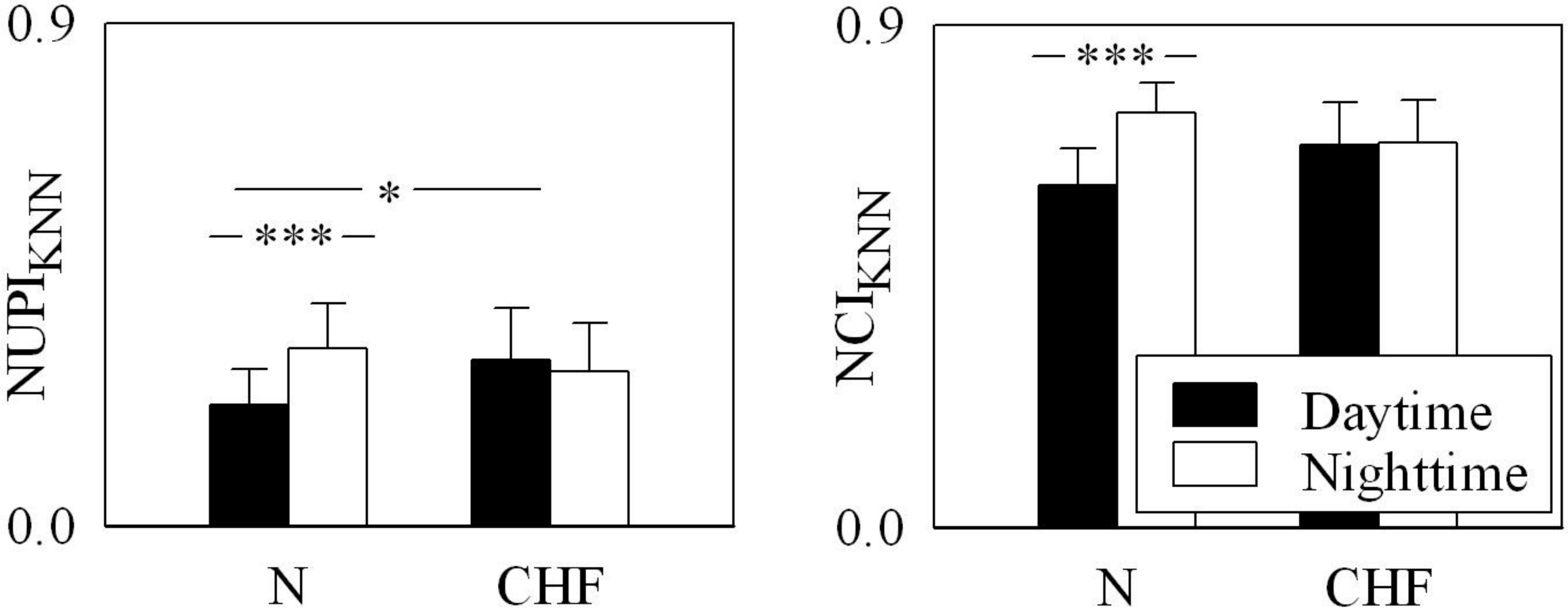
Porta A et al, Chaos, 17, 015117 2007

NCI_{UQ} in N subjects and CHF patients



Porta A et al, Chaos, 17, 015117 2007

$NUPI_{KNN}$ and NCI_{KNN} in N subjects and CHF patients



Conclusions

Complexity of heart period variability can be assessed via predictability-based approaches

These approaches lead to conclusions similar to those drawn using entropy-based methods

Complexity analysis of heart period variability is helpful to distinguish healthy subjects from pathological patients

Complexity analysis of heart period variability does not require controlled experimental conditions to provide meaningful results