Exploiting sensor data in professional road cycling: personalized data-driven approach for frequent fitness monitoring

Leeuw, A.W. de; Heijboer, M.; Verdonck, T.; Knobbe, A.J.; Latre, S.

Data Mining and Knowledge Discovery
https://doi.org/10.1007/s10618-022-00905-5

Exploiting sensor data in professional road cycling: personalized data-driven approach for frequent fitness monitoring

Arie-Willem de Leeuw[1] · Mathieu Heijboer[2] · Tim Verdonck[1] · Arno Knobbe[3] · Steven Latré[1]

Received: 22 April 2022 / Accepted: 29 November 2022

Abstract
We present a personalized approach for frequent fitness monitoring in road cycling solely relying on sensor data collected during bike rides and without the need for maximal effort tests. We use competition and training data of three world-class cyclists of Team Jumbo–Visma to construct personalised heart rate models that relate the heart rate during exercise to the pedal power signal. Our model captures the non-trivial dependency between exertion and corresponding response of the heart rate, which we show can be effectively estimated by an exponential kernel. To construct the daily heart rate models that are required for day-to-day fitness estimation, we aggregate all sessions in the previous week and apply sampling. On average, the explained variance of our models is 0.86, which we demonstrate is more than twice as large as for models that ignore the temporal integration involved in the heart's response to exercise. We show that the fitness of a cyclist can be monitored by tracking developments of parameters of our heart rate models. In particular, we monitor the decay constant of the kernel involved, and also analytically determine virtual aerobic and anaerobic thresholds. We demonstrate that our findings for the virtual anaerobic threshold on average agree with the results of exercise tests. We believe this work is an important step forward in performance optimization by opening up avenues for switching to adaptive training programs that take into account the current physiological state of an athlete.

Responsible editor: Ulf Brefeld and Albrecht Zimmermann

Mathieu Heijboer, Tim Verdonck, Arno Knobbe and Steven Latré have contributed equally to this work.

✉ Arie-Willem de Leeuw
arie-willem.deleeuw@uantwerpen.be

1 University of Antwerp - imec, Sint-Pietersvliet 7, 2000 Antwerp, Belgium

2 Team Jumbo-Visma, Het Zuiderkruis 23, 5215, MV 's-Hertogenbosch, The Netherlands

3 LIACS, Leiden University, Niels Bohrweg 1, 2333, CA Leiden, The Netherlands

Published online: 16 December 2022

Keywords Sport analytics · Road cycling · Applied machine learning · Performance optimization · Sensor data

1 Introduction

The season of professional athletes primarily consists of competing in races, training and recovery periods. Throughout the year, there are usually a few specific moments at which an athlete aims for an optimal performance. The main challenge of the training staff is to construct a training program that allows for a peak performance of the athlete at their moments of choice.

In the past years, more and more information has become available on the optimization of training programs (Mujika et al. 2018). According to the Fitness-Fatigue model (Bannister et al. 1975), the performance of an athlete is related to the sum of a positive contribution that describes the increased fitness from training adaptations and a negative part due to increased fatigue from exercising. Here, the most important challenge is to find the right balance between the physical stress and recovery (Kellmann 2010). As current models are unable to accurately describe this balance, coaches often rely on previous experiences and intuition (Borresen and Lambert 2009). Therefore, it is important to frequently monitor athletes to make sure that the adaptations to the training program are as planned or to pick up early signs of decreased fitness of an athlete (Bourdon et al. 2017).

Due to the recent technological developments, athlete monitoring has become more and more common practice in professional sports (Taylor et al. 2012). The main focus is on measuring changes in fitness or fatigue with respect to modification in the training load. Here, it is important to consider both the *internal* and *external* training load. The external training load is the work that is performed by the athlete, e.g., cycling for an hour at 200 W. While this external load is independent of the individual characteristics of an athlete (Wallace et al. 2009), the internal training load quantifies how the external training load is experienced by the athlete. Common measures for the internal training load are the heart rate or the Session Rating of Perceived Exertion (Foster 1998).

Collecting data about the external and internal training loads alone is not sufficient for proper athlete monitoring. Only after performing sport-specific analyses and adding contextual information in close collaboration with coaches, the data can be properly interpreted and be of added value for optimizing training programs (Thornton et al. 2019). Moreover, although the internal and external training loads are important separately, it is usually the relationship between both quantities that provides the crucial information about the fitness of an athlete (Halson 2014).

There are several complications in finding accurate connections between the external and internal training loads. Most importantly, each individual responds differently to training activities and therefore a personalized approach is preferable (Wallace et al. 2009). However, in most cases there is not enough data to apply an individualized approach and groups of athletes need to be considered together to find significant dependencies. Moreover, for a complete understanding of the relationship between the external and internal training loads it is important to also consider heavy efforts. However, this is very demanding for athletes and thus difficult to include in a well thought-out training program, especially around competition events (Meeusen et al. 2013).

Due to the aforementioned difficulties, detailed athlete monitoring is not possible in many sports. However, professional road cycling is an exception where there are many opportunities for frequent and accurate athlete monitoring. The first reason is that both the rider and their bike are equipped with sensors. As most professional cyclists ride their bike for hours a day and sensors usually detect many different quantities every second, ranging from terrain information to effort-related variables such as speed and produced power, an enormous amount of data is collected about the external training load and the corresponding response of the body. Second, the training program is dominated by cycling activities and therefore it is straightforward to obtain a complete picture of all efforts. This combination makes professional road cycling an ideal sport for investigating frequent athlete monitoring and even opens up the possibility to perform individualized analyses.

Although the abundance of available data for every individual professional cyclist stimulates the application of data science techniques, the scientific literature on data science and athlete monitoring in professional road cycling is limited. Most studies that do consider data-driven approaches in professional road cycling use publicly available data on for example competition results. There are studies that investigate talent identification (Bulck et al. 2021; Janssens et al. 2022), forecast race results (de Spiegeleer; Kholkine et al. 2021) or predict power output during races (Kataoka and Gray 2019). In other works, collaborations with professional cycling teams are established to predict the heart rate (Hilmkil et al. 2018), investigate the effect of different training programs for groups of athletes (Karetnikov 2019), or study the dependence between the performance of a specific rider and the altitude profile of the stage by using an individualized approach (de Leeuw et al. 2020). Recently, sensor data is also used for performance monitoring in cycling. Several machine learning techniques have been applied to construct a recommendation module for increasing the cycling performance (Demosthenous et al. 2022) and automatic handslings performance monitoring is studied during a madison in track cycling (Steyaert et al. 2022). Whereas these studies focus on specific performance measures, e.g., an increased speed with barely affecting the heart rate, we here consider fitness monitoring by extracting physiological characteristics of riders from sensor data. Up to the best of our knowledge, this has not been addressed in professional cycling using a personalized and data-driven approach, which is the type of analysis we consider in this work.

In this paper, we describe the results of a joint project between professional cycling team Team Jumbo-Visma and researchers from the University of Antwerp and Leiden University. This project is initiated by the coaching staff of the professional road cycling team and the ultimate goal is retrieving the fitness of an individual professional cyclist on a daily basis without performing heavy and disrupting exercise tests. By frequently monitoring the fitness of professional road cyclists, coaches can make a transition from static predetermined training programs to more dynamic programs that take into account the current physiological state of the athlete.

In addition to the clear practical relevance for sport practice, this work is also relevant from a scientific perspective, as we will apply several data science techniques to collections of time-series data that will give new insights that are also applicable in other domains. At the core of our approach, i.e., the construction of an accurate model that describes the relationship between the heart rate and the external training load, we encounter three main data science challenges. First, we need some form of aggregation to model the heart rate response to exercising as this depends on the recuperation of previously delivered efforts. Here, we will consider the convolution of the pedal power with a kernel and introduce a *recuperation time* parameter of the athlete. The exact shape of this kernel is not immediately clear as the relationship between the external training load and heart rate response is athlete-specific and non-trivial. The second challenge is about the data. To retrieve the fitness of an athlete, it is important to accurately model the response at small, moderate and high external training loads. As professional cyclists do not touch upon each of these areas during a single ride, we need to combine several training sessions. Therefore, we need to find an optimal balance between a large enough range of the pedal power and a model that is still able to describe daily variations. Moreover, we have imbalanced data as the moderate exercise intensities are most common. As there is ample research about handling imbalanced data sets in regression settings (Krawczyk 2016), this is an interesting test case for obtaining more insights in this research area. The third and final challenge is about extracting the fitness from our heart rate model that is based on data from daily practice. Here, we will connect with exercise tests and demonstrate that with our method aspects of fitness can be monitored.

In summary, our work has the following key contributions:

- The construction of a personalized heart rate model that describes the relationship between heart rate and external training load during exercise.
- Methods for convolution of time series data to describe the athlete-specific response of the heart rate to current and previous delivered efforts.
- A general approach for fitness monitoring in professional road cycling via the recuperation time, heart rate performance curve, as well as virtual aerobic and anaerobic thresholds without performing exercise tests or making adjustments to the training program.
- The presentation of results for world-class professional cyclists.

In the remainder of this article, we start by discussing the state of the art of athlete monitoring in professional road cycling in Sect. 2. Then, we describe the data and our modeling approach in Sect. 3. In Sect. 4, we present results from the heart rate model and the experiments for finding the optimal parameter settings. Moreover, we describe several options for athlete-specific fitness monitoring and we validate our results for the virtual thresholds. Finally, we discuss the results and conclude in Sect. 5.

2 Athlete monitoring in road cycling

The fitness of professional athletes is usually determined by performing an exercise test in a controlled environment. The aim of these tests is finding the key determinants that influence their performance. For endurance athletes, important performance indicators are the VO_2-max, lactate threshold and efficiency (Joyner and Coyle 2008).

In cycling, lactate testing is an important tool for assessing the body capacity for performance in a rather straightforward and inexpensive way. Here, lactate values at fixed exercise intensity (e.g. power) are measured. These values and the manner in which lactate concentration is building up when increasing exercise intensity are good indicators of the performance level of cyclists (San Millán et al. 2009; Wasserman et al. 1981; Karlsson and Jacobs 1982). Typically, at low exercise intensities, the lactate concentration will slowly increase. If the increase of lactate concentration starts to be significant, we are at the so-called first lactate threshold (LT1) or aerobic threshold. Theoretically, this is defined as the maximum exercise intensity that can be maintained indefinitely. If the exercise intensity is further increased, the lactate concentration will build up more rapidly. At the second lactate threshold (LT2) or anaerobic threshold, there is a balance between the amount of lactate that is produced and can be cleared (Binder et al. 2008). These are exercise intensities that can be maintained by the athlete for roughly an hour. Efforts with even higher exercise intensities can only be sustained for short periods.

Although retrieving this information is insightful, these testing procedures are not performed frequently. For athletes, these tests are demanding and this makes it cumbersome to include them into the training programs. Moreover, results found in a laboratory usually differ from performances observed in the field. As a result, professional cyclist are usually monitored by directly using the data that is collected during their rides. Here, the main focus is on the exercise intensity and in particular on the produced power (Leo et al. 2021). A frequently used feature is the so-called power duration curve (Hunter et al. 2019). This curve gives the maximum power that can be maintained by the cyclist for any given time period. To quantify the effects of a particular training period, this curve can be determined before and after a training camp. For example, the coach could observe that the rider is able to produce more power for short time periods, but a lower power for longer efforts.

In addition to the external training load, coaches also monitor the internal training load. Here, the most common measure is the heart rate, since it is relatively cheap to determine accurately. Although the heart rate can be affected by many factors, for professional athletes global changes over a longer period are assumed to be caused by adaptations to training programs (Lamberts et al. 2010; Buchheit 2014). However, there is no unambiguous evidence whether certain changes in heart rate are related to an increased or decreased fitness (Bellenger et al. 2016) and therefore adding contextual information is crucial (Kellmann et al. 2018).

Currently, the monitoring approaches are solely based on either the internal or external training load. However, it is rather the coupling between both loads that is relevant for determining the physical fitness, i.e., the duration of certain exercise intensities are only relevant if information on how these are perceived by the cyclist is included.

In this study, we present an approach for athlete monitoring that combines the advantages of the aforementioned methods, without the presence of the disadvantages, i.e., we will frequently determine the fitness without any additional burden for the cyclist. Recent studies use a non-invasive methodology with a detrended fluctuation analysis of the heart rate variability to determine estimates of the lactate thresholds (Mateo-March et al. 2022). Here, the authors found an agreement with the first lactate threshold and also obtained large correlations with the second lactate threshold. Alternatively, we will present an approach based on daily heart rate models that couple the external training load to the internal training load. Hereby, we cannot only determine virtual aerobic and anaerobic thresholds and monitor their daily fluctuations, but also monitor the fitness via the recuperation time, i.e., the length of the time window in which efforts affect the heart rate.

3 Methods

3.1 Data

We consider three world-class general classification riders of Team Jumbo-Visma, dubbed rider A, B and C. The data consists of a large variety of characteristics that are collected through a cycling computer once a second during all of the rides in a certain time period. For rider A, we consider a period of two and a half years during which the sensor data of 791 sessions is gathered. The data set of rider B consists of 301 sessions collected throughout a single year. Finally, we consider 120 sessions of rider C that were all completed in a time span of four months.

For every session, we have many distinct attributes. Most sessions are on different dates, but there are a few cases where there are multiple sessions on a single day. For example, if a cyclist is participating in a time trial, we can have a session with a reconnaissance of the course, one with the warm-up before the race and a race file. Since there is sufficient time between all sessions, each ride is considered separately.

The main attributes of the data are as follows:

- Power [Watt]
- Cadence [rpm]
- Heart rate [bpm]
- Covered distance [km]
- Duration [sec]
- Speed [km/h]
- Altitude [m]
- GPS coordinates [longitude, latitude].

The other attributes characterize more detailed information about the power that is produced by the cyclist, e.g, the power produced by the left or right leg at different angles of the pedal. These attributes are not considered in this study.

For some rides, the produced power or heart rate is not recorded. During many competition events, there is no heart rate information, as wearing a heart rate chest strap gives some discomfort. Moreover, there are some cases where during (parts of) bike rides, the sensors were not functioning properly. Fortunately, this occurs infrequently and the riders cycle for hours per session. In this abundance of data, we removed the times without heart rate information and the instances where the convolution of the pedal power could not be determined. An overview of the data, including the percentage of missing values, can be found in Table 1.

Table 1 Overview of most important characteristics of the used data

Rider	Number of sessions	Percentage missing values	
		Power (%)	Heart rate (%)
A	791	6.4	9.9
B	301	1.1	1.3
C	120	1.1	0.7

3.2 Modeling approach

We want to determine the fitness of a cyclist by only using data that is collected during competition and training rides. In our approach, we will not consider an entire time period at once as this would incorrectly suggest that the fitness of the cyclist is constant. Alternatively, we will construct our heart rate model iteratively for a shorter period. For each day d, we consider all rides that were within N days prior to d. Typically, $N = 7$ as for professional athletes weekly exercise programs contain the necessary variation for constructing our models. On this subset of the data, we use the heart rate models to determine the fitness of a cyclist on day d. We repeat this procedure for all days to monitor how the fitness of our cyclist is changing throughout the entire period, except for the first N days. Ultimately, we compare the findings of our approach to results during exercise tests in the same time period, as validation of our model.

Our approach consists of two important parts. First, we construct a heart rate model to couple the internal and external training loads. Here, we apply regression techniques to find the relationships between heart rate and variables that characterize the external training load. In the second part, we extract characteristics from our heart rate model to determine information about the fitness of the cyclist. In the remainder of this chapter, we will discuss the details of both steps.

3.3 Heart rate model

From experiments in controlled settings, we know that the heart rate is influenced by many factors. For example, the phenomenon *cardiovascular drift* describes the increase in heart rate for prolonged exercises at constant power (Hamilton et al. 1991), and the heart rate is positively correlated with power output (Grazzi et al. 1999). Therefore, most attributes in our data collection are predictor variables for a heart rate model.

However, in our case the situation is different from these aforementioned experimental studies. In these studies, the dependencies are obtained by only changing a single variable and fixing the values of the others. On the other hand, we are using data that are collected in uncontrolled settings. This implies that many of the attributes are related to each other. For instance, if we straightforwardly consider all rides together for rider A, the Pearson's correlation between the cadence and power output is 0.63. Thus, an approach where all attributes are considered as independent variables would involve many collinearities.

In principle, this is not an issue if the goal is to construct a model with high accuracy. However, to determine the physical fitness of a cyclist, it is crucial to accurately describe the relationship between heart rate and produced power. Therefore, we would obtain incorrect dependencies if we would use a model that includes many correlated attributes and investigate the relationships between power output and heart rate while fixing the value of the other attributes. Hence, we here take a different approach.

3.3.1 Convolved pedal power

As a first step in our methodology, we introduce the *convolved pedal power*. This variable is a weighted average of the produced power in a specific time interval. More specifically, the convolved pedal power at time t for a session s of length l is defined as

$$P^{\mathrm{s}}_{\mathrm{conv}}(t) = \frac{\sum_{i=0}^{w} h(-i) p^{\mathrm{s}}(t-i)}{\sum_{i=0}^{w} h(-i)} := \sum_{i=0}^{w} \tilde{h}(-i) p^{\mathrm{s}}(t-i), \tag{1}$$

where w is the length of the convolution window, $p^{\mathrm{s}} = \{p^{\mathrm{s}}(0), p^{\mathrm{s}}(1), ..., p^{\mathrm{s}}(l)\}$ is the pedal power attribute and *h(x)* is the kernel. Note that we normalised the kernel such that the values of the convolved pedal power can be interpreted as weighted averages during the time window w with $\tilde{h}(-i)$ the corresponding weights. Therefore, the sum of the normalised kernel is equal to 1 and the convolution does not introduce an arbitrary scaling of the pedal power values.

In principle, the kernel function can take any shape. To investigate which kernel functions are relevant, we assume a linear relationship between the heart rate and the convolved pedal power. As explained before, this linear dependency is accurate for a large range of moderate exercise intensities (Grazzi et al. 1999). Although for the construction of the final heart rate model more complex modeling is required, the linear assumption suffices for this purpose.

In this study, we examine two different approaches. First, we consider an empirical kernel where we do not make any assumptions and directly use Eq (1). In this case, we thus find

$$\begin{aligned} \mathrm{HR}(t) &= \mathrm{HR}_0(t) + a \cdot P_{\mathrm{conv}}(t) \\ &= \mathrm{HR}_0(t) + a \cdot \sum_{i=0}^{w} \tilde{h}(-i) p(t-i) \\ &:= \mathrm{HR}_0(t) + a \left[\tilde{h}(0) p(t) + \tilde{h}(-1) p(t-1) + ... + \tilde{h}(-w) p(t-w)\right]. \end{aligned} \tag{2}$$

From this equation, we observe that the specific values of the kernel can be determined by using multiple linear regression with the pedal power at several times as predictor variables. More specifically, we include the pedal power at time t and also at times shifted backwards by an integer value i between 1 and w. Note that this approach resumes many parameters to be optimized and therefore is computationally intensive.

For our second approach, we use the response of the body to change in workloads in controlled environments can be modeled by an exponential kernel (Bunc et al. 1988; Ludwig et al. 2016). Therefore, we will consider the following exponential kernel

$$h(t) = e^{t/\tau}, \tag{3}$$

where τ is the weight of the exponential. Note that the value of τ specifies how previous produced power outputs contribute to the convolved pedal power. For large values of τ, the exponential kernel decreases slowly and the relevant window w in Eq. (1) is relatively large. On the other hand, the kernel decreases faster for small values of τ. In this case, to determine the convolved pedal power at time t, only efforts relatively close to this time are important. In contrast to the first approach, there is only a single parameter that needs to be optimized. Moreover, as we will discuss in the next section, this approach has the advantage that the single parameter τ is interpretable.

3.3.2 Recuperation time

In addition to the mathematical description, there is also a practical interpretation of our exponential kernel approach. As τ is the center of mass of the kernel, we roughly introduce a lag of τ seconds for the response of the heart rate to changing exercise intensity. Moreover, the speed at which the heart rate decays after exercising is related to the physical fitness (Buchheit et al. 2007). In our approach, a rapid heart rate recovery implies that only efforts close to the time for which we model the heart rate are relevant. In this case, only a relative small time window for the convolved pedal power is important for modeling the heart rate. Conversely, large windows are relevant for slow heart rate recovery. Hence, there is a connection between the value of τ and the heart rate recovery. If τ is small, the heart rate recovery is fast and the cyclist on average recuperates fast from previous efforts. This indicates good physical fitness. The opposite holds for large values of τ. Hence, τ is a rough estimator of the recuperation time and has a negative correlation with the physical fitness.

Note that via the parameter τ, we implicitly take into account the effect of all factors different from the power, that influence the heart rate. Thus, the value of τ depends on many factors, such as the altitude on which the rides are executed or weather conditions. For example, distinct value of τ on two specific days can be a consequence of the different altitude at which the bike rides in the previous days were executed. Additionally, the recuperation time is also influenced by the wellness of the cyclist, e.g., the amount of stress or hours slept. Although the effects of all these factors are relevant and ideally we would explicitly model all of them properly, it is rather the combination that is most important for determining the fitness. The strength of our method is that we model the effect of all factors by a single parameter, where taking into account each dependency separately would necessitate the application of a more complicated modeling approach.

3.3.3 Model

The basis of our heart rate model is the convolved pedal power. This variable uses the power and takes into account the time lag between a change in produced power and a corresponding response of the heart rate. Moreover, other influential factors are implicitly taken into account via the shape of the kernel, e.g., the recuperation time of the exponential decaying function. Therefore, it suffices to construct a model where the convolved pedal power is the only independent predictor variable.

For a correct modeling of the relationship between the heart rate and the produced power while cycling, we need to mention that this dependency is rather complex. In accordance with previous work in running (Conconi et al. 1982; Brooke and Hamley 1972), the relationship between the heart rate and power is typically described by an 'S-shaped' curve. At low power, the increase is usually rather small. This is followed by a linear increase for moderate power values and finally the heart rate reaches a plateau if the power is even further increased.[1]

To model this behaviour and especially the relationship at high exercise intensities, we set

$$\mathrm{HR}(t) = \mathrm{HR}_0(t) + a \cdot P_{\mathrm{conv}}(t) + b \cdot [P_{\mathrm{conv}}(t)]^2 + c \cdot [P_{\mathrm{conv}}(t)]^3 \,, \tag{4}$$

where $\mathrm{HR}(t)$ is the heart rate at time t and $\mathrm{HR}_0(t)$ denotes the intercept. Moreover, the three parameters a, b and c are coefficients for the linear, quadratic and cubic dependencies between the heart rate and convolved pedal power, respectively. In addition to the shape of the kernel, we thus have four more parameters in our approach that need to be optimized.

3.3.4 Sampling

The use of sensor data from daily practice leads to another complication. During most rides, a cyclist is only covering a narrow range of exercise intensities. Moreover, most of the time, a cyclist is performing his training sessions at moderate exercise intensities. Therefore, it is usually not possible to obtain reliable information about the heart rate power model for low, moderate and high exercise intensities if only one session is considered. However, later on we will demonstrate that for the determination of the physical fitness of a cyclist, it is important to accurately describe the relationship at all exercise intensities. In particular, we need sufficient data at high exercise intensities for estimating the virtual anaerobic threshold. Therefore, we combine training sessions on subsequent days and construct the heart rate model based on the sensor data of all training sessions that were completed in a small time period. More specifically, for every day d, we consider all training sessions that were completed N days prior to d.

Although this enlarges the range of values for the convolved pedal power, we still encounter that the low and high exercise intensities are highly underrepresented. This is also illustrated in Fig. 1, where we observe that especially the large values

[1] Although this is the most common pattern for healthy persons, different relationship are also observed. For example, there could be a linear relationship for the entire range of exercise intensities or even an upward inflection at large power outputs (Hofmann et al. 1997).

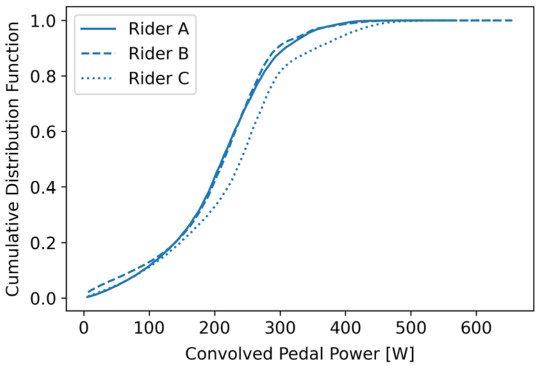


Fig. 1 Cumulative probability density function for the convolved pedal power if we combine all training sessions of the cyclists. As an example, we considered a recuperation time $\tau = 40$ s and considered a convolution window of 120 s. We observe that high values for the convolved pedal power are scarce for all three cyclists

of the convolved pedal power are underrepresented in our entire data collection for each cyclist. For example, we observe that for Rider A around 90% of all values of the convolved pedal power are smaller than 300 W, which is a relatively moderate exercise intensity for this cyclist. The consequence of this skewness is that applying regression to all data is not appropriate to construct a reliable model at large values for the convolved pedal power. In this case, the few points in this area would barely influence the sum of squared errors, which makes the model unreliable for these high exercise intensities.

Here, we address this issue by applying sampling. Our strategy is based on the values of the convolved pedal power. This approach is preferred over sampling on the target variable as this might interfere with the randomness that is present in the target variable. In total, our strategy consists of the three following steps

1. Round the values of the convolved pedal power to the nearest integer.
2. Determine the frequency f_p of each value of the convolved pedal power.
3. Apply undersampling where the probability of a data point to be included in our sampled data set is inversely proportional to f_p.

After performing this procedure, we lower the relative presence of more common values and increase the proportion of underrepresented values. Thereby, we improve the balance in our data set.

3.3.5 Summary

Our approach consists of several parts. To construct a heart rate model for a day d with sufficient data available and the exponential decaying kernel as defined in Eq. 3, we take the steps as described in Algorithm 1.

As the differences in accuracy of our heart rate model are small for minor variations in the shape of the kernel, a different under-sampling occasionally results into other values for the recuperation time and model coefficients. For each day, we therefore execute our algorithm for ten under-sampled strategies with each a different random seed. The day-specific fitness parameters, i.e., the recuperation time and the virtual thresholds that are discussed in the next section, are the averages over these ten results.

Algorithm 1 Heart Rate model

Combine data from the rides within N days prior to d
for $\tau \in [0, 120]$ **do**
  Apply convolution to time series to produce convolved pedal power
  Apply proportional under-sampling strategy on data set
  Split training data set into 10 folds
  for $i \in [1, 10]$ **do**
    Select fold i as test set
    Fit model on remaining 9 folds with optimal parameters $\mathrm{HR}_0(t)$, a,
    b and c of Eq. (4)
    Determine explained variance of the model on the test set
  end for
  Determine explained variance for τ, averaged over 10 folds
end for
Select recuperation time τ that optimizes cross-validated model fit
Construct heart rate model on the entire sampled data set

3.4 Fitness monitoring

As a first step, the heart rate model can be used to obtain a global picture of the fitness of the cyclist. By comparing the relationship between the heart rate and convolved pedal power on different days, we can determine an overall picture of the physiological changes. For example, we can monitor the changes in heart rate at low, moderate and high exercise intensities throughout a training period to examine the physical adaptations to the training programs. Additionally, the model can be compared to a benchmark period in a previous season.

To obtain more concrete information about the physical fitness of the cyclist, we consider specific characteristics of the heart rate convolved pedal power curve. As mentioned before in Sect. 3.3.3, in most practical cases this curve is S-shaped. In this case, there are two points where there is a deflection from the linear behavior. We will determine the coordinates of these points. Then, we connect the deflection points at low and high intensities with the aerobic and anaerobic threshold, respectively. Note that this is analogous to previous works on heart rate performance curves that consider the heart rate response at specific values of the produced power (Hofmann and Pokan 2010), although we here consider the relationship to the convolved pedal power instead.

We have the task of finding the deflection points of the heart rate model as given by Eq. (4). Therefore, we introduce the three parameters P_{min}, P_{lin} and P_{max}. Here, P_{min} and P_{max} are the minimum and maximum values for the convolved pedal power for which the heart rate model is reliable. In Sec. 4.5.1, we will elaborate on the determination of these two parameters in more detail. Moreover, P_{lin} is the average of the minimal and maximal values of the convolved pedal power and therefore typically a moderate aerobic exercise intensity. We define the deflection points to correspond to the points on the S-curve where the perpendicular distance to the straight lines between P_{min} and P_{lin} or P_{lin} and P_{max} has a maximal value (Cheng et al. 1992). This procedure is sketched in Fig. 2.

There are several methods to determine the deflections points of these type of curves. Since the analytical function describing the S-curve is known, we can determine the

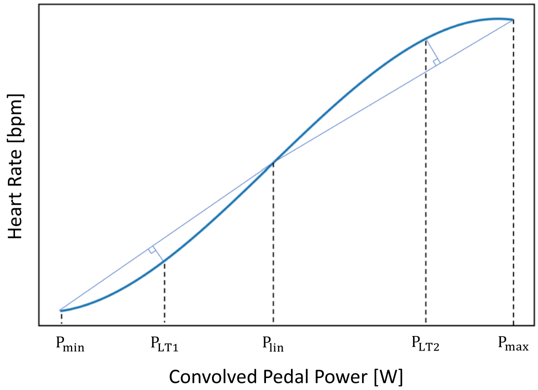


Fig. 2 Illustration of the deflection points P_{LT1} and P_{LT2} in an S-curve for the relationship between the heart rate and the convolved pedal power. The deflection points are defined as the points on the S-curve that have the largest perpendicular distance to the straight lines drawn between P_{min} and P_{lin} or P_{lin} and P_{max}, respectively

coordinates of the deflection points analytically. For this, we observe that for all values of the convolved pedal power between P_{min} and P_{lin}, the deflection point is the only point at which the slope of the S-curve is equal to the slope of the straight line between P_{min} and P_{lin}. As the relationship between the heart rate and convolved pedal power is given by Eq. (4), the first deflection point can be found by solving for P_{LT1},

$$a + 2b \cdot P_{LT1} + 3c \cdot (P_{LT1})^2 = \alpha_{LT1}, \tag{5}$$

where α_{LT1} is the slope of the straight line connecting P_{min} and P_{lin}. More specifically,

$$\alpha_{LT1} = [\mathrm{HR}\,(P_{lin}) - \mathrm{HR}\,(P_{min})] \,/\, [P_{lin} - P_{min}]\,, \tag{6}$$

with $\mathrm{HR}\,(P^*)$ the heart rate at P^*. Hence,

$$P_{LT1} = \frac{-2b \pm \sqrt{4b^2 - 12c \cdot (a - \alpha_{LT1})}}{6c}. \tag{7}$$

Note that there are two solutions. However, there is only one value inside the range $[P_{min}, P_{lin}]$ which corresponds to the value for the convolved pedal power at the first deflection point. By following an analogous reasoning, the second deflection point is the value for the convolved pedal power in the range $[P_{lin}, P_{max}]$ that satisfies

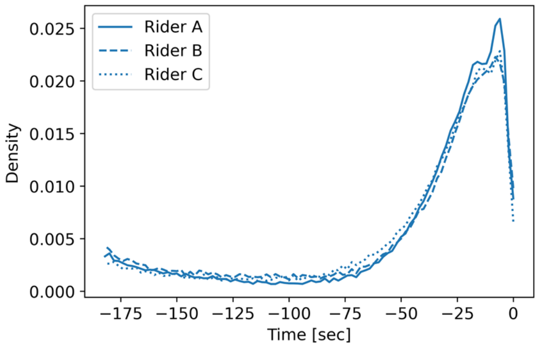
$$P_{LT2} = \frac{-2b \pm \sqrt{4b^2 - 12c \cdot (a - \alpha_{LT2})}}{6c}, \tag{8}$$

with

$$\alpha_{LT2} = [\mathrm{HR}\,(P_{max}) - \mathrm{HR}\,(P_{lin})] \,/\, [P_{max} - P_{lin}]\,. \tag{9}$$

To obtain the values of the heart rate at the deflection points, the values for the convolved pedal power given by Eqs. (7) and (8) can be substituted in Eq. (4).

Fig. 3 The normalised empirical kernel $\tilde{h}$ for all data of each of the three riders, separately



4 Experimental results

In this section, we first discuss the experimental justification of specific choices in our methodology and the parameter settings. Then, we deliberate on the results of our heart rate model and describe the options for fitness monitoring.

4.1 Convolved pedal power

Here, we will compare the two approaches discussed in Sec. 3.3.1 and elaborate on our choice for taking an exponential kernel with recuperation time τ.

To compare both approaches, we consider all data of each rider separately. Per rider, we use 80% and 20% as training and test sets, respectively. The kernel parameters are obtained by applying 10-fold cross validation on the training set. The accuracy of the kernel is assessed by determining the explained variance of the found model on the test set.

In our first approach, we consider the empirical kernel as described by Eq. (2) and thus perform multiple linear regression with a window length w of 180 s. In Fig. 3, we display the empirical kernels for our three riders. Apart from some subtle differences, we find similar behavior for all riders. We observe that the normalised kernel increases for the most recent past. Hereafter, it reaches a maximum value around 7 seconds and then decreases for a long time. Finally, it marginally increases for large negative times. Note that according to Eq. 2, the value of $\tilde{h}(-a)$ is equal to the coefficient for pedal power at time $t - a$ for all $a \in [0, 180]$.

From this experiment, we observe that for all our cyclists the normalised kernel has a maximum value at a time different from zero. This implies that for the riders considered, on average the pedal power that is produced 7 s prior to t is most important for predicting the heart rate at a specific time t. Interestingly, the pedal power at time t is less important. This phenomenon can be explained from a physiological perspective as it will take some time before the body responds to a change in exercise intensity. This clearly shows that for modeling the heart rate at a specific time, it is important to include the past exertions, which in our approach is done by using the convolved pedal power.

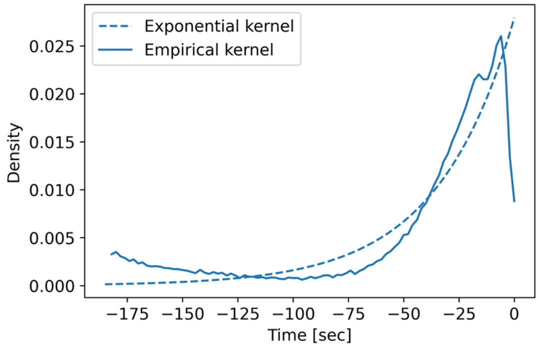


Fig. 4 The normalised kernel $\tilde{h}$ for all data of Rider A for our two approaches. The solid line corresponds to the empirical kernel after performing multiple linear regression for modeling the heart rate at time t with the pedal power at time t and previous times as predictor variables. The exponential decaying function is displayed by the dashed line

Table 2 Comparison of the results for our two approaches for the kernel as discussed in Sect. 3.3.1

Rider	Accuracy (R^2)		Recuperation time (s)
	Empirical kernel	Exponential kernel	
A	0.768	0.758	35
B	0.763	0.756	41
C	0.839	0.834	41

For each rider, we selected all data and determined the rider-specific characteristics of the kernels that optimize the accuracy of the heart rate model. We display the accuracy as explained variance R^2 and also specify the recuperation time of our exponential kernel

For our second approach, we determine the convolved pedal power by using an exponential decaying kernel with a recuperation time τ, as is described by Eq. (3). We again find similar results for our riders. As an illustration, we display the result of both approaches for rider A in Fig. 4. Although the behavior of the normalised kernel is different for the most recent history, we find that on the whole, an exponential function is still a good approximation.

In Table 2, we show the detailed information of both approaches for our three riders. We find that the accuracy of both approaches is very similar. However, compared to the first approach, the exponentially decaying kernel is analytical and has a single interpretable parameter. Additionally, as explained in Sect. 3.3.3, with the exponentially decaying kernel it is also straightforward to model the non-linear behaviour of the relationship between heart rate and produced power. Therefore, we opt for this approach from here on. Note that the exponential function decays quite fast, i.e., the recuperation time is around 40 seconds. Therefore, we can restrict ourselves to a relative small convolution window w of 120 s to save computation time.

4.2 Parameter settings

As explained before, the lack of variation during most training sessions forces us to combine several training sessions to construct an accurate heart rate model for low, moderate and high values of the convolved pedal power. Thus, for a given day d we consider all training session completed in the N days before d. Hereby, we need to find

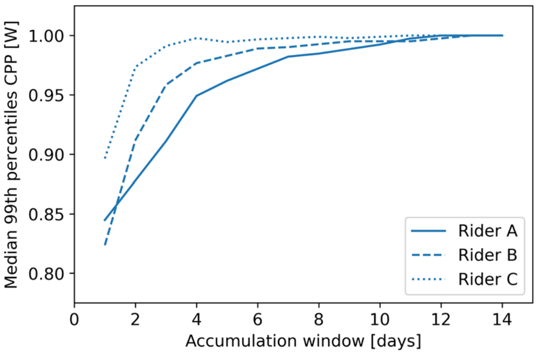


Fig. 5 Median of the 99th percentiles for the distributions of the convolved pedal power (CPP) for all days in our data collection as a function of the accumulation window. For each rider, we normalised the values with respect the value of the median for an accumulation window of 14 days. Here, we consider a recuperation time $\tau = 40$ seconds and a convolution window of 120 s

a proper value for this accumulation window. More specifically, we will discuss how many days prior to d the training session should be included. As discussed previously, we here specifically need to focus on the presence of data points at large values of the convolved pedal power (in other words, large exertions).

To investigate the availability of data at large values of the convolved pedal power, we performed the following experiment. First, we select a certain accumulation window. For a given day d, this accumulation window is defined as the number of the days N prior to d for which the training session data is included in the data set for this day. For example, we only include data of the day before d if we set $N = 1$. After choosing this accumulation window, we select a day d and calculate the convolved pedal power for each data point. Here, we considered $\tau = 40$ s, which is a typical value for the recuperation time that is defined in Eq. (3). Then, we determine the 99th percentile of all values of the convolved pedal power. We repeat this procedure for all days in our entire data collection.

In Fig. 5, we display the median of the 99th percentiles of the convolved pedal power as a function of the accumulation window for our riders. We find that this measure is equal to small values of the convolved pedal power if we consider a small accumulation window. This agrees with our claim that in most training session a cyclist only rides at relatively small values of the convolved pedal power. However, we observe that the value for the median increases if we increase the accumulation window. This nicely demonstrates that the training sessions of cyclists are diverse and larger values of the convolved pedal power are explored if we consider all training sessions in a large enough time period. More specifically, we find that the median of the 99th percentiles of the convolved pedal power rapidly increases up to an accumulation window of 4 days, after which it rises more slowly. Based on the results of Fig. 5, there is no clear-cut value for the accumulation window. However, a slightly different value for the window will not have very large consequences for the results our approach. Therefore, we set the value for the accumulation window based on practical arguments. We take $N = 7$ to allow for a straightforward connection with training programs that are usually planned on a weekly basis and to obtain sufficient data at large values of the convolved pedal power.

As the ultimate goal of this study is to determine the deflection points of the curve that describes the relationship between heart rate and convolved pedal power, we need

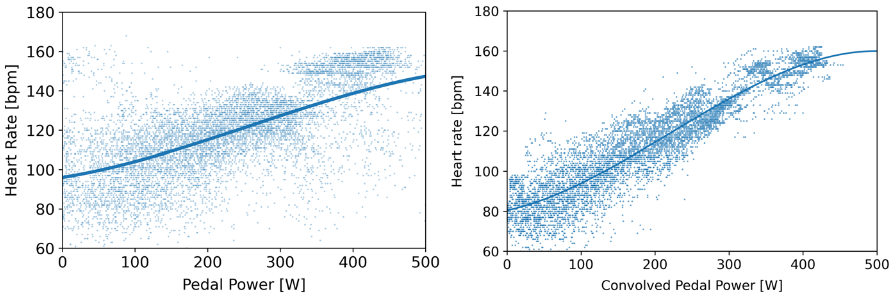


Fig. 6 The heart rate models (solid lines) for rider A as a function of the pedal power (left) or convolved pedal power (right) for a day chosen at random. Each point corresponds to a combination of the heart rate and (convolved) pedal power in the underlying data. The introduction of the convolved pedal power decreases the variance in the data and thereby increases the explained variance R^2 of the heart rate model from 0.42 to 0.86

to increase the relative presence of large values of the convolved pedal power. For $N = 7$, the average number of data points is roughly 60,000 for a day in our data collection. Here, we apply under-sampling and take a sample number of 10,000. In this case, the median of the 95th percentile of the convolved pedal power of the sampled data set is approximately equal to the median of the 99th percentile of the convolved pedal power of the original data set that is displayed in Fig. 5. Hereby, we increase the presence of relevant values for the virtual threshold determination to 5%, which we believe is sufficient to obtain reliable results.

4.3 Heart rate model

The core of our approach is a model that describes the relationship between heart rate and pedal power. We determine this model for each day with sufficient data, i.e., we ignore days for which the total number of data points is less than our sample size of 10 000. Note that this also includes holiday breaks. In total, we construct our model for 1229 days. We have 781, 329 and 119 days for riders A, B and C, respectively.

In Fig. 6, we illustrate the importance of considering the convolved pedal power by considering the heart models for rider A, as a function of the (raw) pedal power (left) or the convolved pedal power (right) for a randomly chosen day. In the left figure, we still find a large variation in heart rate values for each value of the pedal power. On the other hand, the spread in heart rate values is significantly decreased if we consider the convolved pedal power.

This fact is also demonstrated if we consider the fit of the models. In Fig. 7, we show the distribution of the explained variance of the heart rate models for all days in our data collection with sufficient data, as a function of the original and convolved pedal power. Although we here only display the combined results for all riders, our findings are consistent for all three riders. The R^2-score of the models (mean $\pm$ SD) on the test set of the sampled data set is 0.35 ± 0.12 or 0.86 ± 0.07, when we consider the pedal power or convolved pedal power, respectively. By performing an independent t-test, we find that the difference is statistically significant ($p < 0.01$). We calculate

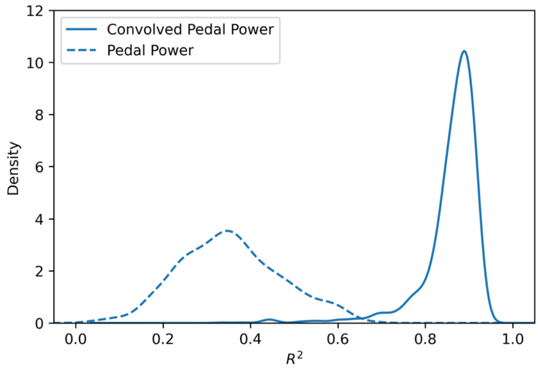


Fig. 7 Distribution of the accuracy (explained variance on the test set) of the heart rate models for the pedal power and convolved pedal power

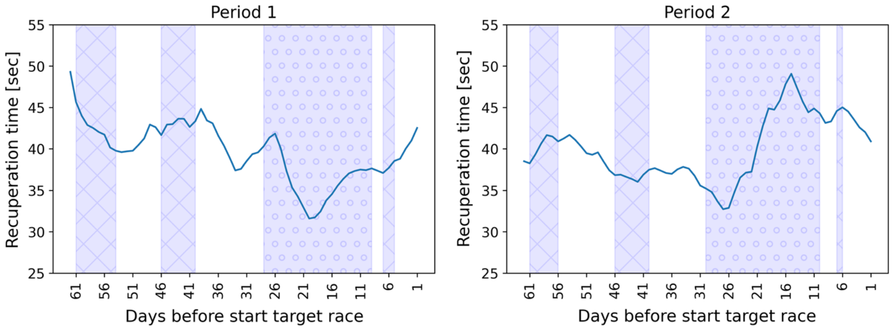


Fig. 8 The recuperation time for rider A during two similar build-up periods for the same focus race in subsequent seasons. We averaged the recuperation times by calculating a rolling mean of the previous week to display the global trends in both phases. The dashed area with circles indicates an altitude camp and the races are denoted by the shaded area with crosses

the effect size of this difference by using Cohen's d (Cohen 1992), and find that the effect size is *large* ($d > 2$). Hence, the introduction of the convolved pedal power significantly increases our understanding of the variance in the heart rate.

4.4 Fitness monitoring

The goal of this study is developing a frequent fitness monitoring system. Within our approach, the main components of this system are the recuperation time, the heart rate performance curves, as well as the virtual aerobic and anaerobic thresholds. In the remainder of this section, we will use the data of rider A to illustrate the usefulness of our monitoring system.

4.4.1 Recuperation time

An example of the recuperation time monitoring is displayed in Fig. 8. To illustrate the global trend, we show the rolling average of daily recuperation times in the previous week. We consider two periods of the same length, that occurred in the same part of the season in two subsequent years, and are the build-up phases for the most important

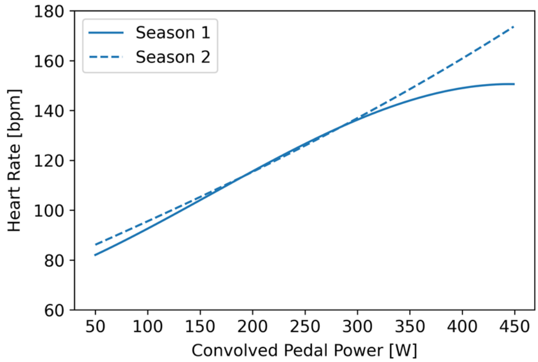


Fig. 9 Heart rate performance curves for rider A at the same date in two subsequent seasons. The solid (dashed) line represents the heart rate performance curve at roughly the middle of the altitude camp in the build-up phase for the main goal of the season, i.e., period 1 (2) in Fig. 8

race of the season for our rider of interest. Moreover, the content of these periods is almost identical with an altitude camp at the same timing, location and the same races that are used as preparation.

Despite the similarity of both periods, we observe that the behavior of the recuperation time is quite different. The main difference occurs during the altitude camp (the largest blue area in each plot), which is the most important part in the build-up towards the target race. At the start and end of the altitude camp, the recuperation times are roughly the same in both periods. However, during the altitude camp, the evolution is quite different. In the first period, a decline in the recuperation time at the beginning of the altitude camp is followed by an increase. On the other hand, the recuperation time increases during almost the entire altitude camp with only a short period of decline in the final stage of the altitude camp. This could be indication that the training load was too large during the altitude camp.

If we compare both seasons, we find a difference of roughly 15 seconds in the recuperation time at 16 days before the start of the target race. Note that these fluctuations are not a consequence of the random noise in data. Namely, the spread in recuperation time on a single day due to the noise, defined as the standard deviation for ten runs with different under-sampling strategies, is typically around 4 s. Therefore, combined with the fact that the rider performed worse in the target race during the second season, this example might indicate that this recuperation time contains valuable information in fitness monitoring.

4.4.2 Heart rate performance curve

With the construction of our heart rate model, we obtained the relationships between heart rate and convolved pedal power for each day in our data collection. Therefore, we can determine the heart rate performance curves and monitor the fitness of our cyclist by studying the development of this curve through time.

In Fig. 9, we show two heart rate performance curves at the same date in two subsequent seasons. Although the run-up to the altitude camp and the goal of both training periods were similar, we find different results. Most importantly, we observe that the heart rate at large values of the convolved pedal power is larger in the second

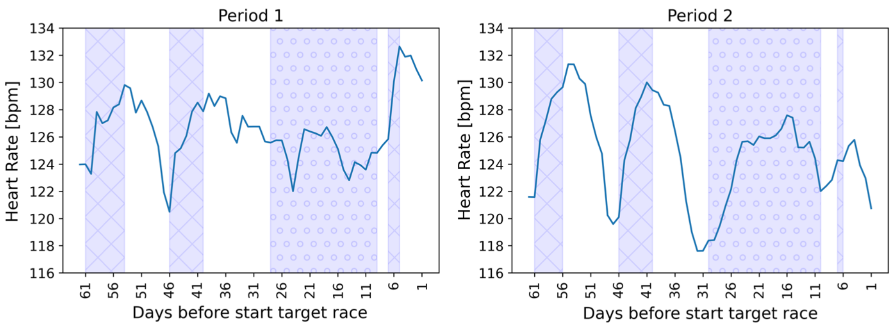


Fig. 10 Heart rate for rider A at a fixed value for the convolved pedal power of 250 W in the build-up phase for the main goal of the season in two subsequent seasons, i.e., period 1 (2) in Fig. 8. The crossed-filled area denotes days on which the rider participated in races and the area filled with circles indicates an altitude camp

altitude camp (dashed line). Interestingly, we observe that one of the curves flattens at high intensities, whereas the other is nearly straight.

Apart from the difference in shape of the heart rate performance curves, we can also extract other useful information from these. In addition to the virtual thresholds, which will be discussed in the next section, it is also insightful to consider intersections at a fixed convolved pedal power. In Fig. 10, we display the course of the heart rate corresponding to a fixed convolved pedal power of 250 W. Although the heart rate values are quite stable, small fluctuations could be examined to study exercising in the aerobic zone. Moreover, we observe that there are differences between both seasons. In particular, the day before the target race, the heart rate at 250 W is almost 10 bpm higher in period 1 than in period 2. This demonstrates that relevant information about the fitness can be derived at moderate exercise intensities. For example, a higher heart rate at the same convolved pedal power could indicate a decreased aerobic fitness (Thorpe et al. 2017), if other possible explanations such as a reduction in training volume can be excluded (Mujika and Padilla 2000a, b).

4.4.3 Virtual aerobic and anaerobic thresholds

As explained in Sect. 3.4, we can analytically obtain the deflection points of the curve describing the relationship between heart rate and convolved pedal power. To determine the coordinates of these points, we need values for the minimal and maximal convolved pedal power, i.e., we have to specify $P_{\min}$ and $P_{\max}$, respectively.

For monitoring the fitness of a cyclist, the precise values of these parameters are not very important. More specifically, a change of the parameter variables only results in an offset of all values. Therefore, slightly different values do not affect the fluctuations of the virtual aerobic and anaerobic thresholds throughout the season, i.e., relative changes. For the purpose discussed in this section, it suffices to choose typical values for these parameters, and we set $P_{\min} = 25$ W and $P_{\max} = 475$ W. In the following section, we will elaborate on a more data-driven procedure to set these

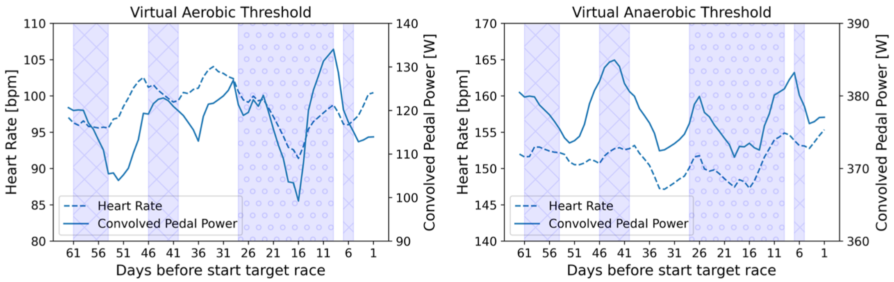


Fig. 11 Heart rate (dashed) and convolved pedal power (solid) at the virtual aerobic (left) and anaerobic (right) thresholds for rider A during the build-up period for the target race in a season. We display the global behavior of the heart rate and convolved pedal power by using a rolling mean with a back looking window of a week. The area with circles indicates an altitude camp and the competition races is denoted by the shaded area with crosses

values and demonstrate that the parameter values need to be altered if a different rider is considered.

In Fig. 11, we show the development of the virtual aerobic (left) and anaerobic (right) thresholds in the build-up towards a target race of our rider. As with the recuperation time, we display the rolling mean with a backward looking window of a week to display the global trend during the preparation phase. Note that also in the case the fluctuations are not due to the noise in the data. Here, the standard deviations for ten different under-sampled data sets are approximately 3 W and 0.7 W for the virtual aerobic and anaerobic thresholds, respectively. Moreover, the standard deviation in the corresponding heart rate values is 0.7 bpm and 0.4 bpm. Therefore, the results of these figures nicely illustrate that actual fluctuations of the thresholds would have been missed if monitoring occurs on a less frequent basis.

4.5 Validation of results

To validate our results for the recuperation time, we would need to compare our findings with additional information about the fatigue and recovery of a rider. Although our findings coincide with the experiences of the coach, e.g., a worse recovery during the altitude camp in period 2 compared to period 1 as is displayed in Fig. 8, more validation would be in order. For example, this can be achieved by using a daily wellness questionnaire (de Leeuw et al. 2021). Unfortunately, this data is not collected and a more detailed validation needs to be addressed in future work.

4.5.1 Virtual anaerobic threshold

The virtual anaerobic threshold can be validated by using the results of physical tests. The tests can be divided into two different categories. First, there are field tests in which the rider completed six blocks of 6 min. In each block, the rider was instructed to ride at different exercise intensities. In the first period of 6 min, the exercise intensity is low. In subsequent blocks, the exercise intensity gradually increased until an all-out 6 min effort in the final block. After each block, the lactate concentration was measured.

Table 3 Comparison of our approach with the results of exercise tests

Rider	Number of physical tests	Average power (virtual) anaerobic threshold	
		Exercise tests	Our approach
A	5	373 ± 11 W	379 ± 10 W
B	3	388 ± 11 W	391 ± 19 W
C	1	423 W	419 ± 12 W

We consider the convolved pedal power at the virtual anaerobic threshold from our approach and the power at the anaerobic threshold obtained in exercise tests. We display the average values and standard deviations

Second, we have the results of two exercise tests that were performed in a controlled laboratory setting. In the tests, the rider starts cycling at a low pedal power. After cycling at this intensity for a fixed time period, the exercise intensity is increased. Typically, the increment of pedal power is 30 W and the rider needs to maintain this intensity for a few minutes, e.g. 3 min. This procedure is repeated until complete exhaustion.

As argued, these tests are invasive, and are only performed rarely. For rider A, we have results of three field tests and two tests that were performed in a laboratory. Rider B and C completed three and one tests, respectively. With this limited data, that is also mostly obtained in the off-season, it is unfeasible to directly compare our results for the virtual anaerobic thresholds with the outcomes of the tests. In this case, the most appropriate validation is to compare average results for the produced power at the virtual thresholds.

To compare the results of our approach with the outcomes of the exercise test, we first need to specify the rider-specific parameters P_{min} and P_{max}. For every rider, we determine the 99th percentile of the convolved pedal power for each day in their data collection and set P_{max} to the largest value. More specifically, P_{max} is equal to 483 W, 497 W and 545 W for rider A, B and C, respectively. Moreover, the presence of low values of the convolved pedal power is similar for our cyclists, and we set $P_{\text{min}} = 0$ W for all riders. In Table 3, we compare our findings with the outcomes of the exercise tests. These results demonstrate that the power at our virtual anaerobic threshold is consistent with the results of the exercise tests.

In addition to the comparing the average values, it is also interesting to look at the changes in the convolved pedal power at the anaerobic threshold throughout the entire period. Therefore, we determine the middle 90% of all values for the convolved power at the anaerobic threshold. By using the physiological characteristics of our riders, we obtain that for rider A the spread is roughly $0.46\,\text{W} \cdot \text{kg}^{-1}$. Moreover, for riders B and C we obtain values of approximately $1.03\,\text{W} \cdot \text{kg}^{-1}$ and $0.48\,\text{W} \cdot \text{kg}^{-1}$, respectively. Within a span of a few months during a season, the variations in the Functional Threshold Power (FTP), i.e., the maximal power that a rider can sustain for an hour, can already be $0.4\,\text{W} \cdot \text{kg}^{-1}$ (Nimmerichter et al. 2010). As the FTP is strongly correlated with the anaerobic threshold (Valenzuela et al. 2018), we find that the fluctuations in our virtual anaerobic threshold are comparable to the results of this study for riders A and B. For rider C, we obtain fluctuations that are larger. Note that these larger values of the fluctuations can be obtained as we consider longer periods that includes race days as well as low-intensity periods within the season, i.e., days after a holiday break.

4.5.2 Virtual aerobic threshold

Unfortunately, a similar validation of our virtual aerobic threshold is not possible. Namely, the physical tests start at exercise intensities at which there is already an elevation in heart rate compared to the heart rate in rest. Thus, the starting exercise intensity of the test is larger than we used for our definition of the virtual aerobic threshold, i.e., the intensity at which there is the first significant increase in heart rate. So, these tests can not be used for validation of the virtual aerobic threshold. However, we know that the values of our virtual aerobic threshold are quite different from the aerobic threshold obtained from exercise tests. For example, we find a value of 106 ± 18 W for our virtual anaerobic threshold for rider A. This is quite different from the 283 W found for the aerobic threshold in the tests performed in the laboratory. The disagreement can be explained by our definition of the virtual anaerobic threshold. We defined this threshold as the exercise intensity at which there is the first significant increase in heart rate and thus can be interpreted as a minimum for low-level exercising. However, professional cyclists are able to maintain low lactate concentrations for low to moderate exercise intensities with already a significant elevation in heart rate. As recreational cyclists have less developed aerobic systems, our virtual aerobic threshold might be a more relevant estimate of the aerobic threshold for these riders.

The advantage of considering elite cyclists is that the structured and frequently collected data are ideal for developing a data-driven method as considered here. The disadvantage is that their training program is very strict and they participate in many races throughout the season. Therefore, it is infeasible that elite cyclists will perform the frequent exercise tests that are necessary for validating the short-term variations in the virtual thresholds. Hence, for further validation of our approach, we need to consider a group of sub-elite cyclists that still collect sufficient data and have more freedom to perform regular exercise test. However, this is beyond the scope of this study and will be addressed in future work.

5 Conclusion

We have presented a methodology for frequently monitoring the fitness of professional cyclists by solely relying on data that is collected in training sessions and competition. Although certain steps still need to be taken before our approach can be used in daily practice, we believe our work is an important step in obtaining frequent and objective information about the fitness of a cyclist without performing additional exercise tests.

In our approach, we introduce the *convolved pedal power* by applying convolution to the pedal power. Hereby, we take into account that the heart rate at a specific time also depends on previously delivered efforts. By performing an experiment with a kernel of arbitrary shape, we obtained that there is lag between a change in exercise intensity and the response in the heart rate. In the literature, this phenomenon is known as the cardiac lag (Grazzi et al. 1999; Jeukendrup and Diemen 1998). For frequent fitness monitoring, the kernel needs to be optimized for each day in our data collection. Therefore, it is computationally expensive to consider an arbitrarily-shaped kernel with many independent parameters. As an alternative, we here worked with an exponentially decaying kernel that only has a single parameter, i.e., the decay constant, that needs to be optimized. Although we have demonstrated that this is a good approximation, we will investigate different alternatives for the kernel in future work.

Subsequently, we constructed a model that couples the external training load via the convolved pedal power to the internal training load by means of the heart rate. To model the complex behavior between these variables, we describe the heart rate as a third-order polynomial in the convolved pedal power. For most of the time, our cyclist rides at low or moderate exercise intensities. Therefore, we have used proportional under-sampling to obtain a model that is also accurate at high exercise intensities. We have performed several experiments to compare different parameter settings in our modeling approach. For a given day, we consider all sessions in the previous 7 days to construct a heart rate model that is valid for all relevant values of the convolved pedal power. Here, we used sampling weights that are inversely proportional to the frequency of occurrence in the original data collection.

The accuracy of our heart rate model expressed in explained variance R^2 is 0.86 ± 0.07. This accuracy is more than two times larger than if we would have considered the pedal power without applying convolution. This demonstrates the importance of introducing the convolved pedal power. The accuracy of our model is comparable to the results of previous studies (Mazzoleni et al. 2016; Lefever et al. 2014). In these works, the constructed heart rate models rely on the pedal power on one or a few specific instances and their accuracy is verified in controlled settings. Here, we model the heart rate at a specific time t by taking into account the pedal power that is produced at all instances prior to t. Moreover, we solely rely on data that is collected without following any protocol. In future research, it is worthwhile to investigate possible extensions of our model. For example, the heart rate might be modeled more accurately if the pedal power and cadence are considered simultaneously (Mazzoleni et al. 2016). Moreover, it would be interesting to explicitly consider altitude effects.

In total, we have presented several personalised methods that can be used to monitor the fitness of a cyclist. First, we considered the decay constant of the kernel that is used for calculating the convolved pedal power. The decay constant specifies the length of a time window in which efforts affect the heart rate and therefore can be interpreted as a recuperation time of our cyclist. We have presented anecdotal results by comparing the course of the recuperation time of two almost identical period in two subsequent season for one of our cyclists. Our findings about the recuperation of the cyclist in these periods are confirmed by the experience of the coaching staff, but a more detailed validation will be addressed in future work.

Second, we considered the heart rate performance curves. The coaching staff can use these curves for monitoring the fitness by for example looking at the development of the heart rate of a fixed exercise intensity throughout a training period. Interestingly, we have observed that there are both heart rate performance curves with upward and downward deflection at large values of the convolved pedal power. As the type of deflection is related to specific functions of the heart (Hofmann et al. 2005; Chwalbinska-Moneta et al. 1996), our methodology opens up new avenues for studying the relationships between training status and heart function in sports.

Finally, we determined virtual aerobic and anaerobic thresholds that are defined as the deflection points of the curve describing the relationship between heart rate and convolved pedal power. With our approach, we can determine exact expression for these thresholds. As is also shown in the context of heart rate performance curves (Hofmann and Pokan 2010), this is a prerequisite for obtaining reliable results. Although our formulas for the virtual thresholds were derived for the most-commonly observed ’S-shaped’ heart rate performance curve, we would like to mention that our formulas are also valid for other shapes, such as an inverted curve that is also observed occasionally here and among subjects in other studies (Hofmann et al. 1997; Lucía et al. 2000).

We validated our findings with the outcomes of exercise tests, although only sparingly available. We have found that our virtual anaerobic threshold on average agrees with the anaerobic threshold found in exercise tests. This agrees with previous studies that found large correlations between the power at the anaerobic threshold and the second deflection point of a heart rate performance curve (Ribeiro et al. 1985; Bunc et al. 1995; Hofmann et al. 1994, 1997). Moreover, we found that our virtual aerobic threshold disagrees with the true aerobic threshold and can rather be seen as minimum intensity for easy exercising. Unfortunately, the relatively small number of exercise tests prevented us from performing a detailed validation of for example the short-term variations in the anaerobic threshold. Therefore, this will be addressed in follow-up studies.

There are also some limitations of our work. As we consider a data-driven approach, our method relies on the collection of sufficient and accurate data. As the precision of power meters can vary considerably (Maier et al. 2017), it is crucial to use power meters that are accurate and regularly calibrate these devices. Since cyclists in professional teams and their coaches are familiar with properly using power meters, this is self-evident for world-class riders as considered in this study. However, this is important for recreational cyclists that are less familiar with the limitations of power meters. Moreover, there could be issues due to switching between different brands. Occasionally, there might be a switching between brands, inspired by switching (power meter) sponsors, but this is never done within a season. As a consequence this should only have little effect. However, we cannot fully guarantee the correctness of our approach when switching power meters. Hence, we would recommend to be consistent in power meter brands. Finally, we need enough data at all relevant exercise intensities to obtain accurate results from our modelling approach. Except from some edge cases, e.g., low-intensity training weeks due to health issues, we found that the data of professional cyclists contains the appropriate information if an aggregation window of a week is considered. However, we expect this needs to be altered for semi-professional or recreational cyclists that ride less often and usually follow less strict training programs.

In addition to the aforementioned directions for future research in the near future, there are also interesting avenues in the long term. Although we only consider general classification cyclists, our approach is generic and can therefore directly be applied to other type of cyclists, or with some small adjustments, to athletes in other sports. Therefore, with our method the fitness of an entire cycling team can be monitored more precisely. On a daily basis, the coach could obtain information about the recuperation time as well as the virtual anaerobic and aerobic thresholds. Moreover, trends in these quantities over a longer time period for example could be visualized in a dashboard. As manual inspection for a large group of cyclist might be too time-consuming in busy periods, specific alerts could be set as early warning signs. For example, the coach might want to have a closer look at the data and have a conversation with a cyclist, if the virtual anaerobic threshold is dropping or the recuperation time is constantly increasing for weeks.

In conclusion, our work opens up new possibilities for addressing the optimization of training programs in sports. By using the fitness monitoring, detailed information about the physiological response of an individual athlete after executing a training session can be obtained. This can be used to identify which kind of training is most effective in each part of the season. Hereby, training programs can be optimized in an athlete-specific manner and made dynamic by adjusting the program with respect to the current fitness status.

Acknowledgements The research leading to these results received funding from ZonMW under project WielerFitheid, winner of the National SportInnovator Prize 2020.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethics and reproducibility Informed consent of the riders was obtained for using the data in this scientific study and publishing the results. The data are property of Team Jumbo-Visma and therefore cannot be shared publicly. However, peers in road cycling with access to similar power and heart rate data can reproduce all steps in our methodology with the descriptions given in the manuscript.

References

Bannister EW, Calvert TW, Savage MV, Bach T (1975) A systems model of training for athletic performance. Aust J Sports Med 7:57–61

Bellenger C, Fuller J, Thomson R, Davison K, Robertson E, Buckley J (2016) Monitoring athletic training status through autonomic heart rate regulation: a systematic review and meta-analysis. Sports Med 46(10):1461–1486. https://doi.org/10.1007/s40279-016-0484-2

Binder RK, Wonisch M, Corra U, Cohen-Solal A, Vanhees L, Saner H, Schmid J-P (2008) Methodological approach to the first and second lactate threshold in incremental cardiopulmonary exercise testing. Eur J Prev Cardiol 15(6):726–734

Borresen J, Lambert MI (2009) The quantification of training load, the training response and the effect on performance. Sports Med 39:779–795. https://doi.org/10.2165/11317780-000000000-00000

Bourdon PC, Cardinale M, Murray A, Gastin P, Kellmann M, Varley MC, Gabbett TJ, Coutts AJ, Burgess DJ, Gregson W, Cable NT (2017) Monitoring athlete training loads: consensus statement. Int J Sports Physiol Perform 12(s2):2–1612170. https://doi.org/10.1123/IJSPP.2017-0208

Brooke J, Hamley E (1972) The heart-rate-physical work curve analysis for the prediction of exhausting work ability. Med Sci Sports 4(1):23–26

Buchheit M (2014) Monitoring training status with hr measures: do all roads lead to rome? Front Physiol 5:73. https://doi.org/10.3389/fphys.2014.00073

Buchheit M, Papelier Y, Laursen PB, Ahmaidi S (2007) Noninvasive assessment of cardiac parasympathetic function: postexercise heart rate recovery or heart rate variability? Am J Physiol Heart Circ Physiol 293(1):8–10

Bulck DV, Weghe AV, Goossens D (2021) Result-based talent identification in road cycling: discovering the next eddy merckx. Ann Oper Res. https://doi.org/10.1007/s10479-021-04280-0

Bunc V, Heller J, Leso J (1988) Kinetics of heart rate responses to exercise. J Sports Sci 6(1):39–48

Bunc V, Hofmann P, Leitner H, Gaisl G (1995) Verification of the heart rate threshold. Eur J Appl Physiol 70(3):263–9. https://doi.org/10.1007/BF00238574

Cheng B, Kuipers H, Snyder AC, Keizer HA, Jeukendrup A, Hesselink M (1992) A new approach for the determination of ventilatory and lactate thresholds. Int J Sports Med 13(7):518–522. https://doi.org/10.1055/s-2007-1021309

Chwalbinska-Moneta J, Krysztofiak H, Ziemba A, Nazar K, Kaciuba-Uściłko H (1996) Threshold increases in plasma growth hormone in relation to plasma catecholamine and blood lactate concentrations during progressive exercise in endurance-trained athletes. Eur J Appl Physiol 73(1):117–120

Cohen J (1992) Statistical power analysis. Curr Dir Psychol Sci 1(3):98–101. https://doi.org/10.1111/1467-8721.ep10768783

Conconi F, Ferrari M, Ziglio PG, Droghetti P, Codecá L (1982) Determination of the anaerobic threshold by a noninvasive field test in runners. J Appl Physiol 52(4):869–73

de Leeuw A-W, van der Zwaard S, van Baar R, Knobbe A (2021) Personalized machine learning approach to injury monitoring in elite volleyball players. Eur J Sport Sci, 1–10

de Leeuw A-W, Heijboer M, Hofmijster M, van der Zwaard S, Knobbe A (2020) Time series regression in professional road cycling. In: Appice A, Tsoumakas G, Manolopoulos Y, Matwin S (eds) Discovery science. Springer, Cham, pp 689–703

De Spiegeleer E. Predicting cycling results using machine learning

Demosthenous G, Kyriakou M, Vassiliades V (2022) Deep reinforcement learning for improving competitive cycling performance. Expert Syst Appl 203:117311

Foster C (1998) Monitoring training in athletes with reference to overtraining syndrome. Med Sci Sports Exerc 30(7):1164–1168

Grazzi G, Alfieri N, Borsetto C, Casoni I, Manfredini F, Mazzoni G, Conconi F (1999) The power output/heart rate relationship in cycling: test standardization and repeatability. Med Sci Sports Exerc 31:1478–83. https://doi.org/10.1097/00005768-199910000-00019

Halson SL (2014) Monitoring training load to understand fatigue in athletes. Sports Med 44:139–147. https://doi.org/10.1007/s40279-014-0253-z

Hamilton MT, González-Alonso J, Montain S, Coyle EF (1991) Fluid replacement and glucose infusion during exercise prevent cardiovascular drift. J Appl Physiol 71(3):871–7

Hilmkil A, Ivarsson O, Johansson M, Kuylenstierna D, van Erp T (2018) Towards machine learning on data from professional cyclists. arXiv preprint arXiv:1808.00198

Hofmann P, Pokan R (2010) Value of the application of the heart rate performance curve in sports. Int J Sports Physiol Perform 5(4):437–47

Hofmann P, Pokan R, Preidler K, Leitner H, Szolar D, Eber B, Schwaberger G (1994) Relationship between heart rate threshold, lactate turn point and myocardial function. Int J Sports Med 15:232–7. https://doi.org/10.1055/s-2007-1021052

Hofmann P, Pokan R, von Duvillard SP, Seibert FJ, Zweiker R, Schmid P (1997) Heart rate performance curve during incremental cycle ergometer exercise in healthy young male subjects. Med Sci Sports Exerc 29(6):762–768

Hofmann P, Wonisch M, Pokan R, Schwaberger G, Smekal G, Duvillard S (2005) ß1-adenoceptor mediated origin of the heart rate performance curve deflection. Med Sci Sports Exerc 37(10):1704–9

Hunter A, Coggan AR, McGregor S (2019) Training and racing with a power meter. VeloPress, Boulder

Janssens B, Bogaert M, Maton M (2022) Predicting the next pogačar: a data analytical approach to detect young professional cycling talents. Ann Oper Res, 1–32

Jeukendrup A, Diemen AV (1998) Heart rate monitoring during training and competition in cyclists. J Sports Sci 16(sup1):91–99

Joyner MJ, Coyle EF (2008) Endurance exercise performance: the physiology of champions. J Physiol 586(1), 35–44. https://doi.org/10.1113/jphysiol.2007.143834

Karetnikov A (2019) Application of data-driven analytics on sport data from a professional bicycle racing team. Eindhoven University of Technology, The Netherlands

Karlsson J, Jacobs I (1982) Onset of blood lactage accumulation during muscular exercise as a threshold concept I theoretical considerations. Int J Sports Med 3(4):190–201. https://doi.org/10.1055/s-2008-1026087

Kataoka Y, Gray P (2019) Real-time power performance prediction in tour de france. In: Brefeld U, Davis J, Van Haaren J, Zimmermann A (eds) Mach Learn Data Min Sports Anal. Springer, Cham, pp 121–130

Kellmann M (2010) Preventing overtraining in athletes in high-intensity sports and stress/recovery monitoring. Scand J Med Sci Sports 20(Suppl 2):95–102. https://doi.org/10.1111/j.1600-0838.2010.01192.x

Kellmann M, Bertollo M, Bosquet L, Brink M, Coutts A, Duffield R, Erlacher D, Halson S, Hecksteden A, Heidari J, Kallus K, Meeusen R, Mujika I, Robazza C, Skorski S, Venter R, Beckmann J (2018) Recovery and performance in sport: consensus statement. Int J Sports Physiol Perform 13(2):240–245. https://doi.org/10.1123/ijspp.2017-0759

Kholkine L, Servotte T, de Leeuw A-W, Schepper TD, Hellinckx P, Verdonck T, Latré S (2021) A learn-to-rank approach for predicting road cycling race outcomes. Front Sports Active Living. https://doi.org/10.3389/fspor.2021.714107

Krawczyk B (2016) Learning from imbalanced data: open challenges and future directions. Progress Artif Intell 5:221–232. https://doi.org/10.1007/s13748-016-0094-0

Lamberts RP, Rietjens GJ, Tijdink HH, Noakes TD, Mi L (2010) Measuring submaximal performance parameters to monitor fatigue and predict cycling performance: a case study of a world-class cyclo-cross cyclist. Eur J Appl Physiol 108:183–190. https://doi.org/10.1007/s00421-009-1291-3

Lefever J, Berckmans D, Aerts J-M (2014) Time-variant modelling of heart rate responses to exercise intensity during road cycling. Eur J Sport Sci 14(sup1), 406–412. https://doi.org/10.1080/17461391.2012.708791. PMID: 24444235

Leo P, Spragg J, Podlogar T, Lawley JS, Mujika I (2021) Power profiling and the power-duration relationship in cycling: a narrative review. Eur J Appl Physiol. https://doi.org/10.1007/s00421-021-04833-y

Lucía A, Carvajal A, Pérez M, Boraita A (2000) Heart rate response during incremental exercise in master runners. Jpn J Physiol 50(1):155–8. https://doi.org/10.2170/jjphysiol.50.155

Ludwig M, Grohganz H, Asteroth A (2016) A convolution model for heart rate prediction in physical exercise. https://doi.org/10.5220/0006030901570164

Maier T, Schmid L, Müller B, Steiner T, Wehrlin JP (2017) Accuracy of cycling power meters against a mathematical model of treadmill cycling. Int J Sports Med 38(6):456–461. https://doi.org/10.1055/s-0043-102945

Mateo-March M, Moya-Ramón M, Javaloyes A, Sánchez-Muñoz C, Clemente-Suárez VJ (2022) Validity of detrended fluctuation analysis of heart rate variability to determine intensity thresholds in elite cyclists. Eur J Sport Sci. https://doi.org/10.1080/17461391.2022.2047228 (**PMID: 35238695**)

Mazzoleni M, Battaglini C, Martin K, Coffman E, Mann B (2016) Modeling and predicting heart rate dynamics across a broad range of transient exercise intensities during cycling. Sports Eng 19:117–127. https://doi.org/10.1007/s12283-015-0193-3

Meeusen R, Duclos M, Foster C, Fry A, Gleeson M, Nieman D, Raglin J, Rietjens G, Steinacker J, Urhausen A (2013) European college of sport science; American college of sports medicine. prevention, diagnosis, and treatment of the overtraining syndrome: joint consensus statement of the european college of sport science and the american college of sports medicine. Med Sci Sports Exer 45(1):186–205

Mujika I, Padilla S (2000) Detraining: loss of training-induced physiological and performance adaptations. Part i: short term insufficient training stimulus. Sports Med. 30(2):79–87. https://doi.org/10.2165/00007256-200030020-00002

Mujika I, Halson S, Burke LM, Balagué G, Farrow D (2018) An integrated, multifactorial approach to periodization for optimal performance in individual and team sports. Int J Sports Physiol Perform 13(5):538–561. https://doi.org/10.1123/ijspp.2018-0093

Mujika I, Padilla S (2000) Detraining: loss of training-induced physiological and performance adaptations. Part ii: Long term insufficient training stimulus. Sports Med 30(3), 145–54. https://doi.org/10.2165/00007256-200030030-00001

Nimmerichter A, Williams C, Bachl N, Eston R (2010) Evaluation of a field test to assess performance in elite cyclists. Int J Sports Med 31:160–6. https://doi.org/10.1055/s-0029-1243222

Ribeiro JP, Fielding RA, Hughes V, Black A, Bochese MA, Knuttgen HG (1985) Heart rate break point may coincide with the anaerobic and not the aerobic threshold. Int J Sports Med 6(4):220–4. https://doi.org/10.1055/s-2008-1025844

San Millán I, Gonzalez-Haro C, Sagasti M (2009) Physiological differences between road cyclists of different categories. a new approach: 733. Med Sci Sports Exercise 41:64–65. https://doi.org/10.1249/01.mss.0000353467.61975.ae

Steyaert M, De Bock J, Verstockt S (2022) Sensor-based performance monitoring in track cycling. In: Brefeld U, Davis J, Van Haaren J, Zimmermann A (eds) Machine learning and data mining for sports analytics. Springer, Cham, pp 167–177

Taylor K, Chapman D, Cronin J, Newton MJ, Gill N (2012) Fatigue monitoring in high performance sport: a survey of current trends. J Aust Strength Condition 20(1):12–23

Thornton HR, Delaney JA, Duthie GM, Dascombe BJ (2019) Developing athlete monitoring systems in team sports: Data analysis and visualization. Int J Sports Physiol Perform 14(6):698–705. https://doi.org/10.1123/ijspp.2018-0169

Thorpe RT, Atkinson G, Drust B, Gregson W (2017) Monitoring fatigue status in elite team-sport athletes: Implications for practice. Int J Sports Physiol Perform 12(Suppl 2):227–234. https://doi.org/10.1123/ijspp.2016-0434

Valenzuela PL, Morales JS, Foster C, Lucia A, de la Villa P (2018) Is the functional threshold power a valid surrogate of the lactate threshold? Int J Sports Physiol Perform 13(10):1293–1298. https://doi.org/10.1123/ijspp.2018-0008

Wallace LK, Slattery KM, Coutts AJ (2009) The ecological validity and application of the session-rpe method for quantifying training loads in swimming. J Strength Condition Res 23(1):33–38

Wasserman K, Whipp BJ, Davis JA (1981) Respiratory physiology of exercise: metabolism, gas exchange, and ventilatory control. Int Rev Physiol 23:149–211