MAT^2

MATerials MATemàtics
Volum 2007, treball no. 8, 24 pp. ISSN: 1887-1097
Publicació electrònica de divulgació del Departament de Matemàtiques
de la Universitat Autònoma de Barcelona
www.mat.uab.cat/matmat

Obtaining the basic response pattern of physiological time series data: a comparison of methods

M. S. Zakynthinaki, J. R. Stirling, M. Sillero, J. Sampedro, I. Refoyo

1 Physiological variables

An understanding of the kinetics of physiological variables such as the heart rate or the rate of change of volume of oxygen uptake is fundamental not only to training methodology and competitive success in sport and exercise, but also to our knowledge of cardiovascular health. A correct and efficient means of interpreting and analyzing the data obtained is of vast importance, as exercise testing is routinely used in both of these areas.

The rate of change of oxygen uptake (defined as $\dot{V}O_2$) is equal to the product of cardiac output and the volume of oxygen extracted from the blood [1, 2]. A breath-by-breath recording of $\dot{V}O_2$ in the form of time series data is possible. Today portable $\dot{V}O_2$ measurement devices are available, including a mask (see Figure 1) and a system that puts the measurements in the form of time series data and stores it in a computer.

The heart rate, on the other hand, is the most commonly used and easiest to obtain, cardiovascular variable. For the recording of heart rate time series data, heart rate monitors are commonly used by athletes [2], as well as by the general public [3, 4, 5]. A chest strap with electrodes (see Figure 2) records the beat to beat interval of the heart and transmits this signal to a wrist receiver (watch). Heart rate monitors are capable of recording beat to

Figure 1: Obtaining heart rate and oxygen uptake data in a maximal oxygen uptake test (first and second authors).

beat intervals and allow one to download this information, in the form of a time series, to a computer, for analysis. There exist numerous popular texts demonstrating the way to analyze and understand the time series recoded with the use of a heart rate monitor. The analysis of the heart rate time series in response to exercise is also commonly used in the field of cardiovascular rehabilitation [3] as well as other areas of health, such as fitness and weight management [4, 5]. There are also many important clinical applications [6].

Figure 2: Heart rate monitors.

2 Noise in physiological time series data

Physiological recordings consisting of breath-by-breath or beat-to-beat data sets, corresponding to measurements of the oxygen uptake or heart rate re-

spectively, show extreme variability. Usually the data oscillates with a high frequency and large amplitude around a smooth curve, which describes the basic response pattern, or underlying dynamics [7]. Such oscillations can be the result of noise due to the measuring device or can be physiological, including ectopic beats [8], or abnormal breathing during exercise including shallow breathing or breath holding [9].

In modelling the kinetics of physiological data, techniques to decrease the noise levels are extensively used [7, 10, 11, 12, 13, 14]. These techniques aim at a reduction of the noise of the signal and the enhancement of the basic (underlying) response pattern, in order to obtain some of the features which best model the way the body adapts to different exercise intensities. Averaging, as a noise reduction and curve smoothing technique is used extensively in the work on modelling of the kinetics of physiological variables such as the heart rate or the rate of change of oxygen uptake. In many studies the time series of physiological data are averaged before the data was modelled. For example in [7] repeated experiments were averaged, both for the individuals and for the whole group of subjects. In [10] a single physiological data set, obtained from breath-by-breath oxygen uptake recordings, was averaged using a 3-rolling breath average. In [14] five sequences of data sets were split, time-aligned and ensemble-averaged to produce a single data set per test. Other examples of data averaging can also be found in [11, 12, 13] as well as in numerous other studies in the literature.

Although a number of studies in the field of exercise physiology address the issue of noise in physiological time series data and conclude that it is random [13, 14, 15, 16, 17, 18], modern research concludes that physiological data is contaminated by nonwhite noise [19]. In [9] it is stated that it is possible that the variability of physiological time series may have biological significance. Recent studies have shown that physiological systems fluctuate in a nonlinear manner, even under resting conditions [20, 21]. It is only under pathological conditions that the physiological signals show uncorrelated randomness similar to white noise [21]. Mathematically it is well known that nonlinear systems such as the physiological ones under question can produce signals which appear noisy [22, 23]. In such systems, the basic response pattern is not necessarily the average; this depends on how the oscillations are distributed.

Averaging over multiple bouts has also been questioned by recent studies in the field of oxygen uptake kinetics: in their conclusions, the authors in [11] state that it is debatable whether averaging is indeed an accurate method

for removing noise and finding the underlying response parameters. From a mathematical point of view, the average of the response patterns of multiple bouts of exercise does not necessarily correspond to the underlying response pattern of a single bout of exercise. A non systematic removal of all the high frequency and high amplitude oscillations found in such time series can result in an averaged system with features not contained in the original non averaged signals [24]. There is no mathematical reason why discrete features found in the averaged signal should also be found in the underlying basic response pattern of the raw, un-averaged data of a single bout of exercise.

3 Averaging and the 3-phase model

It was proposed [25] that the oxygen uptake kinetics in response to a constant exercise intensity is approximately an exponential function of time. Currently three phases are believed to be observable in the time series of the heart rate of an athlete running at constant velocity [7, 14, 26], as well as in the time series of the rate of change of oxygen uptake, $\dot{V}O_2$, of the athlete [7, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 6, 38, 39]; for a mathematical review see [24].

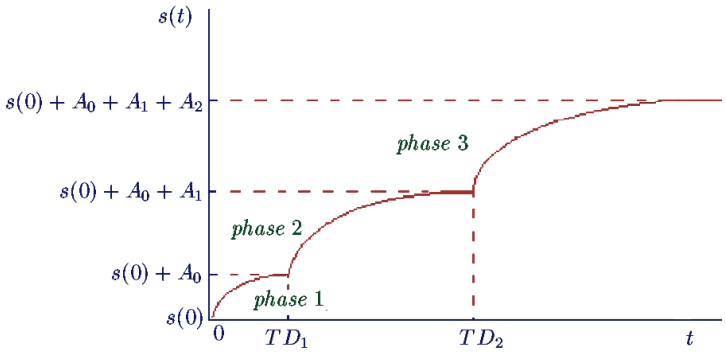


Figure 3: The 3-phase model for the on-transient kinetics of heart rate.

To give a brief introduction to the 3-phase model, let us consider the general case of a physiological variable s given as a function of time $s(t)$, where s can denote either the rate of change of oxygen uptake, or the heart rate, see also [7]. According to the 3-phase model, the on-transient kinetics

of the variable $s(t)$ is modelled using the following triple exponential form:

$$s(t) = s(0) + \begin{cases} A_0\left(1 - e^{-\frac{t}{\tau_0}}\right), & \text{for } 0 \leq t < TD_1 \\ A_0\left(1 - e^{-\frac{t}{\tau_0}}\right) + A_1\left(1 - e^{-\frac{(t-TD_1)}{\tau_1}}\right), & \text{for } TD_1 \leq t < TD_2 \\ A_0\left(1 - e^{-\frac{t}{\tau_0}}\right) + A_1\left(1 - e^{-\frac{(t-TD_1)}{\tau_1}}\right) + A_2\left(1 - e^{-\frac{(t-TD_2)}{\tau_2}}\right), & \text{for } t \geq TD_2 \end{cases} \quad (1)$$

where $s(0)$ is the initial value of the physiological variable s at time $t = 0$, A_0, A_1 and A_2 are the asymptotic values, τ_0, τ_1 and τ_2 are the time constants, and TD_1 and TD_2 are time delays for each exponential term. Figure 3 shows the kinetics of the physiological variable $s(t)$ according to the 3-phase model of equation (1).

We note here that, in the literature, the 3-phase model described in equation (1) is usually expressed in the form:

$$\begin{aligned} s(t) = s(0) &+ A_0\left(1 - e^{-\frac{t}{\tau_0}}\right) && \text{“}\textit{phase}\ 1\text{”} \\ &+ A_1\left(1 - e^{-\frac{(t-TD_1)}{\tau_1}}\right) && \text{“}\textit{phase}\ 2\text{”} \\ &+ A_2\left(1 - e^{-\frac{(t-TD_2)}{\tau_2}}\right) && \text{“}\textit{slow component}\ (\textit{phase}\ 3)\text{”} \end{aligned}$$

It is worth adding that, when the exercise work rate is light or moderate (below the so-called lactate threshold), the kinetics of $s(t)$ is modelled by two exponentials, thus there is no slowly increasing phase 3 [31, 39]. For heavy but not severe exercise intensities [40] phase 3 has been described as an additional distinct component added to the kinetics, to model the delay in reaching a steady state. Exercise intensities that are termed severe are believed to consist of two phases, as the $\dot{V}O_2$ is limited by its maximum value, $\dot{V}O_{2max}$, before a slow component is manifest [40].

Current research in the field of exercise physiology follows the above described 3-phase model (it should be noted here that there is much debate as to the existence of three discrete phases and an alternative model has recently been introduced [24, 41]). No conclusive physiological reason has been provided by the researchers in the field of exercise physiology for the existence of time delays and separate phases, other than data fitting of averaged

data sets. The 3-phase model is described as the best fit to the data, from a statistical point of view, but before such a 3-phase model is fitted to the physiological time series data averaging, as a curve smoothing technique, is extensively used (see section 2). It is worth noticing that, due to the presence of high frequency oscillations in the time series of the raw physiological data, it is only after averaging that the features that are termed the three phases are observable (for examples regarding modelling the $\dot{V}O_2$ time series [10]; in [9] it is clear that it is impossible to clearly observe three separate phases and time delays in the un-averaged signal). The introduction of different phases and time delays in the time series of single bouts of exercise only serves as an improvement (however marginal) in a process of fitting an exponential curve to averaged data.

Regarding the first two phases, after the combination of the two separate exponentials with a time delay was justified as a statistically better fit to the data, physiological reasons have been provided by the researchers [42, 43, 44]. There has been, however, much debate as to the physiological reasons for the third phase, the so-called slow component [32, 33, 36] with most of the proposed mechanisms being refuted. We believe that it is misleading to use a system which has been averaged to support the existence of different phases of physiological behavior.

It is worth noticing that recent research in the field of oxygen uptake kinetics models the 3rd phase (the so-called slow component) in the 3-phase model not as a mono-exponential, but as a number of exponential components (in serial increments) to account for the possible recruitment of additional motor units during prolonged heavy exercise. Findings conclude that it is more plausible that phase 3 is composed of progressive increases in the physiological variable rather than being a single and abrupt increase, as this way the model would be more consistent with a dynamic maintenance of energy balance [11]. As an example, a 10-compartment model has been recently presented [45], which used a sequence of increments, which become smaller and smaller in an effort to follow the underlying response pattern of the physiological time series data. From a mathematical point of view, this process refers to the way a smooth function is approximated using first principles of infinitesimal calculus. It becomes obvious therefore that a more rigorous way of modelling the physiological data is by using a smooth function describing the kinetics of the physiological variable (t) under study.

4 Our aim

We present here three different methods of curve smoothing in order to find the basic response pattern of physiological time series data:

- Ensemble averaging (averaging over repeated bouts of exercise - the most popular curve smoothing technique currently applied to physiological time series data)
- n-rolling averaging (currently in the field of exercise physiology n-rolling averaging is used with $n = 3$ [10])
- Fourier low-pass filtering, with the cut-off frequency chosen by trial and error.

We concentrate on heart rate time series data near maximal, or submaximal exercise intensities (see section 5). We also show the derivation of the basic response pattern of a highly irregular heart rate data set, obtained from rapidly oscillating exercise.

5 Data acquisition and editing

The data presented here correspond to a healthy male subject (age 30, height 1,83m and weight 82 Kgr). The subject's heart rate parameters were: maximum heart rate 188 bpm, minimum heart rate 40 bpm, heart rate reserve 148 bpm. The subject preformed two experiments, each one of which was carried out in accordance with the following protocols:

- Protocol I consisted of five work periods of 3 minute duration at a constant velocity equal to $v\dot{V}O_{2max}$, separated by 3 minutes recovery jogging at a constant velocity equal to 70% of $v\dot{V}O_{2max}$, see [64] for more information regarding the training protocol. The experiment was carried out on a mixture of compound sand and hard footpaths and the first 6 minutes were considered to be a warm-up period.
- Protocol II consisted of 30 minutes of repetitions on a set of short sand hills requiring 30 seconds (at maximal speed) for the ascent and 30 seconds (of easy recovery jogging) to descend.

It should be noted here that a mathematically strict constant (non-zero) velocity exercise bout when starting from rest is never achievable. During the experimental procedure described above, however, care was taken to obtain as good an approximation as possible using split times.

The beat-to-beat heart rate data sets were recorded using a Polar S810i heart rate monitor [65]. Such heart rate monitors have been shown to provide reliable measurements of the electrocardiogram (ECG), when compared with direct ECG measurements [66, 67, 68]. In addition, heart rate monitors have the big advantage that they are portable and can be used for recordings in the field [47]. The beat-to-beat data recordings can, however, be disturbed by any kind of external electromagnetic field. For this reason, care was taken to ensure that the experimental environment was free of any additional electromagnetic signals (such as high tension power lines, mobile phones, engines producing electromagnetic fields etc).

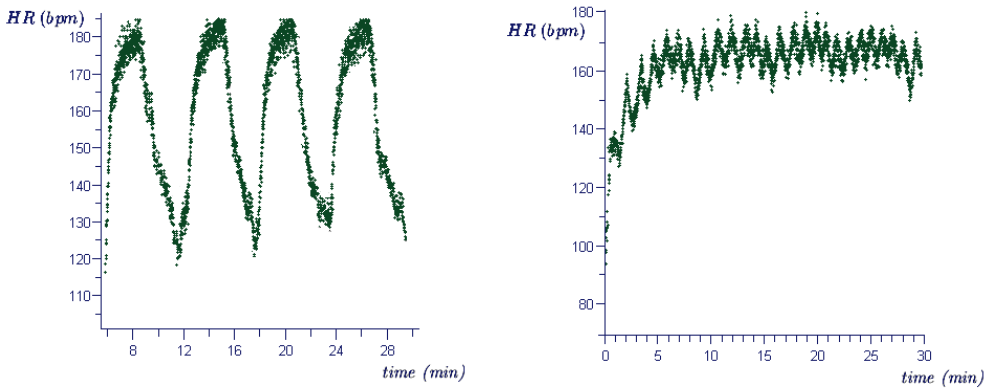
So let us assume that the heart rate of the subject can be expressed as a function of time, $s(t)$. Thus the recorded data set can be thought to consist of the N consecutive sampled values

$$s_k \equiv s(t_k), \quad k = 0, 1, \cdots, N-1$$

where there is $s(t_k) = 0$ for $k < 0$ and $k > N-1$.

Editing of the raw recorded heart rate data set $\{s_k\}_{k=0}^{N-1}$ was only performed to exclude occasional errant data points, as such points were considered not to be reflective of the underlying kinetics. Such occasional deviations from the local mean heart rate can be caused by either technical reasons, physiological reasons occurring in response to various mental or physical impulses, or can be ectopic beats which can occur sporadically due to abnormal origin; for more information of heart rate deviations see [8, 65, 69, 70, 71]. In other studies, data points with values greater than 6 standard deviations from the local mean value were omitted [13, 18, 40]. In what we present here, however, the occasional deviations excluded were the ones occurring mainly due to technical reasons, such as sporadically missed R waves (resulting in a HR interval about two times longer than the normal one), misread T waves (resulting in a sudden too short HR interval), or recorded spurious peak waves instead of the R waves (resulting in a pair of too short and too long intervals). Any other abnormal data points which very rarely appeared in the time series data were also edited, as their values were easily seen to be much larger than 10 standard deviations from the local heart rate.

Figures 4(a) and (b) present the time series of the data recorded during the experiments of protocols I and II described above. In these figures it is easy to observe that the data oscillates around a smooth curve which is a function of both time and intensity.



(a) (b)

Figure 4: Raw heart rate time series data. (a) Protocol I (the first 6 minutes are excluded as they were considered to be a warm-up period) (b) Protocol II.

6 Curve smoothing techniques

6.1 Ensemble averaging

Given m sequences $\{s_k^j\}_{k=0}^{N-1}$, $j = 1, \cdots, m$ consisting of N physiological data points each, the (unweighed) ensemble average of all m data sets is the new sequence $\{\alpha_k\}_{k=0}^{N-1}$ defined from the s_k^j as follows,

$$\alpha_k = \frac{1}{m} \sum_{j=1}^{m} s_k^j \, .$$

6.2 n-Rolling averaging

Given the sequence $\{s_k\}_{k=0}^{N-1}$ of the N physiological data points, the n-rolling (or moving) average is the sequence $\{r_k^n\}_{k=0}^{N-n}$ defined from the s_k by taking

the average of subsequences of n terms,

$$r_k^n = \frac{1}{n} \sum_{j=k}^{k+n-1} s_j.$$

6.3 A few words about Fourier low-pass filtering

Fourier low-pass filtering can be used as an alternative curve smoothing technique, as well as an easy and very effective way for obtaining the basic response pattern from noisy physiological data [46, 47]. Fourier filtering is a very powerful tool and has therefore been widely used for signal smoothing purposes in many applications such as signal processing, as well as a variety of experimental situations [48, 49, 50, 51, 52, 53, 54, 55, 56, 57]. We should note here that methods of spectral analysis have been also widely used in the field of exercise physiology, especially in analyzing the frequencies of heart rate variability [19, 58, 59, 60] and also in other areas of cardiovascular dynamics, where the study was concentrated in finding the frequencies of the physiological response [14, 61, 62, 63]. However, Fourier low-pass filtering, as a curve smoothing method, is never mentioned in these studies. Moreover, in the work related to 3-phase modelling, a spectral analysis is performed after the raw data has been subjected to averaging, see for example [14], where the breath-by-breath raw data were firstly linearly interpolated to yield a time series with the $\dot{V}O_2$ valuated each second and then the data sets were ensemble-averaged before frequency domain analysis was performed.

The discrete Fourier transform of the sequence $\{s_k\}_{k=0}^{N-1}$ of the N data points is given [52, 53] by the set of N points $\{S_n\}_{n=0}^{N-1}$ as follows,

$$S_n \equiv \sum_{k=0}^{N-1} s_k e^{2\pi i k n/N}.$$

The above discrete Fourier transform can be computed in $O(N \log_2 N)$ operations by use of the Fast Fourier Transform algorithm (FFT). We adopted the FFT algorithm from [53]. When N is not an integer power of 2, as is the general case of physiological data, the length of the time series can be artificially increased by adding zero values, a method that is referred to as zero-padding [50, 52, 53].

After application of a discrete Fourier transform, it is possible to achieve the desired smoothing of the data curve by removing the frequency components above a certain cut-off frequency threshold, while retaining the low

frequency components which are carrying the required information. This is equivalent to the implementation of a so-called rectangular window filter [50, 52] that truncates the frequency time series after the M-th term (where $M \leq N-1$).

The inverse Fourier transform gives the 'filtered' sequence $\{s_k^f\}_{k=0}^{N-1}$ that describes the basic response pattern of the data set. Regarding the application of a rectangular window, techniques exist by which the optimal value of the cut-off frequency (M-th frequency term) can be calculated [55]. The value of the cut-off frequency can also be chosen interactively to correctly estimate the noise level. As this value is a parameter in the Fourier curve smoothing procedure, it can be selected so as to provide a filtering/reduction of high frequency oscillations at any desired level.

The application of a rectangular window such as the one described above can, however, result in the introduction of 'ripples' of infinite duration in the resulting smoothed curve [50, 52]. This oscillatory behavior arises from the discontinuity that characterizes the applied low-pass filter (abrupt transition to zero) and is known in the literature as the Gibb's phenomenon. The way to avoid these oscillatory effects is by providing a smooth transition between the pass-band of frequencies and the stop-band. This can be achieved by use of windows that does not contain abrupt discontinuities [50, 52] but decay to zero gradually. A number of such window filters can be found in the literature, for example spline transitions [50, 53].

7 On-transient, constant exercise intensity data sets

The four on-transient responses of protocol I shown in figure 4(a) are characteristic classical responses to an exercise of constant velocity, including the so-called slow component.

We have split and time-aligned the on-transient responses of figure 4(a). Assuming that each of the four data sets is a sequence of N points, figure 5 shows the superposition of the following four sequences, derived from the complete data set of figure 4(a):

- $\{s_k^1\}_{k=0}^{N-1}$, time interval 1 [6.01min, 9min]
- $\{s_k^2\}_{k=0}^{N-1}$, time interval 2 [12.01min,15min]

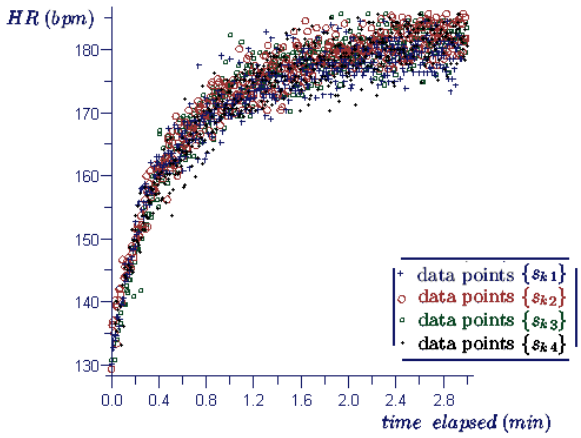


Figure 5: The four on-transient data sets of protocol I superimposed

- $\{s_k^3\}_{k=0}^{N-1}$, time interval 3 [18.01min,21min] and
- $\{s_k^4\}_{k=0}^{N-1}$, time interval 4 [24.01min, 27min]

It is worth noticing that the starting time of the data sets shown in figure 5 were synchronized before the data sets were superimposed.

7.1 Fourier low-pass filtering

Figure 6 presents the smooth curves resulting from Fourier low-pass filtering (sequences $\{s_k^{f-1}\}_{k=0}^{N-1}$, $\{s_k^{f-2}\}_{k=0}^{N-1}$, $\{s_k^{f-3}\}_{k=0}^{N-1}$ and $\{s_k^{f-4}\}_{k=0}^{N-1}$) superimposed on the corresponding on-transient data sequences $\{s_k^1\}_{k=0}^{N-1}$, $\{s_k^2\}_{k=0}^{N-1}$, $\{s_k^3\}_{k=0}^{N-1}$ and $\{s_k^4\}_{k=0}^{N-1}$ respectively. The cut-off frequency for the plots of figure 6 was 0.13 Hz. At the bottom of each figure the time series for the residuals (vertical distance) between the time series data and the time series resulting from low-pass filtering of the data are shown. By observation of the time series of the residuals for these four graphs it is easy to see that there is no obvious pattern which the residual time series follow.

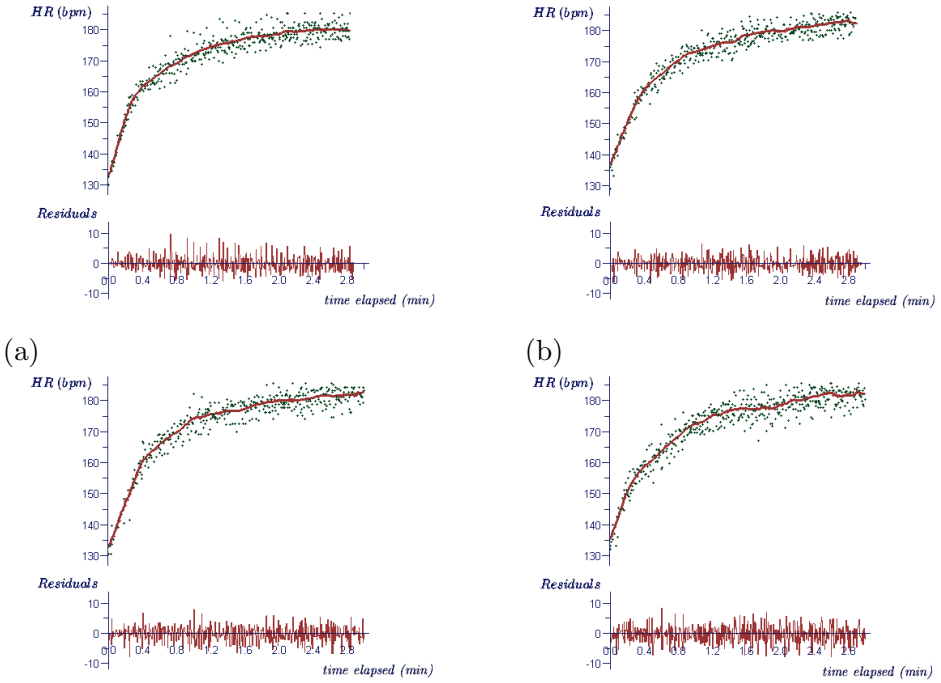


Figure 6: On-transient time series data (green points) and the corresponding smoothed Fourier low-pass curves (red curves). The residuals give the vertical distance between the time series data and the time series resulting from low-pass filtering of the data. a) time interval 1, mean residuals value: 0.046 (b) time interval 2, mean residuals value: 0.008 (c) time interval 3, mean residuals value: 0.018 and (d) time interval 4, mean residuals value: -0.003. Cut-off frequency: 0.13 Hz.

7.2 Ensemble averaging

The averaging procedure was synchronized to the work load changes, see also [7]. As the on-transient data sequences $\{s_k^j\}_{k=0}^{N-1}$, $j = 1, \ldots, 4$ correspond to beat-to-beat data that are not recorded at exactly the same time intervals, the points s_k^j were interpolated to provide data points at time intervals that are the same for each sequence.

Figure 7 presents the sequence $\{\alpha_k\}_{k=0}^{N-1}$ of the ensemble average of the four data sets ($m = 4$, see section 6.1) superimposed on the corresponding on-transient data sequences $\{s_k^1\}_{k=0}^{N-1}$, $\{s_k^2\}_{k=0}^{N-1}$, $\{s_k^3\}_{k=0}^{N-1}$ and $\{s_k^4\}_{k=0}^{N-1}$ respec-

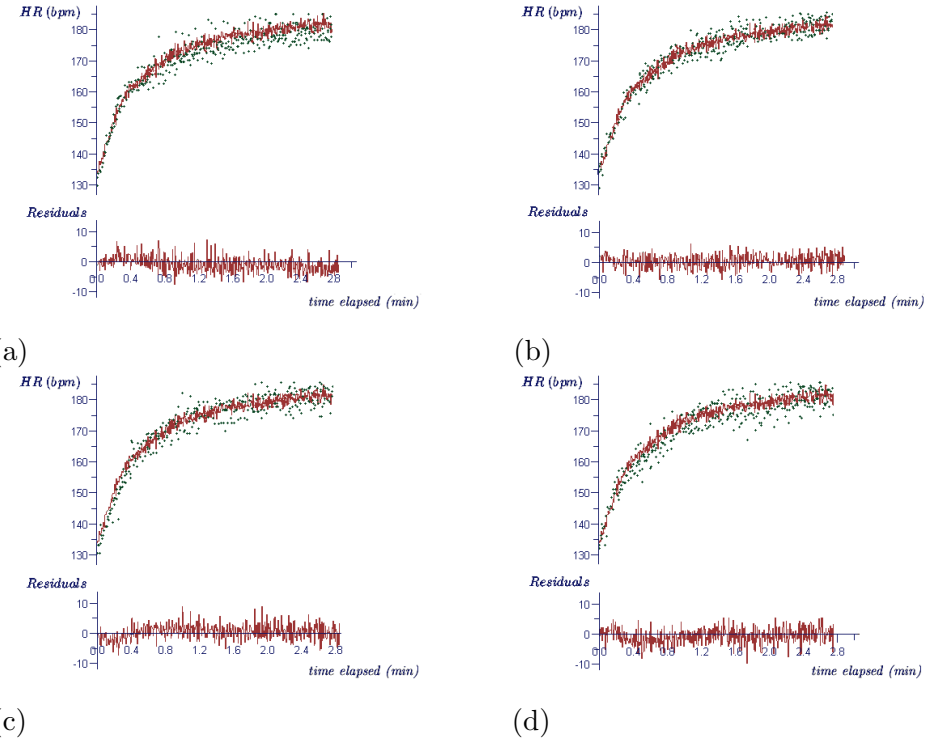


Figure 7: On-transient time series data (green points) and the corresponding ensemble average curves (red curves). The residuals give the vertical distance between the time series data and the time series resulting from ensemble averaging of the data. (a) time interval 1, mean residuals value: -0.804 (b) time interval 2, mean residuals value: 0.597 (c) time interval 3, mean residuals value: 0.728 and (d) time interval 4, mean residuals value: -0.520.

tively. At the bottom of each figure we show the time series for the residuals (vertical distance) between the time series data and the time series resulting from ensemble averaging of the data. For the particular case we present here, there is $N = 450$.

The averaged sequence of data points $\{\alpha_k\}_{k=0}^{N-1}$ is the same for all four figures 7(a)-(d), as it corresponds to the average of all four on-transient responses $\{s_k^1\}_{k=0}^{N-1}$, $\{s_k^2\}_{k=0}^{N-1}$, $\{s_k^3\}_{k=0}^{N-1}$ and $\{s_k^4\}_{k=0}^{N-1}$. As expected, oscillations of high frequency and amplitude are still observable in the time series of the data after averaging. As can be seen in figure 5, the four on-transient data sets are very similar but not identical, even though they correspond to

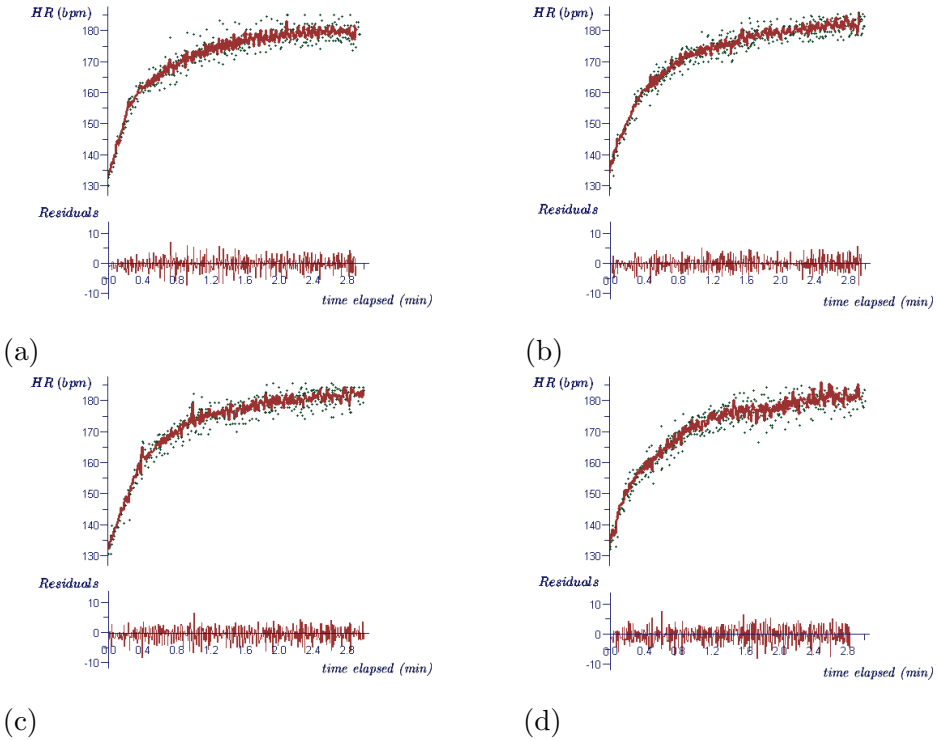


Figure 8: On-transient time series data (green points) and the corresponding 3-rolling averaged curves (red curves). The residuals give the vertical distance between the time series data and the time series resulting from a 3-rolling averaging of the data. a) time interval 1, mean residuals value: -0.103 (b) time interval 2, mean residuals value: -0.110 (c) time interval 3, mean residuals value: -0.114 and (d) time interval 4, mean residuals value: -0.112.

identical velocities and identical work load changes. This is also the case in the literature, see for example [7, 14], where such data sets are however characterized as identical. We would therefore expect that the averaging procedure introduces features not included in the original data sets.

It can be easily observed that the time series of the residuals in figures 7 are not evenly distributed around zero and obvious patterns exist. We observe also that, as expected, the absolute mean value of the residuals is far larger in the case of ensemble averaging, than in the previous case of low-pass filtering.

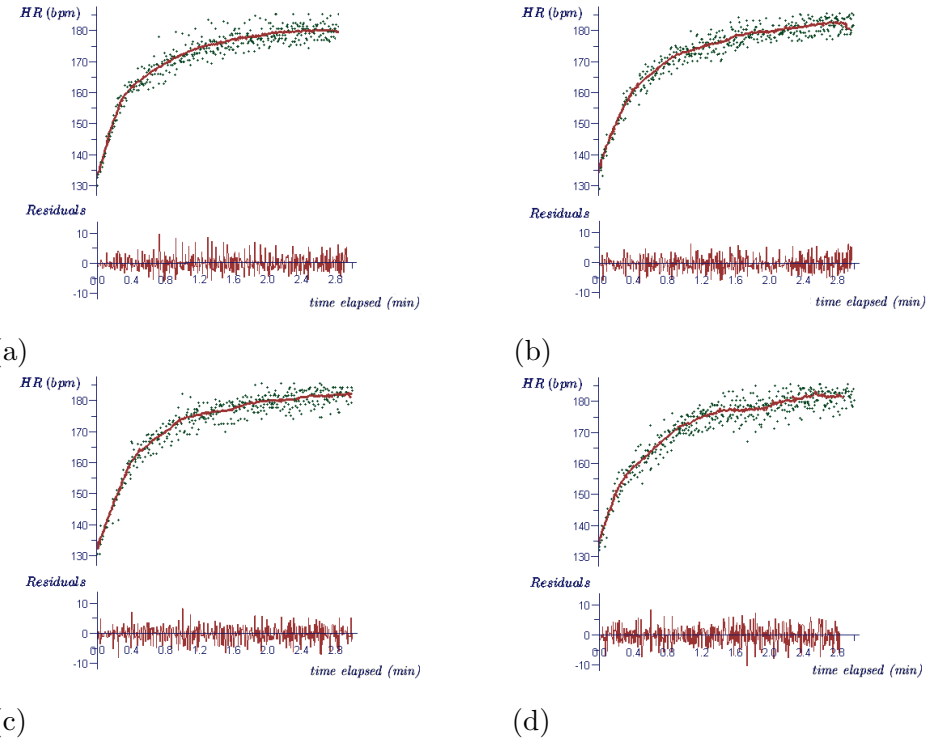


Figure 9: On-transient time series data (green points) and the corresponding 30-rolling averaged curves (red curves). The residuals give the vertical distance between the time series data and the time series resulting from a 30-rolling averaging of the data. a) time interval 1, mean residuals value: 0.021 (b) time interval 2, mean residuals value: -0.003 (c) time interval 3, mean residuals value: -0.006 and (d) time interval 4, mean residuals value: -0.021.

7.3 n-Rolling averaging

Following the rolling averaging methods currently used in the field of exercise physiology, an n-rolling averaging method with $n = 3$ [10] was performed on the data sets of figure 5. Figure 8 shows the superposition of the curves resulting from 3-rolling averaging of the data (sequences $\{r_k^{3-1}\}_{k=0}^{N-1}$, $\{r_k^{3-2}\}_{k=0}^{N-1}$, $\{r_k^{3-3}\}_{k=0}^{N-1}$ and $\{r_k^{3-4}\}_{k=0}^{N-1}$) on the corresponding data points of each interval (sequences $\{s_k^1\}_{k=0}^{N-1}$, $\{s_k^2\}_{k=0}^{N-1}$, $\{s_k^3\}_{k=0}^{N-1}$ and $\{s_k^4\}_{k=0}^{N-1}$). The time series of the residuals (vertical distance) between the 3-rolling averaging data and the raw data points are also shown in these figures. Note that, as can

be easily seen in figure 8, such a 3-rolling method does not provide a smooth curve describing the data, as high frequency oscillations are still present in the 3-rolling averaged data set.

The implementation, however, of an n-rolling averaging procedure with $n > 3$ could result in a greater reduction of the high frequency oscillations and a smooth curve free of such oscillations, especially when used repeatedly. As an indication of the results of an n-rolling averaging procedure with $n > 3$, we present in figure 9 the superposition of the curves resulting from 30-rolling averaging of the data (sequences $\{r_k^{30-1}\}_{k=0}^{N-1}$, $\{r_k^{30-2}\}_{k=0}^{N-1}$, $\{r_k^{30-3}\}_{k=0}^{N-1}$ and $\{r_k^{f30-4}\}_{k=0}^{N-1}$) on the corresponding data points of each interval (sequences $\{s_k^1\}_{k=0}^{N-1}$, $\{s_k^2\}_{k=0}^{N-1}$, $\{s_k^3\}_{k=0}^{N-1}$ and $\{s_k^4\}_{k=0}^{N-1}$). In figure 9 the time series of the residuals (vertical distance) between the 30-rolling averaging data and the raw data points are also shown. No patterns can be observed in the time series of the residuals. In addition, the mean value of the residuals is low, as in the case of Fourier low-pass filtering.

7.4 Fourier low-pass filtering and n-rolling averaging

When the n-rolling averaging method is compared to the low-pass filtering with cut-off frequency 0.13 Hz, there is an optimum area of values for n, which we have observed to be $n \geq 15$, $n \leq 35$. This means that if an n-rolling averaging method is to be implemented instead of a Fourier low-pass filtering, similar results will be obtained if the value of n in a rolling procedure lies within this optimal interval of values.

In a more general case, this optimal interval for the values of n obviously depends on the number of data and the desired degree of smoothing, as it is the case with the optimal value for the cut-off frequency.

8 The basic response pattern of a highly irregular data set

We present the application of Fourier low-pass filtering to the irregular and rapidly oscillating data set of Protocol II of 4(b). Figure 10 shows the results. The cut-off frequency was approximately 0.27 Hz. As expected, we can see that the calculated Fourier curve (thick curves superimposed on the data points in figure 10) provides an excellent description of the underlying dynamics of the time series data.

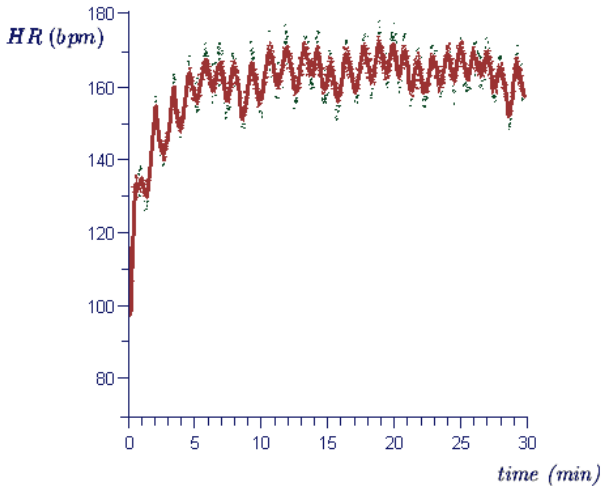


Figure 10: Time series data-protocol II (green points) and the corresponding smoothed Fourier low-pass curve (red curve). Cut-off frequency: 0.27 Hz.

It should be noted here that the application of an n-rolling averaging procedure would yield a similar response pattern as the one shown in figure 10 when the appropriate value of n is chosen interactively.

9 Discussion

An ensemble average will either underestimate, or overestimate, the basic response pattern, depending on whether the amplitude and/or concentration of the high frequency oscillations is larger below, or above, the basic response pattern. When the purpose of the study is to find the underlying dynamics, methods of treating the data such as averaging over repeated bouts of exercise, will cause a problem as they can introduce features not existing in the raw data set. Also much confusion could arise, as averaging over repeated experiments can introduce sharp deviations from the underlying curve. This has important implications when deciding on the smoothness of a model and the underlying physiological reasons for non smooth behavior. A good example of such implications is that the three phases of the so-called 3-phase model can only be clearly seen following averaging over multiple data sets, as in a single raw un-averaged data set the spread of individual time series

data is too large to identify such features [9].

As discussed in section 8 a Fourier low-pass filtering or an n-rolling averaging method is also able to provide basic response patterns of rapidly oscillating data sets, in cases where the ensemble averaging of more that one exercise bouts becomes impossible to implement.

For the case of the data of the rate of change of oxygen uptake ($\dot{V}O_2$), we hypothesize, looking at the published data which has not been averaged [9] that, as with the heart rate data, the curve smoothing methods described here successfully provide a smooth curve describing the underlying response pattern of such data.

This work was supported by the programs Ramón y Cajal 2004 and I3 2006.

References

[1] P. O. Astrand, K. Rodahl, H. A. Dahl, S. B. Stromme, *Text book of Work Physiology: Physiological of Bases of Exercise* (Human Kinetics, 2003).

[2] D. E. Martin, and P. N. Coe, *Better training for distance runners* (Champaign IL: Human Kinetics, 1997).

[3] B. A. Franklin, T. Noakes, O. A. Brusis *Active Cardiac Rehabilitation* (Polar Electro Oy, 2001).

[4] J. Rippe, *Target your Fitness and Weight Management Goals* (Polar Electro Oy, 2000).

[5] K. Kukkonen-Harjula, R. Laukkanen *Precision Weight Management* (Polar Electro Oy, 1998).

[6] K. Wasserman, J. E. Hansen, D. Y. Sue, R. Casaburi, and B. J. Whipp, *Principles of exercise testing and interpretation, including pathophysiology and clinical applications* (Lippincott Williams and Wilkins, 1999).

[7] D. Linnarsson, Acta Physiol. Scand. Suppl. **415**, 1 (1974).

[8] Polar Electro Oy, *Polar precision performance, user guide* (1999).

[9] F. R. Borrani, R. Canadau, G. Y. Millet, S. Perrey, J. D. Fuchslocher, and J. D. Rouillon, J. Appl. Physiol. **90**, 2212 (2001).

[10] S. E. Bearden, and R. J. Moffatt, J. Appl. Physiol. **88**, 1407 (2000).

[11] S. E. Bearden, P. C. Henning, T. A. Bearden, and R. J. Moffatt, Eur. J. Appl. Physiol. **91**, 586 (2004).

[12] T. J. Barstow, and P. A. Mole, J. Appl. Physiol. **71(6)**, 2099 (1991).

[13] I. Lamarra, B. J. Whipp, S. A. Ward, and K. Wasserman, J. Appl. Physiol. **62**, 2003 (1987).

[14] H. C. Xing, J. E. Cochrane, Y. Yamamoto, and R. L. Hughson, J. Appl. Physiol. **71(6)**, 2394 (1991).

[15] S. Koga, Ann. Physiol. Anthropol. **9(2)**, 163 (1990).

[16] C. R. Potter, D. J. Childs, W. Houghton, and N. Armstrong, Eur. J. Appl. Physiol. Occup. Physiol. **80(2)**, 118 (1999).

[17] L. Puente-Maestu, J. B. Abad, R. Godoy, J. M. Pérez-Parra, J. M. Gubillo, and B. J. Whipp, Eur. J. Appl. Physiol. **87**, 535 (2002).

[18] H. B. Rossiter, F. A. Howe, S. A. Ward, J. M. Kowalchuk, J. R. Griffiths, and B. J. Whipp, J. Physiol. **528(2)**, 359 (2000).

[19] Y. Yoshiharu, and R. L. Hughson, J. Appl. Physiol. **71(3)**, 1143 (1991).

[20] V. L. Billat, E. Wesfreid, F. Cottin, C. Kapfer, J. P. Koralsztein, S. Bonneau, and y. Meyer, Int. J. Comp. Sci. in Sport **2**, 16 (2003).

[21] A. L. Goldberger, L. A. N. Amaral, J. M. Hausdorff, P. C. Ivanov, C. K. Peng, and H. E. Stanley, Proc. Natl. Acad. Sci. USA. **99(1)**, 2466 (2002).

[22] H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, 1999).

[23] J. Guckenheimer, and P. Holmes, *Nonlinear oscillations, dynamical systems and bifurcations of vector fields* (Springer-Verlag, 1983).

[24] J. R. Stirling, M. S. Zakynthinaki, and B. Saltin, Bull. Math. Biol. **67**, 989 (2005).

[25] A. Krogh, and J. Lindhard, J. Physiol. (Lond). **47**, 112 (1913).

[26] M. Engelen, J. Porszasz, M. Riley, K. Wasserman, K. Maehara, and T. J. Barstow, J. Appl. Physiol. **81(6)**, 2500 (1996).

[27] T. J. Barstow, and P. A. Mole, J. Appl. Physiol. **63**, 2253 (1987).

[28] R. Casaburi, J. Daly, J. E. Hansen, and R. M. Effros, J. Appl. Physiol. **67**, 1106 (1989).

[29] R. Casaburi, T. J. Barstow, T. Robinson, and K. Wasserman, J. Appl. Physiol. **67**, 547 (1989).

[30] M. MacDonald, P. K. Pedersen, and R. L. Hughson, J. Appl. Physiol. **83(4)**, 1318 (1997).

[31] D. H. Paterson, and B. J. Whipp, J. Physiol. London **443**, 575 (1991).

[32] D. C. Poole, Med. Sci. Sport Exerc. **26(11)**, 1335 (1994).

[33] D. C. Poole, T. J. Barstow, G. A. Gaesser, W. T. Willis, and B. J. Whipp, Med. Sci. Sport Exerc. **26(11)**, 1354 (1994).

[34] T. J. Barstow, Med. Sci. Sport Exerc. **26(11)**, 1327 (1994).

[35] V. L. Billat, Brit. J. Sports Med. **4(2)**, 83 (2000).

[36] G. A. Gaesser, and D. C. Poole, Exerc. Sport Sci. Rev. **24**, 35 (1996).

[37] K. Wasserman, B. J. Whipp, and J. Castagna, J. Appl. Physiol. **36**, 457 (1974).

[38] B. J. Whipp, Med. Sci. Sport Exerc. **26(11)**, 1319 (1994).

[39] B. J. Whipp, and K. Wasserman, J. Appl. Physiol. **33**, 351 (1972).

[40] F. Özyener, H. B. Rossiter, S. A. Ward, and B. J. Whipp, J. Appl. Physiol. **53(3)**, 891 (2001).

[41] J. R. Stirling, M. S. Zakynthinaki, J. Sampedro, I. Refoyo (2007), submitted for publication.

[42] B. Grassi, M. C. Hogan, K. M. Kelley, W. G.Aschenbach, J. J. Hamann, R. K. Evans, R. E. Patillo, amd L. B. Gladden, J. Appl. Physiol. **89**, 1293 (2000).

[43] D. C. Poole, G. A. Gaesser, M. C. Hogan, D. R. Knight, and P. D. Wagner, J. Appl. Physiol. **72**, 805 (1992).

[44] H. B. Rossiter, S. A. Ward, V. L. Doyle, F. A. Howe, J. R. Griffiths, and B. J. Whipp, J. Physiol. (Lond). **518**, 921 (1999).

[45] B. J. Whipp, H.B. Rossiter, and S. A. Ward, Biochem. Soc. Trans. **30(2)**, 237 (2002).

[46] J. R. Stirling, M. S. Zakynthinaki, and V. Billat (2007), submitted for publication.

[47] M. S. Zakynthinaki, J. R. Stirling, Com. Phys. Commun. **176(2)**, 98(2007)

[48] R. Bracewell, *Fourier Transform and its Applications* (McGraw-Hill Education, 1999).

[49] O. E. Brigham, *The Fast Fourier Transform and its Applications* (Prentice Hall, 1988).

[50] S. K. Mitra, *Digital signal processing: a computer - based approach.* (McGraw-Hill, 2005).

[51] P. Z. Peebles, *Probability, random variables, and random signal principles*, (McGraw-Hill, 1987).

[52] J. G. Proakis, and D. G. Manolakis, *Digital signal processing: Principles, algorithms and applications* (MacMillan Publishing Company, 1992).

[53] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C: The art of scientific computing* (Cambridge University Press, 1993).

[54] T. Inouye, T. Harper, and N. C. Rasmussen, Nucl. Instrum. Meth. **67**, 125 (1969).

[55] E. L. Kosarev, and E. Pantos, J. Phys. E: Sci. Instrum. **16**, 537 (1983).

[56] D. Sundararajan, *The discrete Fourier transform: theory, algorithms and applications* (World Scientific Publishing, 2001).

[57] M. S. Zakynthinaki, and J. R. Stirling, Sport Wyczynowy bimonthly periodical **11-12**, 59 (2003).

[58] M. Kobayashi, and T. Musha, IEEE Trans. Biomed. Eng. **29**, 456 (1982).

[59] J. P. Saul, P. Albrecht, R. D. Berger, and R. J. Cohen, Comp. Cardiol. **14**, 419 (1988).

[60] Y. Yoshiharu, R. L. Hughson, and J. Peterson, J. Appl. Physiol. **71(3)**, 1136 (1991).

[61] F. Cottin, Y. Papelier, F. Durbin, J. P. Koralsztein, and V. Billat, Eur. J. Appl. Physiol. **87**, 17 (2002).

[62] D. Essfeld, U. Hoffman, and J. Stegemann, J. Appl. Physiol. Occup. Physiol. **56**, 508 (1987).

[63] J. O. Fortrat, D. Sigaudo, R. L. Hughson, A. Maillet, Y. Yamamoto, and C. Gharib, Autonom. Neurosci.: Basic and Clinical **86**, 192 (2001).

[64] V. L. Billat, B. Flechet, B. Petit, G. Muriaux, and J. P. Koralsztein, Med. Sci. Sports Exerc. **31(1)**, 156 (1999).

[65] Polar Electro Oy, *Polar S810i heart rate monitor, user guide* (2005).

[66] J. Karvonen, J. Chwalbinska-Moneta, and S. Säynäjäkangas, Phys. Sport Med. **12(6)**, 65 (1984).

[67] B. Seaward, R. Sleamaker, T. McAuliffe, and J. Clapp, Biomed. Instrum. Technol. **24(1)**, 37 (1990).

[68] R. Godsen, T. Carroll, and S. Stone, Med. Sci. Sports Exerc. **23(4)**, Suppl. (1991).

[69] O. V. De Vel, IEEE Trans. Biomed. Eng. **33(11)**, 715 (1984).

[70] G. M. Friesen, T. C. Jannett, M. A. Jadallah, S. L. Yates, S. R. Quint, and H. T. Nagle, IEEE Trans. Biomed. Eng. **37(1)**, 85 (1990).

[71] P. S. Hamilton, and W. J. Tompkins, IEEE Trans. Biomed. Eng. **33(12)**, 1157 (1986).

M. S. Zakynthinaki
Departament de matemàtiques
Universitat Autònoma de Barcelona
mzakyn@mat.uab.cat

J. R. Stirling, M. Sillero, J. Sampedro, I. Refoyo
Facultad de Ciencias de la Actividad Física y del Deporte
Universidad Politécnica de Madrid
jr.stirling@upm.es
manuel.sillero@upm.es
javier.sampedro@upm.es
ignacio.refoyo@upm.es

Publicat el 19 de juliol de 2007