Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

School of Statistics and Computer Science

Machine Learning to go *nyoom*

Using Machine Learning to evaluate rowing training and predict
training outcomes or performances

Liam Junkermann

Supervisor: Dr. Lucy Hederman

February 6, 2024

A Final Year Project submitted in partial fulfilment
of the requirements for the degree of
BA(Mod) in Science in Computer Science

Declaration

I hereby declare that this Final Year Project is entirely my own work and that it has not been submitted as an exercise for a degree at this or any other university.

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at `http://www.tcd.ie/calendar`.

I have completed the Online Tutorial on avoiding plagiarism 'Ready Steady Write', located at `http://tcd-ie.libguides.com/plagiarism/ready-steady-write`.

I consent / do not consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

Signed: _____   Date: _____

Abstract

A short summary of the problem investigated, the approach taken and the key findings. This should not be more that around 400 words.

This must be on a separate page.

Lay Abstract

Similar to the actual abstract in terms of the information, but written for a non-specialist. So no jargon, no acronyms. Explain to a member of the general public what this project entailed. Should be no longer than the actual abstract.

This must be on a separate page.

Acknowledgements

Thanks Everyone!

You should acknowledge any help that you have received (for example from technical staff), or input provided by, for example, a company.

Contents

List of Figures

List of Tables

Vocabulary

Heart Rate Metrics

HRV	Heart Rate Variablity
HR	Heart Rate
HR_{max}	Maximum Heart Rate

Training Zones

UT2	Basic Oxygen Utilisation Training
UT1	Oxygen Utilization Training
AT	Anaerobic Threshold Training
TR	Oxygen Transport Training
AN	Anaerobic Capacity Training

Chapter 1

Introduction

This chapter will introduce the general approach and motivation for this project, the expected outcomes, and a structure for the remainder of the report.

1.1 Contents of the Introduction

The introduction presents the nature of the problem under consideration, the context of the problem to the wider field and the scope of the project. The objectives of the project should be clearly stated.

1.2 Contents of the background chapter

The second chapter is typically a literature review, or survey of the state of the art, or a detailed assessment of the context and background for the project. The exact nature of this chapter depends on the topic and/or methods of the project. It is essential that the work of other people is properly cited. This will be discussed in detail in Chapter 2 below. Note that you should use references wherever is appropriate through the report, not just in the literature review chapter.

1.3 The Conclusions chapter

The final chapter should give a short summary of the key methods, results and findings in your project. You should also briefly identify what, if any, future work might be executed to resolve unanswered questions or to advance the study beyond the scope that you identified in Chapter 1.

Chapter 2

Background

This chapter will cover the basic background of rowing, the sports science which guides rowing training, ad how an athletes body responds to training stimulus. Next, a review of performance modelling will explore the development of human performance modelling since the introduction of the basic Bannister model in 1975 [1]. The section will outline how training load and performance can be quantified, and explain the way these approximations are used in various performance models to date.

2.1 An introduction to rowing training fundamentals

Rowing is an Olympic sport, raced across a 2,000 metre course. It is classed as a power-endurance sport, this means training is focused on building aerobic, anaerobic, and power while also developing rowing technique [2]. Most time is spent building endurance, next most time is spent building anerobic capacity, finally building strength and power through Strength and Condition sessions [3]. There are many different approaches to how training is conducted and which energy systems are targeted. This section will discuss the basic training principles which guide training, the way athletes respond to different kinds of training loads, and how performance is evaulated in rowing.

2.1.1 Training Principles

Generally when a coach builds a training plan they have a few factors they can work with: volume, the amount of mileage or total time spent training, sessions intensity, how hard the given session is meant to be, and finally the frequency, or the time spent in different intensity zones. There are various ways to measure intensity, including heart rate, blood lactate concentration, velocity at maximal oxygen uptake (VO2 max), and rate of perceived exertion (RPE) [4]. Rowers tend to use heart rate zones or blood lactate concentration depending on access to the equipment to test blood lacate. Typically when a rower uses

calculated aerobic zones each zone will be a percentage of HR_{max}. Typically these zones are typically defined as follows:

Z1 "Very Light" intensity, 50% - 60% of HR_{max}

Z2 "Light" intensity, 60%-70% of HR_{max}

Z3 "Moderate" intensity, 70%-80% of HR_{max}

Z4 "Hard" intensity, 80%-90% of HR_{max}

Z5 "Maximum" intesity, 90%-100% of HR_{max}

The exact definition of these zones varies in the literature, as does the method to determine HR_{max}. However, most high level athletes will have completed some kind of stress test to determine their HR_{max}in order to train more effectively on their prescribed zones. A rower who uses lactate based training zones might use the following zones:

T1 basic oxygen utilization training (UT2) [lactate = 0-2 mmol/L]

T2 oxygen utilization training (UT1) [lactate = 2-3.5 mmol/L]

T3 anaerobic threshold training (AT) [lactate = 3.5-4.5 mmol/L]

T4 oxygen transport training (TR) [lactate = 4.5-6 mmol/L]

T5 anaerobic capacity training (AN) [lactate $\geq$ 6 mmol/L] [5]

Depending on how rigorous the testing protocol was, Heart Rate zones may be calculated for each zone, these may vary from the aerobic zones calculated from HR_{max}.

The most basic zone approximation approach uses three zones based around certain physiological thresholds, like, lacate thresholds (LT_1 and LT_2) and ventilatory thresholds. Cyclists may use critical power to determine these three basic zones, although this practice has not become popular in rowing training. The zones become simply, low-intensity, moderate-intensity, and high-intensity.

There are a few different approaches for distributing intensity for endurance training. The three main methods are: polarised training, sweet spot or threshold training, and pyramidal training. This guides the final factor a coach considers when building a general training plan, frequency. For the purposes of comparing polarized training (POL), threshold training (THR), and pyramidal training (PYR), the more basic three zones of intensity will be used. The breakdown per zone for each training method is as follows:

Polarised Training Far more time spent in the low-intensity zone [6].

> **Low-Intensity** 75%-85% of total training volume

> **Medium-Intensity** 5%-10% of total training volume

High-Intensity 5%-10% of total training volume

Threshold Training More time spent in the medium-intensity zone [6].

Low-Intensity 45%-55% of total training volume

Medium-Intensity 35%-55% of total training volume

High-Intensity 15%-20% of total training volume

Pyramidal Training Most time spent in low-intensity zone with progressively less time spent in higher zones [7].

Low-Intensity 75%-85% of total training volume

Medium-Intensity 15%-20% of total training volume

High-Intensity 5%-10% of total training volume

This report will not compare the effectiveness of different training distributions. Different distributions tend to be used by different sports, or depending on which energy system is being targeted. The use of polarized training is most common in rowing [4].

2.1.2 Energy Systems

An understanding of the aerobic and anerobic Systems

2.1.3 Physiological Response to Training

Adapations and that

2.1.4 Further training related sections

Discuss the need to include subsections for over reaching/training, detraining, tapering, etc.

2.1.5 Heart Rate Variablity

An explanation of HRV and how its used in training and recovery, include correlation with training load here, can include data from Churchill (2014) [8] here.

2.1.6 Performance

How it is measured, what feeds into it.

2.1.7 Summary

Does what it says on the box.

2.2 A Review of Performance Modelling

A review of models that already exist, including history of modelling, starting with Banister *et al.* [1], and including Edelmann-nusser *et al.* [9] and their comments on the efficacy of a linear systems model on an inherently non-linear biological behaviour. Essentially go through the steps done by Churchill (2014) [8].

2.2.1 Quantifying Training Load (Fatigue)

RPE

2.2.2 TRIMP

2.2.3 Quantifying Performance

erg score, otw results, telemetry

2.2.4 Impulse-Response Models

Limitations to the Impulse-Response Model

2.2.5 Alternative Models

Artifcial Neural Network (ANN) approaches.

Chapter 3

Data Collection and Management

Discussing approaches to data collection, key decisions in generating the model, eg. building a model which can be run with minimal effort by the users.

Chapter 4

The Model

4.1 Model Considerations

What led to the final model, difficulties, and data Considerations

4.2 The Implementation

Details about the model, including some examples of outputs (on my own data?)

Chapter 5

Discussion

Evaluation of methods and model, and a discussion of potential further step.

Chapter 6

Conclusion

Did we do what we set out to do.

Bibliography

[1] E. W. Banister, T. W. Calvert, M. V. Savage, and T. Bach, "A systems model of the effects of training on physical performance," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-6, no. 2, pp. 94–102, Feb. 1976. DOI: `10.1109/tsmc.1976.5409179`. [Online]. Available: `https://doi.org/10.1109/tsmc.1976.5409179`.

[2] J. Mäestu, J. Jürimäe, and T. Jürimäe, "Monitoring of performance and training in rowing," *Sports Medicine*, vol. 35, no. 7, pp. 597–617, Jul. 2005. DOI: `10.2165/00007256-200535070-00005`.

[3] K. S. Seiler and G. Ø. Kjerland, "Quantifying training intensity distribution in elite endurance athletes: is there evidence for an "optimal" distribution?" *Scandinavian Journal of Medicine & Science in Sports*, vol. 16, no. 1, pp. 49–56, 2006, ISSN: 0905-7188. DOI: `10.1111/j.1600-0838.2004.00418.x`.

[4] M. A. Rosenblat, A. S. Perrotta, and B. Vicenzino, "Polarized vs. threshold training intensity distribution on endurance sport performance: A systematic review and meta-analysis of randomized controlled trials," *Journal of Strength and Conditioning Research*, vol. 33, no. 12, pp. 3491–3500, Dec. 2019. DOI: `10.1519/jsc.0000000000002618`.

[5] A. Das, U. S. Kaniganti, S. J. Shenoy, P. Majumdar, and A. K. Syamal, "Monitoring training load, muscle damage, and body composition changes of elite indian rowers during a periodized training program," *Journal of Science in Sport and Exercise*, vol. 5, no. 4, pp. 348–359, Nov. 2022. DOI: `10.1007/s42978-022-00197-7`.

[6] S. Seiler and G. Ø. Kjerland, "Quantifying training intensity distribution in elite endurance athletes: Is there evidence for an "optimal" distribution?" *Scandinavian Journal of Medicine and Science in Sports*, vol. 16, no. 1, pp. 49–56, Oct. 2004. DOI: `10.1111/j.1600-0838.2004.00418.x`.

[7] S. Selles-Perez, J. Fernández-Sáez, and R. Cejuela, "Polarized and pyramidal training intensity distribution: Relationship with a half-ironman distance triathlon competition," *Journal of Sports Medicine and Science*, vol. 18, no. 4, pp. 708–715, Nov. 2019.

[8] T. Churchill, "Modelling athletic training and performance," Ph.D. dissertation, University of Canberra, Australian Capital Territory, Australia, 2014.

[9] J. Edelmann-nusser, A. Hohmann, and B. Henneberg, "Modeling and prediction of competitive performance in swimming upon neural networks," *European Journal of Sport Science*, vol. 2, no. 2, pp. 1–10, Apr. 2002. DOI: 10.1080/17461390200072201. [Online]. Available: https://doi.org/10.1080/17461390200072201.

Appendix A1

Appendix

You may use appendices to include relevant background information, such as calibration certificates, derivations of key equations or presentation of a particular data reduction method. You should not use the appendices to dump large amounts of additional results or data which are not properly discussed. If these results are really relevant, then they should appear in the main body of the report.

A1.1   Appendix numbering

Appendices are numbered sequentially, A1, A2, A3... The sections, figures and tables within appendices are numbered in the same way as in the main text. For example, the first figure in Appendix A1 would be Figure A1.1. Equations continue the numbering from the main text.