# Fast Detrending of Unevenly Sampled Series with Application to HRV

Valeria Villani, Antonio Fasano

Università Campus Bio-Medico di Roma, Rome, Italy

## Abstract

*Detrending RR series is a common processing step prior to HRV analysis. In the classical approaches RR series, which are inherently unevenly sampled, are interpolated and uniformly resampled, thus introducing errors in subsequent HRV analysis. In this paper, we propose a novel approach to detrending unevenly sampled series and apply it to RR series. The approach is based on the notion of weighted quadratic variation, which is a suitable measure of variability for unevenly sampled series. Detrending is performed by solving a constrained convex optimization problem that exploits the weighted quadratic variation. Numerical results confirm the effectiveness of the approach. The algorithm is simple and favorable in terms of computational complexity, which is linear in the size of the series to detrend. This makes it suitable for long-term HRV analysis. To the best of the authors' knowledge, it is the fastest algorithm for detrending RR series.*

## 1. Introduction

Heart rate variability (HRV) is a non invasive measure of the autonomic nervous system. Variations in the duration of heartbeat intervals are consequence of changes in the so-called sympathovagal balance, which is the balance between the sympathetic and the parasympathetic branches of the autonomic nervous system [1]. HRV is assessed from RR interval time series, or tachograms, by computing some time or frequency-domain metrics [2]. HRV analysis is usually preceded by a detrending step of the RR series. This is mainly due to the fact that most frequency-domain metrics require power spectral density estimation and slowly-varying trends introduce nonstationarity, which adversely affects subsequent analysis [1]. Further, the analysis of slow trends in short-term HRV is considered a dubious measure [1].

Classical approaches to RR series detrending consist of subtracting the trend estimated using first or higher order polynomials [3, 4]. More recently, a technique based on smoothness priors was proposed in [5]. The major limitation of these approaches lies in the fact that they apply to uniformly sampled series, despite RR series are inherently unevenly sampled. Thus, linear or cubic spline interpolation and resampling are required prior to detrending [2, 5]. However, It has been shown that the interpolation-resampling process introduces significant errors in the spectral analysis of HRV [6, 7]. In particular, it overestimates the total power in the LF and HF bands, and the LF/HF ratio, the most used frequency-domain measures of HRV [2, 6]. Very recently, an algorithm for detrending RR unevenly sampled series has been proposed in [8].

In this paper we propose a novel approach to detrending *unevenly* sampled series and apply it to RR series. We avoid the need of interpolation-resampling process. The proposed algorithm is based on the notion of *weighted quadratic variation reduction*. The algorithm is simple and proves to be effective in detrending RR series. Moreover, it is remarkably fast, thus lending itself to the detrending of 24-hour RR series for long-term analysis [1]. Finally, detrending using the proposed algorithm can be followed by spectral analysis using the Lomb-Scargle periodogram [2], which is tailored to unevenly sampled data.

The paper is organized as follows. The rationale behind the approach is described in Section 2. The detrending algorithm is derived in Section 3. Sections 4 and 5 follow with simulation results and conclusions.

## 2. Rationale

Trends are low[1] "variability" components of RR series. Thus, provided that we introduce a suitable measure of "variability", a trend can be estimated searching for the low-variability component closest, in some sense, to the RR series. Then, the estimated trend can be subtracted from the RR series. To make this idea precise, we propose to measure the variability of a generic vector through the following

**Definition 1.** Given a vector $\boldsymbol{x} = [x_1 \cdots x_n]^{\mathrm{T}} \in \mathbb{R}^n$, with $n \geq 2$, and the set $\mathcal{W} = \{w_1, \ldots, w_{n-1}\}$, with $w_k > 0$, the *weighted quadratic variation* of $\boldsymbol{x}$ is defined as

$$[\boldsymbol{x}]_{\mathcal{W}} \doteq \sum_{k=1}^{n-1} w_k^2 \left(x_k - x_{k+1}\right)^2 \tag{1}$$

[1] Low with respect to RR series "variability".

and is denoted by $[\boldsymbol{x}]_{\mathcal{W}}$.

When $w_k = 1$, for $k = 1, \ldots, n-1$, the weighted quadratic variation (WQV) is known in the literature as the quadratic variation and is a well-known property used in the analysis of stochastic processes [9]. In this context, we consider WQV as a function of deterministic or random vectors.

Introducing the $(n-1) \times n$ matrix

$$\boldsymbol{D}_{\mathcal{W}} = \begin{bmatrix} w_1 & -w_1 & 0 & \cdots & 0 \\ 0 & w_2 & -w_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & w_{n-1} & -w_{n-1} \end{bmatrix}, \quad (2)$$

the WQV of $\boldsymbol{x}$ can be equivalently expressed as

$$[\boldsymbol{x}]_{\mathcal{W}} = \|\boldsymbol{D}_{\mathcal{W}}\boldsymbol{x}\|^2, \quad (3)$$

where $\|\cdot\|$ denotes the Euclidean norm.

The WQV is a *consistent* measure of variability and its use is motivated by this property: for vectors affected by additive noise, on average it does not decrease and is an increasing function of noise variances, *regardless* of the set $\mathcal{W}$ of weights. This property relies on the following

**Proposition 2.** *Let $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{\nu}$, where $\boldsymbol{x}_0 \in \mathbb{R}^n$ is a deterministic vector and $\boldsymbol{\nu} \in \mathbb{R}^n$ is an absolutely continuous random vector with zero mean and covariance matrix $\boldsymbol{K}_\nu = \mathbb{E}\left\{\boldsymbol{\nu}\boldsymbol{\nu}^T\right\}$. Then, for any choice of the set $\mathcal{W}$*

$$\mathbb{E}\left\{[\boldsymbol{x}]_{\mathcal{W}}\right\} = [\boldsymbol{x}_0]_{\mathcal{W}} + \operatorname{tr}\left(\boldsymbol{D}_{\mathcal{W}}\boldsymbol{K}_\nu\boldsymbol{D}_{\mathcal{W}}^T\right) \geq [\boldsymbol{x}_0]_{\mathcal{W}} \quad (4)$$

*where the inequality is strict almost surely*[2].

The WQV is well suited for measuring the variability of unevenly sampled series. This is due to the degrees of freedom provided by the set of weights $\mathcal{W}$. A suitable choice of it allows us to introduce into the WQV information about the – possibly uneven – sampling grid.

Indeed, let $\boldsymbol{x} = [x_1, \ldots, x_n]^T$ be a vector collecting samples of a series taken at (non-uniform) times $t_1, \ldots, t_n$, with $t_i < t_{i+1}$. By setting weights to

$$w_k = \frac{1}{t_{k+1} - t_k}, \qquad k = 1, \ldots, n-1 \quad (5)$$

$[\boldsymbol{x}]_{\mathcal{W}}$ becomes a measure of variability that takes into account both values and sampling times. Moreover, it is appropriate for non-uniformly sampled series, because each summand in (1) is normalized by the time interval between consecutive samples. Indeed, even though each term $(x_k - x_{k+1})^2$ might be large, its contribution to the WQV reduces as the time interval $t_{k+1} - t_k$ increases. This is what we expect from a measure of variability for unevenly sampled data.

[2] A property holds almost surely if it is true with probability one.

Now, let us apply these results to RR series

$$RR_k = R_{k+1} - R_k, \qquad k = 1, \ldots, n \quad (6)$$

where $R_k$ is the time lag of the $k$th R-peak. The corresponding time instants are

$$t_k = \sum_{i=1}^{k-1} RR_i, \qquad k = 1, \ldots, n \quad (7)$$

with $t_1 = 0$. Thus, taking into account both (5) and (7), the weights for RR series become

$$w_k = \frac{1}{RR_k}, \qquad k = 1, \ldots, n-1. \quad (8)$$

The next section is devoted to the development of an efficient algorithm for detrending RR series that exploits the concept of WQV.

## 3. Trend estimation and removal

In this section, we denote by $\widetilde{\boldsymbol{\mathcal{R}}}$ the vector collecting $n$ samples of an RR series, which is affected by a slowly varying trend, by $\boldsymbol{x}$ the vector of estimated trend, and by $\boldsymbol{\mathcal{R}} = \widetilde{\boldsymbol{\mathcal{R}}} - \boldsymbol{x}$ the corresponding RR series detrended. Following the line of reasoning presented in the previous section, the trend can be estimated searching for a component of the series that has reduced variability, with respect to the measured one. This amounts to searching for a component that is "close" to the observed series, but has *reduced* WQV. We propose to estimate the trend $\boldsymbol{x}$ by solving the following optimization problem

$$\begin{cases} \underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{minimize}} & \left\|\boldsymbol{x} - \widetilde{\boldsymbol{\mathcal{R}}}\right\|^2 \\ \text{subject to} & [\boldsymbol{x}]_{\mathcal{W}} \leq \rho \end{cases} \quad (9)$$

where weights are set according to (8) and $\rho$ is a nonnegative constant that controls the WQV of the estimated trend. Note that we do not need to know in advance the appropriate value for $\rho$ in any particular problem, as will be clear later.

Let us consider (9) in more detail. It is a convex optimization problem, since both the objective function and the inequality constraint are convex. As a consequence, any locally optimal point is also globally optimal and Karush-Kuhn-Tucker (KKT) conditions provide necessary and sufficient conditions for optimality [10]. Moreover, since the objective function is strictly convex and the problem is feasible, the solution exists and is unique. It is possible to prove that the solution to (9) is given by

$$\boldsymbol{x} = \left(\boldsymbol{I} + \lambda\boldsymbol{D}_{\mathcal{W}}^T\boldsymbol{D}_{\mathcal{W}}\right)^{-1}\widetilde{\boldsymbol{\mathcal{R}}} \quad (10)$$

where $\boldsymbol{I}$ denotes the $n \times n$ identity matrix, and $\lambda$ is a non-negative parameter determined by

$$[\boldsymbol{x}]_{\mathcal{W}} = \left\| \boldsymbol{D}_{\mathcal{W}} \left( \boldsymbol{I} + \lambda \boldsymbol{D}_{\mathcal{W}}^T \boldsymbol{D}_{\mathcal{W}} \right)^{-1} \widetilde{\boldsymbol{\mathcal{R}}} \right\|^2 = \rho. \quad (11)$$

Note that in (10) the matrix inverse exists for any $\lambda \geq 0$, since $\boldsymbol{D}_{\mathcal{W}}^T \boldsymbol{D}_{\mathcal{W}}$ is positive semidefinite. It is interesting that the solution to (9) is a linear operator acting on $\widetilde{\boldsymbol{z}}$. The parameter $\lambda$ controls the WQV of the solution $\boldsymbol{x}$, i.e., the degree of variability of the estimated trend. Moreover, since (11) establishes a one-to-one correspondence between $\rho$ and $\lambda$, this last can be used in place of $\rho$. As $\lambda$ ranges from 0 to $+\infty$, solution (10) captures components of the measured series $\widetilde{\boldsymbol{\mathcal{R}}}$ with decreasing WQV. It is possible to prove that when $\lambda \to +\infty$ such components reduce to a constant vector, whereas when $\lambda$ is finite more complex trends of the series are captured.

Once estimated, the trend can be removed from the RR series by subtraction

$$\boldsymbol{\mathcal{R}} = \widetilde{\boldsymbol{\mathcal{R}}} - \boldsymbol{x} = \left[ \boldsymbol{I} - \left( \boldsymbol{I} + \lambda \boldsymbol{D}^{\mathrm{T}} \boldsymbol{D} \right)^{-1} \right] \widetilde{\boldsymbol{\mathcal{R}}} \quad (12)$$

In the following, we will refer to (10) or (12) as detrending by *WQV reduction* (WQVR).

## 3.1. Computational issues

Formula (10) involves matrix inversion, which has complexity $O(n^3)$. Thus, computational aspects become important, especially when the size of the vector to detrend is large. However, it is possible to prove that (10) can be performed efficiently with complexity $O(n)$, i.e., *linear* in the size of the RR series to detrend. To give an idea of the speed, an efficient implementation of (10) in MATLAB (ver. 7.11), running on a PC equipped with a 2.3 GHz Core i5 processor, takes about 0.82 s to detrend an RR series of $10^7$ double precision floating point samples, which corresponds to a tachogram of about 100 days. To the best of the authors' knowledge, this is the fastest algorithm for detrending RR series.

## 4. Simulation results

We carried out a comparative analysis of the proposed approach with competing algorithms, both on real and synthetic RR series. For real time series, the quality of detrending is evaluated by visual inspection, since the trend affecting the RR series is not known. Synthetic RR series allows us to numerically assess the performance.

Regarding real data, in the top panel of Figure 1 we report a portion of the RR series $\widetilde{\boldsymbol{\mathcal{R}}}$ extracted from the ECG record nsrdb/16273 from PhysioNet [11]. The lower panel shows the corresponding series $\boldsymbol{\mathcal{R}}$ detrended using

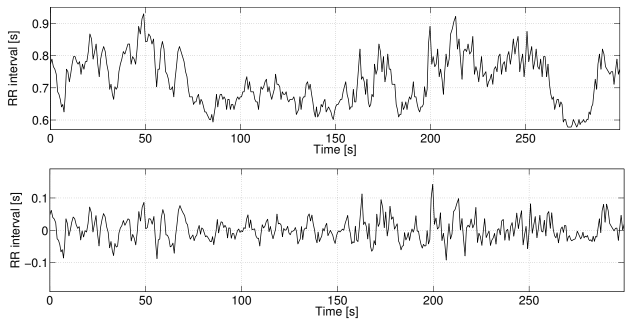


Figure 1. RR series from real data (top) and the same detrended using the proposed algorithm (bottom).

WQVR, with $\lambda$ roughly set to 15. A visual comparison highlights that the proposed approach managed to effectively detrend the RR series, even though the value of $\lambda$ was not optimized.

To assess the performance, we considered synthetic RR series (trend-free) affected by simulated trends. The synthetic RR series, denoted by $\boldsymbol{\mathcal{R}}^{(0)}$, was generated according to the model described in [12], with LF/HF ratio equal to 0.5, heart rate with mean 60 bpm and standard deviation 5 bpm, and duration 4.5 minutes. Then, $\boldsymbol{\mathcal{R}}^{(0)}$ was corrupted by synthetic trends $\boldsymbol{\xi}_i$ rendered as white Gaussian noise ($\mu = 0$, $\sigma^2 = 1$) low-pass filtered with bandwidth 0.05 Hz, unevenly sampled at the time of the series $\boldsymbol{\mathcal{R}}^{(0)}$. The resulting series is denoted by $\widetilde{\boldsymbol{\mathcal{R}}}_i = \boldsymbol{\mathcal{R}}^{(0)} + \boldsymbol{\xi}_i$. We generated 300 independent realizations of $\boldsymbol{\xi}_i$.

We compared our algorithm with: i) the smoothness priors (SP) approach [5], which is one of the most used, if not the most used, techniques for RR series detrending; ii) the smoothing by Gaussian process priors (SGP) method [8], which can handle unevenly sampled series. SP can handle only uniformly sampled data. It was implemented using the code provided in [5]. For SGP we used the MATLAB script available on the website reported in [8]. The quality of detrending was assessed through the following errors

$$\varepsilon_i(\mathrm{alg}_k) = \left\| \boldsymbol{\mathcal{R}}_i^{(\mathrm{alg}_k)} - \boldsymbol{\mathcal{R}}_0 \right\|^2, \quad i = 1, \ldots, 300 \quad (13)$$

where $\boldsymbol{\mathcal{R}}_i^{(\mathrm{alg}_k)}$ is the series detrended using algorithm $\mathrm{alg}_k \in \{\mathrm{WQVR}, \mathrm{SP}, \mathrm{SGP}\}$. Note that in the case of SP, which works only on uniformly sampled data, the corresponding detrended series $\boldsymbol{\mathcal{R}}_i^{(\mathrm{SP})}$ was unevenly resampled at the same time of $\boldsymbol{\mathcal{R}}^{(0)}$, to allow a meaningful comparison in (13).

Table 1 reports the mean and variance of the error (13) computed over 300 realizations $\boldsymbol{\xi}_i$ of the simulated trend. From the table, it emerges that WQVR outperforms both SP and SGP in terms of error variance, since $\mu_\epsilon$ has comparable values. In particular, the advantage over SGP is very large. Moreover, WQVR has the benefit of being

| | $\mu_\varepsilon$ | $\sigma^2_\varepsilon$ |
|---|---|---|
| WQVR | $2.55 \times 10^2$ | $2.12 \times 10^{-4}$ |
| SP | $2.58 \times 10^2$ | $1.61 \times 10^{-3}$ |
| SGP | $2.55 \times 10^2$ | 9.56 |

Table 1. Mean and variance of error (13) for unevenly sampled RR series ($\mu_\epsilon$ in s$^2$, $\sigma^2_\epsilon$ in s$^4$).

| | $n = 10^3$ | $n = 10^5$ | $n = 10^7$ |
|---|---|---|---|
| WQVR | $9.50 \times 10^{-5}$ | $8.20 \times 10^{-3}$ | $8.20 \times 10^{-1}$ |
| SP | 1.24 | – | – |
| SGP | $1.10 \times 10^{-3}$ | $1.47 \times 10^{-1}$ | $6.90 \times 10^2$ |

Table 2. Average execution time in seconds.

faster than SGP and much faster than SP. Indeed, while WQVR has complexity $O(n)$, i.e., *linear* in the size of the series to detrend, SGP and SP have complexity $O(n^3)$, since they use Cholesky decomposition [8] and matrix inversion [5], respectively. Note that SGP is faster than SP since it exploits some MATLAB features. Further, WQVR uses much less memory resources since, differently from SGP and SP, it does not store any matrix. Furthermore, in contrast to SP, WQVR does not require interpolation and resampling steps.

Finally, we measured the execution time required by WQVR, SP and SGP to detrend an RR series of double precision floating point samples with size $n$ equal to $10^3$, $10^5$, and $10^7$. The algorithms were tested under MATLAB (ver. 7.11), running on a PC equipped with a 2.3 GHz Core i5 processor. The average execution time is reported in Table 2: for WQVR it grows linearly with $n$, differently from SGP and SP. In particular, it was practically impossible to compute execution times for SP for $n \geq 10^4$. Note that in the case of SP the time required by interpolation and resampling was excluded. The results of Table 2 confirm that the proposed approach is remarkably fast.

## 5. Conclusions

In this work we considered the problem of detrending RR series. Customarily, this is achieved by interpolating the RR series and then resampling to obtain an evenly sampled series, which is then detrended. The interpolation-resampling step introduces errors in subsequent HRV analysis. To tackle this problem, we proposed a novel approach to detrending unevenly sampled RR series. The approach is based on the notion of *weighted quadratic variation reduction* and is suited for subsequent analysis through the Lomb-Scargle periodogram, since it does not require interpolation and resampling. Simulation results confirm the effectiveness of the approach, which is moreover very fast. This makes it suitable for long-term HRV analysis. To the best of the authors' knowledge, it is the fastest algorithm for detrending RR series.

## References


[1] Task Force of The European Society of Cardiology & The North American Society of Pacing and Electrophysiology. Heart rate variability - Standards of measurement, physiological interpretation, and clinical use. Eur Heart J 1996; 17:354–381.

[2] Clifford GD, Azuaje F, McSharry P (eds.). Advanced Methods and Tools for ECG Data Analysis. Artech House, Inc., 2006.

[3] Mitov I. A method for assessment and processing of biomedical signals containing trend and periodic components. Med Eng Phys 1998;20:660–668.

[4] Litvack DA, Oberlander TF, Carney Laurel H. SJP. Time and frequency domain methods for heart rate variability analysis: A methodological comparison. Psychophysiology 1995;32:492–504.

[5] Tarvainen M, Ranta-aho P, Karjalainen P. An advanced detrending method with application to HRV analysis. IEEE Trans Biomed Eng feb. 2002;49(2):172–175.

[6] Clifford GD, Tarassenko L. Quantifying errors in spectral estimates of HRV due to beat replacement and resampling. IEEE Trans Biomed Eng 2005;52(4):630–638.

[7] Moody GB. Sepctal analyisis of heart rate without resampling. Computers in Cardiology 1993;20:715–718.

[8] Eleuteri A, Fisher AC, Groves D, Dewhurst CJ. An efficient time-varying filter for detrending and bandwidth limiting the heart rate variability tachogram without resampling: MATLAB open-source code and internet web-based implementation. Comp Math Meth Med 2012;.

[9] Shreve SE. Stochastic Calculus for Finance II: Continuous-Time Models. Springer Science+Business Media, Inc., 2004.

[10] Boyd S, Vandenberghe L. Convex Optimization. Cambridge University Press, March 2004.

[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;101(23):e215–e220.

[12] McSharry PE, Clifford GD, Tarassenko L, Smith LA. A dynamical model for generating synthetic electrocardiogram signals. IEEE Trans Biomed Eng 2003;50(3):289–294.



Address for correspondence:

Antonio Fasano
Università Campus Bio-Medico di Roma
via Álvaro del Portillo, 21
00128 Rome, Italy
E-mail: a.fasano@unicampus.it