Exploration of Heart Rate Variability for the Prediction of Performance in Youth Footballers

Kacper Korzeniewski[1], Jakub S. Gąsior[2], Magdalena Mikielewicz[1], Maciej Rosoł[1], Robert Makuch[3], Marcel Młyńczak[1]

[1]Warsaw University of Technology, Warsaw, Poland
[2]Medical University of Warsaw, Warsaw, Poland
[3]Kazimierz Pulaski University of Technology and Humanities in Radom, Radom, Poland

Abstract

This study aimed to evaluate the prediction quality of short-term resting heart rate variability (HRV) parameters for estimating endurance performance in youth footballers.

Male players (n = 222, mean age 13 ± 2.3 years) underwent 5-minute RR intervals (RRi) measurement in a supine position using a heart rate monitor. RRi stationarity was confirmed using the Augmented Dickey-Fuller test (α = 0.05). Time-domain and nonlinear HRV parameters were calculated. The data was split 75:25 for training and testing. Gradient Boosting (GBR), AdaBoost (ABR), and Random Forest (RFR) were used to predict VO2max from the Yo-Yo Intermittent Recovery Test. Models were validated using mean absolute error (MAE) and root mean square error (RMSE). Key HRV parameters were ranked using explainable artificial intelligence (XAI) tools.

78% of RRi were stationary. The GBR model achieved MAE: 4.2, RMSE: 5.1; ABR achieved MAE: 3.7, RMSE: 4.4; and RFR achieved MAE: 3.8, RMSE: 4.6. Key features included mean RRi, DFAα2, and SD1/SD2.

Machine learning models was used to evaluate the quality of endurance performance prediction. They enabled to identify significant HRV parameters that can be further analyzed in this regard.

1. Introduction

One of the principal criteria employed in the assessment of a footballer's performance is endurance. It has been demonstrated in studies that there is a strong correlation between a player's maximal aerobic capacity and their ability to perform sprints, as well as the total distance covered during a match [1].

A common method for assessing aerobic performance is the Yo-Yo Intermittent Recovery Level 2 (IR2) test, which is considered to be valid, reliable, specific, and easily integrated into testing routines. The test, which lasts between five and 15 minutes, gauges an athlete's capacity to repeatedly engage in high-intensity exercise bouts that demand a substantial anaerobic energy contribution [2]. During the Yo-Yo IR2, players achieve their maximal oxygen uptake (VO2max), which is the upper limit of the cardiorespiratory system, beyond which further increases in workload do not result in a proportional increase in oxygen consumption.

Nevertheless, it is not always feasible or advisable to conduct exhaustive tests, even with young athletes. When the evaluation of multiple athletes is required concurrently, the process can be both time-consuming and labour-intensive. In such instances, a substantial number of personnel may be necessary to effectively manage the tests. As an alternative approach, the assessment of cardiac autonomic function at rest through heart rate variability (HRV) has been demonstrated to have a strong correlation with aerobic capacity [3].

The ability to accurately estimate VO2max without the necessity of performing the Yo-Yo IR2 test has the potential to significantly enhance the way coaches design training programmes.

Considering the growing popularity of machine learning (ML) tools in recent years, these techniques have also been applied to predict VO2max, offering a more efficient approach to athlete performance evaluation [4].

The objective of this study was, therefore, to investigate the quality of VO2max prediction by models based on HRV parameters obtained from resting. Additionally, we assessed the importance of HRV features included in the models for VO2max prediction.

2. Materials and methods

2.1. Study population

A total of 222 youth male footballers in between 8 and 18 years old participated in the study. The key characteristics of the study population are presented in

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.136

Table 1.

The study was conducted in accordance with the principles of the Declaration of Helsinki for research involving human subjects. Ethical approval was obtained from the appropriate ethics committee (approval number: [KB/55/N02/2019]). All participants and their legal guardians were fully informed about the purpose, procedures, potential benefits, and risks of the study. Written informed consent was obtained from the parents or legal guardians of the participants prior to the study.

Table 1. Descriptive statistic of study population.

Parameters	Mean ± Std	Min / Max
Age	13 ± 2.3	8 / 18
Height [cm]	160.9 ± 15.0	128.1 / 190.4
Weight [kg]	51.5 ± 13.6	23.9 / 85.8
BMI [kg/m^2]	19.6 ± 2.6	14.6 / 30.7
VO2max [ml/kg/min]	45.5 ± 5.4	38.8 / 62.3

2.2. Study design

The study involved two stages of assessment for each athlete. The first stage consisted of a 5-minute heart rate (HR) measurement under controlled resting conditions. To ensure HR stabilization, participants were instructed to lie in a supine position, remaining still and refraining from speaking or moving. Based on data collected in this stage, the HRV parameters used for modeling were calculated.

In the second stage, athletes performed the Yo-Yo IR2 test, followed by the calculation of their VO2max. This stage of study was only conducted to obtain the reference value of VO2max.

2.3. RRi data acquisition

The Polar Team2 [5] device was used to collect data during the protocol. This device allows HR and HRV to be recorded and data to be transmitted via Bluetooth in real time. The Polar transmitter was placed on the athlete's chest during the study, on an elastic band equipped with electrodes.

Registered RR intervals (RRi) were reviewed by a medical staff to identify artifacts based on graphical presentation of raw RRi. Technical and physiological artifacts were identified and corrected by using dedicated computer program.

2.3. Modeling

Based on the collected dataset, we aimed to investigate the quality of VO2max prediction using HRV parameters calculated from RR intervals (RRi) captured during relaxation, and to assess the importance of the included features. The first step involved checking the stationarity of the time series using the Augmented Dickey-Fuller test (α=0.05) [6].

Next, we prepared the input dataset of features for machine learning (ML) models. In this case, both linear and non-linear HRV parameters were calculated. The HRV parameters included: Mean RRi, SDNN, RMSSD, pNN50, SD1, SD2, SD1/SD2, DFA alpha 1, DFA alpha 2, Approximate Entropy and Sample Entropy.

The dataset was split into training and testing sets in a 75:25 ratio and standardized to prepare for the ML models. We utilized three commonly used regression algorithms: Gradient Boosting Regressor (GBR), AdaBoost Regressor (ABR), and Random Forest Regressor (RFR) to estimate endurance performance (VO2max) from the Yo-Yo IR2 test.

To find the best sets of hyperparameters, the tuning was performed on the training set for each algorithm using the grid-search with 10-fold cross-validation technique. In each iteration of the validation, metrics such as mean absolute error (MAE) and root mean squared error (RMSE) were calculated. The best sets of hyperparameters for each model were determined based on the lowest MAE score.

Finally, the models with the best hyperparameters were trained on the full training dataset, validated on the test set, and their performance metrics were calculated.

2.4. Explainable AI

To evaluate the significance of individual features used in ML modelling, explainable artificial intelligence (XAI) techniques were applied. For this analysis, the Dalex Python package was employed [7]. Model-level variable importance based on drop-out loss values were calculated on the test set and visualized on plot.

Based on the results from each model, a cumulative ranking of key HRV features was prepared. The top 5 most important features for each model were assigned scores from 1 to 5 (with 5 being the most significant). These scores were then summed and presented as the final result.

3. Results

Based on results of Augmented Dickey-Fuller test the 78% of collected RRi were stationary. The stationarity of the signals justified the use of time domain HRV parameters during modeling.

For the GBR model, the best hyperparameters were a learning rate of 0.1, a maximum depth of 6, and number of estimators 100. The model's performance on the test data resulted in a MAE of 4.2, a RMSE of 5.1 and mean absolute percentage error (MAPE) of 9.4%

The ABR model performed slightly better. The best set of hyperparameters is a learning rate of 0.01, a square loss function, and 225 estimators. For this configuration model yielded an MAE of 3.7, an RMSE of 4.4 and MAPE of 8.1% on the test data, reflecting improved prediction compared to GBR.

For the RFR model, the optimal configuration of hyperparameters included 125 estimators and the use of absolute error as the criterion. The model achieved an MAE of 3.8, an RMSE of 4.6 and MAPE of 8.4%, positioning it between the GBR and ABR models in terms of predicting the V02max.

The visualization of feature importance for each model was presented in Figure 1.

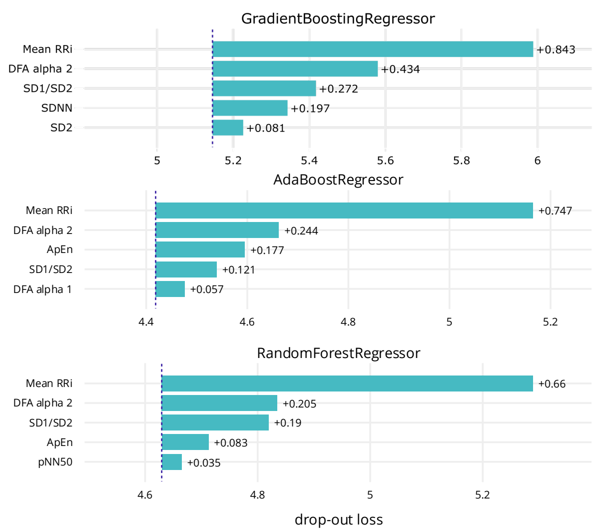

Figure 1. Variable importance analysis.

The summarizing table showing the ranking of the most important HRV features based on machine learning modeling approach was presented in Table 2.

Table 2. Results of XAI analysis.

HRV Parameters	Score*
Mean RRi	15
DFAα2: detrended fluctuation analysis (DFA), which describes long-term fluctuations	12
SD1/SD2: ratio of the Poincaré plot standard deviation perpendicular the line of identity to the Poincaré plot standard deviation along the line of identity	8
ApEn: approximate entropy, which measures the regularity and complexity of a time series	5
SDNN: standard deviation of NN intervals	2
DFAα1: DFA, which describes short-term fluctuations	1
SD2: Poincaré plot standard deviation along the line of identity	1
pNN50: percentage of successive RR intervals that differ by more than 50 ms	1

* The sum of scores given to the most important features of ML models in variable importance analysis.

4. Discussion

In this study, we explored the use of machine learning models to predict endurance performance (VO2max) in youth male footballers based on HRV parameters measured during rest.

Our findings demonstrate that it is feasible to predict VO2max (estimated during Yo-Yo IR2) using short-term RR interval (RRi) signals. Furthermore, we identified key HRV parameters that significantly contribute to predicting endurance performance, with mean RRi, DFAα2, and SD1/SD2 emerging as the most important predictors.

These results are consistent with existing physiological knowledge, particularly the observation that well-trained individuals typically exhibit lower resting heart rates [8], making mean RRi a logical and crucial predictor. The identification of DFAα2 and SD1/SD2 further supports the relevance of nonlinear HRV metrics in assessing cardiovascular adaptability and endurance capacity in young athletes.

However, there are limitations to our study that we aim to address in future research. Increasing the size of the dataset could enhance the performance and generalizability of the machine learning models. Additionally, expanding the range of HRV parameters analyzed could help identify other significant features that may improve prediction accuracy. We also plan to investigate the combination of HRV metrics with physical parameters, such as body composition or training history, to further refine VO2max predictions and provide more comprehensive insights into youth athletes' endurance capabilities.

References

[1] B. Karsten, E. Larumbe-Zabala, G. Kandemir, T. Hazir, A. Klose, and F. Naclerio, "The Effects of a 6-Week Strength Training on Critical Velocity, Anaerobic Running Distance, 30-M Sprint and Yo-Yo Intermittent Running Test Performances in Male Soccer Players," PLoS ONE, vol. 11, no. 3, p. e0151448, Mar. 2016.

[2] J. Bangsbo, F. M. Iaia, and P. Krustrup, "The Yo-Yo Intermittent Recovery Test: A Useful Tool for Evaluation of Physical Performance in Intermittent Sports," Sports Medicine, vol. 38, no. 1, pp. 37–51, 2008

[3] L. A. Pereira et al., “Relationship Between Resting Heart Rate Variability and Intermittent Endurance Performance in Novice Soccer Players,” Research Quarterly for Exercise and Sport, vol. 90, no. 3, pp. 355–361, Jul. 2019.
[4] M. Rosoł, M. Petelczyc, J. S. Gąsior, and M. Młyńczak, “Prediction of peak oxygen consumption using cardiorespiratory parameters from warmup and submaximal stage of treadmill cardiopulmonary exercise test,” PLoS ONE, vol. 19, no. 1, p. e0291706, Jan. 2024.
[5] “Polar for business: from personal trainers to large-scale organisations” *Polar.com*, 2023. https://www.polar.com/uk-en/business (accessed Sep. 23, 2024).
[6] R. Mushtaq, “Augmented Dickey Fuller Test,” SSRN Journal, 2011.
[7] H. Baniecki, W. Kretowicz, P. Piatyszek, J. Wisniewski, and P. Biecek, “dalex: Responsible Machine Learning with Interactive Explainability and Fairness in Python,” Journal of Machine Learning Research, vol. 22, pp. 1–7, 2021.
[8] A. D’Souza et al., “Exercise training reduces resting heart rate via downregulation of the funny channel HCN4,” Nat Commun, vol. 5, no. 1, p. 3775, May 2014.

Address for correspondence:

Kacper Korzeniewski
Faculty of Mechatronics, Institute of Metrology and Biomedical Engineering, Warsaw University of Technology, 8 St. Andrzeja Boboli Street, Warsaw Poland
kacper.korzeniewski.dokt@pw.edu.pl