



HAL
open science

Classification and detection of atypical training sessions in cycling.

Tom Chassard, Steven Golovkine, Alice Meignié, Juliana Antero,
Jean-François Toussaint

► **To cite this version:**

Tom Chassard, Steven Golovkine, Alice Meignié, Juliana Antero, Jean-François Toussaint. Classification and detection of atypical training sessions in cycling.. ECSS congress 2022, Aug 2022, Séville, Spain. hal-04032765

HAL Id: hal-04032765

<https://hal.science/hal-04032765v1>

Submitted on 16 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Tom Chassard^{1,2}, Steven Golovkine³, Alice Meignie^{1,2},
Juliana Antero^{1,2}, Jean-François Toussaint^{1,2,4}

¹Institute for Research in Medicine and Epidemiology of Sports (IRMES – URP 7329), University of Paris Cité, Paris, France. ²National Institute of Sports, Expertise and Performance (INSEP), Paris, France. ³University of Limerick, Limerick, Ireland. ⁴CIMS, AP-HP, Hôtel-Dieu, Paris, France

Context

This study is part of the Empow'Her project, which aims to investigate the specificities of female athletes and their potential impact on performance. This field of research offers many prospects, as female athletes are largely under-represented in studies and scientific literature in the field of sports science, even though initial results are beginning to appear¹.

This project is following up different sports such as rowing, skiing, gymnastics, football, as well as cycling, of which we will present a first analysis. Cycling has been extensively studied in recent years^{2,3,4} but it remains difficult to find a method to analyze the relationship between internal and external load.

References : ¹Meignie & al, 2021, ²Lucia & al, 2001, ³Faria & al, 2005, ⁴Grappe & al, 2009

Objectives

1. Automatically classify training sessions by type of effort made by the cyclist
2. Characterize training sessions by the relationship between external load and internal response
3. To analyze environmental factors leading to atypical relation between external load and internal response

Data

Daily 6-month follow up of 2 female professional cyclists

Training data:

- Power (Watts)
- Heart Rate (BPM)
- Cadence (Rotations per minute)
- Temperature (°C)
- Latitude and Longitude (degrees)
- Elevation (meters)

Waking up questionnaire:

- Menstrual status
- Sleep Duration and Quality
- Feeling
- Energy

1) Training Classification

To compare training sessions based on a similar external load, we try to extract characteristics from the time series of the training sessions. We therefore decide to characterize the training session by quantities (e.g. mean, standard deviation, maximum of a set of quantitative data such as bike cadence, elevation, duration, the percentage of time spent in each of the 7 Coggan zones⁵).

PCA was then applied and results were compared with true session labels given by the coaches themselves. Without true labels, clusters are hard to define.

A more complex method, using Variational Recurrent AutoEncoder^{6,7} to study latent representation of temporal series were tested, but cycling temporal series seem complex for the current capacity of these models.

⁵ Coggan & al, 2010, ⁶Costa & al, 2013, ⁷Post & al, 2020

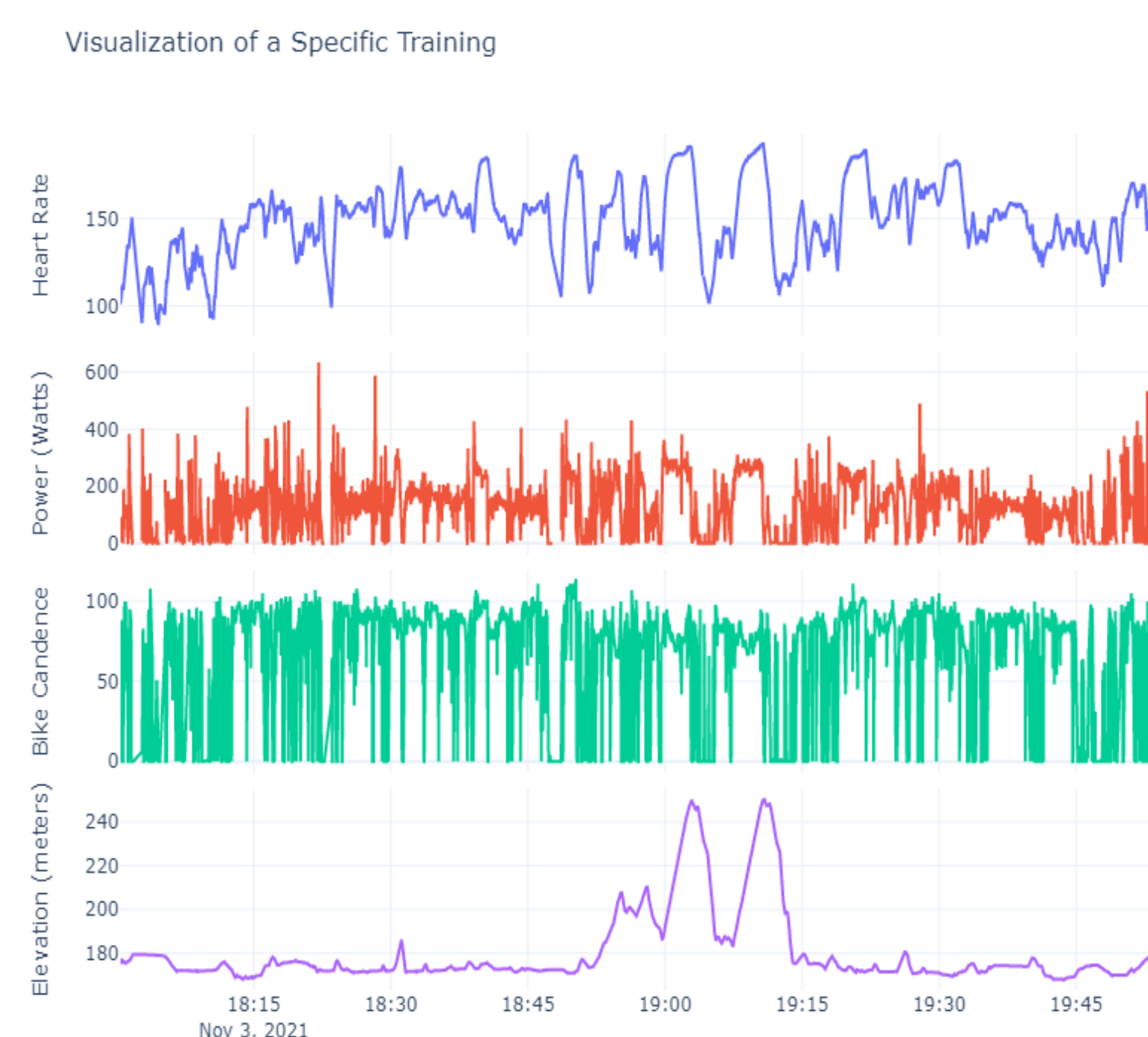


Figure 1: Visualisation of training data

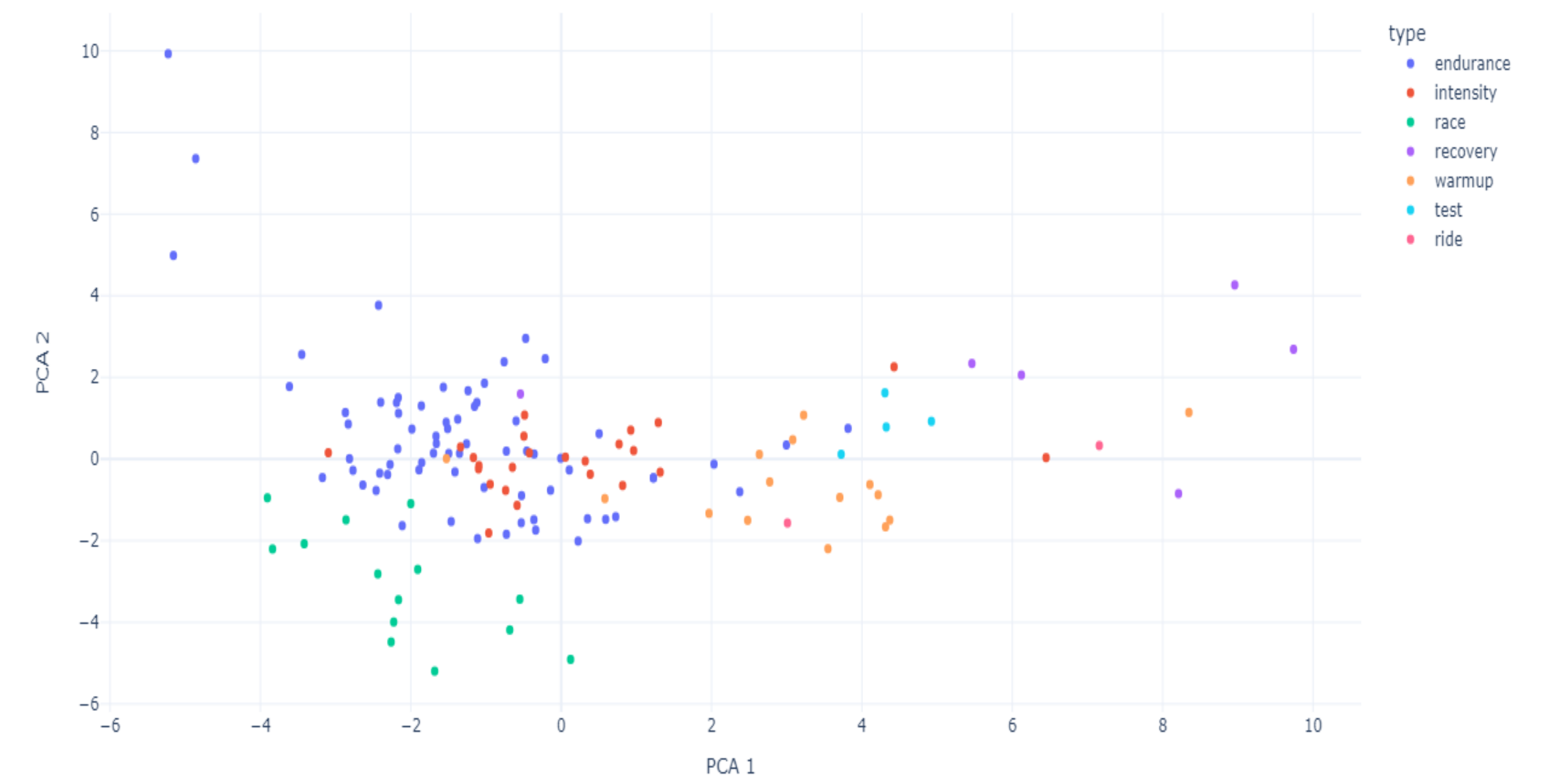


Figure 2: Plot of PCA on feature extracted from training data, coloured by true labels

2) Training Characterization

Instead of representing the training session as a multivariate time series, we choose here to treat it as a set of "power/heart rate" pairs, to study the variations in the relationship of this pair.

To get rid of the biases that arise when using "classical" regression that predicts the mean, which "hides" some of the complexity of this external load/internal load relationship, we use quantile regression⁷ to determine typical individual heart rate intervals for each power output.

We chose the 10% and 90% quantiles to isolate atypical heart rate behaviour in response to power. Quantile regression was implemented using a neural network, with one input layer, two dense layers and two outputs layers (one for each quantile).

⁷Koenker & al, 2001

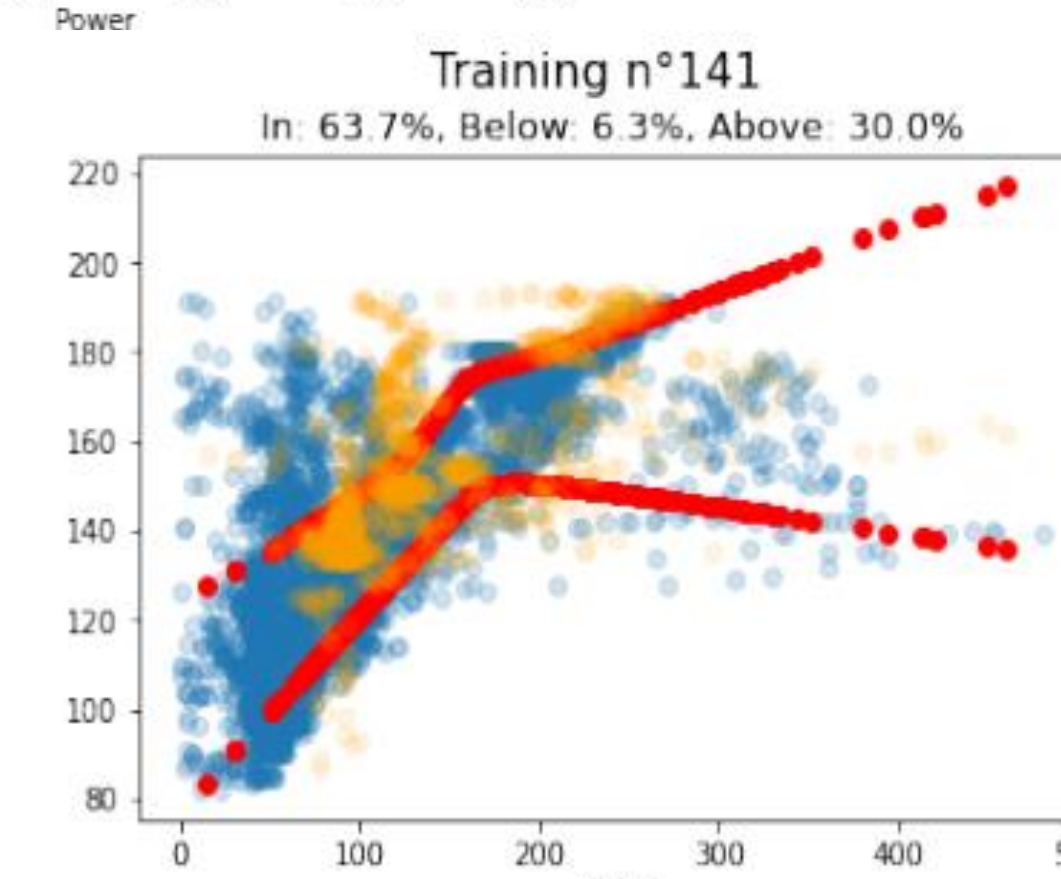
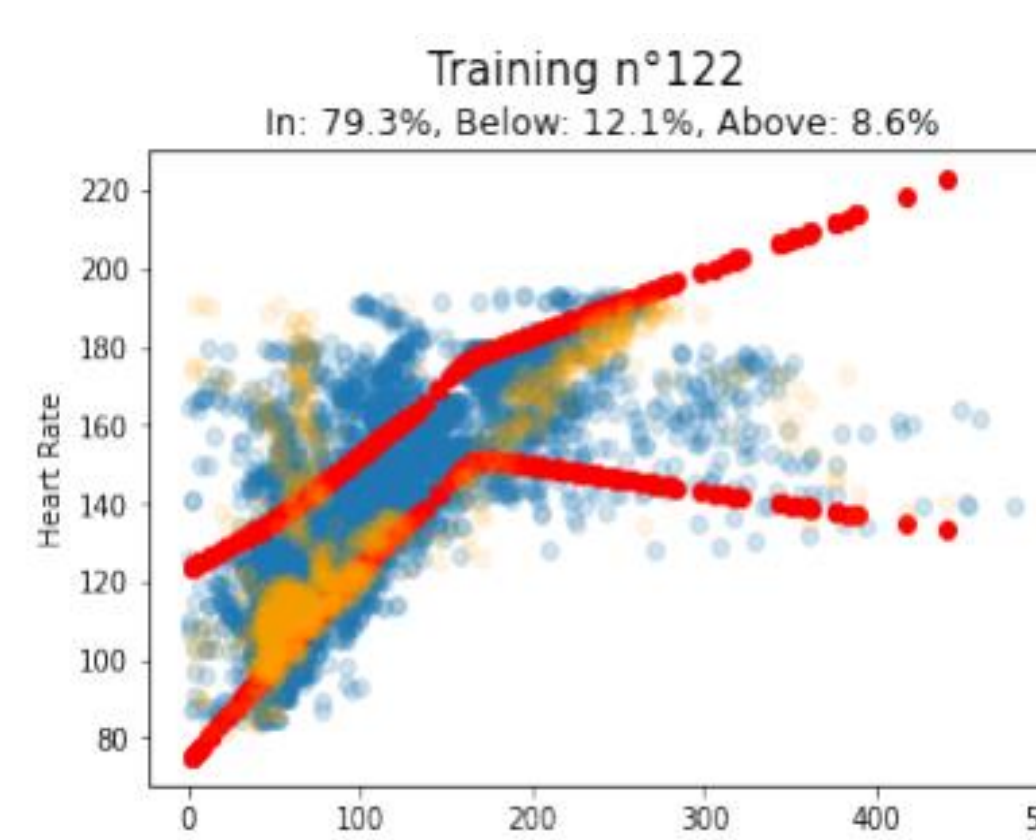


Figure 3: Example of training with data mainly in the specific Heart Rate interval (above), below this interval (left) and above this interval (right).

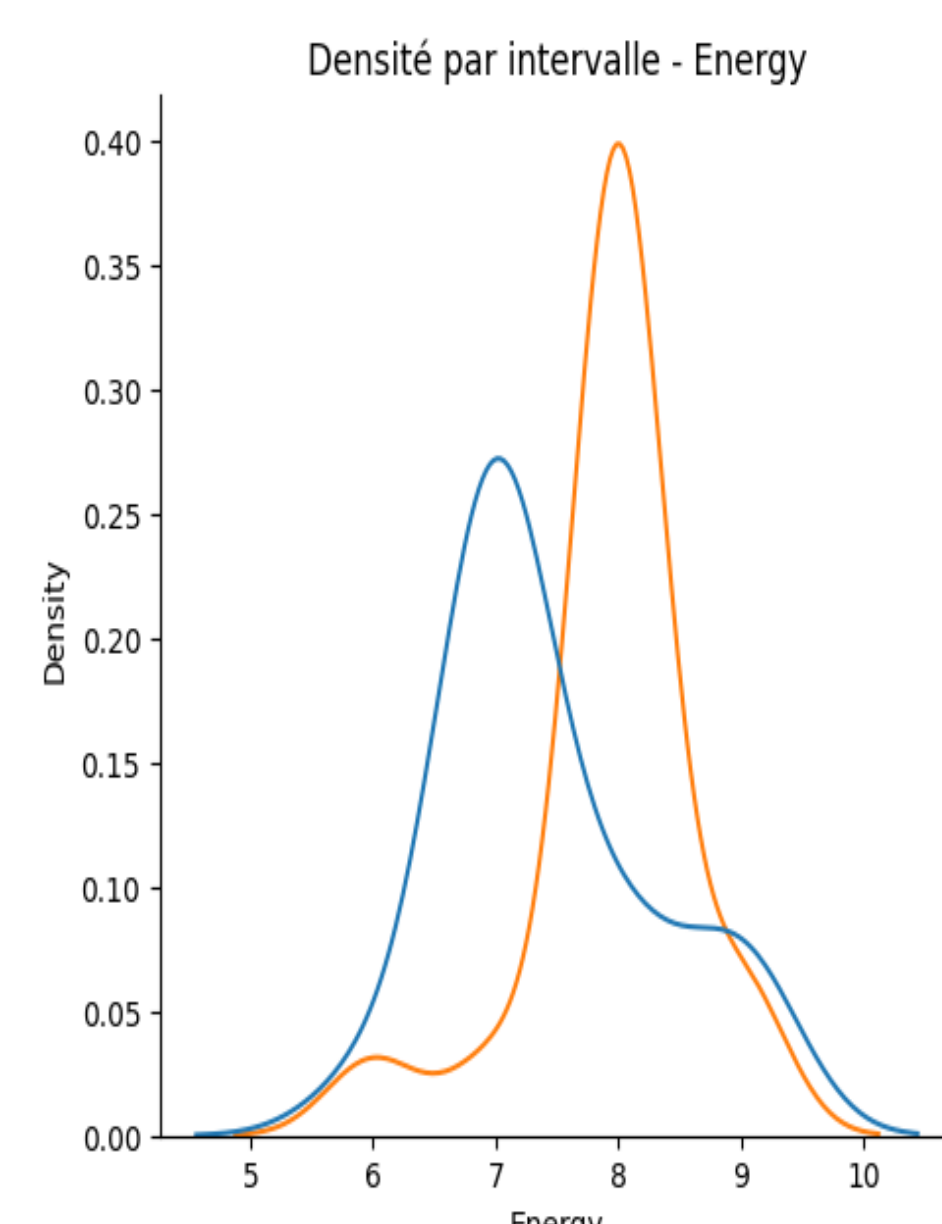


Figure 4: Density difference of Energy variable when majority of point belong to below (orange) or above (blue) interval

3) Results

We highlighted very unbalanced training sessions, where the below-in-above ratio is far from the expected 10-80-10 ratio.

To identify the variables that lead to these atypical training sessions, we selected those specific sessions, where the heart rate is predominantly below the usual threshold or where the heart rate is predominantly above the standard interval.

We examine the density of variables in the daily waking questionnaires in both groups, and apply the Mann-Whitney method to measure the difference between the studied variables.

We found that on days when the heart rate is predominantly below the characteristic range, the reported waking energy is statistically half a point higher (scale 0-10), than on days when the heart rate is predominantly above the standard power response.

Conclusion and Discussion

- The classification of training sessions is complex, but by combining the knowledge of cycling experts and statistical methods, it seems possible to add knowledge to this field.
- Quantile regression allows us to identify training sessions that behave differently from what is expected.
- Significant changes in well-being variables may provide explanations for differences in internal response during training sessions.
- Our classification method still needs more data and more athletes for statistical power.
- A method that takes into account the continuous nature of the heart rate will improve the model.