A Generalised Hydraulic Performance Model For Intermittent Exercise

WESTERN SYDNEY
UNIVERSITY

Fabian Clemens Weigend

School of Health Sciences
and
School of Computer, Data and Mathematical Sciences

A thesis submitted for the degree of

Doctor of Philosophy

2022

Statement of Authentication

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

Abstract

Data science advances in sports commonly involve 'big data', i.e., large sport-related data sets. However, such big data sets are not always available, necessitating specialised models that apply to relatively few observations. One important area of sport science research that features small data sets is the study of energy recovery during intermittent exercise. In this area, models are typically fitted to data collected from exhaustive exercise test protocols, which athletes can perform only a few times. Recent findings highlight that established recovery models, such as the so-called work-balance models, are too simple to adequately fit observed trends in the data. These models summarise the available energy capacities of an athlete during exercise in a single variable, which is referred to as work balance.

In this thesis we revisit a so-called hydraulic performance model and hypothesise that it is able to address the recently highlighted shortcomings of work-balance models. However, current literature has not fully validated the original hydraulic model, because it depends on physiological measures that cannot be acquired at the required precision or quantity.

We introduce a generalised interpretation and formalisation of the original hydraulic model that removes its ties to concrete physiological measures. We use evolutionary computation to fit its parameters to an athlete. In this way, we investigate a new hydraulic model that requires the same few data points as work-balance models, but promises to predict recovery dynamics more accurately.

To compare the hydraulic model to established work-balance models, we retrospectively apply them to data compiled from previously published studies. The hydraulic model outperforms established work-balance models on all defined metrics, even those that penalise models featuring higher numbers of parameters. However, the more accurate energy recovery predictions of the hydraulic model come at the cost of inaccurate predictions of metabolic responses during exercise, such as oxygen uptake.

In conclusion, while the new hydraulic performance model should not be used to predict metabolic responses during exercise, it promises to be a powerful tool for predicting energy recovery. This work carefully positions the new hydraulic model among existing models, with its benefits and limitations. The results render the

new hydraulic model a powerful alternative to address the shortcomings of established work-balance models and incentivise further investigation. Data and code are published as open source.

Acknowledgements

List of Figures

List of Tables

Abbreviations

A_nA anaerobic alactic energy source.

A_nL anaerobic lactic energy source.

Ae aerobic energy source.

AnF anaerobic fast energy source.

AnS anaerobic slow energy source.

An anaerobic energy sources.

D_{CP} the difference between CP and power output.

LF limited fast energy source.

LS limited slow energy source.

M_U maximal flow from the unlimited energy source.

M_{LF} maximal flow from the limited fast energy source.

M_{LS} maximal flow from the limited slow energy source.

O oxidative energy source.

P_U flow from U.

P_{peak} peak power output.

P_{rec} fixed recovery power output.

P_{work} fixed work power output.

P_{LS} flow from LS.

P constant power output.

T_{rec} recovery time.

U unlimited energy source.

$W'_{\text{bal}-\text{int}}$ work-balance model by Skiba et al. (2012).

$W'_{\text{bal}-\text{ode}}$ work-balance model by Skiba et al. (2015).

W'_{bart} work-balance model with $\mathcal{T}_t$ fitted by Bartram et al. (2018).

W'_{skib} work-balance model with $\mathcal{T}_t$ fitted by Skiba et al. (2015).

W'_{weig} work-balance model with $\mathcal{T}_t$ fitted by Weigend et al. (2022a).

W' finite energy reserve for work above critical power.

W_{lim} total achievable workload.

$\mathcal{T}_{\text{bart},t}$ $\mathcal{T}_t$ by Bartram et al. (2018).

$\mathcal{T}_{\text{skib},t}$ $\mathcal{T}_t$ by Skiba et al. (2015).

$\mathcal{T}_{\text{weig},t}$ $\mathcal{T}_t$ by Weigend et al. (2022a).

$\dot{V}_{\text{O}_2\text{max}}$ maximal oxygen uptake.

$\dot{V}_{\text{O}_2\text{peak}}$ peak oxygen uptake.

$\dot{V}_{\text{O}_2}$ oxygen uptake.

γ distance from the bottom to the limited hydraulic model tank
(LS of hydraulic$_{\text{weig}}$, or $A_n L$ of the M-M model).

AIC_{c} the small-sample version of the Akaike Information Criterion.

AWC anaerobic work capacity.

CP critical power.

hydraulic$_{\text{2t}}$ two-tank hydraulic.

hydraulic$_{\text{weig}}$ generalised hydraulic model by Weigend et al. (2021).

ϕ distance from the bottom to the unlimited hydraulic model tank
(U of hydraulic$_{\text{weig}}$, or O of the M-M model).

ψ distance from the top to An.

θ distance from the top to the limited hydraulic model tank
(LS of hydraulic$_{\text{weig}}$, or $A_n L$ of the M-M model).

g the level of depletion of a hydraulic model tank
(LS of hydraulic$_{\text{weig}}$, or $A_n L$ of the M-M model).

h the level of depletion of a hydraulic model tank
    (An of hydraulic$_{2t}$, LF of hydraulic$_{weig}$, or A_nA of the M-M model).

p_{Ae} flow from Ae to An.

M-M-S Margaria-Morton-Sundström.

M-M Margaria-Morton.

TTE time to exhaustion.

TTF time to task failure.

AT anaerobic threshold.

AutoML Automated Machine Learning.

LAT lactate threshold (moderate-heavy boundary).

MAE mean absolute error.

MAP maximum aerobic power.

MOEA/D Multi-Objective Evolutionary Algorithm with Decomposition.

MSE mean squared error.

MVC maximal voluntary contraction.

NRMSE normalised root mean squared error.

ODE ordinary differential equations.

RB recovery bout.

RMSE root mean squared error.

SD standard deviation.

SEE standard error of estimation.

WB1 work bout 1.

WB1→RB→WB2 recovery protocol as defined in Section 2.3.3.1.

WB2 work bout 2.

Contents

Chapter 1

Introduction

Emerging technologies that enable the real-time monitoring of athletes in training and competition have fostered interest in methods to predict capabilities of athletes. Predictive models for how much an athlete 'has left in the tank' enable the investigation of pacing strategies (Behncke, 1997; Sundström et al., 2014; de Jong et al., 2017) and to optimise the outcome of a competition (Hoogkamer et al., 2018). They can be described as a 'digital athlete': a computer-based model for enhancing training programming or strategy optimisation. A foundation for such advances is research in performance modelling, which can be understood as the mathematical abstraction of exercise physiology.

Many mathematical models of the human body exist and, in a simplified view, they are either simple and imprecise or complex but difficult to use. Typically, the application of performance models requires fitting to exhausting exercise tests, which athletes can only perform a few times. Therefore, an advantage of simple models is that they can be applied to few data points. However, because these models are simple, they are either limited to specific applications or their predictions can be imprecise (Skiba and Clarke, 2021). This is dangerous because an overestimation of capabilities may cause athletes to overexert themselves and risk injury. Conversely, an underestimation can result in athletes not exercising at their full potential.

More accurate predictions can be expected from complex models that, for example, consider how efficiently chemical processes of the human body generate required energy. But these models require more data to be fitted, and some relevant chemical processes are impossible to monitor during exercise (Morton, 2006; Sundström, 2016; Lidar et al., 2021). Therefore, complex models are difficult to use because they require extensive exercise tests or measurements that are impossible to obtain.

This work developed and investigated a combination of a simple and a complex model. The combined model is a simplified form of a so-called three-tank hydraulic model. The new hydraulic model can be applied to commonly collected data with

an optimisation approach and, at the same time, it is complex enough to make more accurate predictions than competing currently established models.

1.1 Main contributions

- This work developed a hydraulic performance model that supports athletes in understanding and predicting their exercise capabilities.

- The model has been published as a new pathway to approximate energy expenditure and recovery of an athlete at the *Genetic and Evolutionary Computation Conference* (GECCO) (Weigend et al., 2021).

- An extensive comparison of energy recovery predictions of established models and our new model has been published in *Annals of Operational Research* (ANOR) (Weigend et al., 2022a).

- Further elaborations of benefits and limitations of the new model have been published as a preprint as Weigend et al. (2022b). The submitted paper focuses on oxygen uptake predictions during exercise.

- Two open-source python packages have been developed and published to make our tools, models, and data more easily available to the community. These packages are `pypermod`[1] and `threecomphyd`[2] and their function and content are described throughout this work.

1.2 Layout

This introduction is followed by a literature review in Chapter 2, which covers currently established types of performance models, their applications, and their shortcomings. Subsequently, Chapter 3 details the problem statement of this work and the arising research questions. Chapter 4 is dedicated to a comprehensive mathematical definition of the new model. Chapter 5 describes how the new model is fitted to data from athletes. Chapter 6 compares predictions of the new model to established models and previously published data. Chapter 7 scrutinises limitations of metabolic predictions of the new model. Chapter 8 outlines future research directions and concludes this work.

---

[1] https://github.com/faweigend/pypermod
[2] https://github.com/faweigend/three_comp_hyd

Chapter 2

Literature Review

This chapter discusses currently established mathematical models to describe responses of the human body to the stimulus of exercise. Section 2.1 begins by defining how the terms 'performance' and 'energy' are used. Then, advances in predicting human performance as well as energy expenditure and energy recovery are discussed in Sections 2.2 and 2.3. Hydraulic models are introduced in Section 2.4 and related to other established models for energy expenditure and recovery.

2.1 Exercise performance

The definition of athlete performance depends on the perspective. For a soccer player, it might be the amount of successful passes, ball contact time or contribution to scored goals. However, in endurance sports such sports specific performance measures are often not considered and performance is defined in its most general measure: as an amount of expended energy. In units, energy is measured in joules (J) and power in watts (W). The following relationship holds:

$$\text{power (W)} = \frac{\text{energy (J)}}{\text{time (s)}}.$$
(2.1)

Winter et al. (2016) also considered that, in sports, the understanding of power can vary according to the perspective. As an example, a nutritionist may use power as the rate of chemical energy transfer from food and quantify energy expenditure in kilo calories per second (Whaley et al., 2006; Winter et al., 2016). This work uses the above in Equation (2.1) stated units and defines performance as an amount of expended energy.

2.2 Energy expenditure

The analysis of energy expenditure and how much energy an athlete has available lies at the core of performance modelling. This section begins with early findings and the prominent critical power model.

2.2.1 Critical power model

Some early findings and observations in human performance analysis were published by Hill (1925). He was the first to suggest a hyperbolic relationship between a constant power output and time until exhaustion. He based this suggestion on observations of isolated frog muscles, which then were related to humans with velocity versus time plots of swimming and running world records over various distances.

Monod and Scherrer (1965) formalised the findings of Hill (1925) as a mathematical model. They were the first to define the term critical power (CP) of a muscle. On the basis of their observations, they proposed that when work is performed at a constant rate above CP, the relation of the total achievable workload (W_{lim}) and time to exhaustion (TTE) is the result of a linear function

$$W_{lim} = a + \text{CP} \cdot \text{TTE}. \tag{2.2}$$

Monod and Scherrer (1965) define a as the muscle's energy reserve. Given this equation, a power output above CP requires a muscle to draw energy from a to match power demands. Because a is a finite capacity, it will ultimately be depleted when TTE is reached and the muscle will be unable to perform at the required rate.

Moritani et al. (1981) confirmed that these ideas also apply to more whole-body exercise tests on a cycle ergometer and related the estimated CP to the anaerobic threshold (AT). According to Beaver et al. (1986), AT marks the onset point from which anaerobic energy production is needed in addition to aerobic energy production to meet energy demands. Moritani et al. (1981) argued that the ideas of CP and AT are very similar, because both mark the border between a sustainable (aerobic) and a limited (anaerobic) energy production, and observed a high correlation between both variables. This way they fostered the understanding of a in Equation (2.2) as the anaerobic work capacity (AWC). This notion implies that work below CP can be accomplished by aerobic energy sources and work above CP must additionally draw from anaerobic capacities. A corresponding update of a in Equation (2.2) results in:

$$W_{lim} = \text{AWC} + \text{CP} \cdot \text{TTE}. \tag{2.3}$$

Given that Equation (2.2) and Equation (2.3) assume a constant work rate above CP, W_{lim} equals the product of constant power output (P) and TTE:

$$W_{lim} = P \cdot \text{TTE}. \tag{2.4}$$

Using Equation (2.4), Equation (2.3) can be reformulated as:

$$\text{TTE} = \frac{\text{AWC}}{P - \text{CP}}. \tag{2.5}$$

Whipp et al. (1982) then denoted AWC as a finite energy reserve for work above critical power (W') and both terms were used interchangeably. When solved for P, that results in the equation:

$$P = \frac{W'}{\text{TTE}} + \text{CP}. \tag{2.6}$$

Hill (1993) referred to Equation (2.5) as the nonlinear power-time relationship, to Equation (2.6) as the linear power-1/time relationship, and to Equation (2.3) as the linear work-time relationship. Current literature refers to the critical power model in one of these three forms or in all of them (Clarke and Skiba, 2013; Mattioni Maturana et al., 2018; Sreedhara et al., 2019; Jones et al., 2019).

2.2.1.1 Assumptions of the critical power model

Hill (1993) and Morton (2006) summarised the assumptions of the critical power model as the following:

1. An individual's power output is a function of two energy sources: aerobic (using oxidative metabolism) and anaerobic (non-oxidative metabolism).

2. Aerobic energy is unlimited in capacity but its conversion rate into power output is limited (CP).

3. Anaerobic energy is limited in capacity (W') but its conversion rate is unlimited.

4. Exhaustion occurs when all of the W' is depleted.

However, more recent publications suggest that the assumed clear separation between aerobic and anaerobic energy sources is incorrect and AWC and W' are not the same. Dekerle et al. (2006) compared the estimated W' using Equation (2.6) to their understanding of AWC: as the time integral above the power output expected from the measured oxygen uptake. The conclusion was that W' underestimates the accumulation of expended anaerobic energy. In addition, Noordhof et al. (2013) summarised that CP and W' should not be considered separate aerobic and anaerobic entities but interrelated components. Poole et al. (2016) suggested the conceptualisation of W' as a buffer 'to resist exercise intolerance during supra-CP exercise, where the source of the buffer will vary depending on the conditions' (p. 2331).

As already shown by varying definitions for W' and AWC, all of the above four assumptions by Hill (1993) are incorrect from a strict physiological perspective (Morton, 2006; Clarke and Skiba, 2013; Poole et al., 2016). Moreover, in the review by Hill (1993) in which he summarised these assumptions, he had already emphasised that the attractiveness of this model lies in its simplicity. The simplicity of the critical power model allows a straightforward procedure to obtain predictions and to fit the model to athletes.

2.2.1.2 Time to exhaustion and time to task failure

In the light of the above stated assumptions of the critical power model, the distinction between time to time to exhaustion (TTE) and time to task failure (TTF) becomes important. The fourth assumption of the critical power model states that exhaustion occurs when all of the W' is depleted. Therefore, the time from the beginning of exercise until W' is depleted is being referred to as TTE (Hill, 1993; Morton, 2006; Sreedhara et al., 2019). However, literature reconsidered the term "exhaustion" because it may not appropriately describe the physiological state after short extreme-intensity efforts (Hunter et al., 2004).

As an example for this reconsideration, Jones et al. (2010) referred to the critical power model as the power to TTE relationship. Later, Jones and Vanhatalo (2017) write "exhaustion" in quotes and, more recently, Jones et al. (2019) avoided the term exhaustion and refer to it as the time to the limit of exercise tolerance where the limit coincides with the exhaustion of W'. Not mentioning exhaustion at all, McCrary et al. (2018) and Muniz-Pumares et al. (2019) used the same critical power model equations as introduced in the previous sections, but they used TTF instead of TTE.

On the contrary, other recent work, for example (Sreedhara et al., 2019; Caen et al., 2021; Skiba and Clarke, 2021), kept the semantics of the early works and refer to the depletion of W' as exhaustion and use the term TTE. This work is in large parts based upon publications by Morton (2006); Caen et al. (2021); Skiba and Clarke (2021). Therefore, we follow their notation and hereby define that we use TTE to refer to the time until W' is depleted or predicted to be to be depleted.

2.2.1.3 Fitting the critical power model

A well-known hyperbolic representation of the critical power model is depicted in Figure 2.1. The nonlinear power-time relationship from Equation (2.6) is plotted as a TTE versus P plot. The resulting hyperbolic curve has one asymptote at TTE $= 0$ and one at $P = $ CP.

CP and W' of an athlete are traditionally derived from a series of exhaustive exercise tests at distinct constant exercise intensities (Hill, 1993; Clarke and Skiba, 2013). For each of these tests, the athlete has to exercise at a constant power output until volitional exhaustion is reached. Constant power output and TTE are recorded and, as an example, represent one dot in Figure 2.1. Henceforth, such a test is referred to as a TTE test. Hill (1993) suggested conducting three to five TTE tests at distinct constant power outputs. Then, the nonlinear power-time relationship, the linear power-1/time relationship, or the linear work-time relationship can be fitted to all test results such that they minimise the prediction error.
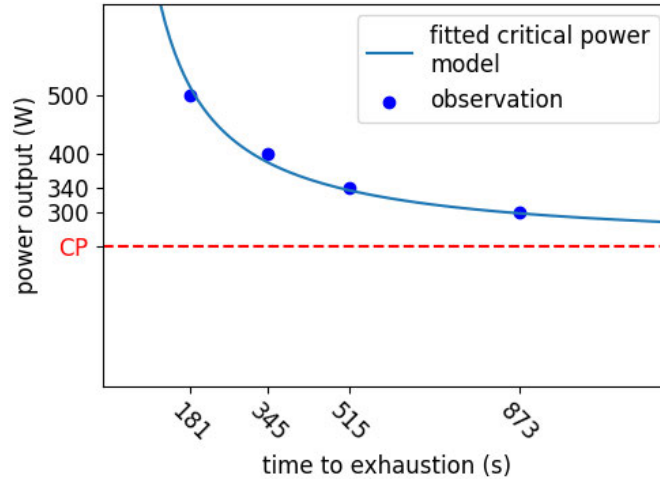
Figure 2.1: The nonlinear power-time relationship of constant power output and time to exhaustion. The red dashed line marks the CP asymptote. Example TTE test results are depicted as dots.

Because TTE tests require an athlete to exercise until volitional exhaustion, test results are subject to a natural variability. Therefore, the fitted CP and W' of all three forms of the critical power model are rarely identical, even if they are applied to the same test results (Hill, 1993; Jones et al., 2019). Research has continuously evaluated how to obtain accurate and reliable estimates of CP and W'. Jones et al. (2019) summarised that TTE trials should be between 2 min and 15 min with the shortest trial 2 min to 3 min and the longest 10 min to 15 min. Accuracy of fitted parameters should be determined by relative standard error of estimation (SEE). The SEE associated with fitted CP should be $< 5\%$ of CP and associated with W' $< 10\%$ of W' (Jones et al., 2019). The 'individual best fit' for an athlete should be the critical power model relationship (nonlinear power-time, linear power-1/time, or linear work-time) that results in the smallest sum of relative SEE associated with CP and W' (Jones et al., 2019). Using these 'individual best fit' CP and W', reliable predictions for other power outputs in the severe exercise intensity domain can be made.

Vanhatalo et al. (2007) proposed a procedure to approximate CP and W' without the need for multiple TTE tests. They proposed the so-called 3-min all-out cycling test for which the athlete has to pedal at maximal effort against a fixed resistance on a cycle ergometer. This test is especially strenuous because the athlete is required to not pace themselves. From the start of the test, the power output reaches its peak within the first seconds and then declines as the athlete depletes their W'. By approximately the last 30 s of the test, power output stabilises around a value similar to CP. Vanhatalo et al. (2007) compared to CP and W' estimated with the conventional 5 TTE tests. The study found that the power outputs at the

end of the 3-min all-out test were not significantly different from CP and that the expended energy above CP during their tests was not significantly different from W'. The 3-min all-out is attractive because it requires only one test to determine the power- duration parameter estimates, but Muniz-Pumares et al. (2019) noted that some studies have reported significant differences in W' or an overestimation of CP. The 'gold standard' remains a series of TTE tests (Muniz-Pumares et al., 2019). Therefore, one has to differentiate between CP and the 'end test power' of the 3-min all-out cycling test and as well as between W' the 'work above end test power'.

As previously stated, Hill (1993) and Clarke and Skiba (2013) emphasised that the attractiveness of the critical power model lies in its simplicity. It should not be employed if highly accurate predictions are required. In fact, because of its imprecise assumptions, subsequent publications have suggested that no efforts shorter than 2 min or longer than 30 min should be considered for prediction (Housh et al., 1989; Vanhatalo et al., 2011; Clarke and Skiba, 2013). Nevertheless, its straightforward application and its elegant abstraction have led to improved understanding and prediction of performance dynamics (Vanhatalo et al., 2011; Poole et al., 2016; Jones et al., 2019).

2.2.2 Extensions of the critical power model

This section introduces alternative concepts, relates them to the assumptions of CP and W' and discusses where shortcomings are addressed. Another often referenced model, which is also mentioned in reviews by Morton (2006); Sreedhara et al. (2019), was proposed by Ward-Smith (1985) for the prediction of sprint performances between 100 m and 10 000 m. Similarly to the assumptions of critical power based models, Ward-Smith (1985) differentiated between aerobic and anaerobic energy systems. He defined the maximum rate of energy release that can be maintained by oxidative metabolism as R. Of special interest in comparison with the critical power model is the equation for power available during a sprint at time t:

$$P(t) = P_{\max} \cdot e^{-\lambda t} + R \cdot (1 - e^{-\lambda t}). \tag{2.7}$$

Ward-Smith (1985) defined $P_{\max}$ as the maximum power available from the anaerobic metabolism and λ was used as a parameter to govern the variation of power over time. Depending on λ, the available anaerobic power decreases at an exponential rate, while available power from aerobic sources increases to the maximum of R at the same exponential rate. Thus, a decrease in $P_{\max}$ represents drainage of an additional finite energy storage that diminishes over time, while R defines the rate of energy that can be maintained indefinitely. Similarities to W' and CP are apparent.

Differences lie in the limited use of W' and the initial increment of anaerobic energy contribution from the beginning of exercise. Ward-Smith (1985) incorporated this model into concepts of running resistance and thermal energy generation and reported conforming predictions for actual running data.

Peronnet and Thibault (1989) extended the work by Ward-Smith (1985) to account for longer running distances and even a full marathon. While Ward-Smith (1985) proposed R as the maximal sustainable energy rate, Peronnet and Thibault (1989) replaced it with the term maximum aerobic power (MAP) and suggested that this power output is not maintainable over an indefinite time.

Relating the assumptions of Ward-Smith (1985) and Peronnet and Thibault (1989) to those of the critical power model, two of its major inaccuracies were addressed. As observable in Figure 2.1, one of the asymptotes of the hyperbola lies at TTE = 0 s, suggesting that immediately accessible power is infinite. Infinite power output is biologically impossible. Ward-Smith (1985) accounts for this by incorporating $P_{\max}$ in a way that limits the available power at times close to $t = 0$ s. Peronnet and Thibault (1989) accounted for the assumption that CP can be maintained for indefinite time, that is, the second asymptote at $P = $ CP. Their diminishing MAP over time results in a more realistic conception (Peronnet and Thibault, 1989).

Although these models result in higher prediction accuracy, they come at the price of flexibility. Because Ward-Smith (1985) designed his concept to model sprint performances, the assumption of this and of comparable models is that the athlete always performs at the maximal possible power output to finish the task. When examining the origin of the critical power model, this may seem comparable, but the assumed linear use of W' also allows an individual to predict performance for efforts with varying power outputs.

A valuable extension that maintains this flexibility was proposed by Morton (1996), who extended the critical power model by an additional parameter and named it the three parameter model. He added the parameter k to the known Equation (2.6), which resulted in

$$\text{TTE} = \frac{W'}{P - \text{CP}} + k. \tag{2.8}$$

The limitation $k < 0$ applies and thus k moves the x-asymptote of the hyperbola into the negative values. This causes the model to have a y-axis intercept, which can be regarded as the maximal instantaneous power. Similarly to the $P_{\max}$ defined by Ward-Smith (1985), this caps the power output for TTE = 0 s. Still, this model shows that inaccuracies for short durations and adaptions have been proposed by Vinetti et al. (2019). However, the three parameter model is well covered by reviews Clarke and Skiba (2013); Jones and Vanhatalo (2017); Mattioni Maturana

et al. (2018); Sreedhara et al. (2019) and provides a meaningful addition without increasing complexity significantly.

An alternative model, which has recently gained more attention (Gorostiaga et al., 2022; Drake et al., 2022), is the so-called power-law model, which was popularised by Riegel (1981, 1977). This model assumes that the relationship of a constant power output P to TTE is

$$P = \text{TTE } S^{E-1}. \tag{2.9}$$

The parameter S is required to be greater than 0 and affects the scaling of the power-duration curve. As described by Drake et al. (2022), it can be called "speed parameter" and it is the power that a cyclist can sustain for exactly one second. The parameter E is called "endurance parameter", is required to be between 0 and 1, and affects the slope of the power-duration curve. The higher the value of E, the faster the curve decays. A cyclist who specialises on sprints will have a large value for S and a small E. On the contrary, a cyclist who specialises on endurance will have a small value for S and a large E (Drake et al., 2022).

While other reviews report the critical power model as the 'gold standard' (Jones and Vanhatalo, 2017; Muniz-Pumares et al., 2019), it was highlighted in an extensive review by Drake et al. (2022) that the power-law model results in a better fit and more realistic predictions when retrospectively applied to data of previous publications. We will follow the current debate around the critical power model with interest and will consider alternatives as future work (Gorostiaga et al., 2022). However, in this work we will mainly focus on energy recovery predictions during intermittent exercise. As discussed in the following section the established models in this domain require an estimation of critical power.

2.3 Energy recovery and intermittent exercise

While the critical power model predicts energy expenditure at high intensities, it does not consider the recovery of W' after exercise has ended or during exercise at low intensities. Formally, exercise protocols that alternate between intensities below CP and above CP can be defined as intermittent exercise. Predictions for intermittent exercise are important for athletes, such as cyclists, whose power demands vary between having to climb steep ascents and recovering while riding downhill. To predict the performance capabilities of athletes during intermittent exercise, models need to predict the recovery of W' during phases of exercise below the CP intensity.
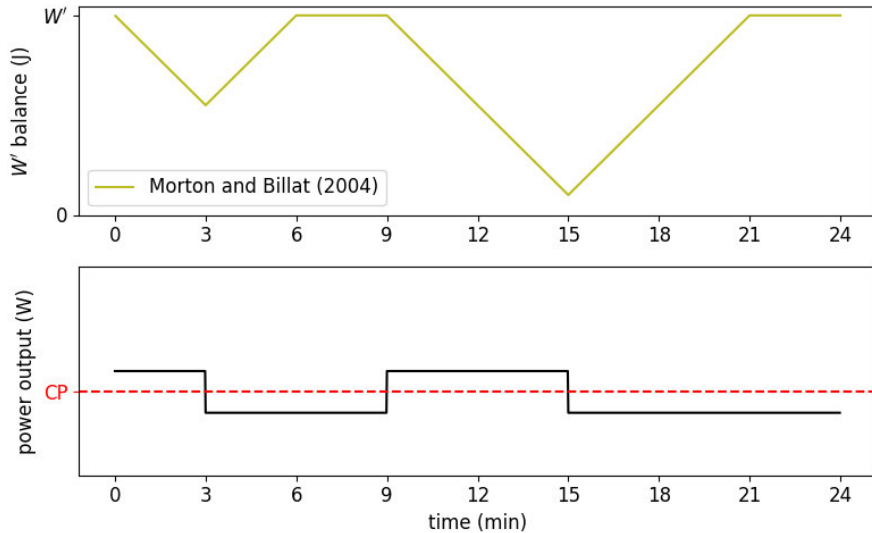
Figure 2.2: Utilisation and recovery mechanics of W' according to Morton and Billat (2004). Exercise intensity and CP are depicted in the lower plot. The upper plot shows predicted utilisation and recovery of W' balance.

2.3.1 Linear recovery

Morton and Billat (2004) were among the first to extend the CP model to consider intermittent power outputs. They proposed that recovery kinetics are similar to the assumed linear utilisation kinetics of the critical power model. According to the model by Morton and Billat (2004), W' recovers linear dependent on the difference from CP. An example is depicted in Figure 2.2. The bottom plot shows the power demand on the athlete, stated as power output. The exercise features a 3 min and a 6 min work bout above CP and recovery bouts below CP. While the athlete works at a rate above CP, available W' capacity is drained according to the assumptions of the critical power model. Remaining W' capacity may be denoted as W' balance. In the example in Figure 2.2, CP is at 200 W, work-bout intensity at 250 W, and recovery intensity at 150 W. Chidnok et al. (2012) confirmed the validity of this model and related W' recovery to prolonged exercise tolerance in intermittent exercise. Chidnok et al. (2012) suggested that the consideration of recovery aids the creation of effective interval training programs. However, the assumption of a linear recovery turned out to be too simple.

2.3.2 Work-balance models

Further refinements of W' recovery predictions were introduced by Ferguson et al. (2010), who suggested a curvelinear recovery. Skiba et al. (2012) then confirmed this observation and proposed an exponential recovery model: the work-balance model. Since then, various forms of the work-balance model have developed into

one of the most widely investigated approaches to predict recovery of W' during intermittent exercise (Jones and Vanhatalo, 2017; Sreedhara et al., 2019; Skiba and Clarke, 2021). Since the first publication by Skiba et al. (2012), an updated form was introduced by Skiba et al. (2015) and another alternative form was proposed by Bartram et al. (2018). Work-balance models have been used to search for optimal drafting strategies in running (Hoogkamer et al., 2018) or to predict phases of perceived exhaustion during cycling exercise (Skiba et al., 2014). Each of the available work-balance models is introduced in the following sections.

2.3.2.1 The integral model

Skiba et al. (2012) formulated their model with a convolution integral. Henceforth, we will refer to it as work-balance model by Skiba et al. (2012) ($W'_{\text{bal-int}}$). The same primary author reported in Skiba and Clarke (2021) that their original definition of $W'_{\text{bal-int}}$ was mathematically imprecise and updated their equation to:

$$W'_{\text{bal-int}}(t) = W' - \int_0^t e^{\frac{-(t-u)}{\tau_{W'}}} W'_{\text{exp}}(u)\mathrm{d}u, \tag{2.10}$$

which denotes how the model estimates remaining W' balance at a given time point t. $W'_{\text{exp}}(u)$ represents energy that was spent in one time step u and is estimated as:

$$W'_{\text{exp}}(u) = \begin{cases} 0, & P(u) \leq \text{CP} \\ \int (P(u) - \text{CP})\mathrm{d}u, & P(u) > \text{CP} \end{cases}. \tag{2.11}$$

This equation shows that the model does not distinguish between the severe and extreme intensity exercise domains because it only distinguishes between above CP or below. According to Equation (2.10), for every u between $0\,\text{s}$ and t, spent energy at time step u is subtracted from the total available W' capacity. Before the subtraction, spent energy is multiplied by $e^{\frac{-(t-u)}{\tau_{W'}}}$, which approaches 0 as the difference between t and u increases. Thus, energy that was spent at u recovers exponentially as t increases, and $\tau_{W'}$ affects the speed of recovery. Skiba et al. (2012) fitted an exponential function to obtain $\tau_{W'}$ to observed recovery ratios during various recovery intensities. Their fitted function is:

$$\tau_{W'} = 546 \cdot e^{(-0.01 \cdot D_{\text{CP}})} + 316, \tag{2.12}$$

where the difference between CP and power output (D_{CP}) uses the average power output during all phases of recovery.

Together, Equations (2.10) to (2.12) define the $W'_{\text{bal-int}}$ model and Figure 2.3 depicts an example of how it predicts energy expenditure and recovery. Predictions of the previously introduced model by Morton and Billat (2004) have been added for comparison. The step size for the estimations was $1\,\text{s}$. During the first $3\,\text{min}$ of
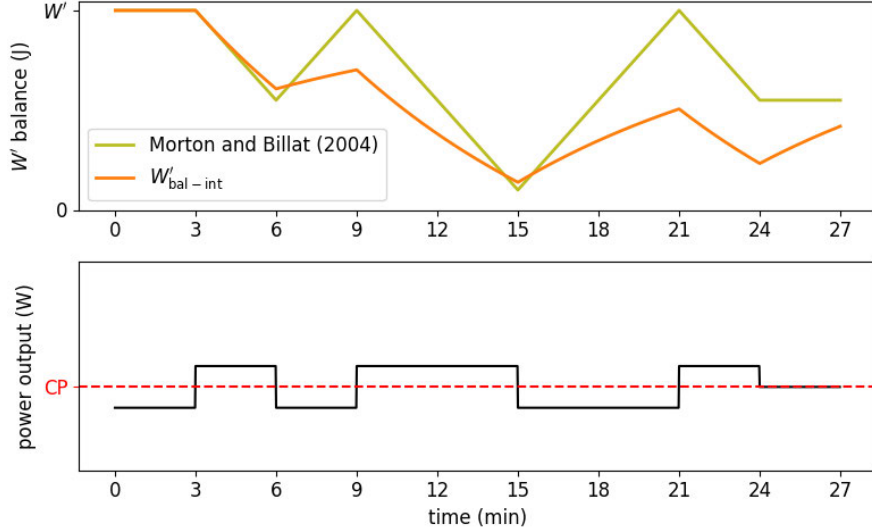
Figure 2.3: Example energy expenditure and recovery predictions of the work-balance model for intermittent exercise by Skiba et al. (2012) ($W'_{\text{bal}-\text{int}}$) and the model by Morton and Billat (2004).

exercise, the power output was below CP, so no additional energy had to be drawn from the W'.

At 181 s (3 min and 1 s) the power output switched to 250 W, which was 50 W above CP. Therefore, with a step size of 1 s, $W'_{\text{exp}}(181)$ was 50 W $\cdot$ 1 s = 50 J and $W'_{\text{bal}-\text{int}}(181) = W' - 50$ J. For the next step $t = 182$ s, the power output was the same, so $W'_{\text{exp}}(182) = 50$ J. However, $W'_{\text{bal}-\text{int}}(182)$ did not drop by an additional 50 J, because of the integral in Equation (2.10). Instead, the sum of $W'_{\text{exp}}(181) \cdot e^{\frac{-(182-181)}{\tau_{W'}}}$ and $W'_{\text{exp}}(182)$ were subtracted from W', which resulted in slightly less than 100 J. The energy that was spent in the 181st second recovered slightly in the 182nd second. This dynamic is a key difference of $W'_{\text{bal}-\text{int}}$ in comparison with the critical power model and the previous model by Morton and Billat (2004). As observable in Figure 2.3, energy expenditure of $W'_{\text{bal}-\text{int}}$ is not linear at a constant power output.

Between 6 min and 9 min in Figure 2.3, power output dropped below CP again. The model by Morton and Billat (2004) predicted linear recovery, the $W'_{\text{bal}-\text{int}}$ model predicted a slower and curvelinear recovery. For all u during this time $W'_{\text{exp}}(u) = 0$ J, so no additional energy was drawn from W'. The predicted recovery by $W'_{\text{bal}-\text{int}}$ is the exponential decrease of previously accumulated W'_{exp} that approached 0 J as t increased.

Another key difference between the model by Morton and Billat (2004) and $W'_{\text{bal}-\text{int}}$ is observable in the final 3 min of Figure 2.3. The critical power model as in Equation (2.6) and the model by Morton and Billat (2004) assume that CP is the maximal power output that can be maintained for an indefinite amount of time.

13

During power output equal to CP, no energy was drawn from W', but according to Morton and Billat (2004) no recovery was possible. However, while $W'_{\mathrm{bal-int}}$ also predicted that no additional energy was drawn from W' it assumed that previously spent energy recovered. $W'_{\mathrm{bal-int}}$ predicts recovery during exercise at CP. The $W'_{\mathrm{bal-int}}$ model assumes that the maximal power output that can be maintained for an indefinite amount of time is higher than CP.

Skiba and Clarke (2021) acknowledged that $W'_{\mathrm{bal-int}}$ makes such logically inconsistent predictions in extreme cases. The model uses CP as the threshold for sustainable exercise but predicts recovery during exercise at CP, and therefore violates its own assumption. As a result, Skiba and Clarke (2021) recommended as best practice to use $W'_{\mathrm{bal-int}}$ predictions only for exercise in which power outputs above CP are limited to short bursts. In Skiba et al. (2015), another form of the model was proposed that addresses some of these issues.

2.3.2.2 Ordinary Differential Equations models

Subsequent to $W'_{\mathrm{bal-int}}$, Skiba et al. (2015) derived a model for expenditure and recovery of W' from first principles. The formal derivation is detailed in Appendix 1 of Skiba et al. (2015). They conceptualised available W' balance as liquid in a tank that is emptied or refilled according to energy expenditure or recovery. These dynamics were specified as ordinary differential equations; henceforth, their model is referred to as work-balance model by Skiba et al. (2015) ($W'_{\mathrm{bal-ode}}$). In Skiba and Clarke (2021), they also specified a discrete version of $W'_{\mathrm{bal-ode}}$, which is introduced in the following section. Henceforth, the by $W'_{\mathrm{bal-ode}}$ predicted remaining capacity of W' at a discrete time point t is denoted as $W'_{\mathrm{bal-ode},t}$. P_t refers to the power output at a discrete time step t. Δt is the difference between the discrete time step $t-1$ and t in seconds. The discrete $W'_{\mathrm{bal-ode}}$ is defined as:

$$W'_{\mathrm{bal-ode},t} = \begin{cases} W'_{\mathrm{bal-ode},t-1} - (P_t - \mathrm{CP})\Delta t, & \text{for } P_t \geq \mathrm{CP} \\ W' - (W' - W'_{\mathrm{bal-ode},t-1}) \cdot e^{\frac{-\Delta t}{\mathcal{T}_t}}, & \text{for } P_t < \mathrm{CP}. \end{cases} \quad (2.13)$$

During a constant power output above or at CP, $W'_{\mathrm{bal-ode},t}$ decreases linearly as t increases, as the critical power model predicts. During power outputs below CP, $W'_{\mathrm{bal-ode},t}$ increases exponentially with W' as its asymtote. The value $\mathcal{T}_t$ affects recovery speed and its estimation varies between distinct $W'_{\mathrm{bal-ode}}$ models. At a discrete time step t, the $\mathcal{T}_t$ by Skiba et al. (2015) ($\mathcal{T}_{\mathrm{skib},t}$) is estimated as

$$\mathcal{T}_{\mathrm{skib},t} = \frac{W'}{D_{\mathrm{CP},t}}, \quad (2.14)$$

where $D_{\mathrm{CP},t}$ represents the difference between P_t and CP. Henceforth, Equation (2.13) with Equation (2.14) is referred to as work-balance model with $\mathcal{T}_t$ fitted by Skiba et al. (2015) (W'_{skib}).
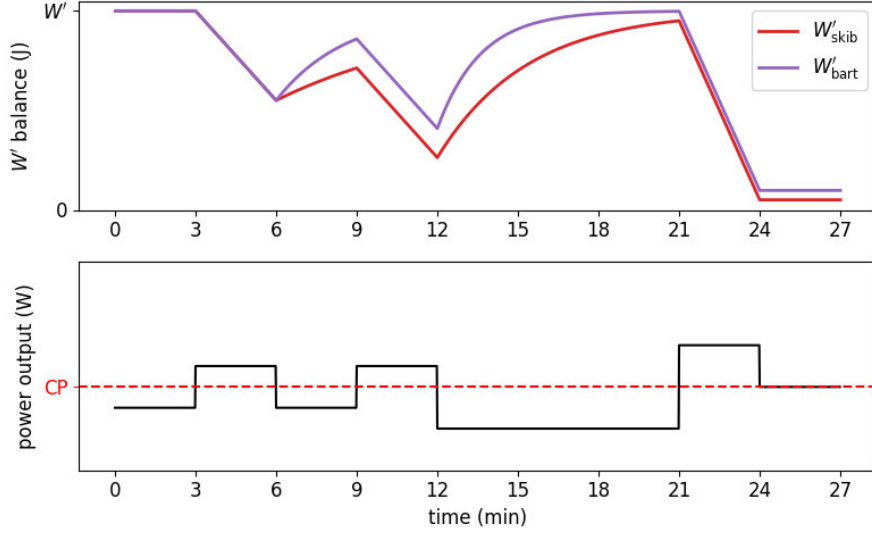
Figure 2.4: Predicted energy expenditure and recovery by the work-balance models by Skiba et al. (2015) or by Bartram et al. (2018). W' expenditure during exercise above CP resembles the assumed linear depletion of the critical power model. Differences in predicted curvelinear recovery kinetics are clearly visible.

Figure 2.4 depicts an example for W'_{skib} predictions. During the time steps of the first 3 min, P_t was below CP and $W'_{\text{bal}-\text{ode},t}$ (i.e., the predicted available W' balance) remained at its maximum. Then, the power output increased above CP for the next 3 min. The available W' balance decreased by $P_t - \text{CP}$ per second. Between 6 min and 9 min, P_t dropped below CP again, W'_{skib} simulated recovery and $W'_{\text{bal}-\text{ode},t}$ rose exponentially with W' as its asymptote. Speed of recovery was affected by $\mathcal{T}_{\text{skib},t}$ from Equation (2.14), which took the difference between P_t and CP into account. The effect of the difference between P_t and CP on recovery dynamics is observable by comparing recovery between 6 min and 9 min to recovery between 12 min and 21 min. During the second recovery bout, P_t was lower and therefore the slope of the exponential recovery was steeper. During the last 3 min, P_t was exactly equal to CP and thus $W'_{\text{bal}-\text{ode},t}$ did not change. If $W'_{\text{bal}-\text{ode},t}$ were to reach 0 J, exhaustion would be predicted. In the example in Figure 2.4, the athlete was predicted to be close to exhaustion, but some energy capacities remained.

Bartram et al. (2018) investigated the recovery speed of W' of elite athletes and observed faster recovery rates than Skiba et al. (2015). They proposed another $\mathcal{T}_t$ for Equation (2.13) to predict quicker recovery ratios. The $\mathcal{T}_t$ by Bartram et al. (2018) ($\mathcal{T}_{\text{bart},t}$) was defined as

$$\mathcal{T}_{\text{bart},t} = 2287.2 \cdot D_{\text{CP},t}^{-0.688}. \tag{2.15}$$

Henceforth, Equation (2.13) with Equation (2.15) is referred to as work-balance model with $\mathcal{T}_t$ fitted by Bartram et al. (2018) (W'_{bart}). Predictions of W'_{bart} are

| WB1 | → | RB | → | WB2 |

| intensity: P_work | intensity: P_rec | intensity: P_work |
| time: until exhaustion | time: T_rec | time: until exhaustion |
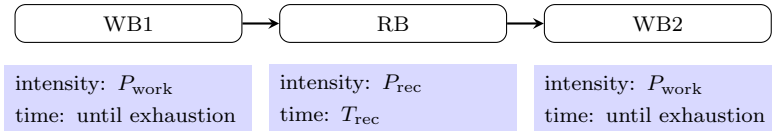
Figure 2.5: A schematic of the protocol to estimate recovery ratios. An exhaustive work bout (WB1) at a set intensity (P_work) is prescribed. Immediately after exhaustion is reached, a recovery bout (RB) follows at a lower recovery intensity (P_rec) for a set duration (T_rec). Then, a second exhaustive work bout (WB2) is conducted and the ratio of the time to exhaustion of WB2 to the one of WB1 represents the amount that was recovered during RB.

depicted next to those of W'_skib in the example from Figure 2.4. It is observable that W'_bart predicted quicker recovery dynamics.

Research into energy recovery modelling during intermittent exercise is an evolving field with promising applications. Skiba et al. (2014) applied their $W'_\text{bal−int}$ to cycling data of triathletes and deemed the model useful for predicting the fatigue state of an athlete. Hoogkamer et al. (2018) used the W'_skib model to investigate optimal pacing strategies for a sub-2-hour marathon. Reviews by Jones and Vanhatalo (2017); Skiba and Clarke (2021) highlighted promising possibilities of work-balance models for monitoring capabilities of an athlete during exercise but also stated that these models require future refinement.

2.3.3 Refined energy recovery predictions

Recent findings have suggested that the currently assumed monoexponential recovery of discussed $W'_\text{bal−int}$ and $W'_\text{bal−ode}$ models still overly simplify energy recovery dynamics, and that model modifications that account for characteristics of prior exhaustive exercise (Caen et al., 2019), as well as biexponential recovery dynamics (Caen et al., 2021), might improve recovery predictions. To discuss these suggested improvements in more detail, the next subsection introduces the test protocol with which Caen et al. (2019, 2021) measured energy recovery.

2.3.3.1 Procedure for measuring recovery

Recovery ratios can be computed from exercise protocols involving two exhaustive work bouts named work bout 1 (WB1) and work bout 2 (WB2) interspersed with a recovery bout (RB). A schematic of the protocol is depicted in Figure 2.5. First, the athlete exercises at a fixed work power output (P_work) above CP until exhaustion. Immediately after exhaustion is reached, exercise intensity switches to a lower fixed recovery power output (P_rec) below CP. After a set recovery time (T_rec) at that recovery intensity, the second work bout until exhaustion at P_work starts. The time to exhaustion of WB2 is expected to be shorter than the one of WB1 because of the
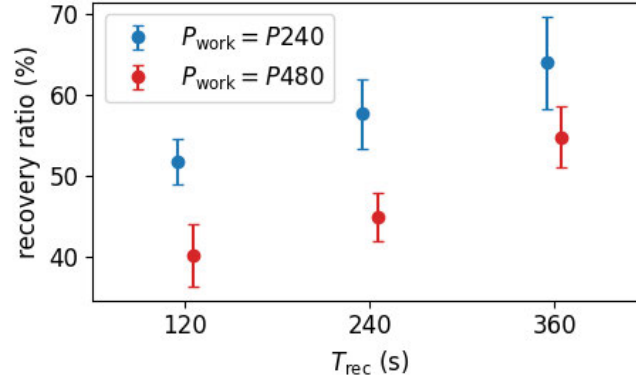
Figure 2.6: This figure displays observed recovery ratios by Caen et al. (2019) for distinct WB1→RB→WB2 test conditions. $P240$ and $P480$ are P_{work} intensities and denote the power output predicted to lead to exhaustion after 240 s or 480 s. Reported observations indicate that recovery is quicker after an exhaustive work bout at a higher power output.

limited recovery time in between the work bouts. Because it is assumed that W' is completely depleted at the end of WB1, the ratio of the second time to exhaustion to the first represents the amount of W' recovered. Thus, the time to exhaustion of WB2 divided by the time to exhaustion of WB1 multiplied by 100 results in a recovery ratio in percent. Henceforth, this protocol is referred to as the recovery protocol as defined in Section 2.3.3.1 (WB1→RB→WB2). Caen et al. (2019, 2021) used this protocol for their reported findings.

2.3.3.2 Effects of prior exhaustive exercise

Caen et al. (2019) found that recovery kinetics depend on previous work rates in addition to time and difference to CP. Their results are depicted in Figure 2.6. Caen et al. (2019) obtained their measurements from 11 athletes and 15 to 17 performance tests per athlete. As a first step, they estimated CP and W' of an athlete by fitting the critical power model to results of three to five TTE tests. From CP and W', they estimated the constant power outputs that were predicted to lead to exhaustion after 240 s ($P240$) and 480 s ($P480$). These were the P_{work} intensities for their WB1→RB→WB2 protocol. Their P_{rec} was either 33 % of CP (CP_{33}) or 66 % of CP (CP_{66}). Their investigated T_{rec} were 120 s, 240 s, and 360 s.

Caen et al. (2019) reported means and standard deviations of combined CP_{33} and CP_{66} data of all athletes. These are depicted in Figure 2.6. They observed that athletes recovered faster when exhausted after a work rate of $P240$ in comparison with $P480$. $W'_{\text{bal-int}}$ and $W'_{\text{bal-ode}}$ models are insensitive to P_{work} in the WB1→RB→WB2 protocol. Equation (2.12), Equation (2.14), and Equation (2.15) define recovery speed of these models and are only affected by $D_{\text{CP},t}$. Caen et al.
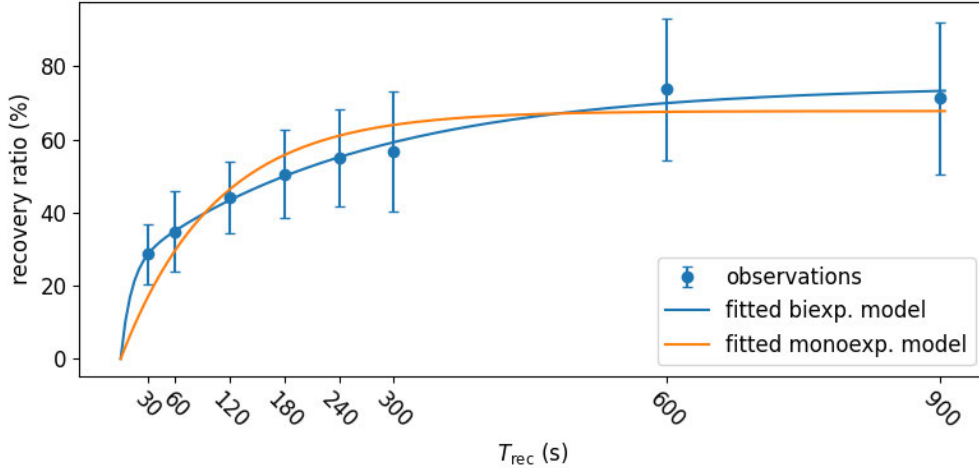
Figure 2.7: This figure displays observed recovery ratios by Caen et al. (2021) for distinct WB1→RB→WB2 test conditions. P_{work} and P_{rec} were constant but T_{rec} ranged from 30 s to 900 s. Caen et al. (2021) reported that their fitted biexponential model explained their observations more accurately than a monoexponential one.

(2019) concluded that modifications to work-balance models that take work-bout characteristics into account would improve prediction quality.

2.3.3.3 Biexponential recovery

In a subsequent study, Caen et al. (2021) showed that their energy recovery observations were better explained with a biexponential model that implements a steeper slope during the beginning of recovery. Caen et al. (2021) obtained recovery ratios of 20 participants for a range of 8 T_{rec} conditions with constant P_{work} and P_{rec}. They observed that group averaged recovery ratios rose quickly during shorter T_{rec} and slower during the longer T_{rec}. When fitting a biexponential and a monoexponential model to their recovery data, they observed that the biexponential model made more accurate predictions. Figure 2.7 summarises reported findings that Caen et al. (2021) presented in Figure 3 of their publication. In addition, Skiba et al. (2012); Skiba and Clarke (2021) proposed an alternative biexponential version of their work-balance models but such biexponential models have yet to be applied in practice.

Work-balance models with these proposed modifications would feature additional parameters, which introduces challenges in fitting them to small data sets. As demonstrated by the WB1→RB→WB2 protocol, recovery dynamics are derived from exhaustive exercise tests and therefore data for model fitting and validation are sparse (Vanhatalo et al., 2011; Sreedhara et al., 2019). Indeed, the search for models that optimally balance complexity with applicability to few data points is a primary challenge of energy recovery modelling. The review by Sreedhara et al.
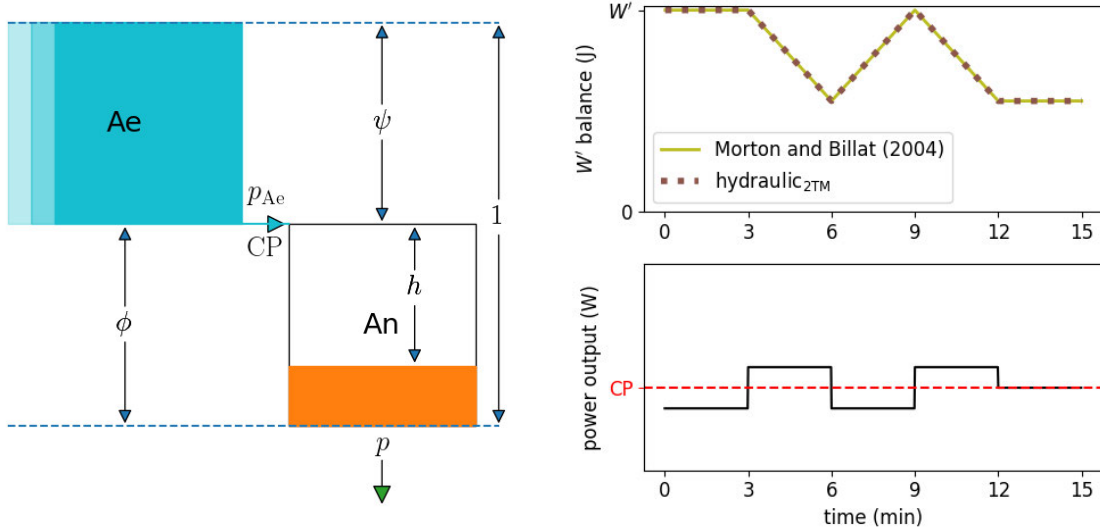
Figure 2.8: Left: A schematic of a hydraulic two-tank model. The left tank represents the aerobic energy source (Ae) and is of infinite capacity, which is indicated by the fading colour to the left. The right tank represents anaerobic energy sources (An), which in this particular example is considered to be of capacity W'. A pipe connects the bottom of Ae to An, allows flow from Ae to An (p_{Ae}), and has the maximal flow capacity CP. A tap p is attached to the bottom of An and h indicates the level of depletion of An. Tank positions are defined by ϕ and ψ. Right: The hydraulic two-tank model predicts energy expenditure and recovery according to the assumptions of Morton and Billat (2004).

(2019) discussed work-balance models in detail and highlighted the search for models that refine energy recovery predictions as future research opportunities.

2.4   Hydraulic models

Hydraulic models offer an alternative to work-balance models for predicting energy expenditure and recovery dynamics during exercise. They represent human bioenergetic responses to exercise as liquid flow within a system of pipes and tanks. This section begins with an analogy between hydraulic and work-balance models as an introduction.

2.4.1   Hydraulic related to work-balance

Morton (2006) discussed a hydraulic model with two tanks, which adheres to the assumptions of the critical power model defined in Section 2.2.1.1. Henceforth, this is referred to as the two-tank hydraulic (hydraulic$_{2t}$) model. A schematic of the hydraulic$_{2t}$ model is depicted on the left in Figure 2.8. The first of the listed assumptions of the critical power model in Section 2.2.1.1 is that power output is a function of two energy sources: aerobic and anaerobic. Each energy source is

represented by one tank in Figure 2.8. The second assumption says that aerobic energy is unlimited; therefore, the aerobic tank is of infinite size, indicated by the fading colour to the left. The third assumption is that anaerobic energy is limited in its capacity.

As discussed above, it is controversial whether W' equates the anaerobic work capacity. For example, Sreedhara et al. (2019) avoided this connotation, naming the anaerobic tank 'limited capacity' when they described the hydraulic$_{2t}$ model. However, because the present example uses the hydraulic$_{2t}$ model as an analogy for the assumptions of the critical power model, it is justified to interpret the limited tank as anaerobic source with the capacity W' in this instance. The fill-state of An can drop and rise by h and remaining liquid represents available W' balance.

A pipe (p_{Ae}) connects the bottom of the aerobic tank to the anaerobic tank. Flow from this pipe represents a sustainable energy contribution because the aerobic tank is of infinite capacity. The second assumption in Section 2.2.1.1 states that the conversion of aerobic energy into power output is limited by CP. Therefore, the pipe has a maximal flow capacity of CP. At the bottom of the An tank is a tap p. Flow from this tap represents power output.

As depicted in Figure 2.8, liquid flow within the example hydraulic$_{2t}$ model resembles the relation of power output, CP, and W', as assumed by the critical power model and the assumed linear recovery by Morton and Billat (2004). At the beginning of the simulated exercise in Figure 2.8, the anaerobic tank was filled. During the first 3 min, power output was below CP and thus flow from the tap p could be matched by flow from the aerobic tank. Then, power output rose above CP, maximal flow from the aerobic tank was reached and the liquid level of the anaerobic tank dropped by the difference between flow from p and CP every time step. From the 6th minute power output was below CP again, allowing flow from the aerobic tank to refill the anaerobic tank. Predicted W' balance rose by the difference between flow from p and CP every time step. During the last 3 min, power output was exactly at CP, the flow from p matched the maximal flow capacity from the aerobic tank and no changes to the fill level of the anaerobic tank occurred.

Changes to tank positions allow hydraulic models to make more sophisticated predictions and to resemble bioenergetic responses to exercise in more detail. For example, as discussed earlier Ward-Smith (1985) suggested in his Equation (2.7) that power output during the first seconds of exercise is nearly entirely provided by anaerobic sources and the contribution of aerobic sources increases exponentially from the onset of exercise. If in Figure 2.8 the parameter ψ was smaller, then liquid in the filled anaerobic tank would pressure against the pipe exit of p_{Ae}. In such a scenario, when p is opened, the fill level of the anaerobic tank would drop first and the flow from the aerobic tank would increase according to decreasing liquid pressure

against the pipe exit. Thus, the hydraulic$_{2t}$ model would predict an increase in the contribution from aerobic energy sources as suggested by Ward-Smith (1985).

Other more complex hydraulic models make use of such liquid pressure dynamics and even predict body responses to power demands with three or four tanks (Morton, 2006; Sundström, 2016). In addition, the first hydraulic model was more complex than the hydraulic$_{2t}$ model and is introduced in the following section.

2.4.2   M-M model

The first hydraulic model was published by Margaria (1976) and later further elaborated by Morton (2006). Morton mathematically defined its dynamics and published it as the Margaria-Morton (M-M) model. An example of the M-M model with the notation of Morton (2006) is depicted in Figure 2.9. Like Margaria (1976), Morton (2006) labelled the left infinitely big tank as oxidative energy source (O) and described it as the aerobic energy source. Because it is a sustainable energy source, it is a tank of infinite capacity. This is indicated by the fading colours to the left in Figure 2.9. The limited tank in the middle is labelled anaerobic alactic energy source (A_nA) and Morton (2006) stated it represents the alactic anaerobic energy source. The third tank is also limited, labelled anaerobic lactic energy source (A_nL), and represents the lactic anaerobic source (Morton, 1986b, 1990, 2006).

The depicted situation in Figure 2.9 is that T was opened and the fill level of the middle tank A_nA dropped by h. The more liquid flows out of the middle tank, the less liquid pressures against R_1, and the more flows from O through R_1 into A_nA. This way R_1 contributes to flow out of T. In the depicted situation in Figure 2.9, T was opened so wide that $h > \theta$ and flow from A_nL began to contribute as well. Its fill level has dropped by g.

Because the pipe R_1 enables flow from the aerobic source O into the middle tank, Morton (2006) referred to flow through R_1 as oxygen uptake ($\dot{V}_{O_2}$). He also defined that the maximal flow through pipe R_1 represents the maximal oxygen uptake ($\dot{V}_{O_2max}$). Morton (1986b) fitted differential equations of his model to collected $\dot{V}_{O_2}$ dynamics of participants during exercise on a treadmill. Fitted equations could explain the observations well. However, Morton (2006) concluded it remained to be seen what predictions of the M-M model would conform to reality because the model parameters (e.g. A_nL tank capacity and its height defined by θ and γ) are extremely difficult or impossible to obtain from individual athletes.

The M-M model in Figure 2.9, defines tank positions with the values θ, γ and ϕ. Because each tank represents a concrete bioenergetic energy source, Morton (1986b, 1990, 2006) developed several constraints on θ, γ and ϕ to find a realistic arrangement of tanks for his model:
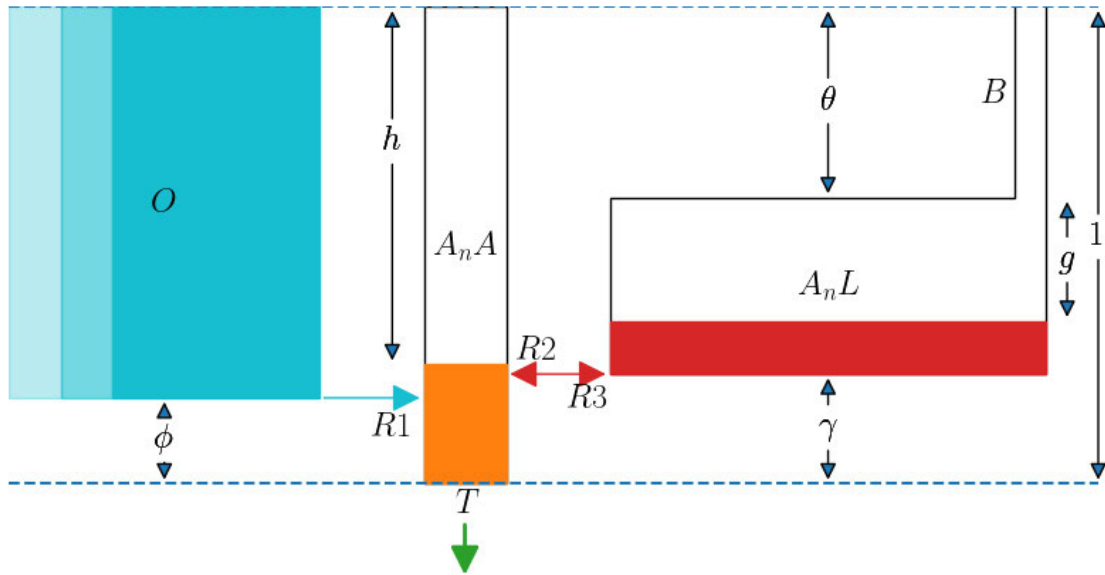
Figure 2.9: The M-M model as published by Morton (2006). Hydraulic models approximate human bioenergetic responses to exercise as liquid flow within a system of pipes and tanks. The left infinitely big tank O represents the oxidative or aerobic energy source. The limited tank in the middle (A_nA) represents the alactic anaerobic phosphagens. The right limited tank (A_nL) represents an anaerobic lactic energy source. A tap (T) is attached to A_nA, which can be opened and closed according to energy demand. The tube B accounts for early lactic acid occurrence. The pipes R_1, R_2, R_3 enable flow between the tanks and have maximal flow capacities.

- Pipe R_2/R_3 has to be above R_1 ($\gamma > \phi$) because athletes can deplete their glycogen stores when exercising below $\dot{V}_{O_2max}$.

- The onset of flow through R_2 represents the lactate threshold (moderate-heavy boundary) (LAT), i.e., the commencement of increased lactic acid production. Therefore, the top of tank A_nA has to have some distance to the top of the entire system ($\theta > 0$) and should be at approximately 40 % of the height of O.

- During constant severe intensity exercise, $\dot{V}_{O_2}$ rises asymptotic to a maximal value. When exercise stops, oxygen consumption does not decline immediately. Therefore, R_1 cannot be at the top or bottom of the middle tank ($0 < \phi < 1$).

From these constraints, Morton (1990) argued that the only realistic configuration of the three component hydraulic model is the one depicted in Figure 2.9.

2.4.3 Extensions of the M-M model

Despite the lack of clear verification, the M-M model was developed even further by Behncke (1993). He increased complexity by introducing adaptions to account for fatigue effects during runs of 10 km or longer. Then Behncke (1997) applied his form of the M-M model to world records in competitive running. Behncke (1997)

described the many constraints of the hydraulic model to make calculations 'cumbersome' and highlighted that naive conclusions about metabolic capacities according to tank fill levels are not justified. Further, he proposed that even more adaptions to the hydraulics model were necessary to precisely simulate energy flows, and finally suggested abandoning the hydraulics model instead of adding more dynamic capacity constraints Behncke (1997).

Sundström et al. (2015) later argued that Behncke (1997) modelled switches between energy systems as too rapid and proposed another hydraulic model with four tanks. He used this model to simulate constant power output exercises as well as all-out efforts and compared the values to empirical findings by Gastin et al. (1995). Sundström et al. (2015) reported good resemblance but also stated that their results could not be regarded as a validation. All parameters of their simulations were chosen to represent a general trained athlete according to reported values by Gastin et al. (1995); Karatzaferi et al. (2001); Watt et al. (2002). However, for a validation parameters would have to be collected from an individual and results matched to its performance. Indeed, the Margaria-Morton-Sundström (M-M-S) model by Sundström (2016) has yet to be validated experimentally.

2.4.4 Validation of the M-M model

Sundström et al. (2014) investigated predictions of the M-M model further in theoretical elaborations. They compared predicted optimal pacing strategies of the M-M model to those of a critical power model for intermittent exercise on an artificial course and reported that the M-M model made more realistic predictions.

Lidar et al. (2021) developed an approach to apply the M-M to real athlete data by fitting parameters with an optimisation approach to available $\dot{V}_{O_2}$ data. The study estimated accumulated anaerobic energy expenditure. They fitted two versions of the M-M and two versions of the $hydraulic_{2t}$ model to collected data during treadmill roller-skiing time trials. According to Lidar et al. (2021), the $hydraulic_{2t}$ model provided the highest validity and reliability for data it was fitted to and for predictions on unknown data. Further, Lidar et al. (2021) reported that optimal parameters for the fitted M-M were likely outside the physiologically reasonable ranges.

Therefore, hydraulic models promise more realistic predictions but a challenge remains in determining to what extent model predictions conform to reality. Applications by Morton (1986a); Behncke (1997); Sundström et al. (2014); Lidar et al. (2021) could only approximate parameters with optimisation approaches. Behncke (1997); Lidar et al. (2021) concluded from their approximations that the M-M model could not fully capture bioenergetic responses of the human body to exercise.

2.5 Summary

The discussed mathematical models of human performance come with many simplifying assumptions. Exercise performance is reduced to measures of expended and recovered energy. Hydraulic models represent complex metabolic responses during exercise as liquid flow within systems of tanks. However, these simplifications have resulted in useful models that serve as evidence-based tools for performance prediction and exercise prescription.

As summarised by Morton (2006) for hydraulic models or by Skiba and Clarke (2021) for work-balance models: one-size-fits-all performance models do not exist. For simplicity, let us imagine a spectrum where simple and applicable performance models are on the left and complex and theoretical models are on the right.

One could say the critical power model is on the far left, because it requires just a few performance tests to estimate its two parameters CP and W'. $W'_{\text{bal-int}}$ and $W'_{\text{bal-ode}}$ models would be on the left as well, because they only need CP and W' of an athlete to make predictions. The elegant abstraction and simplicity of these models allow for a relatively easy application to individual athletes. However, recent research has highlighted they might be too simple and energy recovery predictions may require further refinement (Caen et al., 2019, 2021).

In contrast, the hydraulic M-M or M-M-S models would be on the far right of the spectrum. They are complex and require eight or more parameters to be adjusted according to in-depth knowledge about bioenergetic capacities of an athlete(Morton, 2006; Lidar et al., 2021). Their intricate systems of tanks, pipes and liquid pressure promise more realistic predictions, but a challenge remains in investigating how realistic their predictions are.

Chapter 3

Problem Statement and Research Questions

3.1 Problem Statement

As research in performance modelling advances, existing models are challenged, verified, refuted or refined according to new studies and observations. Work-balance models are also subject to such advances and, as discussed in Section 2.3.2, while they are promising tools to monitor the energy capacities of an athlete, modifications to improve energy recovery predictions have been proposed (Caen et al., 2019, 2021).

The foundation of the problem statement of this work is two thought experiments. Both show that the M-M model discussed in Section 2.4.2 already addresses the improvements to energy recovery predictions of work-balance models discussed in Section 2.3.2.

3.1.1 Thought experiment 1

Caen et al. (2019) suggested a quicker recovery after short exhaustive exercise at a high power output is than after long exhaustive exercise at a lower power output. Work-balance models are insensitive to the characteristics of prior exhaustive exercise and modifications are required to address the suggestions by Caen et al. (2019).

In contrast, the M-M model is sensitive to these characteristics, as shown in the example depicted in Figure 3.1. Chosen parameters for this example serve as an illustration and are not based on real experiments. On the left, the tap of the M-M model was opened for 60 s according to 600 W. In this example, this can be considered a large power output over a short time. As a result, the liquid level in the middle tank dropped fast, and limited by the maximal flow capacity of the pipe from the right tank, not much liquid could flow from the right tank into the middle tank.
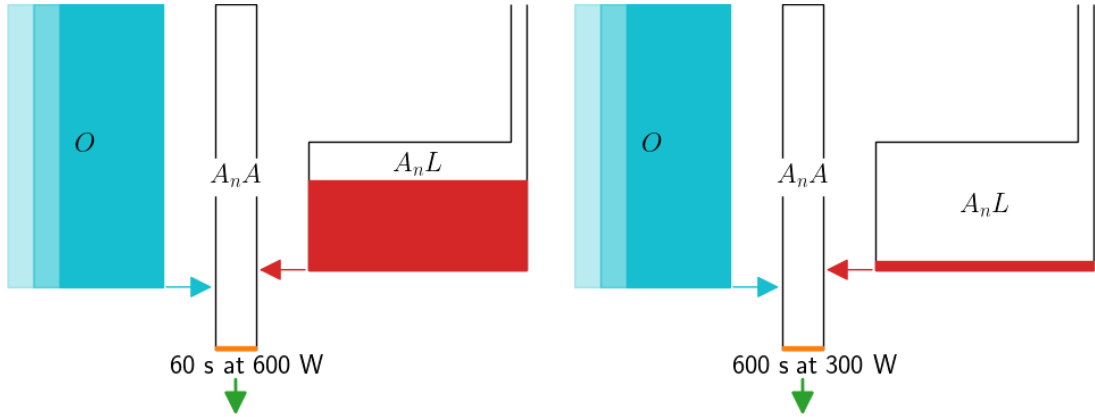
Figure 3.1: The M-M model predicts energy capacities after 60 s at a power output of 600 W (left) and 600 s at 300 W (right). Chosen parameters for this example serve as an illustration and are not based on real experiments. Left: The maximal flow capacity of the pipe from A_nL limited how much liquid could flow into A_nA during 60 s of rapid energy expenditure. Right: During 600 s of slower energy expenditure more liquid from A_nL could flow into A_nA.

On the right in Figure 3.1, the tap was opened for 600 s according to 300 W, such that the fill level of the middle tank dropped more slowly. Thus, there was more time for liquid to flow from the right into the middle tank. In both depicted cases, the models predicted that the athlete was exhausted because the middle tank was depleted and liquid flow out of the tap did not meet the demand any more.

In this situation, if the taps of both models were closed again, flow from the O tank would refill both limited tanks and the models would mimic recovery. The model on the left would recovery more quickly because less liquid has flown out of the system. As a result, the M-M model predicted quicker recovery after short exhaustive exercise at a high power output than after long exhaustive exercise at a lower power output. The M-M models is capable of addressing the suggested improvements of Caen et al. (2019).

3.1.2 Thought experiment 2

Skiba et al. (2012), and more recently, Caen et al. (2021); Skiba and Clarke (2021) have already noted that predicting energy recovery as a biexponential function might improve energy recovery predictions. However, established $W'_{\text{bal}-\text{ode}}$ and $W'_{\text{bal}-\text{int}}$ models predict monoexponential recovery dynamics.

Predictions of the hydraulic M-M model are closer to biexponential dynamics because it understands power output and recovery as a function of three interacting tanks and liquid flow. With rising or falling fill levels, liquid pressure increases or decreases exponentially. The model features two tanks that are connected to the middle tank; therefore, flow into and out of the middle tank is a biexponential

function. For example, assuming the tap is fully closed in the situation depicted in Figure 2.9, Equation 15 of Morton (1986b) defines the level of depletion at that time t as

$$h = k_1' \, \exp\left(r_1 t\right) + k_2' \, \exp\left(r_2 t\right) + \frac{c'}{b} + \theta. \tag{3.1}$$

This is a biexponential equation where $k_1', r_1, k_2', r_2, c', b, \theta$ are constants, which Morton (1986b) determined from boundary conditions. Therefore, the M-M model can predict energy recovery as a biexponential function and is capable of addressing the suggested improvements of Caen et al. (2021).

3.1.3 The problem of using the M-M model

Both thought experiments suggest that the M-M model can predict energy recovery with the proposed improvements by Caen et al. (2019, 2021); Skiba and Clarke (2021). By addressing these shortcomings, the M-M model would be able to make more accurate predictions than work-balance models for remaining energy capacities of an athlete during intermittent exercise.

However, using the M-M model for intermittent exercise predictions is not straightforward. As discussed in Section 2.4.2, the M-M model was designed to predict oxidative, lactic, and alactic energy production in response to power output. Its parameters are ascribed to concrete metabolic responses during exercise and Behncke (1997) reported it to be cumbersome to apply the model to individual athletes. Lidar et al. (2021) concluded that the M-M model cannot fully capture the dynamics of the human bioenergetics system and Behncke (1997) suggested abandoning the hydraulic model. At this stage, the M-M model cannot be used as an alternative to work-balance models for intermittent exercise predictions of an athlete.

3.1.4 A new hydraulic model

This work proposes a new hydraulic model, which combines the advantages of work-balance models and the M-M model for predicting energy expenditure and recovery during intermittent exercise. The advantage of work-balance models is that they only require CP and W' of an athlete to make predictions. We propose a three-tank hydraulic model that also only requires CP and W'. That becomes possible by defining a generalised three-tank hydraulic model where the parameters are not ascribed to concrete bioenergetic entities.

Our generalised hydraulic model has three tanks, like the M-M model, and therefore it retains the benefits outlined in the thought experiments above. Because its parameters are not limited by physiological constraints, it allows the development of an optimisation approach that freely adjusts parameters until the model predicts

energy expenditure according to the critical power model and recovery ratios according to previously published observations. In the following chapters, the new model will be defined, applied, and validated to answer three main research questions.

3.2   Research Questions

Chapter 4 mathematically defines the new generalised model. Then, the following chapters answer three research questions:

1. **Can the new hydraulic model be a feasible alternative to work-balance models?** The suggested advantages of the new hydraulic model make a difference only if it can be applied to intermittent exercise with data similar to those required by work-balance models and makes predictions with at least comparable accuracy. Chapter 5 presents a proof-of-concept to confirm this. We developed a procedure with which the generalised hydraulic model could successfully predict both energy expenditure and recovery kinetics for one example case. The procedure was designed to use CP and W' as inputs and to enable the generalised hydraulic model to be an alternative to work-balance models.

2. **How accurate are intermittent exercise predictions?** After confirming the feasibility of the hydraulic model in a proof-of-concept, it requires validation in more detail. The above outlined thought experiments suggest that the generalised hydraulic model can predict energy recovery more accurately than work-balance models. In Chapter 6, previously published data were used to compare model predictions and to confirm more accurate predictions of the generalised hydraulic model.

3. **How accurate are predictions for metabolic responses during exercise?** The original M-M model was designed to predict energy production of oxidative, lactic, and alactic metabolic pathways during exercise. Our generalised hydraulic model resembles the M-M model but does not ascribe parameters to human metabolism. Chapter 7 scrutinises predictions for metabolic responses of our generalised hydraulic model and compares them to collected data from study participants.

Chapter 4

The New Hydraulic Model

We published our generalised hydraulic model in Weigend et al. (2021). Henceforth, the model is referred to as generalised hydraulic model by Weigend et al. (2021) (hydraulic$_{\text{weig}}$). Hydraulic$_{\text{weig}}$ is an abstract version of the M-M model for which all bionergetic or metabolic context of its parameters has been erased. This chapter formalises the model in its entirety.

4.1 Generalisation of the M-M model

Behncke (1997); Lidar et al. (2021) fitted the M-M model to metabolic or bioenregetic empirical measures. Behncke (1997) concluded that the naive interpretation of its parameters is not justified and Lidar et al. (2021) that the M-M model was unlikely to fully capture the human bioenergetic system.

To prevent these issues caused by concrete physiological interpretation, the parameters of hydraulic$_{\text{weig}}$ are defined as being more abstract entities than those of the M-M model. The sole remaining similarity between the M-M model and hydraulic$_{\text{weig}}$ is that hydraulic$_{\text{weig}}$ also quantifies performance as a function of three energy sources, which are represented as tanks. However, hydraulic$_{\text{weig}}$ ignores relations to oxidative, lactic or alactic energy sources. It re-imagines the left infinitely big tank as a generic unlimited energy source (U), the middle tank as limited fast energy source (LF) and the right tank as limited slow energy source (LS)[1].

An example with these generalised labels is depicted in Figure 4.1. Maximal flow capacities of pipes are labelled maximal flow from the unlimited energy source (M_U), maximal flow from the limited fast energy source (M_{LF}) and maximal flow from the limited slow energy source (M_{LS}). Tank positions and sizes are defined by θ, γ and ϕ. Hydraulic$_{\text{weig}}$ does not feature tube B (depicted in Figure 2.9). Tube B

---

[1]Initially, in Weigend et al. (2021, 2022a) we referred to the three tanks as aerobic energy source (Ae), anaerobic fast energy source (AnF) and anaerobic slow energy source (AnS). Later, these labels were changed to U, LF, and LS, due to discrepancies between the measured and predicted oxygen uptake in athletes. These details are reported in Appendix A.1 and Chapter 7
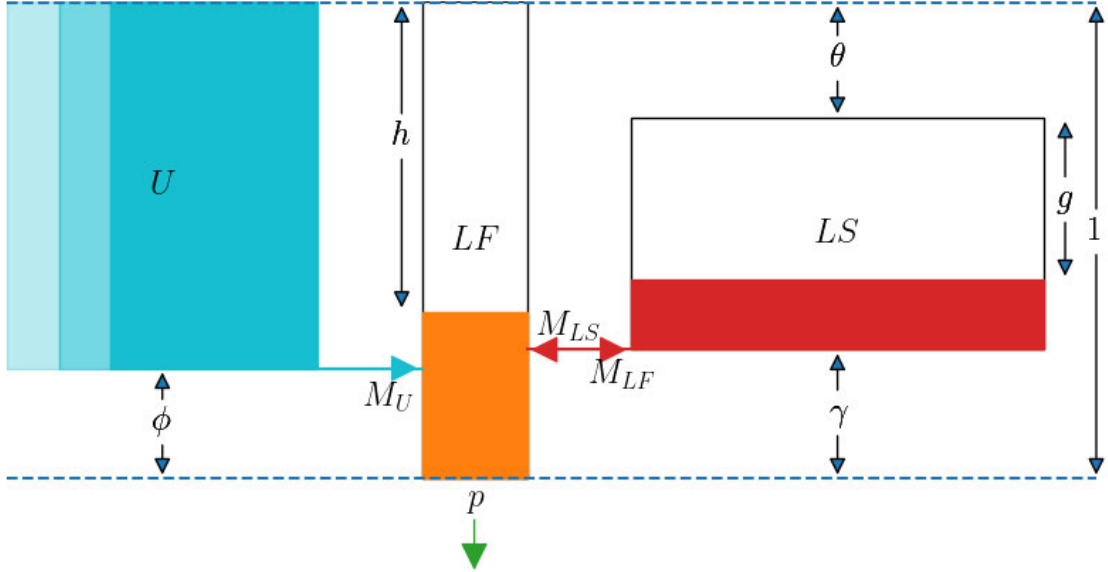
Figure 4.1: Representative example of hydraulic$_\text{weig}$, as published in Weigend et al. (2021). In comparison with the M-M model in Figure 2.9, tube B was removed and the tanks renamed as unlimited energy source (U), limited fast energy source (LF) and limited slow energy source (LS). The maximal flow capacities between tanks are labelled M_U, M_{LS}, M_{LF}.

was included by Margaria (1976); Morton (1986a) to account for early lactate levels in blood at the onset of exercise.

4.2 A configuration

Hydraulic$_\text{weig}$ features eight adjustable parameters. As depicted in Figure 4.1, the parameters θ, γ and ϕ denote tank positions and sizes, LF, LS are adjustable tank capacities measured in J, and M_U, M_{LF} and M_{LS} are maximal flow capacities measured in W. A configuration for hydraulic$_\text{weig}$ entails the positions, sizes and capacities of each tank and is therefore defined as

$$[LF,\ LS,\ M_U,\ M_{LF},\ M_{LS},\ \theta,\ \gamma,\ \phi]. \tag{4.1}$$

The parameters LF, LS, M_U, M_{LF} and M_{LS} are in the interval $[0, \infty)$. The bounds of tank positions θ, γ and ϕ are $[0, 1]$. The only constraint on tank positions is that LS must have a positive cross-sectional area; thus, $\theta + \gamma < 1$ must be satisfied. In summary, a configuration defines the parameters with which the liquid flow within the three tanks of hydraulic$_\text{weig}$ is computed.

4.3 Liquid flow phases

Liquid flow between tanks can be described in nine distinct phases. In this work, these phases are named after the tanks involved (U for flow from U and L for flow

from LS) and flow characteristics (l for limited, f for full, r for recovery). For example, U_l means that only liquid from U flows into LF and the flow is limited; that is, flow through the pipe is not at maximal capacity because of liquid pressure against the pipe exit. This phase U_l occurs when the fill level of LF is above the pipe exit of U and above the top of LS, and LS is full.

As a second example, Figure 4.1 depicts the phase $U_l L_l$. Liquid from U and liquid from LS flows into LF at limited capacity. Flow from U is limited because the fill level of LF is above the pipe exit of U. Liquid flows from LS into LF because the fill level of LF is below the top of LS ($h > \theta$). Flow is limited because the fill level of LF is above the pipe exit of LS ($h \leq 1 - \gamma$).

The phase $U_l L_r$ is an example of liquid flow during recovery. Liquid flows from U into LF at limited capacity and liquid from LF re-flows into LS. During this phase, the fill level of LF is above the pipe exit of U and above the fill level of LS ($h \leq \theta + g$). LS is not full ($g > 0$).

All possible phases are named U_l, U_f, $U_l L_l$, $U_l L_r$, $U_l L_f$, $U_f L_l$, $U_f L_r$, $U_f L_f$ and L_r. Each phase features a distinct differential equation that describes flow out of the tap at the bottom of LF. Not all three-tank hydraulic models require all of these liquid flow phases to be defined. Which phases are required depends on the configuration type.

4.4 Configuration types

As described above, a configuration of hydraulic$_{\text{weig}}$ assigns values to tank capacities LF, LS, to maximal flow capacities M_U, M_{LF}, M_{LS}, and tank positions ϕ, θ, and γ. Configurations can be categorised in types according to the liquid flow phases that occur in them. The overview in Figure 4.2 depicts all 20 possible configuration types named from A to T. Table 4.1 summarises what liquid flow phases can occur with each type.

For example, all configuration types in the top row of Figure 4.2 (types A-D) feature $\phi = 1$, such that the pipe exit of the left unlimited tank U is at the top of the middle limited tank LF. For these, the fill level of LF cannot rise above the pipe exit of U. Therefore, flow from the tank U cannot be limited by pressure from liquid above the pipe exit. As a result, phases that involve U_l cannot occur with types A-D.

The original hydraulic model by Margaria (1976) was of configuration type T (see Figure 1.29 in Margaria (1976)). Morton (1986b) argued that the model by Margaria (1976) is inconsistent with physiological facts, because flow from the left unlimited tank U can be at maximal capacity, while no liquid flows from the right tank LS. With the original interpretation that flow from U represents oxygen uptake, this
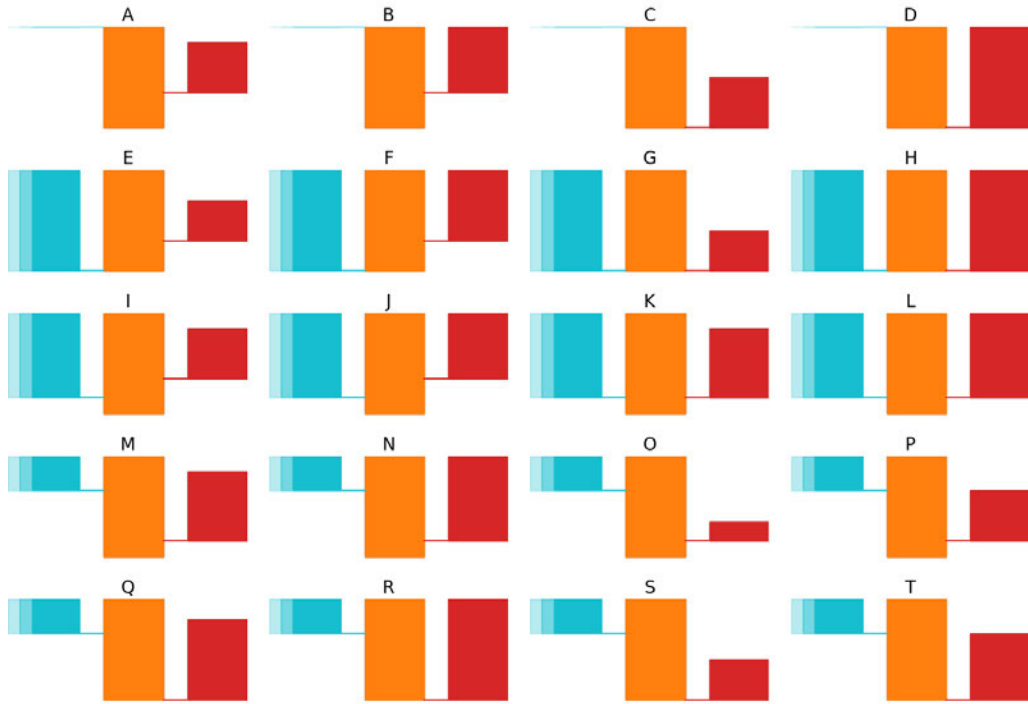
Figure 4.2: All 20 possible configuration types of a three-tank hydraulic model with adjustable θ, γ, and ϕ

Table 4.1: A summary of the liquid flow phases (rows) that can occur in each configuration type (columns).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
U_l					○		○		○		○		○		○	○	○		○	○
U_f	○		○												○				○	
U_lL_l					○	○	○	○	○	○	○	○	○	○			○	○		
U_lL_r					○	○	○	○	○	○	○	○	○	○			○	○		
U_lL_f					○	○			○	○										
U_fL_l	○	○	○	○									○	○	○	○	○	○	○	○
U_fL_r	○	○	○	○									○	○	○	○	○	○	○	○
U_fL_f	○	○							○	○	○	○	○	○	○	○				
L_r	○	○	○	○																

32

would mean that the model assumes that an athlete can work at a sustainable rate while being at their maximal oxygen uptake. As pointed out by Morton (1986b), this is physiologically not possible. Instead, he proposed that the M-M model was either of configuration type I,K,M or Q (A,B,C,D in Figure 2 in Morton (1986b)). Subsequently, Morton (1990) presented the in Section 2.4.2 discussed physiological constraints and concluded that only type I was valid (Figure 1 in Morton (1990)). Behncke (1993); Sundström et al. (2015); Lidar et al. (2021) each presented different reasons for why configuration type I is inconsistent with physiological facts and they introduced changes to the M-M model that resulted in new configuration types.

In contrast, for hydraulic$_{\text{weig}}$, we removed the bioenergetic and metabolic context of the M-M model and therefore removed all physiological constraints on its parameters. As a result, all 20 configuration types are possible and all liquid flow phases can occur.

4.5 Model Formalisation

The following sections formally define all possible liquid flow phases of hydraulic$_{\text{weig}}$. Section 4.5.1 defines these as ordinary differential equations (ODE) in accordance with how they are presented in Morton (1986b). Section 4.5.2 defines them as discretised ODEs in accordance with Sundström et al. (2014). The latter formalisation is practical for step-by-step simulations. We added considerations for extreme cases when step sizes of are large.

4.5.1 Ordinary differential equations

Morton (1986b) defined ODEs for the liquid flow phases of configuration types I,K,M and Q. However, he only described phases that are applicable to all four types, outlining the remaining ones without a rigorous solution. This section describes the liquid flow equations of hydraulic$_{\text{weig}}$ for all flow phases and relates them to Morton (1986b) where possible. Specifically, our additions to the work of Morton (1986b) are the formalisation of the phases U_f, L_r, L_f, as well as full solutions for the phases $U_l L_f$, $U_f L_l$, $U_f L_r$, $U_f L_f$.

4.5.1.1 Phase U_l

This phase is characterised by a limited flow from U into LF, while no liquid flows from LS into LF. This occurs if the fill level of LF is above or at the pipe exit of U ($h <= 1 - \phi$), and at the same time, LS is full ($g = 0$) and the fill level of LF is above or at the top of LS ($h <= \theta$). Using the notation of hydraulic$_{\text{weig}}$, Morton

(1986b) defined in (his) Equation 2 that during this phase, a constant power output P; that is, a constant liquid flow from tap p at the bottom of LF, is estimated as

$$P = \frac{M_U}{1 - \phi} h(t) + LF \ h'(t). \tag{4.2}$$

This is a first-order linear ODE. To estimate tank fill levels at a given time point t, an applicable equation for $h(t)$ is required. Equation (4.2) has as its general solution

$$h(t) = \frac{P(1 - \phi)}{M_U} + c_1 \ \exp\left[\frac{M_U}{LF(\phi - 1)} t\right], \tag{4.3}$$

where c_1 is a constant.

In his examples, Morton (1986b) derived c_1 with the assumption that $h(t) = 0$ at $t = 0$. The present work defines a more general case where $h(t) = h_{\text{start}}$ for $t = 0$. With the condition that h_{start} does not violate any of the above stated conditions for phase U_l ($h_{\text{start}} <= 1 - \phi$ and $h_{\text{start}} <= \theta$), the constant is then estimated as

$$c_1 = h_{\text{start}} - \frac{P(1 - \phi)}{M_U}. \tag{4.4}$$

Substituting Equation (4.4) into Equation (4.3) results in an applicable form of $h(t)$ to determine the fill level of LF at any time during phase U_l, given an initial h_{start} and a constant power demand P.

4.5.1.2 Phase U_f

This phase is characterised by flow from U into LF at full capacity while no liquid flows from LS into LF. During this phase, the fill level of LF is below or at the pipe exit of U ($h \leq 1 - \phi$). At the same time LS is filled ($g = 0$) and the fill level of LF is above or at the top of LS ($h \leq \theta$). Morton (1986b) did not define this phase because it did not occur in the hydraulic model configurations he investigated. For this phase, a constant power demand P is defined as

$$P = M_U + LF \ h'(t). \tag{4.5}$$

The depletion or refilling of LF is linear; therefore, $h'(t)$ is a constant. From a given $h(t) = h_{\text{start}}$ at $t = 0$, the fill level at any time during this phase can be estimated as

$$h(t) = h_{\text{start}} + \frac{P - M_U}{LF} t. \tag{4.6}$$

4.5.1.3 Phase $U_l L_l$

During this phase, liquid flows from U and from LS into LF. Both flows are limited by pressure from liquid in LF above the pipe exits of U and LS. This occurs if the fill level of LF is above or at pipe exit of U ($1 - \phi \leq h$) and above or at pipe

exit of LS ($\theta \leq h$). Further, the fill level of LS must be above the fill level of LF ($g + \theta \leq h$). This phase corresponds to the phases that Morton (1986b) defined as A2 and 2R1. In accordance with his Equation 8, a constant liquid flow out of tap p is defined as

$$P = \frac{M_U}{1 - \phi} h(t) + LF \ h'(t) + LS \ g'(t) \tag{4.7}$$

with

$$LS \ g'(t) = M_{LS} \frac{h(t) - g(t) - \theta}{1 - \theta - \gamma}, \tag{4.8}$$

i.e.,

$$h(t) = \frac{LS(1 - \theta - \gamma)}{M_{LS}} g'(t) + g(t) + \theta. \tag{4.9}$$

As already outlined by Morton (1986b), to derive an applicable form of $g(t)$ and $h(t)$, one can substitute Equation (4.9), together with its derivative $h'(t)$, into Equation (4.7). This yields a second-order differential equation of the form

$$g''(t) + \left[\frac{M_U}{LF(1 - \phi)} + \frac{M_{LS}(LF + LS)}{LF \ LS(1 - \theta - \gamma)} \right] g'(t) + \frac{M_U \ M_{LS}}{LF \ LS(1 - \phi)(1 - \theta - \gamma)} g(t)$$
$$= \frac{M_{LS}(P(1 - \phi) - M_U \ \theta)}{LF \ LS(1 - \phi)(1 - \theta - \gamma)}, \tag{4.10}$$

which corresponds to Equation 11 in Morton (1986b). By summarising constants into a, b, and c, the above equation is of the form

$$g''(t) + a \ g'(t) + b \ g(t) = c \tag{4.11}$$

and has as its general solution

$$g(t) = c_1 \ \exp\left[\frac{1}{2} \left(-\sqrt{a^2 - 4b} - a \right) t \right] + c_2 \ \exp\left[\frac{1}{2} \left(\sqrt{a^2 - 4b} - a \right) t \right] + \frac{c}{b}. \tag{4.12}$$

As per Morton (1986b) in Equation 12 of his paper, the constants in the exponents are summarised as r_1 and r_2. This results in

$$g(t) = c_1 \ \exp\left(r_1 \ t \right) + c_2 \ \exp\left(r_2 \ t \right) + \frac{c}{b}. \tag{4.13}$$

From here, the present work diverts from the procedure outlined in Morton (1986b). Morton (1986b) derived c_1 and c_2 by setting the initial conditions $g(0) = g'(0) = 0$, which assumes that LS is filled at the beginning of the simulation, and $h(0) = g(0) + \theta$. To apply the equations to simulations with more broadly defined initial conditions, this work instead determines c_1 and c_2 from set initial fill levels h_{start} and g_{start}. By substituting the derivative of Equation (4.13) into Equation (4.9) and with $h(t) = h_{\text{start}}$ and $g(t) = g_{\text{start}}$ at $t = 0$, this arrangement yields

$$c_2 = \frac{M_{LS}(h_{\text{start}} - g_{\text{start}} - \theta)}{r_2 \ LS(1 - \theta - \gamma)} - \frac{r_1 \ c_1}{r_2}. \tag{4.14}$$

35

Now substituting into Equation (4.13) and solving for c_1 yields

$$c_1 = \left[g_{\text{start}} - \frac{M_{LS}(h_{\text{start}} - g_{\text{start}} - \theta)}{r_2 \ LS(1 - \theta - \gamma)} - \frac{c}{b} \right] \frac{1}{1 - \frac{r_1}{r_2}}, \tag{4.15}$$

which results in an applicable version of Equation (4.13) for $g(t)$.

For an applicable version of $h(t)$, Equation (4.13) and its derivative $g'(t)$ are substituted into Equation (4.9):

$$h(t) = \left[\frac{LS(1 - \theta - \gamma)}{M_{LS}} c_1 r_1 + c_1 \right] \exp(r_1 \ t) + \left[\frac{LS(1 - \theta - \gamma)}{M_{LS}} c_2 r_2 + c_2 \right] \exp(r_2 \ t) + \frac{c}{b} + \theta. \tag{4.16}$$

4.5.1.4 Phase $U_l L_r$

This phase is characterised by limited liquid flow from U into LF and flow from LF into LS, which causes the liquid level in LS to rise. This occurs if the fill level of LF is above or at the pipe exit of U ($h \geq 1 - \phi$), the fill level of LS is above or at the fill level of LF ($\theta + g \geq h$), and LS is not full ($g > 0$). Because of the re-flow into LS, the change in the fill level of LS is negative. Morton (1986b) defined this as

$$LS \ g'(t) = -M_{LF} \frac{g(t) + \theta - h(t)}{1 - \gamma}. \tag{4.17}$$

To derive applicable forms of the equations $h(t)$ and $g(t)$, this work follows a similar procedure to the previous Section 4.5.1.3. First, Equation (4.17) is rearranged to

$$h(t) = \frac{LS(1 - \gamma)}{M_{LF}} g'(t) + g(t) + \theta. \tag{4.18}$$

Then, Equation (4.18) and its derivative are substituted into Equation (4.7), yielding the second-order differential equation

$$g''(t) + \left[\frac{M_U}{LF(1 - \phi)} + \frac{M_{LF}(LF + LS)}{LF \ LS(1 - \gamma)} \right] g'(t) + \frac{M_U M_{LF}}{LF \ LS(1 - \phi)(1 - \gamma)} g(t)$$
$$= \frac{M_{LF}(P(1 - \phi) - M_U \ \theta)}{LF \ LS(1 - \phi)(1 - \gamma)}. \tag{4.19}$$

With constants summarised into x, y and z, it is of the form

$$g''(t) + x \ g'(t) + y \ g(t) = z. \tag{4.20}$$

With the same procedure as Equation (4.11) to Equation (4.13), previously in Section 4.5.1.3, a general solution can be derived as

$$g(t) = c_1 \ \exp(r_1 t) + c_2 \ \exp(r_2 t) + \frac{z}{y}, \tag{4.21}$$

where now c_1, c_2, r_1 and r_2 are constants of different values than in Section 4.5.1.3.

Again, c_1 and c_2 can be derived with initial fill levels h_{start} and g_{start}. The derivative of Equation (4.21) $g'(t)$ is substituted into Equation (4.18), and with $t = 0$ and $h(t) = h_{\text{start}}$ and $g(t) = g_{\text{start}}$, this arrangement results in

$$c_2 = \frac{M_{LF}(h_{\text{start}} - g_{\text{start}} - \theta)}{LS(1 - \gamma)r_2} - \frac{r_1}{r_2}c_1. \qquad (4.22)$$

This is then substituted into Equation (4.21), yielding

$$c_1 = \left[g_{\text{start}} - \frac{M_{LF}(h_{\text{start}} - g_{\text{start}} - \theta)}{LS(1 - \gamma)r_2} - \frac{c}{b} \right] \frac{1}{1 - \frac{r_1}{r_2}}, \qquad (4.23)$$

which is a solution for c_1 and c_2 and is therefore an applicable version of $g(t)$ with Equation (4.21).

The applicable version of $h(t)$ is derived with the same procedure as in Section 4.5.1.3 and of the same from as Equation (4.16) but with the updated c_1, c_2, r_1 and r_2 from this subsection.

4.5.1.5 Phase $U_l L_f$

This phase is characterised by limited liquid flow from U into LF and flow from LS into LF at full capacity. This phase occurs if the fill level of LF is above or at the pipe exit of U ($h \leq 1 - \phi$) and below or at the pipe exit of LS ($h \geq 1 - \gamma$). During this phase, the flow from LS is only affected by the pressure of the remaining liquid in LS. The change in the fill level of LS was defined by Morton (1986b) as

$$LS\, g'(t) = M_{LS}\frac{1 - \theta - \gamma - g(t)}{1 - \theta - \gamma}. \qquad (4.24)$$

This is a first-order differential equation with the general solution

$$g(t) = (1 - \theta - \gamma) + c_1\, \exp\left[-\frac{M_{LS}\, t}{LS(1 - \theta - \gamma)} \right], \qquad (4.25)$$

where c_1 is a constant. With a given $g(t) = g_{\text{start}}$ at $t = 0$, this equation can be solved for c_1 and therefore is an applicable form of $g(t)$.

To obtain an applicable form of $h(t)$, the derivative of Equation (4.25) is substituted into Equation (4.7), which yields

$$h'(t) = -\frac{M_U}{LF(1 - \phi)}h(t) + \frac{M_{LS}}{LF(1 - \theta - \gamma)}\, c_1\, \exp\left[-\frac{M_{LS}\, t}{LS(1 - \theta - \gamma)} \right] + \frac{P}{LF}, \quad (4.26)$$

i.e.

$$h'(t) = a\, h(t) + b\, \exp(k\, t) + c, \qquad (4.27)$$

where a, b, c, k are constants. This is a first-order differential equation with the general solution

$$h(t) = -\frac{b\, \exp(k\, t)}{a - k} + c_2\, \exp(a\, t) - \frac{c}{a}. \qquad (4.28)$$

Using a known initial fill level $h(t) = h_{\text{start}}$ at $t = 0$, this equation can be solved for c_2. Then, with this given constant c_2, we have an applicable form of $h(t)$.

4.5.1.6 Phase $U_f L_l$

This phase is characterised by liquid flow from U into LF at full capacity and limited flow from LS into LF. During this phase, the fill level of LF is below or at the pipe exit of U ($h \geq 1 - \phi$), above or at the pipe exit of LS, and below or at the fill level of LS ($g + \theta \leq h \leq 1 - \gamma$). Because flow from U is at full capacity, flow from U is a constant of value M_U and total power output of the model is estimated as

$$P = M_U + LF\ h'(t) + LS\ g'(t). \tag{4.29}$$

Because flow from LS is limited Equation (4.8) for L_l from Section 4.5.1.3 above applies:

$$LS\ g'(t) = M_{LS}\frac{h(t) - g(t) - \theta}{1 - \theta - \gamma};$$

that is,

$$h(t) = \frac{LS(1 - \gamma)}{M_{LS}}g'(t) + g(t) + \theta,$$

which has the first derivative

$$h'(t) = \frac{LS(1 - \theta - \gamma)}{M_{LS}}g''(t) + g'(t). \tag{4.30}$$

To obtain an applicable version of $g(t)$, Equation (4.30) is substituted into Equation (4.29), yielding the second-order differential equation

$$g''(t) + \frac{M_{LS}(LF + LS)}{LF\ LS(1 - \theta - \gamma)}g'(t) = \frac{M_{LS}(P - M_U)}{LF\ LS(1 - \theta - \gamma)}, \tag{4.31}$$

which has as its general solution

$$g(t) = \frac{P - M_U}{LF + LS}\ t + c_1\ \exp\left[-\frac{M_{LS}(LF + LS)}{LF\ LS(1 - \theta - \gamma)}\ t\right]\frac{LF\ LS(1 - \theta - \gamma)}{M_{LS}(LF + LS)} + c_2, \tag{4.32}$$

where c_1 and c_2 are constants. The derivative of Equation (4.32) is

$$g'(t) = \frac{P - M_U}{LF + LS} - c_1\ \exp\left[-\frac{M_{LS}(LF + LS)}{LF\ LS(1 - \theta - \gamma)}\ t\right]. \tag{4.33}$$

With a given $g(t) = g_{\text{start}}$ and $h(t) = h_{\text{start}}$ at $t = 0$, Equation (4.33) can be solved for c_1 by substituting it into Equation (4.8), which yields

$$c_1 = \frac{P - M_U}{LF + LS} - \frac{M_{LS}(h_{\text{start}} - g_{\text{start}} - \theta)}{LS(1 - \theta - \gamma)}. \tag{4.34}$$

Then, substituting Equation (4.34) into Equation (4.32) allows a solution for c_2

$$c_2 = g_{\text{start}} - \frac{LF\ LS(P - M_U)(1 - \theta - \gamma)}{M_{LS}(LF + LS)^2} + \frac{LF(h_{\text{start}} - g_{\text{start}} - \theta)}{LF + LS}. \tag{4.35}$$

The solutions for c_1 and c_2 result in an applicable form of $g(t)$ and $g'(t)$ under the condition that at $t = 0$ and $g(t) = g_{\text{start}}$ and $h(t) = h_{\text{start}}$. To obtain an applicable form of $h(t)$, we substitute Equation (4.32) and Equation (4.33) with Equation (4.34) and Equation (4.35) into Equation (4.9).

4.5.1.7   Phase $U_f L_r$

This phase is characterised by liquid flow from U into LF at full capacity and flow from LF into LS, which causes the fill level of LS to rise. During this phase, the fill level of LF is below or at the pipe exit of U ($h \geq 1 - \phi$) and above or at the fill level of LS ($h \leq \theta + g$). As in Section 4.5.1.4, this phase uses Equation (4.17) for re-flow into LS, which is

$$LS \; g'(t) = -M_{LF}\frac{g(t) + \theta - h(t)}{1 - \gamma},$$

i.e.,

$$h(t) = \frac{LS(1 - \gamma)}{M_{LF}}g'(t) + g(t) + \theta.$$

Thus, the procedure to obtain applicable forms of $h(t)$ and $g(t)$ is similar to the procedure in Section 4.5.1.4. The difference lies in the contribution from U, which is now a constant. As such, total power output is determined by Equation (4.29). We substitute the first derivative $h'(t)$ of the above equation $h(t)$ into Equation (4.29) and yield the second-order differential equation

$$g''(t) + \frac{M_{LF}(LF + LS)}{LF \; LS(1 - \gamma)}g'(t) = \frac{M_{LF}(P - M_U)}{LF \; LS(1 - \gamma)}. \tag{4.36}$$

The general solution is

$$g(t) = -\frac{LF \; LS(1 - \gamma)}{M_{LF}(LF + LS)}c_1 \; \exp\left[-\frac{M_{LF}(LF + LS)}{LF \; LS(1 - \gamma)}t\right] + \frac{P - M_U}{LF + LS}t + c_2, \tag{4.37}$$

where c_1 and c_2 are constants.

To estimate these constants, the first derivative $g'(t)$ of Equation (4.37) is substituted into Equation (4.18), and solved for c_1 with defined $h(t) = h_{\text{start}}$ and $g(t) = g_{\text{start}}$ at $t = 0$:

$$c_1 = -M_{LF}\frac{g_{\text{start}} + \theta - h_{\text{start}}}{LS(1 - \gamma)} - \frac{P - M_U}{LF + LS}. \tag{4.38}$$

This solution for c_1 then allows a solution to Equation (4.37) for c_2:

$$c_2 = g_{\text{start}} - \frac{LF \; LS(P_{work} - M_U)(1 - \gamma)}{M_{LF}(LF + LS)^2} - \frac{LF(g_{\text{start}} + theta - h_{\text{start}})}{LF + LS}. \tag{4.39}$$

These solutions for c_1 and c_2 result in an applicable form of $g(t)$ as Equation (4.37). For an applicable form of $h(t)$, Equation (4.37) and its first derivative $g'(t)$ are substituted into Equation (4.18).

4.5.1.8 Phase $U_f L_f$

This phase is characterised by liquid flow from U into LF at full capacity and flow from LS into LF at full capacity. This phase occurs if the fill level of LF is below or at the pipe exit of U and below or at the pipe exit of LS ($h \geq 1 - \phi$ and $h \geq 1 - \gamma$). During this phase, LS is only affected by the pressure of the remaining liquid in LS. Thus, the same Equation (4.24) as in Section 4.5.1.5 is used to estimate the change in the fill level of LS:

$$LS \, g'(t) = M_{LS} \frac{1 - \theta - \gamma - g(t)}{1 - \theta - \gamma}$$

with its general solution

$$g(t) = (1 - \theta - \gamma) + c_1 \, \exp\left[-\frac{M_{LS} \, t}{LS(1 - \theta - \gamma)}\right],$$

where c_1 is a constant that can be derived from a given $g(t) = g_{\text{start}}$ at $t = 0$.

Thus, we have an applicable form of $g(t)$. To obtain an applicable form of $h(t)$, the derivative $g'(t)$ of the above equation for $g(t)$ is substituted into Equation (4.29). This yields

$$h'(t) = \frac{P - M_U}{LF} + \frac{M_{LS}}{LF(1 - \theta - \gamma)} c_1 \, \exp\left[-\frac{M_{LS}}{LS(1 - \theta - \gamma)}t\right], \tag{4.40}$$

which then is integrated to obtain

$$h(t) = \frac{P - M_U}{LF}t - \frac{LS}{LF}c_1 \, \exp\left[-\frac{M_{LS}}{LS(1 - \theta - \gamma)}t\right] + c_2, \tag{4.41}$$

where c_1 is the constant derived with $g(t) = g_{\text{start}}$ at $t = 0$ above, and c_2 can be derived with $h(t) = h_{\text{start}}$ at $t = 0$.

4.5.1.9 Phase L_r

This phase is characterised by flow from U through LF directly into LS. During this phase the fill level of LS rises but LF remains full. This phase only occurs if the pipe exit of U is at the top of LF ($\phi = 1$), LS is not yet full ($g > 0$), and the sum of flow from the tap (P) plus flow into LS is less than the maximal flow capacity from U ($h' \leq 0$). During this phase, $h(t) = 0$ always applies and Equation (4.17) changes to

$$LS \, g'(t) = -M_{LF} \frac{g(t) + \theta}{1 - \gamma}, \tag{4.42}$$

which is a first-order differential equation with its general solution

$$g(t) = c_1 \, \exp\left[-\frac{M_{LS}}{LS(1 - \gamma)}t\right] - \theta, \tag{4.43}$$

where c_1 is a constant that can be derived from $g(t) = g_{\text{start}}$ at $t = 0$.

With this, all possible liquid flow phases are defined and solved. The separation in distinct phases makes using the hydraulic model cumbersome because exact transition points between phases have to be found. For example, when a model of configuration type M predicts liquid flow dynamics during a TTE trial at a high intensity, it will start with phase U_l. Then, when $h(t) = \theta$, it will switch into the phase $U_l L_l$. As $h(t)$ drops further to $h(t) = \phi$, the model will switch into the phase $U_f L_l$. Then, $h(t)$ will reach γ and the model will commence the final phase, $U_f L_f$, where $h(t)$ will reach 0.

Each switch between phases requires solving a mathematical problem. For example, to find the time when $U_l L_l$ switches to $U_f L_l$ requires a solution to Equation (4.9) with $h(t) = \phi$ for t. That t is when phase $U_l L_l$ ends. For this example, it is called t_{end}. The t_{end} is used to estimate $h(t_{\mathrm{end}})$ and $g(t_{\mathrm{end}})$ with Equation (4.16) and Equation (4.13). The resulting values are used as h_{start} and g_{start} for the subsequent $U_f L_l$.

The following section formalises liquid flow phases in an alternative way as discretised equations.

4.5.2 Discretised equations

Typically, digital data describes power outputs of an athlete in discrete time steps; for example, measured power output per second on a cycle ergometer. Further, real-athlete data are subject to natural variation. Even during a TTE test, true power outputs may vary slightly from second to second. With the goal of investigating hydraulic$_{\mathrm{weig}}$ on real-athlete data, this section summarises liquid flow equations as discretised equations that are convenient for application with computer programs.

Sundström et al. (2014) also presented discretised liquid flow equations. For this work, we updated the notation and presentation to match the generalised assumptions of hydraulic$_{\mathrm{weig}}$. Further, we extend the discretised equations with extreme case handling in Section 4.5.2.1 to avoid estimation errors for large step sizes or extreme parameter combinations. Henceforth, estimations assume discrete time steps and the time difference between two time steps t and $t + 1$ is denoted as Δt. Power demand P is assumed to be constant for Δt between two steps.

For the estimations of fill levels and flows at time step t, first the previous fill level of LF defined by h_{t-1} drops according to the current power demand P_t. This results in the intermediate level $h_{P,t}$, which is estimated as

$$h_{P,t} = h_{t-1} + \frac{P_t}{LF} \cdot \Delta t. \qquad (4.44)$$

Then, the entire system of hydraulic$_{\text{weig}}$ reacts to the new intermediate fill level of LF and flows are estimated. Flow from U (P_U) at time point t is estimated as

$$P_{U,t} = \begin{cases} M_U \cdot \frac{h_{P,t}}{1-\phi}, & \text{if } 0 \leq h_{P,t} \leq (1-\phi). \\ M_U, & \text{otherwise.} \end{cases} \qquad (4.45)$$

The flow capacity M_U is scaled with the ratio of the fill level of LF to $(1-\phi)$. This means that maximal flow is reached as soon as $h_{P,t} \geq (1-\phi)$. Because the size of U is infinite, liquid will never flow back into U and thus the interval of $P_{U,t}$ is $[0, M_U]$.

Estimations of flow from LS (P_{LS}) require more definition. Because liquid can refill LS or flow out of it, the interval of P_{LS} is $[-M_{LF}, M_{LS}]$. The height of LS is henceforth referred to as g_{max}. It amounts to

$$g_{\text{max}} = 1 - \theta - \gamma. \qquad (4.46)$$

To introduce all possible flows for P_{LS}, calculations are organised in categories. The full Equation (4.50) for P_{LS} is the combination of Equation (4.47), Equation (4.48), and Equation (4.49). The first category describes cases during which no flow between LS and LF occurs. Thus, P_{LS} at these time points t equals 0:

$$P_{LS,t} = \begin{cases} 0, & \text{if } h_{P,t} \leq \theta \\ & \text{and } g_{t-1} = 0. \\ 0, & \text{if } h_{P,t} \geq (1-\gamma) \\ & \text{and } g_{t-1} = g_{\text{max}}. \\ 0, & \text{if } h_{P,t} = (g_{t-1} + \theta). \end{cases} \qquad (4.47)$$

In the first case, the tank LS is full and the fill level of LF is above the top of tank LS. In the second case, the fill level of LF is below the bottom end of LS and LS is empty. Finally, in the third case, the fill level of LF is exactly equal with the fill level of LS, causing equilibrium between both.

The second category includes cases during which liquid flows out of LS into LF:

$$P_{LS,t} = \begin{cases} M_{LS} \cdot \frac{h_{P,t} - (g_{t-1} + \theta)}{g_{\text{max}}}, & \text{if } h_{P,t} > (g_{t-1} + \theta) \\ & \text{and } h_{P,t} < (1-\gamma). \\ M_{LS} \cdot \frac{g_{\text{max}} - g_{t-1}}{g_{\text{max}}}, & \text{if } h_{P,t} \geq (1-\gamma) \\ & \text{and } g_{t-1} < g_{\text{max}}. \end{cases} \qquad (4.48)$$

If the fill level of LF is below the fill level of LS and above the bottom end of LS, the maximal possible flow is scaled according to the ratio of the difference between fill levels and the total height of LS. If the fill level of LF is below the bottom end of LS and LS is not empty, the maximal flow is scaled according to the amount of the remaining liquid, to consider the pressure of remaining liquid in the tank.

The third category describes the flow back from LF into LS.

$$P_{LS,t} = \begin{cases} M_{LF} \cdot \frac{h_{P,t}-(g_{t-1}+\theta)}{1-\gamma}, & \text{if } h_{P,t} < (g_{t-1}+\theta) \\ & \text{and } g_{t-1} > 0. \end{cases} \tag{4.49}$$

Here the fill level of LF is above the fill level of LS and LS is not full, which causes liquid to flow back into LS. The maximal flow M_{LF} from LF into LS is scaled according to the ratio between the difference of fill levels and the height of LS. Because $h_{P,t}$ is smaller than $g_{t-1}+\theta$, the result will be negative. This indicates, that a re-flow into LS occurs.

As the result, the full equation for $P_{LS,t}$ is the combination of Equation (4.47), Equation (4.48) and Equation (4.49):

$$P_{LS,t} = \begin{cases} 0, & \text{if } h_{P,t} \leq \theta \\ & \text{and } g_{t-1} = 0. \\ 0, & \text{if } h_{P,t} \geq (1-\gamma) \\ & \text{and } g_{t-1} = g_{\max}. \\ 0, & \text{if } h_{P,t} = (g_{t-1}+\theta). \\ M_{LS} \cdot \frac{h_{P,t}-(g_{t-1}+\theta)}{g_{\max}}, & \text{if } h_{P,t} > (g_{t-1}+\theta) \\ & \text{and } h_{P,t} < (1-\gamma). \\ M_{LS} \cdot \frac{g_{\max}-g_{t-1}}{g_{\max}}, & \text{if } h_{P,t} \geq (1-\gamma) \\ & \text{and } g_{t-1} < g_{\max}. \\ M_{LF} \cdot \frac{h_{P,t}-(g_{t-1}+\theta)}{1-\gamma}, & \text{if } h_{P,t} < (g_{t-1}+\theta) \\ & \text{and } g_{t-1} > 0. \end{cases} \tag{4.50}$$

Then, the effect of both flows $P_{U,t}$ and $P_{LS,t}$ on the tank-fill levels of this time step are estimated as

$$h_t = h_{P,t} + \frac{P_{LS,t} + P_{U,t}}{LF} \cdot \Delta t \tag{4.51}$$

and

$$g_t = g_{t-1} + \frac{P_{LS,t}}{LS} \cdot \Delta t. \tag{4.52}$$

Together, these equations allow us to estimate tank fill levels and liquid flows for a time step t with power demand P_t.

4.5.2.1 Handling extreme cases

Large values for Δt, or extreme parameter values, can cause flows to be overestimated. To handle these cases, limitations to the flow P_{LS} are applied:

It can happen that P_{LS} becomes larger than the remaining capacity of LS, or a negative P_{LS} refills more liquid into LS than the tank has capacity for. In case not enough liquid remains in LS, $P_{LS,t}$ is capped at the remaining amount by

$$P_{LS,t} = \min\left(P_{LS,t}, \frac{(g_{\max} - g_{t-1}) \cdot LS}{\Delta t}\right). \tag{4.53}$$

Accordingly, if $P_{LS,t} \cdot \Delta t$ amounts to more re-flow into LS than the available capacity, it is capped to fill LS just to the top:

$$P_{LS,t} = \max \left(P_{LS,t} \, , \, \frac{-g_{t-1} \cdot LS}{\Delta t} \right). \tag{4.54}$$

Further, extreme values can cause flows that overshoot an equilibrium between fill levels of LF and LS. The difference in liquid between the fill levels of LF and LS at time step t (D_t) is estimated as

$$D_t = \frac{h_{P,t} - (g_{t-1} + \theta)}{\frac{1}{LS} + \frac{1}{LF}}. \tag{4.55}$$

This difference is positive if the fill level of LF is above or at the fill level of LS and negative otherwise. Because both $P_{LS,t}$ and D_t can be negative or positive, the limitation applies in the following manner:

$$P_{LS,t} = \begin{cases} \max \left(P_{LS,t} \, , \, \frac{D_t}{\Delta t} \right), & \text{if } P_{LS,t} < 0. \\ \min \left(P_{LS,t} \, , \, \frac{D_t}{\Delta t} \right), & \text{if } P_{LS,t} > 0. \end{cases} \tag{4.56}$$

Applying these limits, computer simulations of hydraulic$_{\text{weig}}$ with discrete time steps are robust, even with extreme values.

4.6   Summary

This chapter introduced and fully defined the hydraulic$_{\text{weig}}$ model. A so-called model configuration assigns values to the eight adjustable parameters of hydraulic$_{\text{weig}}$. These parameters affect the way in which the model predicts energy expenditure and recovery as liquid flow. Configurations were categorised into configuration types depending on what liquid flow phases can occur with them.

Section 4.5.1 formalised liquid flow phases of hydraulic$_{\text{weig}}$ in accordance with how Morton (1990) defined the M-M model. Section 4.5.2 formalised liquid flow in discretised equations, which are convenient for step-by-step computer simulations and align with how Sundström et al. (2014) defined the M-M model. We extended the discretised equations with extreme case handling to avoid estimation errors for large step sizes or extreme parameter combinations.

All equations of Sections 4.5.1 and 4.5.2 were implemented and published in Weigend et al. (2021), and in the python package `threecomphyd`[2]. The package provides scripts to run model simulations, animations, and to fit the model to real-athlete data with an optimisation approach introduced in the next chapter.

---

[2]`https://github.com/faweigend/three_comp_hyd`

Chapter 5

Fitting the Model

The first research question in Section 3.2 asked whether hydraulic$_\text{weig}$ could be a feasible alternative to work-balance models. This chapter provides a proof-of-concept that hydraulic$_\text{weig}$ can be fitted to the same input data as work-balance models, such that it is capable of making accurate predictions for energy expenditure and recovery of an athlete during intermittent exercise. Furthermore, the presented proof-of-concept confirms the first thought experiment presented in Section 3.1.1. Specifically, it confirmed that hydraulic$_\text{weig}$ can address the proposed improvements for work-balance models by Caen et al. (2019).

To achieve this, we developed a pathway to fit hydraulic$_\text{weig}$ to CP and W' of an athlete and to recovery observations published by Caen et al. (2019). The developed pathway was published in Weigend et al. (2021). The present chapter introduces the procedure and adds more detail to our publication. Specifically, we expand further on the quality of fitting results in Section 5.3 and how our findings relate to the overarching research questions of this thesis. Furthermore, we present convergence plots and more example Pareto fronts in Appendix A.2.

5.1 Objective functions

Currently established work-balance models use the CP and W' of an athlete to predict remaining energy capacities of an athlete during intermittent exercise (Skiba and Clarke, 2021). Generally, intermittent exercise can be separated into phases of energy expenditure during high power outputs and phases of energy recovery during low power outputs. Therefore, to use hydraulic$_\text{weig}$ as an alternative to work-balance models, a pathway is needed that uses CP and W' to optimise the parameters of hydraulic$_\text{weig}$ for predicting energy expenditure and recovery during intermittent exercise. Specifically, fitting hydraulic$_\text{weig}$ requires finding optimal values for LS, LF, M_U, M_{LS}, M_{LF}, θ, γ and ϕ to make the model predict energy expenditure and recovery of an athlete.

This separates the task of fitting hydraulic$_\text{weig}$ to an athlete into two objectives. The first objective is to predict energy expenditure as accurately as possible; the second is to predict energy recovery as accurately as possible. To achieve this, the present work defines the following two objective functions that assign a measure of prediction error to hydraulic$_\text{weig}$ predictions for energy expenditure and recovery. Both functions require the inputs CP and W' and a hydraulic$_\text{weig}$ configuration, consisting of values for LS, LF, M_U, M_{LS}, M_{LF}, θ, γ and ϕ. Both functions return a single value as a measure of the prediction error. The smaller the returned value, the higher the prediction accuracy.

5.1.1 Energy expenditure

CP and W' are parameters of the critical power model. Work-balance models use the critical power model for energy expenditure predictions (Skiba and Clarke, 2021). Therefore, the energy expenditure objective function of this work compares the TTE predictions of hydraulic$_\text{weig}$ to TTE predictions of the critical power model. The lower the error measurement returned, the more accurately does a hydraulic$_\text{weig}$ model with the given input configuration resemble the predictions of the critical power model with the given input CP and W'.

The returned prediction error measurement is estimated from TTE test predictions. With the given input values for CP and W', the objective function uses Equation (2.6) to estimate 12 power outputs that lead to exhaustion after 120 s, 130 s, 140 s, 150 s, 170 s, 190 s, 210 s, 250 s, 310 s, 400 s, 600 s and 1200 s. This range features more intensities for shorter times, such that the hydraulic$_\text{weig}$ model is optimised to resemble the critical power model especially closely between 2 min and 5 min. During intermittent exercise, such ranges of energy expenditure are more likely to occur.

Then, the hydraulic$_\text{weig}$ model with the given configuration simulates constant exercise at these power outputs. At the start of a simulation, all tanks are filled. Then the tap of hydraulic$_\text{weig}$ is opened according to the constant power output to be tested. When liquid flow from the tap cannot match demand, exhaustion is reached, and the predicted time to exhaustion is compared with the expected time.

These tests yield 12 differences between the expected TTEs according to the critical power model and the predicted TTEs of a hydraulic$_\text{weig}$ model with the given input configuration. The energy expenditure objective function then returns the normalised root mean squared error (NRMSE) of these differences as the error measurement.

Table 5.1: WB1→RB→WB2 test conditions and observed recovery ratios extracted from Figure 3 of Caen et al. (2019).

P_{work} (W)	P_{rec} (W)	T_{rec} (s)	recovery ratio (%)
$P240$	CP_{33}	120	55
$P240$	CP_{33}	240	61
$P240$	CP_{33}	360	70.5
$P240$	CP_{66}	120	49
$P240$	CP_{66}	240	55
$P240$	CP_{66}	360	58
$P480$	CP_{33}	120	42
$P480$	CP_{33}	240	52
$P480$	CP_{33}	360	59.5
$P480$	CP_{66}	120	38
$P480$	CP_{66}	240	37.5
$P480$	CP_{66}	360	50

5.1.2 Energy recovery

The recovery dynamics of work-balance models $W'_{\text{bal}-\text{int}}$, W'_{skib} and W'_{bart} were fitted to small groups of tested participants (7 participants for $W'_{\text{bal}-\text{int}}$; 10 participants for W'_{skib}; 4 participants for W'_{bart}). The energy recovery objective function of this work uses a similar approach for the hydraulic$_{\text{weig}}$ model.

Caen et al. (2019) reported averaged recovery ratios of 11 participants during 12 distinct conditions for the earlier introduced WB1→RB→WB2 test protocol. The present work chose the data of Caen et al. (2019) due to the diversity of test conditions. We extracted the ratios summarised in Table 5.1 from Figure 3 of Caen et al. (2019). The conditions $P240$ and $P480$ are intensities predicted by the critical power model to lead to exhaustion after 240 s and 480 s respectively. CP_{33} and CP_{66} are 33 % and 66 % of CP. Recovery ratios in % denote how much of W' was recovered during a WB1→RB→WB2 test with the given P_{work}, P_{rec}, and T_{rec} conditions.

The energy recovery objective function of this work estimates $P240$, $P480$, CP_{33}, and CP_{66} from the CP and W' given as inputs to the function. Using these, it simulates the 12 WB1→RB→WB2 tests with a hydraulic$_{\text{weig}}$ model with the given configuration and compares the predicted recovery ratios to the observations of Table 5.1. For this simulation, the tap of hydraulic$_{\text{weig}}$ is opened according to P_{work} until the model predicts exhaustion. Then the tap is opened according to P_{rec} for T_{rec}. Subsequently, the tap is opened according to P_{work} again until the model predicts exhaustion for the second time. The ratio of both times to exhaustion is

the ratio of recovery.

This procedure yields 12 differences in predicted to expected recovery ratios. The energy recovery objective function returns the NRMSE of these differences as the error measurement. The smaller the error, the more accurate the energy recovery predictions for Table 5.1 by a hydraulic$_{\text{weig}}$ model with the given configuration.

5.2 The evolutionary algorithm

As discussed earlier, a configuration of hydraulic$_{\text{weig}}$ consists of eight real-valued parameters for LS, LF, M_U, M_{LS}, M_{LF}, θ, γ and ϕ. Given CP and W' of an athlete, an optimal configuration entails the values that yield the lowest possible NRMSE from the two objective functions defined above. In Weigend et al. (2021), we showed (in a proof-of-concept) that a successful strategy to approximate an optimal configuration in this search space is evolutionary computation (Eiben and Smith, 2015; Bäck and Schwefel, 1993; Biscani and Izzo, 2020).

Specifically, we fitted optimal configurations with the established Multi-Objective Evolutionary Algorithm with Decomposition (MOEA/D) approach as implemented in Pygmo[1](Qingfu Zhang and Hui Li, 2007; Biscani and Izzo, 2020). A common underlying idea behind evolutionary computation approaches, such as MOEA/D, is that a population of individuals is evaluated using objective functions (Eiben and Smith, 2015). In the case of this work, individuals are hydraulic$_{\text{weig}}$ configurations; their performance is measured with the energy expenditure and recovery objective functions defined above. Such a performance evaluation allows to implement a 'survival-of-the-fittest' approach. The well-performing hydraulic$_{\text{weig}}$ configurations of a population are selected as parents. They pass on their mixed and slightly mutated parameters to a new generation of offspring individuals. Over multiple generations of survival of the fittest, recombination, and mutation, the evolving population contains increasingly optimised hydraulic$_{\text{weig}}$ configurations for the objective functions.

MOEA/D was chosen because of its popularity and robust results in our initial experiments. Other global optimisers could be equally suitable. For example, Lidar et al. (2021) chose the non-linear grey-box parameter estimation solver *nlgreyest* of MATLAB[2]. Many more appropriate algorithms could be explored, but the engineering of a problem-specific and fine-tuned fitting algorithm was not the main purpose of this work.

Nevertheless, global optimisation algorithms, such as MOEA/D or *nlgreyest* of MATLAB[3], are tools for a wide range of applications. Therefore, even if fine-tuning

---

[1]https://esa.github.io/pygmo2/
[2]https://au.mathworks.com/products/matlab.html
[3]https://au.mathworks.com/products/matlab.html

Table 5.2: Initial guesses for hydraulic$_{\text{weig}}$ configuration parameters and bounds applied to them.

parameter	normal distribution	lower bound	upper bound
LF	$\mathcal{N}(0.3, 0.12) \cdot W'$	1	500 000
LS	$\mathcal{N}(1, 0.4) \cdot W'$	1	500 000
M_U	$\mathcal{N}(1, 0.4) \cdot \mathrm{CP}$	1	5 000
M_{LS}	$\mathcal{N}(1, 0.4) \cdot \mathrm{CP}$	1	5 000
M_{LF}	$\mathcal{N}(0.1, 0.04) \cdot \mathrm{CP}$	1	5 000
θ	$\mathcal{N}(0.25, 0.1)$	0	1
γ	$\mathcal{N}(0.25, 0.1)$	0	1
ϕ	$\mathcal{N}(0.5, 0.2)$	0	1

the fitting algorithm was not the focus of this work, it was necessary to customise settings to some extent to fit hydraulic$_{\text{weig}}$. Here, it was vital to develop a fitting procedure that was reliable at obtaining accurate results, with speed and efficiency as secondary concerns. The following subsections describe our algorithm design choices and customised MOEA/D settings.

5.2.1 Initial random population

The evolutionary optimisation procedure of MOEA/D starts with a population of random hydraulic$_{\text{weig}}$ configurations. This work defines that the random parameters of these initial configurations are chosen the from normal distributions summarised in Table 5.2.

Initial test runs revealed, that W' was an appropriate starting point for LF and LS, from where the evolutionary algorithm could approximate the optimal values for these parameters. In addition, LF was generally smaller in optimised solutions than LS. Therefore, LF was set to be randomly chosen from a normal distribution with mean 0.3 and standard deviation 0.12, multiplied by W' and LS from a normal distribution with mean 1 and standard deviation 0.4, also multiplied by W'.

As can be seen in entries for M_U, M_{LS}, M_{LF} in Table 5.2, CP was an appropriate starting point for maximal flow capacities. M_{LF} was generally relatively small compared with M_U and M_{LS} in optimised configurations. The values of θ and γ were set to be randomly chosen from a normal distribution with mean 0.25 and standard deviation of 0.1, so that LS generally had a positive size. If $(\theta + \gamma) < 1$ was violated, the configuration was dismissed.

Normal distributions retain a small probability that extreme values will occur. These extreme values can be useful for exploring the search space, but this work limited them by upper and lower bounds to ensure the computations were feasible. Also these boundaries are summarised in Table 5.2. The lower bounds were set at 1 J for tank capacities, 1 W for maximal flow capacities, and 0 for tank positions θ, γ, ϕ.
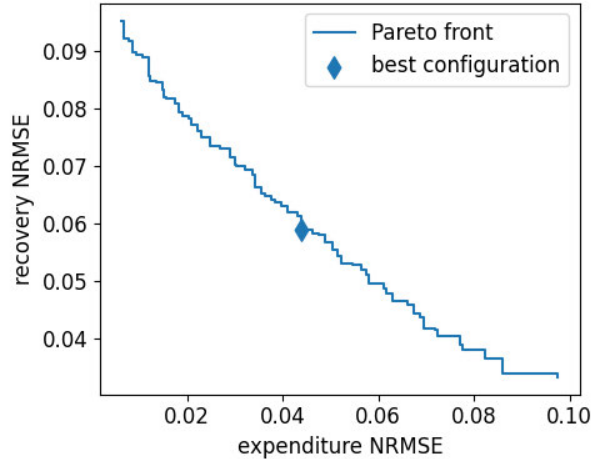
Figure 5.1: An example Pareto front of hydraulic$_{\text{weig}}$ configurations. This work defines the configuration with the best trade-off as the one with minimal distance to point $(0, 0)$. It is marked as a blue diamond.

The upper bounds were set to exceed the average CP and W' of elite athletes more than ten times; for example, Bartram et al. (2018) reported World Championships winners in a track cycling endurance discipline had an averaged CP of $393\,\text{W} \pm 14\,\text{W}$ and W' of $23\,300\,\text{J} \pm 3\,300\,\text{J}$. Therefore, the chosen upper bounds far exceed what a human body is capable of, and none of the optimised fittings throughout this research of this work came close to them.

Starting with an initial random population of hydraulic$_{\text{weig}}$ configurations, the evolutionary optimisation algorithm optimised configurations within the defined parameter boundaries.

5.2.2 Best trade-off

MOEA/D was used to optimise the initial random population of hydraulic$_{\text{weig}}$ configurations to minimise two objectives: NRMSE of the energy expenditure objective function and NRMSE of the recovery objective function. During initial experiments, it became apparent that it was not possible to configure a hydraulic$_{\text{weig}}$ model such that it achieved the optimal NRMSE of 0 for both objectives simultaneously. Some prediction error for energy expenditure must be accepted to minimise energy recovery error and vice versa. An optimal trade-off had to be defined.

A hydraulic$_{\text{weig}}$ configuration is on the so-called Pareto front if it has been optimised so far that the NRMSE of one objective function cannot be reduced without increasing the other. Figure 5.1 displays an example Pareto front. NRMSE of the energy expenditure objective function is on the x-axis, and that of the energy recovery objective function on the y-axis. A configuration on the top left on the front achieves a low NRMSE of < 0.02 from the energy expenditure objective function

but a relatively high NRMSE > 0.09 for energy recovery. For a configuration on the bottom right on the front, it is the other way around.

As defined above, objective functions return a NRMSE of 12 differences, which makes NRMSE values directly comparable. Both energy expenditure and energy recovery predictions are equally important for simulating intermittent exercise. Therefore, the best trade-off balances them equally weighted. This work defined that the optimal trade-off is maintained by the configuration with minimal Euclidean distance to the point $(0, 0)$, marked as a blue diamond in the example in Figure 5.1. The hydraulic$_{\text{weig}}$ configuration with minimal distance is referred to as the 'best-fit solution'.

5.2.3 Asynchronous islands

Because of the initial random population, and the random factors that affect the optimisation steps of the evolutionary fitting procedure, the algorithm occasionally became stuck in local optima during experiments. This caused all individuals of a population to approximate a solution that was less accurate than the best-fit solutions of other runs of the same algorithm. As a result, we observed considerable variance in the best-fit solutions of consecutive algorithm runs on the same problem. To ensure that each run of the algorithm resulted in best-fit solutions of consistently high quality, MOEA/D was coupled with the asynchronous islands functionality of Pygmo (Biscani and Izzo, 2020).

Evolving an optimal solution using the asynchronous islands functionality means that several instances (one for each island) of the evolutionary algorithm were run at the same time, but were isolated from each other (Biscani and Izzo, 2020). After a set number of evolution steps (generations), solutions from each of the island populations travelled in between islands. Then each algorithm continued to evolve its population, which now contained a few migrant solutions from the populations of the other algorithms. This step of evolving for a set number of generations and then exchanging solutions is henceforth referred to as a cycle.

The exchange of intermediate solutions between multiple algorithms diversifies their populations and makes it less likely that they become stuck in a local optimum. The following subsection demonstrates the robustness of the algorithm over 10 consecutive runs and across distinct parameter settings. An example fitting showing the evolving fronts of 7 islands is depicted in Appendix A.2.

5.2.4 Algorithm settings

As outlined earlier, comparison of hydraulic$_{\text{weig}}$ with work-balance models requires a reliable fitting algorithm, but fine-tuning the fitting for speed was only of secondary

Table 5.3: Used Pygmo default settings for MOEA/D and asynchronous islands (Biscani and Izzo, 2020).

algorithm settings	default
weight_generation	"grid"
decomposition	"tchebycheff"
neighbors	20
CR	1
F	0.5
eta_m	20
realb	0.9
limit	2
preserve_diversity	True
seed	random
migration type	"p2p"
topology	"fully connected"

importance. Therefore, future work will need to improve the algorithm settings further.

Most algorithm settings of MOEA/D and the asynchronous islands functionality remained at the default that Pygmo provided; these are summarised in Table 5.3. These settings affect algorithm specifics, such as the recombination of parents for a new generation, random mutation, or migration between islands. We report them for documentation purposes; their elaboration and fine-tuning would exceed the scope of this work.

Nevertheless, value ranges for four settings were investigated to ensure best-fit solutions of consistent quality. This work used a grid search approach to investigate combinations of settings for cycles (10,40,80), generations (10,20,30), population size (32,64), and islands (7,14,21). These value ranges were chosen from initial experiments. The group averages published by Caen et al. (2019) were CP of 248 W and W' of 18 200 J. Using these as the input, the algorithm was run 10 times with each of the possible combinations of settings for cycles, generations, population size and islands.

Table 5.4 summarises a total of 540 fittings of hydraulic$_\text{weig}$ to CP of 248 W and W' of 18 200 J. Each row represents one algorithm setting combination and summarises the results of 10 runs by listing the best (min), average (mean), and worst (max) distances to (0,0) of the resulting 10 best-fit solutions and the configuration of the solution with minimal distance.

The parameters of the optimised hydraulic$_\text{weig}$ configurations in Table 5.4 display values for LF that were typically within 2 000 J of W'. LS could be up to three times larger than LF and the larger it was, the more likely that M_{LS} was below 100 W. M_U had values very close to CP. In comparison, M_{LF} was smaller, with

Table 5.4: Hydraulic$_{\text{weig}}$ fitted to CP of 248 W and W' of 18 200 J. The results of distinct settings for MOEA/D with the asynchronous islands functionality are summarised. Each combination of algorithm settings was run 10 times. The smaller the minimal distance to (0,0), the more accurate the hydraulic$_{\text{weig}}$ configuration, according to the objective functions defined in Section 5.1. Distances are shown as the best (min), average (mean), and worst (max). Furthermore, the configuration that achieved the minimal distance (best hydraulic$_{\text{weig}}$ configuration) is shown. The algorithm setting combination with the lowest mean distance is highlighted in grey.

algorithm settings				distance to (0,0)			best hydraulic$_{\text{weig}}$ configuration							
gens	cycles	pop	islands	min	mean	max	LF(J)	LS(J)	M_U(W)	M_{LS}(W)	M_{LF}(W)	θ	γ	ϕ
10	10	32	7	0.0777	0.0923	0.1156	14 450	29 026	247	108	20	0.51	0.04	0.35
10	10	32	14	0.0759	0.0834	0.1047	18 706	37 041	250	90	16	0.64	0.04	0.2
10	10	32	21	0.0725	0.0793	0.0853	17 679	44 187	247	112	17	0.7	0.01	0.26
10	10	64	7	0.0734	0.0822	0.1055	15 584	39 288	247	116	19	0.65	0.01	0.32
10	10	64	14	0.0737	0.0757	0.0833	15 992	38 930	247	101	19	0.62	0.04	0.3
10	10	64	21	0.0731	0.0747	0.0783	17 651	45 559	247	103	17	0.7	0.03	0.25
10	40	32	7	0.073	0.0772	0.0849	17 413	51 320	248	99	16	0.7	0.07	0.26
10	40	32	14	0.0732	0.0761	0.0921	17 255	76 750	247	96	17	0.74	0.1	0.27
10	40	32	21	0.0724	0.0734	0.0748	16 736	44 441	246	109	19	0.68	0.02	0.29
10	40	64	7	0.0722	0.0735	0.0776	17 108	47 178	247	111	18	0.7	0.02	0.28
10	40	64	14	0.0724	0.0731	0.0743	18 269	53 241	247	107	17	0.74	0.02	0.25
10	40	64	21	0.0721	0.0726	0.073	17 457	46 236	247	105	19	0.69	0.02	0.26
10	80	32	7	0.0733	0.0764	0.0873	16 221	36 450	247	111	19	0.62	0.02	0.31
10	80	32	14	0.0725	0.075	0.0796	16 776	43 518	247	110	18	0.69	0.01	0.29
10	80	32	21	0.0726	0.074	0.0796	16 923	41 780	247	110	18	0.67	0.01	0.29
10	80	64	7	0.0725	0.0734	0.0778	16 749	41 432	247	109	19	0.66	0.01	0.28
10	80	64	14	0.0723	0.0733	0.0777	17 848	48 139	247	107	17	0.72	0.01	0.25
10	80	64	21	0.0724	0.0726	0.0731	16 250	42 734	247	111	20	0.66	0.03	0.3
20	10	32	7	0.0729	0.0802	0.1038	16 374	48 532	247	103	20	0.68	0.05	0.28
20	10	32	14	0.0736	0.0769	0.0832	16 898	35 465	247	108	18	0.62	0.01	0.28
20	10	32	21	0.0734	0.0758	0.0781	15 573	77 767	247	94	18	0.69	0.16	0.32
20	10	64	7	0.0727	0.0755	0.0795	16 911	45 622	247	110	18	0.7	0.02	0.27
20	10	64	14	0.0726	0.073	0.0734	17 275	71 979	248	99	16	0.75	0.09	0.27
20	10	64	21	0.0722	0.0733	0.0752	18 264	49 095	247	110	17	0.73	0.01	0.24
20	40	32	7	0.0731	0.0772	0.0993	15 940	40 369	247	105	19	0.63	0.05	0.31
20	40	32	14	0.0728	0.0744	0.0769	16 458	41 725	247	108	17	0.66	0.03	0.3
20	40	32	21	0.0731	0.074	0.0748	15 481	37 680	247	112	19	0.62	0.03	0.33
20	40	64	7	0.0722	0.0729	0.0761	17 626	58 034	247	96	17	0.71	0.07	0.26
20	40	64	14	0.0721	0.0725	0.0731	18 128	45 966	247	106	17	0.71	0.02	0.24
20	40	64	21	0.0718	0.0723	0.0725	18 042	46 718	247	107	17	0.72	0.02	0.25
20	80	32	7	0.0724	0.074	0.0764	17 569	49 817	247	105	18	0.72	0.03	0.26
20	80	32	14	0.0727	0.0733	0.0767	19 167	53 078	248	104	15	0.76	0.01	0.21
20	80	32	21	0.0727	0.0732	0.0745	16 688	48 099	247	105	18	0.69	0.05	0.27
20	80	64	7	0.0722	0.0726	0.0731	16 682	44 357	247	112	19	0.68	0.02	0.29
20	80	64	14	0.0721	0.0725	0.0728	18 053	50 179	247	106	16	0.73	0.03	0.25
20	80	64	21	0.0722	0.0724	0.0728	16 777	45 583	247	108	19	0.69	0.02	0.28
30	10	32	7	0.0747	0.0809	0.093	14 619	57 423	247	94	19	0.6	0.18	0.35
30	10	32	14	0.0728	0.075	0.078	16 950	57 304	247	96	18	0.69	0.09	0.27
30	10	32	21	0.0729	0.0742	0.0756	16 606	42 383	246	112	19	0.67	0.01	0.29
30	10	64	7	0.0721	0.0734	0.076	18 292	48 369	247	108	16	0.73	0.02	0.25
30	10	64	14	0.072	0.0727	0.0733	18 024	53 422	247	106	18	0.74	0.02	0.25
30	10	64	21	0.072	0.0727	0.0738	17 285	45 790	247	111	18	0.7	0.01	0.27
30	40	32	7	0.0728	0.0747	0.0795	18 208	54 965	248	97	16	0.72	0.06	0.24
30	40	32	14	0.0727	0.0734	0.0742	16 770	42 517	246	112	18	0.67	0.01	0.3
30	40	32	21	0.0725	0.0729	0.0733	16 577	41 662	247	115	19	0.68	0.01	0.3
30	40	64	7	0.0718	0.0726	0.0743	16 770	44 997	247	112	18	0.69	0.02	0.29
30	40	64	14	0.0719	0.0722	0.0728	17 571	43 403	247	108	17	0.69	0.01	0.26
30	40	64	21	0.0719	0.0723	0.0729	16 574	42 663	247	112	19	0.67	0.01	0.29
30	80	32	7	0.072	0.0742	0.0803	17 037	41 622	247	110	19	0.67	0.01	0.28
30	80	32	14	0.0728	0.0729	0.0731	17 431	73 419	248	98	17	0.76	0.08	0.26
30	80	32	21	0.0721	0.0729	0.0743	17 574	43 129	247	109	17	0.69	0.01	0.26
30	80	64	7	0.0722	0.0725	0.0729	16 748	48 486	247	104	19	0.69	0.05	0.28
30	80	64	14	0.072	0.0723	0.0726	16 313	47 535	247	106	19	0.68	0.05	0.3
30	80	64	21	0.0719	0.0723	0.0726	16 381	53 690	247	105	18	0.7	0.07	0.3

values between 15 W and 22 W. An optimal value for θ seemed to revolve around 0.7, for γ around 0.05, and for ϕ around 0.28. Each configuration was obtained with distinct algorithm settings. Of these, some where more suitable than others.

We can observe that the 10 consecutive runs with settings featuring 20+ generations, 40+ cycles, and 14+ islands, were of especially consistent quality with differences < 0.005 between the minimum and maximum distance to (0,0). The setting combination featuring 30 generations, 40 cycles, 64 population size, and 14 islands had the (on average) most accurate best-fit solutions. This is highlighted in grey in Table 5.4. Based on these results, we recommend to run the outlined algorithm with these settings; this was done for the remainder of this work.

5.2.5 Computation time

The described evolutionary algorithm is a time-demanding fitting process. As an example, fitting a configuration with the default parameters from Table 5.3 and 30 generations, 40 cycles, 64 population size, and 7 islands to CP of 248 W and W' of 18 200 J took on average 8 h 46 min and 16 s on 7 cores of an Intel® Xeon® CPU E5-2650 v4 @ 2.20GHz.

To decrease the computation time, we implemented a so-called early stopping approach. If during an algorithm run the current best-fit solution did not improve for more than 10 cycles, computations were stopped, and the best-fit solution of the last cycle was returned as the result. This saved computation time by preventing the algorithm from continuing to optimise solutions that had already converged to an optimum. As an example, early stopping was triggered for 8 of 10 fittings to CP 285 W and W' 21 296 J and the average computation time was 5 h 27 min and 41 s. This is an improvement from the average 8 h 46 min and 16 s of the previous example.

For this work, the main objective was to design a reliable fitting procedure that produced fittings with minimal error. Therefore, more fine-tuning of algorithm settings to improve computation times should be the subject of future work.

5.3 Quality of fitting results

The above sections introduced an evolutionary algorithm and objective functions to fit a hydraulic$_{\text{weig}}$ configuration to given CP and W', such that the hydraulic$_{\text{weig}}$ model predicts energy expenditure and recovery, the former according to the critical power model and the latter according to observations from Caen et al. (2019). To show that this procedure produces reliable results, 10 independent fittings and their predictions were analysed in detail. Further we provide additional convergence plots in Appendix A.2.

Table 5.5: A summary of 10 independent fitting results to CP of 248 W and W' of 18 200 J. Each row is a hydraulic$_{\text{weig}}$ configuration obtained from an independent evolutionary algorithm run with the settings discussed in Section 5.2.4. The columns expNRMSE and recNRMSE are NRMSE errors of the energy expenditure and energy recovery objective functions. The entries in the bottom row are the mean and standard deviation of the respective column.

LF(J)	LS(J)	M_U(W)	M_{LS}(W)	M_{LF}(W)	θ	γ	ϕ	expNRMSE	recNRMSE	dist (0,0)
17 457	46 236	247	105	19	0.69	0.02	0.26	0.0426	0.0582	0.0721
17 571	43 403	247	108	17	0.69	0.01	0.26	0.0395	0.0601	0.0719
18 053	50 179	247	106	16	0.73	0.03	0.25	0.0361	0.0624	0.0721
16 684	44 888	247	113	18	0.69	0.02	0.29	0.0429	0.0576	0.0718
16 852	42 910	247	109	19	0.68	0.02	0.28	0.0432	0.0578	0.0721
16 606	39 680	247	112	18	0.66	0.01	0.29	0.0425	0.0584	0.0722
16 899	41 130	247	109	19	0.66	0.01	0.28	0.0423	0.0587	0.0724
17 532	45 987	247	111	18	0.71	0.01	0.27	0.04	0.0597	0.0718
16 312	38 982	247	112	19	0.64	0.01	0.3	0.0449	0.0565	0.0722
16 891	45 521	247	116	17	0.7	0.01	0.29	0.0407	0.0595	0.0721
17 086	43 892	247	110	18	0.69	0.01	0.28	0.0415	0.0589	0.0721
±512	±3 221	±0	±3	±1	±0.02	±0.01	±0.01	±0.0023	±0.0016	±0.0002

5.3.1 Ten independent fittings

Table 5.5 summarises the parameters and errors of 10 hydraulic$_{\text{weig}}$ configurations fitted to the reported group averages of CP of 248 W and W' of 18 200 J by Caen et al. (2019). The fittings were obtained from independent consecutive runs of the presented evolutionary algorithm with the default settings summarised in Table 5.3 and the recommended settings of 30 generations, 40 cycles, 64 population size, and 14 islands. The bottom row of Table 5.5 summarises the mean and standard error of the respective column. NRMSE errors of the energy expenditure and energy recovery objective functions are summarised in expNRMSE and recNRMSE columns. The last column details the resulting Euclidian distance to (0,0).

The configuration parameters were of magnitudes comparable to those previously reported in Table 5.4. LF had values close to W'. LS was more than two times larger than LF. For every configuration, M_U was 1 W below CP, M_{LS} was (on average) 110 W, and M_{LF} was the smallest maximal flow capacity, with an average of 18 W.

Examining the configuration types introduced in Section 4.4, all fitted configurations summarised in Table 5.5 were of type M. However, their γ, as well as distances between ϕ and $1-\theta$, were small. Therefore, although the liquid flow phases $U_l L_l$ and $U_f L_f$ did occur in these configurations, they had limited influence on model predictions. This applied to all fitted configurations. Hence, the 10 independent fittings in Table 5.5 are consistent in their configuration types and liquid flow dynamics.
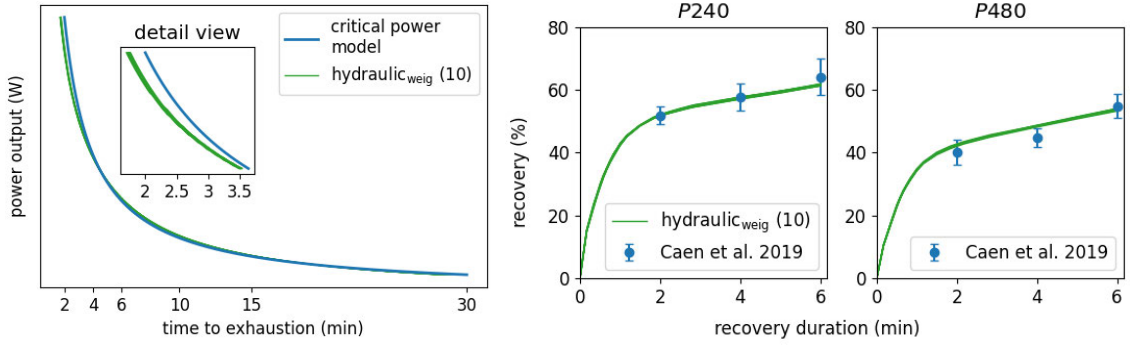
Figure 5.2: Energy expenditure and recovery dynamics of 10 hydraulic$_{\text{weig}}$ models fitted to CP of 248 W and W' of 18 200 J. Predictions of all 10 hydraulic$_{\text{weig}}$ models are plotted in green. **Left:** Predicted power output to TTE relationship in comparison to the critical power model. **Right:** Predicted recovery dynamics in comparison to published observations by Caen et al. (2019).

As an additional observation, the NRMSE values in Table 5.5 indicate that the best possible trade-off between energy expenditure and energy recovery error was not an equal balance. The 10 configurations favoured a smaller error for energy expenditure at the cost of a larger error for energy recovery predictions. The resulting Euclidean distances to (0,0) varied only marginally with a standard deviation of 0.0002.

In conclusion, all 10 configurations in Table 5.5 had parameters of magnitudes similar to previously observed results, were of the same configuration type, and made a very similar trade-off between energy expenditure and recovery error. This strongly indicates that the outlined evolutionary algorithm produces fittings of consistent quality in consecutive runs. However, the fittings are not exactly the same and the standard deviations reported in Table 5.5 show that some variations remained. In the following, we show that variations are too small to have a considerable effect on model predictions.

5.3.2 Impact of variations on predictions

To investigate the effect of the remaining variations between independent fittings, Figure 5.2 depicts a comparison of predictions of the critical power model and observations of Caen et al. (2019) with predictions of hydraulic$_{\text{weig}}$ models with the 10 configurations of Table 5.5.

As observable on the left in Figure 5.2, overall the by the critical power concept suggested hyperbolic relationship of constant power output and TTE was closely matched by all 10 hydraulic$_{\text{weig}}$ models. However, deviations in the predictions for high power outputs were observable. The hydraulic$_{\text{weig}}$ models predicted up to 20 s shorter TTEs at power outputs that led to exhaustion earlier than 3.5 min. This

was the case for all hydraulic$_{weig}$ models, and variations in their configurations had little influence on prediction results. These shorter TTE predictions at high power outputs were part of the necessary trade-off between energy expenditure and energy recovery predictions.

Depicted on the right in Figure 5.2 are the recovery ratios predicted by hydraulic$_{weig}$ models, and the observations reported by Caen et al. (2019). As summarised in Section 2.3.3.2, constant power outputs that were predicted to lead to exhaustion after 240 s (P240) and 480 s (P480) were the P_{work} intensities for their WB1→RB→WB2 protocol (Caen et al., 2019). Their P_{rec} was either 33 % of CP or 66 % of CP. Their investigated T_{rec} were 2 min, 4 min, and 6 min. Caen et al. (2019) reported means and standard deviations of the observed recovery ratios.

The 10 hydraulic$_{weig}$ models simulated the test protocol by Caen et al. (2019). As observable on the right in Figure 5.2, except for recovery after 4 min from exhaustive exercise at P480, the predictions of all models were close to the reported means and within their standard deviation ranges.

Despite the variations in their configurations, the predictions of the hydraulic$_{weig}$ models were nearly identical and all were equally able to account for the characteristics of prior exhaustive exercise. These results show that hydraulic$_{weig}$ models are capable of addressing the improvements for energy recovery predictions suggested by Caen et al. (2019). Therefore, they confirm the first thought experiment presented in Section 3.1.1.

5.4   Summary

This chapter completed a proof-of-concept that hydraulic$_{weig}$ models can be used as an alternative to work-balance models. The fitting of hydraulic$_{weig}$ was formalised as a two-objective evolutionary optimisation procedure that has CP and W' as inputs, and returns a hydraulic$_{weig}$ configuration. The returned configuration optimises two objectives: to predict energy expenditure similar to the critical power model and to predict energy recovery similar to the recovery ratios published by Caen et al. (2019).

Using published group averages for CP and W' of Caen et al. (2019) for our proof-of-concept, we showed that the fitting algorithm produces consistent results. Fitted hydraulic$_{weig}$ models were capable of addressing the by Caen et al. (2019) proposed improvements for energy recovery predictions and thus confirmed the first thought experiment from Section 3.1.1. TTE predictions at high intensities that lead to exhaustion within less than 3.5 min were underpredicted by up to 20 s; however, overall the hyperbolic critical power model was closely matched.

The presented proof-of-concept motivated further investigations as a validation strategy for hydraulic$_{\mathrm{weig}}$. The next chapter validates and scrutinises energy recovery predictions of fitted hydraulic$_{\mathrm{weig}}$ models on more data. The evolutionary fitting approach of this chapter is part of our python package `threecomphy`[4] and was published in Weigend et al. (2021).

---

[4]`https://github.com/faweigend/three_comp_hyd`

Chapter 6

Validation and Comparison

The hydraulic$_{\text{weig}}$ model allows the fitting of its parameters using an optimisation approach that only requires CP and W' as inputs. In the previous chapter, a hydraulic$_{\text{weig}}$ model could be fitted to predict both energy expenditure and recovery kinetics for one example case. This chapter further contributes to the investigation and validation of the hydraulic$_{\text{weig}}$ model and compares it to work-balance models and to data compiled from previous studies.

This chapter addresses the second research question in Section 3.2 and confirms both thought experiments in Sections 3.1.1 and 3.1.2. The findings have been accepted for publication in *Annals of Operational Research* (ANOR). We published the preprint as Weigend et al. (2022a).

6.1   Model comparison on published data

Thought experiments of Sections 3.1.1 and 3.1.2 and the proof-of-concept in the previous Chapter 5 suggest that hydraulic$_{\text{weig}}$ can address improvements for energy recovery predictions of work-balance models proposed by Caen et al. (2019, 2021). To validate this, we compared hydraulic$_{\text{weig}}$ energy recovery predictions to those of three work-balance models on published data of five studies (Weigend et al., 2022a). This comparison was possible because the evolutionary fitting procedure for hydraulic$_{\text{weig}}$ only requires CP and W' as inputs, which are measurements commonly reported in previous studies. We compiled observed recovery ratios from previous studies and hypothesised that the hydraulic$_{\text{weig}}$ model would predict them overall more accurately than the work-balance models.

This section summarises the methodology and results of our Weigend et al. (2022a) comparison study and is structured as follows: Sections 6.1.1 and 6.1.2 introduce the work-balance models with which the hydraulic$_{\text{weig}}$ predictions were compared. Section 6.1.3 lists the previously published studies on the recovery from exercise, from which data were extracted to compare models on. Section 6.1.4

describes the metrics used to assess the model goodness-of-fits and prediction capabilities. Extracted data points and prediction results are detailed in Section 6.1.5. Finally, Section 6.1.6 summarises model performance on all data and goodness-of-fit metric scores.

6.1.1 Work-balance models to compare with

As outlined in Section 2.3.2, the initial $W'_{\text{bal}-\text{int}}$ model by Skiba et al. (2012) was later updated in Skiba et al. (2015). Substantial differences between these versions exist. As shown by Skiba and Clarke (2021), the original model by Skiba et al. (2012) contradicts the assumption of the critical power model that W' linearly depletes. This contradiction interferes with the WB1→RB→WB2 protocol with which most recovery ratios were estimated in this work as well as in several previous studies, such as Caen et al. (2019, 2021). Therefore, we did not compare the $W'_{\text{bal}-\text{int}}$ model, instead focusing on the updated version W'_{skib} and the W'_{bart} models. Henceforth, W'_{skib} and W'_{bart} are also referred to as $W'_{\text{bal}-\text{ode}}$ models. We also introduced a third work-balance model, which also falls into the category of $W'_{\text{bal}-\text{ode}}$ models.

6.1.2 An additional work-balance model

W'_{skib}, W'_{bart}, and the hydraulic$_{\text{weig}}$ model of this work were each created by fitting to different sets of recovery observations. Objective metrics to compare model quality; for example, the Akaike Information Criterion (Burnham and Anderson, 2004), require that models are fitted to the same data. Therefore, to allow a more comprehensive comparison, we added a third $W'_{\text{bal}-\text{ode}}$ model with a new $\mathcal{T}_t$ by Weigend et al. (2022a) ($\mathcal{T}_{\text{weig},t}$). We derived this $\mathcal{T}_{\text{weig},t}$ with a procedure as close as possible to those of Skiba et al. (2012) and Bartram et al. (2018).

As the first step, a constant value for $\mathcal{T}_t$ in Equation (2.13) was fitted to each recovery ratio and recovery time combination from Table 1 of the Appendix of Weigend et al. (2021). For these observations power output was constant for every discrete time step t during recovery and thus $\mathcal{T}_t$ was constant with the same value for all t. We used the standard Broyden-Fletcher-Goldfarb-Shanno algorithm implementation of SciPy (SciPy 1.0 Contributors et al., 2020), with 200 as the initial guess, to fit a constant $\mathcal{T}_t$, which enabled Equation (2.13) to best reproduce the observed recovery ratio. This resulted in twelve pairs of fitted constant $\mathcal{T}_t$s to constant recovery intensities.

As the next step, we then fitted an exponential function to these twelve pairs using the non-linear least squares implementation of SciPy (SciPy 1.0 Contributors et al., 2020). With the recovery intensity as $D_{\text{CP},t}$, the function was of the form $\mathcal{T}_t = a \cdot e^{D_{\text{CP},t} \cdot b} + c$. With the values of Skiba et al. (2012) as the initial guess

(546, -0.01, 316), the resulting optimal constants were $a = 1\,274.45, b = -0.0308$, and $c = 266.65$. Thus, given any $D_{\mathrm{CP},t}$ at a discrete time step t, $\mathcal{T}_{\mathrm{weig}_t}$ can be estimated as

$$\mathcal{T}_{\mathrm{weig},t} = 1\,274.45 \cdot e^{-0.0308 \cdot D_{\mathrm{CP},t}} + 266.65. \qquad (6.1)$$

Unfortunately, this fitted equation failed to fit the data satisfactorily ($R^2 = 0.14$). Nevertheless, we used it because it was developed using a procedure that closely resembled those used to estimate $\mathcal{T}_{\mathrm{skib},t}$ and $\mathcal{T}_{\mathrm{bart},t}$. The introduction of $\mathcal{T}_{\mathrm{weig},t}$ is valuable because it allows the application of the Akaike Information Criterion metric, which requires compared models to be fitted to the same data points. Henceforth, Equation (2.13) with Equation (6.1) is referred to as work-balance model with $\mathcal{T}_t$ fitted by Weigend et al. (2022a) (W'_{weig}).

6.1.3 Data extraction for model comparisons

We extracted data from previous studies that investigated energy recovery dynamics and used these to compare and evaluate recovery ratio predictions of all models. The studies for comparison were identified from Table 1 of the comprehensive review by Chorley and Lamb (2020). From these studies, we retained those that featured appropriate data, except those that met the following exclusion criteria:

- Featured a mode of exercise other than cycling. Cycle ergometers measure power output directly. Power outputs during modes of exercise such as running or swimming are not directly comparable because they are estimated using different methods or are approximated; e.g., (Morton and Billat, 2004) focused only on speed instead of power.

- The observations were made under extreme conditions (e.g., hypoxia or altitude).

- Insufficient information was reported to simulate the prescribed protocol in and/or to infer a recovery ratio of W' in percent; e.g., the integral version of the work-balance model by Skiba et al. (2012) assumes recovery during high-intensity exercise such that recovery ratios cannot be straightforwardly inferred.

- The prescribed protocol leaves doubt if reported recovery ratios are comparable to the WB1→RB→WB2 protocol (e.g, repeated ramp tests until exhaustion, $50\,\%$ W' depletion followed by a 3-min all-out test, or knee-extension maximal voluntary contraction (MVC) test during recovery).

Five studies were included for comparison, four of which were obtained from the review by Chorley and Lamb (2020): (Bartram et al., 2018), (Chidnok et al., 2012), (Ferguson et al., 2010), and (Caen et al., 2019). After the summary of Chorley and Lamb (2020) was published, Caen et al. (2021) published a study that investigated the W' reconstitution dynamics in even more detail and which was thus added to the list.

The data in the listed studies were presented in diverse ways, such that modifications were made to some of the data to enable model comparison. The study by Caen et al. (2019) did not report distinct mean values for every investigated condition, such that we derived approximate values in Weigend et al. (2021) to fit our hydraulic$_{\text{weig}}$ to their conditions. Hence, the data for comparison are the values from Weigend et al. (2021). Further, the study by Bartram et al. (2018) fitted their own W'_{bart} model, where $\mathcal{T}_t$ was defined according to Equation (2.15). Therefore, we used W'_{bart} model predictions for prescribed intensities of Bartram et al. (2018) as the observations against which the other models were compared.

The study by Chidnok et al. (2012) reported times to exhaustion from their intermittent exercise protocol instead of recovery ratios. Power output during recovery was constant in their tests. Therefore, to derive recovery ratio estimations that are comparable with the WB1→RB→WB2 procedure, we fitted a constant value for $\mathcal{T}_t$ of the $W'_{\text{bal}-\text{ode}}$ model to each of their prescribed protocols and times to exhaustion. These constant values for $\mathcal{T}_t$ were fitted with the Brent method implementation by SciPy (SciPy 1.0 Contributors et al., 2020) to find a local minimum in the interval between [100, 1 000]. We then used WB1 → RB → WB2 recovery ratio estimations of $W'_{\text{bal}-\text{ode}}$ models with fitted constant $\mathcal{T}_t$ as the observations with which to compare W'_{skib}, W'_{bart}, W'_{weig}, and hydraulic$_{\text{weig}}$.

6.1.4 The metrics of goodness of fit

The metrics of goodness of fit used to compare the models were root mean squared error (RMSE), mean absolute error (MAE), and the small-sample version of the Akaike Information Criterion (AIC$_{\text{c}}$). Chai and Draxler (2014) discussed RMSE and MAE as widely adopted metrics for assessing model prediction capabilities. We compared predictive accuracy by comparing RMSE and MAE on data to which competing models were not fitted. Lower values for RMSE and MAE were interpreted as more accurate predictions.

To statistically compare prediction error distributions between models, we used a bootstrap hypothesis test (Efron and Tibshirani, 1993; Good, 2000). We did so because only small data sets were available and we could not assume normal distributed prediction errors with equal variances for every compared model. The null hypothesis of our bootstrap test was that prediction error distributions of two compared

models are the same. Because we used two prediction error metrics (RMSE and MAE) we investigated the null hypothesis on both. We used the absolute difference between RMSE and also between MAE of compared groups as our test statistics. With the null hypothesis that error distributions are the same, we could bootstrap new samples by randomly selecting with replacement from all pooled observations. We created a distribution of test statistics from $1\,000\,000$ bootstrap samples to reliably approximate the p-value of our observed test statistic at high precision. We rejected the null hypothesis if the p-value $< .05$.

We also compared models with the $\mathrm{AIC_c}$, which was first proposed by Sugiura (1978). The $\mathrm{AIC_c}$ is a model selection tool used to investigate the balance between model complexity and explanatory capability (Burnham and Anderson, 2004). $\mathrm{AIC_c}$ penalises the number of parameters of the model and thus provides insight into the balance between model complexity and goodness of fit. The lower the $\mathrm{AIC_c}$ score, the better this balance is met. The $\mathrm{AIC_c}$ was calculated as

$$\mathrm{AIC}_c = n \cdot \ln(\mathrm{MSE}) + 2k + \frac{2k \cdot (k+1)}{n-k-1}, \tag{6.2}$$

where MSE is the mean squared error, n is the number of data points and k is the number of parameters of the model. Models have to be fitted to and applied to the same data in order to obtain comparable $\mathrm{AIC_c}$ scores. Therefore, only W'_{weig} and hydraulic$_{\mathrm{weig}}$ were comparable with this criterion in this work.

Altogether, the hypothesis that the more complex hydraulic$_{\mathrm{weig}}$ model fits the data better than the established $W'_{\mathrm{bal-ode}}$ models will be supported if the overall RMSE, MAE and $\mathrm{AIC_c}$ scores are lower for hydraulic$_{\mathrm{weig}}$ than for other models, and if prediction error distributions are significantly different to those of other models.

6.1.5 Results

This section presents the extracted data and the prediction results of $W'_{\mathrm{bal-ode}}$, and hydraulic$_{\mathrm{weig}}$ models for each listed previous study. We refer to extracted data from studies by the last name of the first author; for example, the extracted data from Bartram et al. (2018) are referred to as 'Bartram data set'. All studies collected their data through performance tests that required athletes to exercise until volitional exhaustion. Such tests are affected by circumstances that are difficult to measure and control (e.g., motivation, nutrition and state-of-mind). Therefore, recovery ratio observations are noisy and the extracted group averages were accompanied by large standard deviations. These uncertainties prevented us from drawing conclusions about model quality on averages of individual data sets and instead necessitated to perform the comparison between models across all available data. We begin by presenting the extracted data and model predictions of individual data sets throughout

63
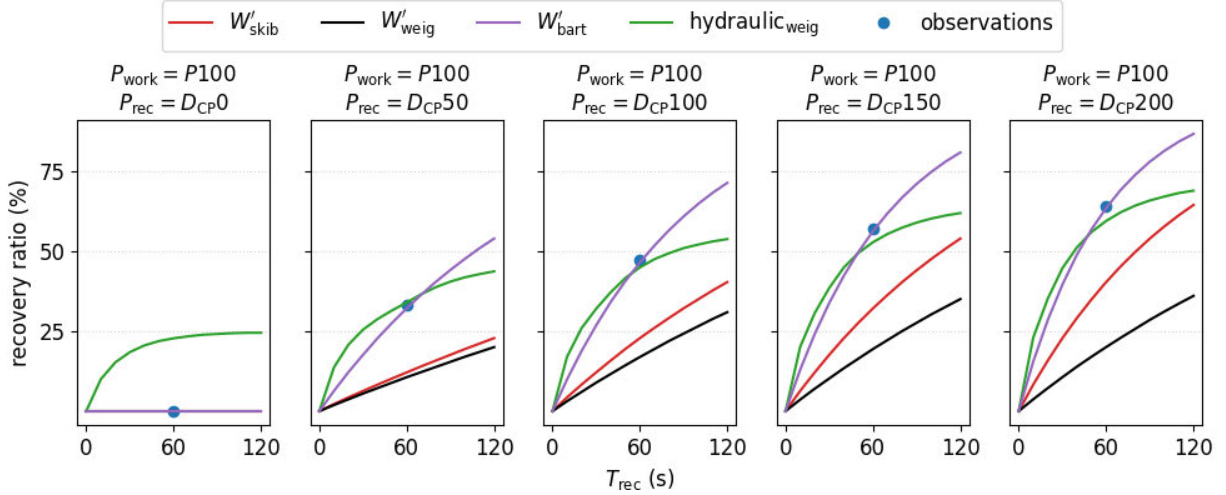
Figure 6.1: Comparison of model predictions with the WB1→RB→WB2 protocol. Depicted are the recovery dynamics around 60 s at various D_{CP} recovery intensities after a preceding exhaustive exercise at the intensity that was predicted to lead to exhaustion after 100 s ($P100$). Chosen intensities and time frames stem from the protocol prescribed by Bartram et al. (2018) and predictions of W'_{bart} were used as the observations with which to compare models.

Table 6.1: The left part of the table summarises extracted data and conditions from Bartram et al. (2018). The right part of the table displays model predictions. Predictions of W'_{bart} were taken as the observed recovery ratios.

Parameters from Bartram et al. (2018)					recovery ratios (%)			
CP(W)	W'(J)	P_{work}(W)	P_{rec}(W)	T_{rec}(s)	observed	W'_{skib}	W'_{weig}	hydraulic$_{\text{weig}}$
393	23 300	626	393	60	0.0	0.0	0.0	22.7
393	23 300	626	343	60	33.0	12.1	10.6	33.9
393	23 300	626	293	60	47.0	22.8	16.9	44.8
393	23 300	626	243	60	57.0	32.1	19.4	52.7
393	23 300	626	193	60	64.0	40.3	20.0	59.3

Sections 6.1.5.1 to 6.1.5.5. This is followed by summarising all prediction errors and resulting RMSE, MAE, and AIC$_{\text{c}}$ scores in the final Section 6.1.6.

6.1.5.1   Bartram data set

The protocol prescribed by Bartram et al. (2018) consisted of three work bouts interspersed with two 60-s recovery bouts. The first two work bouts each lasted for 30 s, and the final one until volitional exhaustion. Work bout exercise intensity (P_{work}) was set to $P100$; that is, the intensity that was predicted to lead to exhaustion after 100 s. The recovery bout intensity (P_{rec}) was set to differences to CP (D_{CP}) of 200 (i.e., CP - 200 W), or 150, 100, 50, or 0. The group averaged CP and W' for the four world-class cyclists featured in Bartram et al. (2018) were 393 W and 23 300 J. Altogether, these input values resulted in an estimated $P100$ exhaustive intensity of
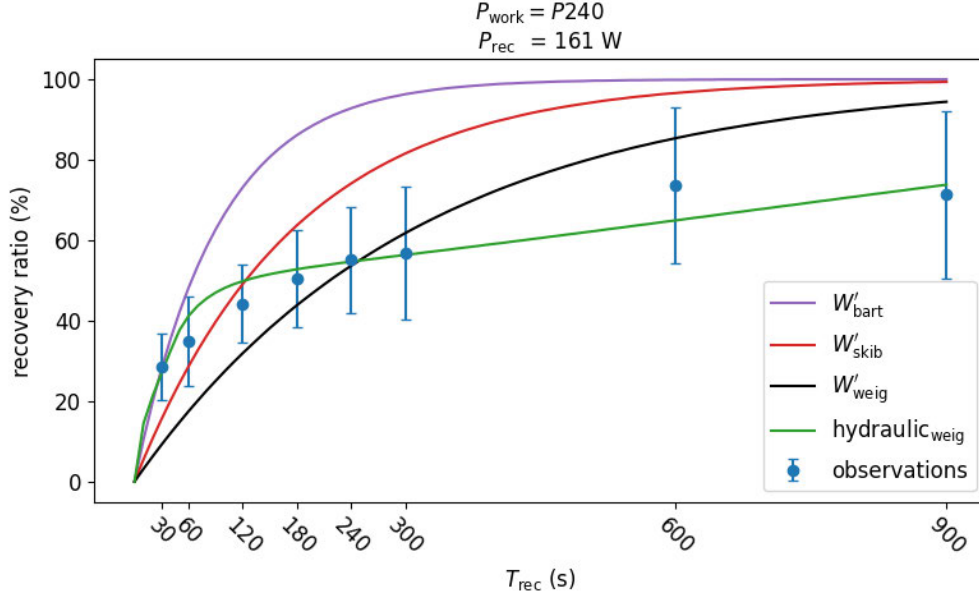
Figure 6.2: Comparison of model predictions with published observations by Caen et al. (2021). After an exhaustive exercise bout at $P240$, recovery dynamics at an intensity of 161 W were simulated using the defined WB1→RB→WB2 recovery estimation protocol. The published observed recovery ratios by Caen et al. (2021) are depicted in blue.

626 W and recovery intensities $D_{CP}0$ of 393 W, $D_{CP}50$ of 343 W, $D_{CP}100$ of 293 W, $D_{CP}150$ of 243 W, and $D_{CP}200$ of 193 W, respectively.

The resulting recovery predictions of $W'_{bal-ode}$ and hydraulic$_{weig}$ models are summarised in Figure 6.1 and Table 6.1. The W'_{bart} model was not compared because it was the model that Bartram et al. (2018) fitted to their observations. We used it to create the observations against which the other models were compared. The fitted hydraulic$_{weig}$ configuration to CP and W' by Bartram et al. (2018) was: [23 112 J, 65 845 J, 392 W, 149 W, 24 W, 0.73, 0.01, 0.24] in the order of Equation (4.1).

Figure 6.1 and Table 6.1 show that in all cases except $D_{CP}0$ the recovery ratios predicted by hydraulic$_{weig}$ model were closest to the ones observed by Bartram et al. (2018), followed by W'_{skib} and then W'_{weig}. On the contrary, the hydraulic model is the only model to predict recovery at $D_{CP}0$.

6.1.5.2 Caen data set

The protocol by Caen et al. (2021) investigated the recovery dynamics following exhaustive exercise at $P240$ (published average of 349 W). They prescribed a recovery intensity of 161 W on average, which was determined by selecting 90 % of the power at the gas exchange threshold (Binder et al., 2008) of their participants. The average CP of their participants was 269 W, and the average W' was 19 200 J. The reported

Table 6.2: The left part of the table summarises extracted data and conditions from Caen et al. (2021). The right part of the table displays model predictions.

Parameters from Caen et al. (2021)					recovery ratios (%)				
CP(W)	W'(J)	P_{work}(W)	P_{rec}(W)	T_{rec}(s)	observed	W'_{bart}	W'_{skib}	W'_{weig}	hydraulic$_{\mathrm{weig}}$
269	19 200	349	161	30	28.6	28.0	15.5	9.2	26.9
269	19 200	349	161	60	34.8	48.2	28.7	17.5	41.2
269	19 200	349	161	120	44.2	73.2	49.1	31.9	49.8
269	19 200	349	161	180	50.5	86.1	63.7	43.8	52.8
269	19 200	349	161	240	55.1	92.8	74.1	53.6	54.7
269	19 200	349	161	300	56.8	96.3	81.5	61.8	56.3
269	19 200	349	161	600	73.7	99.9	96.6	85.4	64.9
269	19 200	349	161	900	71.3	100.0	99.4	94.4	73.8

observed recovery ratios were 28.6 % ± 8.2 % after 30 s, 34.8 % ± 11.1 % after 60 s, 44.2 % ± 9.7 % after 120 s, 50.5 % ± 12.1 % after 180 s, 55.1 % ± 13.3 % after 240 s, 56.8 % ± 16.4 % after 300 s, 73.7 % ± 19.3 % after 600 s, and 71.3 % ± 20.8 % after 900 s.

The simulation parameters and results of the defined recovery estimation protocol are summarised in Figure 6.2 and Table 6.2. Fitting hydraulic$_{\mathrm{weig}}$ to CP and W' group averages resulted in the configuration [17 631 J, 46 246 J, 267 W, 118 W, 21 W, 0.68, 0.01, 0.29] in the order of Equation (4.1). The recovery ratios predicted by the hydraulic$_{\mathrm{weig}}$ model matched the observed values better compared with all the other models. Nevertheless, some lack of fit for the hydraulic$_{\mathrm{weig}}$ model was observed: the model overpredicted the recovery ratios at early time points and underpredicted those at longer time points except for the last one. W'_{skib}, W'_{bart}, and W'_{weig} model predictions consistently overestimated recovery for longer recovery times.

6.1.5.3   Chidnok data set

Chidnok et al. (2012) prescribed a protocol that alternated between 60-s work bouts and 30-s recovery bouts until the athlete reached exhaustion. With their protocol, the work-bout intensity P_{work} was set to $P240$. The protocol prescribed four trials each with a different recovery intensity P_{rec} (20 W as the 'low' recovery intensity, 95 W as 'medium', 173 W as 'high', and 270 W as the 'severe' recovery intensity). The participants had an average CP of 241 W and W' of 21 100 J. Their recorded times to exhaustion were 1 224 s ± 497 s with 'low' recovery intensity, 759 s ± 243 s with 'medium', 557 s ± 90 s with 'high', and 329 s ± 29 s with 'severe'.

As described in Section 6.1.3, to compare observations of Chidnok et al. (2012) to WB1→RB→WB2 protocol estimations, a constant value for $\mathcal{T}_t$ for the $W'_{\mathrm{bal-ode}}$ model was fitted to the protocol by Chidnok et al. (2012) for each of their recovery conditions. The resulting $\mathcal{T}_t$ values were 165.19 s for the 'low' recovery intensity
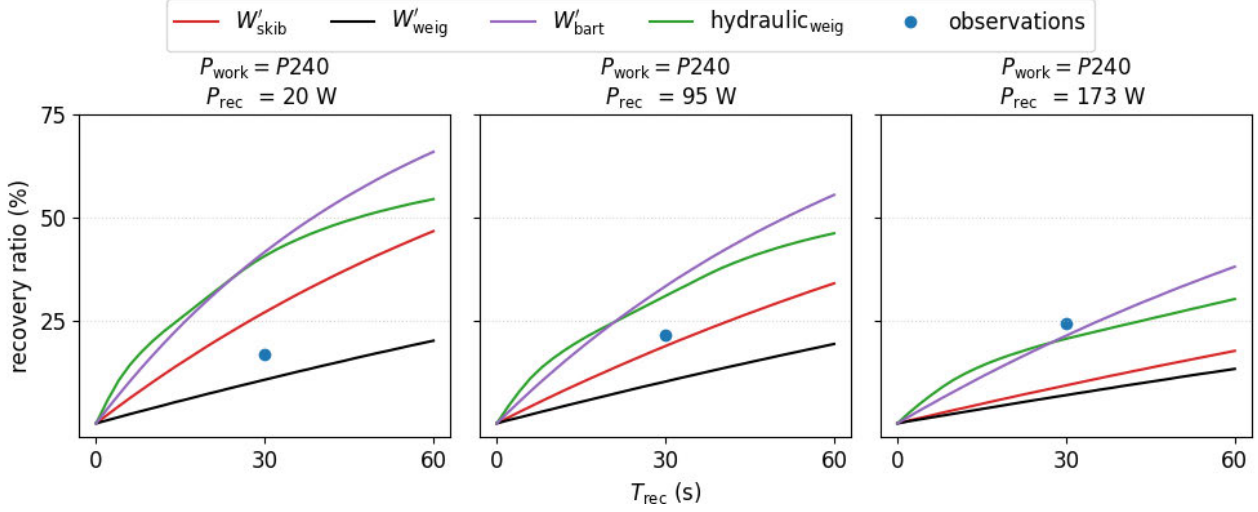
Figure 6.3: Predicted recovery dynamics of compared models up to 60 s after a preceding exhaustive exercise at $P240$ and at three different recovery intensities (20 W, 95 W, and 173 W). Observations were predicted recovery ratios of $W'_{\text{bal-ode}}$ models with a constant $\mathcal{T}$ fitted to reported times to exhaustion by Chidnok et al. (2012).

Table 6.3: The left part of the table summarises extracted data and conditions from Chidnok et al. (2012). The right part of the table displays model predictions.

Parameters from Chidnok et al. (2012)					recovery ratios (%)				
CP(W)	W'(J)	P_{work}(W)	P_{rec}(W)	T_{rec}(s)	observed	W'_{bart}	W'_{skib}	W'_{weig}	hydraulic$_{\text{weig}}$
241	21 100	329	20	30	16.6	41.6	27.0	10.6	40.6
241	21 100	329	95	30	21.4	33.3	18.8	10.1	30.9
241	21 100	329	173	30	24.4	21.3	9.2	6.8	20.5
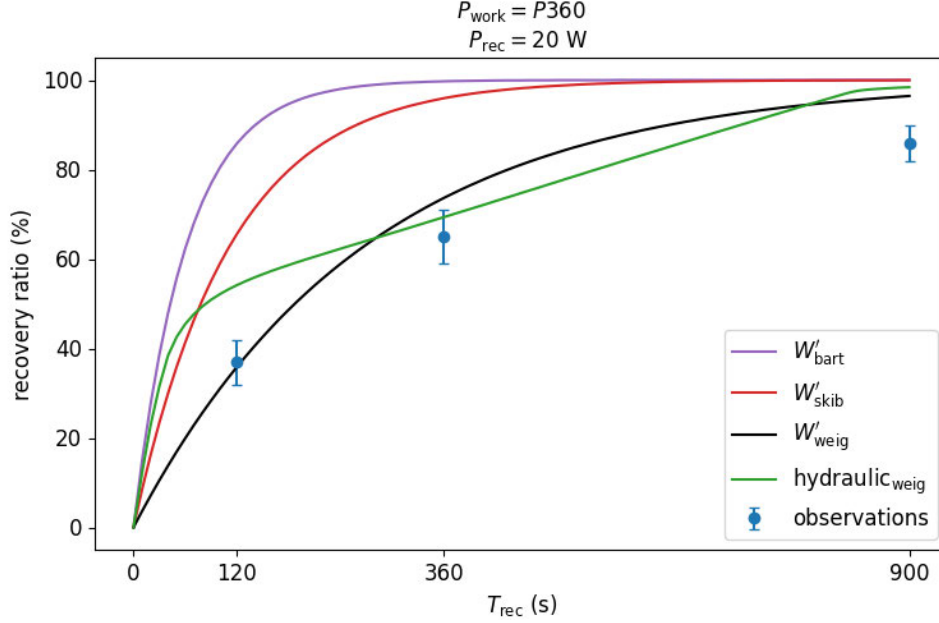
Figure 6.4: A comparison of predicted recovery dynamics after an exhaustive exercise bout at $P360$ and at a recovery intensity of 20 W. Recovery ratios are estimated with the WB1→RB→WB2 protocol, which resembles the prescribed protocol by Ferguson et al. (2010). Published observations by Ferguson et al. (2010) are depicted in blue.

protocol, a $\mathcal{T}_t$ of 124.81 s for 'medium', and a $\mathcal{T}_t$ of 107.45 s for 'high'. The 'severe' recovery intensity was left out because 270 W lies above the average CP of 241 W. In this case, no recovery should occur if the assumptions of $W'_{\text{bal}-\text{ode}}$ model hold true. Chidnok et al. (2012) prescribed recovery bouts of 30 s and WB1→RB→WB2 protocol estimations with corresponding fitted $\mathcal{T}_t$s and with a T_{rec} of 30 s were 24.6 % at the 'low' intensity, 21.7 % at 'medium', and 16.7 % at the 'high' recovery intensity.

The fitted hydraulic$_{\text{weig}}$ configuration to CP and W' by Chidnok et al. (2012) was: [18 919 J, 48 052 J, 240 W, 115 W, 19 W, 0.68, 0.05, 0.31], in the order of Equation (4.1). Predictions of all models and extracted conditions for the recovery estimation protocol are summarised in Figure 6.3 and Table 6.3. In the case of the 'low' recovery intensity predictions of the W'_{skib} and W'_{weig} models were the most accurate. In the case of the 'medium' recovery intensity the W'_{skib} model was the most accurate, and in the remaining 'high' condition hydraulic$_{\text{weig}}$ and W'_{bart} model predictions were closest to the data. None of the models made predictions that were close to all three observations.

6.1.5.4 Ferguson data set

Ferguson et al. (2010) prescribed a protocol with an initial time to exhaustion bout at the intensity that was predicted to lead to exhaustion after 360 s ($P360$), followed by a recovery at 20 W for 2 min, 6 min, or 15 min. After recovery, exercise intensity was then increased back to one of three possible high-intensity work rates. Thus,

Table 6.4: The left part of the table summarises extracted data and conditions from Ferguson et al. (2010). The right part of the table displays model predictions.

Parameters from Ferguson et al. (2010)					recovery ratios (%)				
CP(W)	W'(J)	P_{work}(W)	P_{rec}(W)	T_{rec}(W)	observed	W'_{bart}	W'_{skib}	W'_{weig}	hydraulic$_{\mathrm{weig}}$
212	21 600	269	20	120	37.0	85.8	65.6	35.9	54.2
212	21 600	269	20	360	65.0	99.7	95.9	73.6	69.4
212	21 600	269	20	900	86.0	100.0	100.0	96.4	98.4

each participant performed nine tests in total with three different constant work rates after three different recovery times. The CP model was fitted to these three times to exhaustion after each recovery period to determine changes in CP and W'. Ferguson et al. (2010) published their group averages for CP as 212 W, W' as 21 600 J, the $P360$ as 269 W, and the observed recovery ratios after 2 min as (37 % ± 5 %), 6 min (65 % ± 6 %), and 15 min (86 % ± 4 %).

Extracted parameters for the recovery intensity protocol and model prediction results are summarised in Figure 6.4 and Table 6.3, together with reported means by Ferguson et al. (2010). The fitted hydraulic$_{\mathrm{weig}}$ configuration to CP and W' group averages by Ferguson et al. (2010) was [18 730 J, 81 031 J, 212 W, 94 W, 19 W, 0.63, 0.21, 0.34], in the order of Equation (4.1). In this set-up W'_{weig} predictions were overall closest to the published observations. Hydraulic$_{\mathrm{weig}}$ overestimated the recovery after 120 s and after 900 s. W'_{skib} and W'_{bart} overestimated recovery in every instance.

6.1.5.5 Weigend data set

The values for this data set are those of Table 5.1, or Table 1 in the Appendix of Weigend et al. (2021). Reported measures recreate the depicted means in Figure 3 of Caen et al. (2019). They consist of three recovery ratios for four conditions each: preceding exhausting exercise at $P240$ or $P480$, followed by recovery at 33 % of CP or 66 % of CP. The participants of Caen et al. (2019) had an average CP of 248 W and W' of 18 200 J, which results in a $P240$ of 285 W, a $P480$ of 323 W, 33 % of CP as 81 W, and 66 % of CP as 163 W.

Extracted parameters for the recovery ratio estimation protocol and model predictions are summarised in Figure 6.5 and Table 6.5. The best fit hydraulic$_{\mathrm{weig}}$ configuration to CP and W' group averages was [18 042 J, 46 718 J, 247 W, 107 W, 17 W, 0.72, 0.02, 0.25], ordered as in Equation (4.1). As described earlier, the recovery ratio values by Weigend et al. (2021) were used to fit the $\mathcal{T}_{\mathrm{weig},t}$ for the W'_{weig} model and are used in the evolutionary fitting process for hydraulic$_{\mathrm{weig}}$ to fit recovery dynamics. Therefore, both W'_{weig} and hydraulic$_{\mathrm{weig}}$ were not scrutinised for predictive accuracy on this data set. Their predicted recovery ratios were recorded
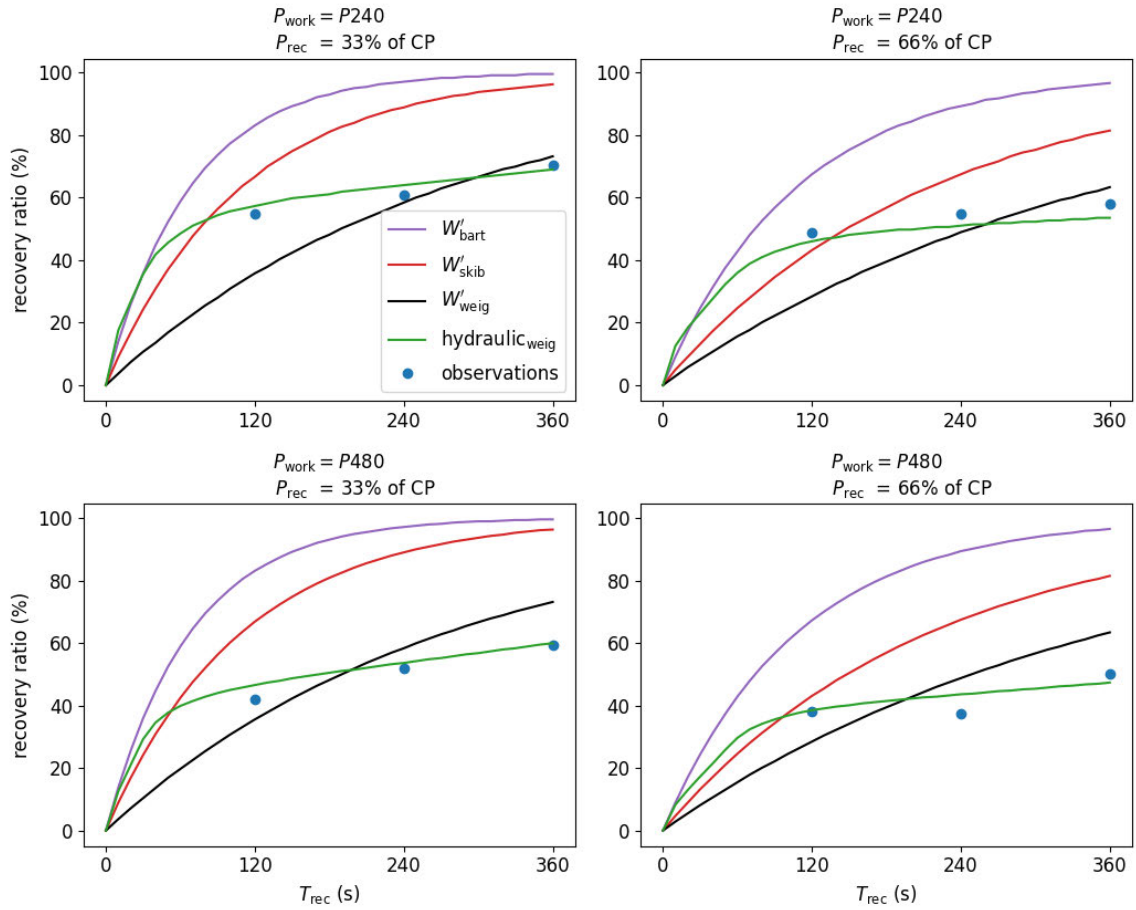
Figure 6.5: Predicted recovery dynamics in comparison with measures that we derived from observations of Caen et al. (2019). We derived three recovery ratios for four conditions each: preceding exhausting exercise at $P240$ or $P480$ followed by recovery at 33 % of CP or 66 % of CP. Depicted observations are the values from Table 5.1 and approximate Figure 3 of the publication by Caen et al. (2019). W'_{weig} and hydraulic$_{weig}$ were fitted to these observations.

Table 6.5: The left part of the table summarises extracted data and conditions that Weigend et al. (2021) derived from Caen et al. (2019). The right part of the table displays model predictions. Both W'_{weig} and hydraulic$_{\text{weig}}$ were fitted to these observations. Their predicted recovery ratios were recorded for the AIC$_{\text{c}}$ goodness of fit estimation metric in Section 6.1.4.

Parameters from Weigend et al. (2021)					recovery ratios(%)				
CP(W)	W'(J)	P_{work}(W)	P_{rec}(W)	T_{rec}(s)	observed	W'_{bart}	W'_{skib}	W'_{weig}	hydraulic$_{\text{weig}}$
248	18 200	323	81	120	55.0	83.1	66.7	35.5	58.3
248	18 200	323	81	240	61.0	97.1	89.0	58.3	65.0
248	18 200	323	81	360	70.5	99.5	96.3	73.1	70.1
248	18 200	323	163	120	49.0	67.2	42.9	28.4	46.5
248	18 200	323	163	240	55.0	89.2	67.4	48.7	51.5
248	18 200	323	163	360	58.0	96.5	81.4	63.3	54.2
248	18 200	285	81	120	42.0	83.0	66.8	35.5	46.8
248	18 200	285	81	240	52.0	97.1	89.0	58.3	54.0
248	18 200	285	81	360	59.5	99.5	96.3	73.1	60.4
248	18 200	285	163	120	38.0	67.2	42.9	28.4	38.5
248	18 200	285	163	240	37.5	89.3	67.4	48.7	43.6
248	18 200	285	163	360	50.0	96.5	81.4	63.3	47.4

for the AIC$_{\text{c}}$ goodness of fit estimation metric in covered in the next subsection. Out of the remaining two models predictions of W'_{skib} were closer to the observations but both overpredict in nearly all instances.

6.1.6 Summary of metrics of goodness of fit

Table 6.6 summarises the prediction errors of the competing models and the resulting metric scores on our investigated data sets. RMSE and MAE were defined as the metrics to assess predictive accuracy. Their MAE scores were 24.87 with a standard deviation of absolute errors (SD) of 14.35 for W'_{bart} and 7.11 (SD = 6.83) for hydraulic$_{\text{weig}}$ (p < .001 for the difference in MAEs, bootstrap hypothesis test). The RMSE scores on Caen, Chidnok, and Ferguson data sets were 28.46 for W'_{bart} and 9.69 for hydraulic$_{\text{weig}}$. Also the bootstrap hypothesis test with the absolute difference in RMSEs as its test statistic resulted in p < .001.

Both remaining models W'_{skib} and W'_{weig} could be compared with hydraulic$_{\text{weig}}$ on the Bartram, Caen, Chidnok, and Ferguson data sets. The hydraulic$_{\text{weig}}$ featured the lowest MAE with 7.07, the lowest SD with 7.17, and lowest RMSE with 9.94. W'_{skib} predictions were significantly different to hydraulic$_{\text{weig}}$ (p < .001 with the MAE test statistic and p = .001 with RMSE). W'_{weig} predictions were significantly different to hydraulic$_{\text{weig}}$ (p = .019 with the MAE test statistic and p = .031 with RMSE).

AIC$_{\text{c}}$ was chosen as the metric to assess which model provides the best trade-off between predictive capabilities and complexity. Models must be fitted to and tested on the same data for AIC$_{\text{c}}$ scores to be comparable. Hence, as reflected in

Table 6.6: Summary of the model prediction errors and estimated metric scores. The first two columns summarise the prediction errors used to compare predictive accuracy of W'_{bart} and hydraulic$_{\text{weig}}$ via MAE, standard deviation of absolute errors (SD), and RMSE. Prediction accuracy had to be assessed using data to which models were not fitted, such that we had to exclude the Bartram and Weigend data sets. In the subsequent three columns, the Weigend data set was excluded because W'_{weig} and hydraulic$_{\text{weig}}$ were fitted to it. Finally, the AIC$_c$ metric requires models to be fitted to and be evaluated on the same data. Therefore, we compared AIC$_c$ scores estimated from prediction errors of W'_{weig} and hydraulic$_{\text{weig}}$ on all data sets. We approximated p-values for absolute differences in MAE and RMSE with a bootstrap hypothesis test and considered $p < .05$ as significant. For every metric, a lower score means a better result.

		data for prediction scores 1 (%)		data for prediction scores 2 (%)			data for AIC$_c$ scores (%)	
		W'_{bart}	hydraulic$_{\text{weig}}$	W'_{skib}	W'_{weig}	hydraulic$_{\text{weig}}$	W'_{weig}	hydraulic$_{\text{weig}}$
Bartram	0			0.0	0.0	22.7	0.0	22.7
	1			−20.9	−22.4	0.9	−22.4	0.9
	2			−24.2	−30.1	−2.2	−30.1	−2.2
	3			−24.9	−37.6	−4.3	−37.6	−4.3
	4			−23.7	−44.0	−4.7	−44.0	−4.7
Caen	0	−0.6	−1.7	−13.1	−19.4	−1.7	−19.4	−1.7
	1	13.4	6.4	−6.1	−17.3	6.4	−17.3	6.4
	2	29.0	5.6	4.9	−12.3	5.6	−12.3	5.6
	3	35.6	2.3	13.2	−6.7	2.3	−6.7	2.3
	4	37.7	−0.4	19.0	−1.5	−0.4	−1.5	−0.4
	5	39.5	−0.5	24.7	5.0	−0.5	5.0	−0.5
	6	26.2	−8.8	22.9	11.7	−8.8	11.7	−8.8
	7	28.7	2.5	28.1	23.1	2.5	23.1	2.5
Chid.	0	25.0	24.0	10.4	−6.0	24.0	−6.0	24.0
	1	11.9	9.5	−2.6	−11.3	9.5	−11.3	9.5
	2	−3.1	−3.9	−15.2	−17.6	−3.9	−17.6	−3.9
Ferg.	0	48.8	17.2	28.6	−1.1	17.2	−1.1	17.2
	1	34.7	4.4	30.9	8.6	4.4	8.6	4.4
	2	14.0	12.4	14.0	10.4	12.4	10.4	12.4
Weigend	0						−19.5	3.3
	1						−2.7	4.0
	2						2.6	−0.4
	3						−20.6	−2.5
	4						−6.3	−3.5
	5						5.3	−3.8
	6						−6.5	4.8
	7						6.3	2.0
	8						13.6	0.9
	9						−9.6	0.5
	10						11.2	6.1
	11						13.3	−2.6
MAE		24.87[*]	7.11	17.23[*]	15.06[*]	7.07		
± SD		±14.35	±6.83	±9.34	±12.19	±7.17		
RMSE		28.46[*]	9.69	19.48[*]	19.17[*]	9.94		
AIC							181.03	151.85

[*] significantly different to hydraulic$_{\text{weig}}$ predictions

the last two columns of Table 6.6, W'_{weig} and hydraulic$_{\text{weig}}$ could be compared on the combined data points of all covered data sets. With a k of 3 for W'_{weig} and a k of 8 for hydraulic$_{\text{weig}}$ the resulting scores were 151.85 for hydraulic$_{\text{weig}}$ and 181.03 for W'_{weig}. The hydraulic$_{\text{weig}}$ achieved the lower AIC$_c$ score.

6.2 Discussion

The in this chapter presented study compared the prediction capabilities and goodness-of-fit of hydraulic$_{\text{weig}}$ to that of work-balance models. We hypothesised that the hydraulic model would more accurately predict the observed recovery ratios from past studies. The hydraulic$_{\text{weig}}$ model outperformed the W'_{skib}, W'_{bart}, and W'_{weig} models with respect to objective RMSE, MAE, and AIC$_c$ metrics on data from five studies. Our findings therefore support the hypothesis. This section discusses the results in more detail, interprets them in the context of findings from the previous literature, and relates them to the thought experiments of Chapter 3 of this work.

6.2.1 Multiple vs. isolated data sets

As presented in Section 6.1.6, the standard deviations of absolute prediction errors as well as the overall MAE and RMSE, were considerably lower for the hydraulic$_{\text{weig}}$ than for the $W'_{\text{bal-ode}}$ models. But when averaging the prediction errors on isolated data sets listed in Table 6.6, hydraulic$_{\text{weig}}$ only made more accurate predictions than its competitors on the Bartram and Caen data sets. For the Bartram data set, the MAE of hydraulic$_{\text{weig}}$ was 6.96, compared with 18.74 for W'_{skib}, and 26.82 for W'_{weig} respectively. For the Caen data set, the MAE of hydraulic$_{\text{weig}}$ was the lowest with 3.52. On the remaining Chidnok data set it was W'_{skib} that achieved the lowest MAE with 9.4 and on the Ferguson data set it was W'_{weig} with 6.7.

As highlighted by Skiba and Clarke (2021) and Sreedhara et al. (2019), $W'_{\text{bal-ode}}$ models are meant to be applied to any athlete on a wide range of possible conditions. A lower MAE score for W'_{weig} on the Ferguson data set means that W'_{weig} predicted recovery ratios more closely for the particular group (six recreational active men) under the particular test conditions that Ferguson tested. However, to determine the usefulness of a model for predicting performance for high-intensity intermittent exercise in a more general sense, models should be evaluated on a multitude of scenarios. After combining all data sets, hydraulic$_{\text{weig}}$ achieved the overall lowest MAE score, which means that hydraulic$_{\text{weig}}$ could predict recovery ratios overall more accurately for a range of groups and settings.

6.2.2 Confirmation of proposed improvements

The observed less consistent prediction quality across data sets of the $W'_{\text{bal-ode}}$ models confirms the findings by Caen et al. (2019), who proposed that the predictive capabilities of work-balance models may improve with modifications that account for intensity and duration of prior exhaustive exercise. As an example, out of all compared studies in this work, Bartram et al. (2018) prescribed the highest work bout intensity for their experimental setup ($P_{\text{work}} = P100$). Considering the suggestion by Caen et al. (2019) that a shorter time to exhaustion at a high intensity allows a quicker recovery, it seems reasonable that the W'_{bart} model estimated the fastest recovery kinetics out of all recovery models.

Conversely, the Caen et al. (2019) study prescribed the lowest work bout intensity out of all compared studies ($P_{\text{work}} = P480$). Their observed recovery ratios are summarised in the Weigend data set and were slower than the W'_{bart} predictions. This observation again matches the assumption that a longer exhaustive exercise at a lower intensity requires a longer recovery.

Despite the differences in observed recovery rates, the $W'_{\text{bal-ode}}$ models allow for only a single recovery rate no matter the nature of the prior exercise. To illustrate this, we conducted simulations to depict the influence of prior exercise intensity on the recovery ratios predicted by the $W'_{\text{bal-ode}}$ and hydraulic$_{\text{weig}}$. Figure 6.6 depicts four simulations. All simulations shared the same test setup except for differing P_{work} intensities. The simulation on the left had $P_{\text{work}} = P100$, as prescribed by Bartram et al. (2018), the simulation on the right featured the lowest $P_{\text{work}} = P480$, as found in the Weigend data set. From left to right, P_{work} of the simulations decreased step wise. Bartram et al. (2018) investigated recovery after 60 s and therefore the W'_{bart} prediction after 60 s is marked as the observation on the left. Recovery ratios with $P_{\text{work}} = P480$ and $P_{\text{rec}} = 33$ % of CP of the Weigend data set are marked as observations on the right. The recovery ratios predicted by the $W'_{\text{bal-ode}}$ models were the same for each P_{work} and their predictions were therefore unable to fit all observations equally well. In contrast, the hydraulic model could account for such characteristics.

In conclusion, the comparison on multiple data sets and conditions further confirms the findings by Caen et al. (2019). The characteristics of prior exhaustive exercise have an effect on recovery ratios. The presented compiled data and the simulation in Figure 6.6, suggest that hydraulic$_{\text{weig}}$ made more accurate predictions because it could account for these characteristics and thus addressed the improvements proposed by Caen et al. (2019). Specifically, we suggest that standard deviations of MAE, as well as overall MAE and RMSE scores of hydraulic$_{\text{weig}}$ model, were
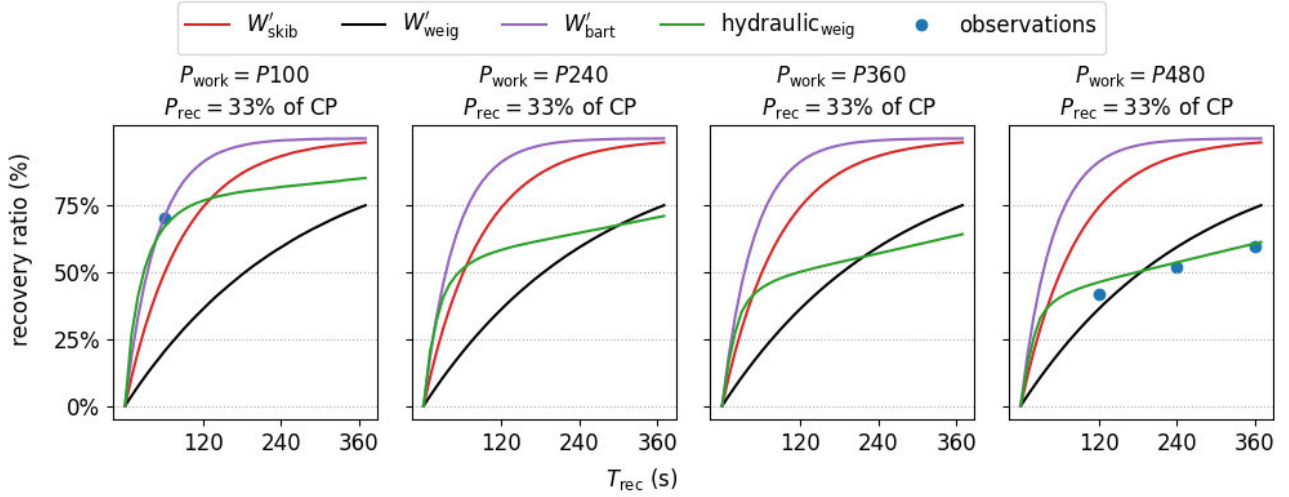
Figure 6.6: Simulated recovery ratios using the hydraulic$_{\text{weig}}$ and $W'_{\text{bal−ode}}$ models in response to prior exercise of differing intensities. The plots show that $W'_{\text{bal−ode}}$ models are insensitive to the properties of prior exhausting exercise; that is, their predictions were not affected by P_{work}. In contrast, the hydraulic$_{\text{weig}}$ was sensitive to the prior exercise properties. Performance models $W'_{\text{bal−ode}}$ were configured with a CP = 393 W, $W' = 23\,300$ J and hydraulic$_{\text{weig}}$ featured the configuration [23 112 J, 65 845 J, 392 W, 149 W, 24 W, 0.73, 0.01, 0.24]. All simulations differed only in P_{work}, which decreased from $P100$ to $P480$ in the simulations depicted from left to right. $P_{\text{work}} = P100$, prescribed by Bartram et al. (2018), was the highest intensity out of compared studies. They investigated recovery after 60 s, therefore the W'_{bart} model prediction after 60 s is marked as the observation. $P_{\text{work}} = P480$ was the lowest prescribed intensity out of compared studies and recovery ratios for $P_{\text{work}} = P480$, $P_{\text{rec}} = 33$ % of CP from the Weigend data set were marked as observations on the right.

smaller than those of $W'_{\text{bal}-\text{ode}}$ models because the hydraulic model could account for characteristics of prior exhaustive exercise.

6.2.3 Confirmation of thought experiments

Furthermore, the above findings confirm the thought experiments presented in Sections 3.1.1 and 3.1.2. The first thought experiment states that the hydraulic$_{\text{weig}}$ model can address the findings by Caen et al. (2019) because of the interactions between the three tanks it uses to model energy recovery. For example, using the notation in Figure 4.1, during high-intensity exercise, the liquid level in LF would rapidly decrease and the contribution of LS would be less than during low-intensity exercise when the liquid level in LF would decrease more slowly. Differences in fill states of LS affected recovery estimations and the above findings confirmed that it enabled hydraulic$_{\text{weig}}$ to predict rapid recovery after high-intensity exercise and a slower recovery after exercise at a lower intensity in a variety of situations.

The second thought experiment in Section 3.1.2 suggests that hydraulic$_{\text{weig}}$ is capable of modelling energy recovery as a bi-exponential function. Caen et al. (2021) showed that their observations were well explained with a bi-exponential model that implements a steeper slope during the beginning of recovery. Indeed, the observed recovery ratios of the Caen data set increased rapidly from 0 s to 120 s and then continued to rise more slowly at longer durations (Figure 6.2). Also, the first work-balance model paper by Skiba et al. (2012) proposed an alternative bi-exponential version of their $W'_{\text{bal}-\text{int}}$ model with two $\mathcal{T}_t$s. However, bi-exponential work-balance models have yet to be applied in practice. As for hydraulic models, the comparison in Figure 6.2 confirmed, that hydraulic$_{\text{weig}}$ is capable of capturing the bi-exponential nature of energy recovery and it therefore addresses proposed work-balance model improvements.

6.3 Summary

Our model comparison on data from five previous studies allows a more holistic view on recovery dynamics and confirms limitations of work-balance models that have been suggested in previous literature (Caen et al., 2019, 2021). The results of this chapter imply that more complex models, such as hydraulic$_{\text{weig}}$, can improve energy recovery predictions.

The findings have been reported in Weigend et al. (2022a), where we proposed that further efforts to merge and compare data are significant steps to advance the research of energy recovery modelling. We further proposed that the predictive capabilities of hydraulic models look strong and that they have to be considered as a possible future direction to advance energy recovery modelling.

However, hydraulic$_{weig}$ represents a different pathway for how hydraulic models can be used. Hydraulic$_{weig}$ was specifically designed and fitted to predict energy recovery during intermittent exercise. But it was also derived from the M-M model, which necessitates that hydraulic$_{weig}$ predictions are also investigated in the context that the M-M model was designed for (e.g., oxygen uptake). Therefore, the next chapter further highlights the differences between the hydraulic$_{weig}$ and M-M model and scrutinises metabolic predictions.

Chapter 7

Metabolic Prediction Limitations

As discussed in Chapter 4, hydraulic$_{\mathrm{weig}}$ is different from the original M-M model because all bioenergetic and metabolic contexts were erased. As presented in the previous Chapter 6, the removal of physiological constraints allowed hydraulic$_{\mathrm{weig}}$ to serve as a promising alternative to work-balance models for energy recovery predictions during high-intensity intermittent exercise.

However, because the origin of hydraulic$_{\mathrm{weig}}$ lies in the M-M model, it is tempting to interpret parameters of fitted hydraulic$_{\mathrm{weig}}$ models in the metabolic contexts ascribed by the M-M model. This chapter confirms that the hydraulic$_{\mathrm{weig}}$ model cannot be used to predict metabolic responses during exercise on the example of collected oxygen uptake measurements.

This chapter addresses the third research question in Section 3.2. It emphasises the distinction between hydraulic$_{\mathrm{weig}}$ and the M-M model and discusses the limitations of hydraulic$_{\mathrm{weig}}$ should it be regarded for metabolic predictions. The findings of this chapter have been published as a preprint as Weigend et al. (2022b).

7.1   Oxygen uptake predictions

According to the definitions of Morton (2006), the left tank of the M-M model represents the aerobic contribution, and therefore the flow from the left tank represents oxygen uptake ($\dot{V}_{\mathrm{O_2}}$). Morton (2006) detailed $\dot{V}_{\mathrm{O_2}}$ predictions of the M-M model in his work and $\dot{V}_{\mathrm{O_2}}$ can be measured via breath-by-breath indirect calorimetry over the course of an exercise trial. Therefore, this flow from the left tank is a suitable measurement to highlight the differences between hydraulic$_{\mathrm{weig}}$ and the M-M model on collected data from participants.

To elaborate how $\dot{V}_{\mathrm{O_2}}$ predictions of hydraulic models were computed, we use the example in Figure 7.1, which displays snapshots of a hydraulic$_{\mathrm{weig}}$ model that simulates an exercise with constant intensity.

At the beginning of exercise at second 0, all tanks were filled. Then, the tap at the bottom of the middle tank was opened according to the energy demand.

(a) 10 s (b) 20 s (c) 80 s

Figure 7.1: Snapshots of a hydraulic$_{\text{weig}}$ model simulating exercise at a constant power output of 400 W. Depicted are fill levels of tanks and the flows between them after 10 s, 20 s and 80 s.

Liquid flowed out, and the fill level of the middle tank dropped accordingly. As a consequence, liquid pressure against the pipe exit of the left tank dropped. Thus, the liquid flow from the left tank increased.

In the example from Figure 7.1, the tap was opened according to 400 W and after 10 s, the fill level of the middle tank dropped halfway to the top of the right tank. Flow from the left tank increased to an equivalent of 47.3 W. As observable in the snapshot at 20 s, when the fill level of the middle tank dropped below the top of the right tank, liquid from the right tank also flowed into the middle tank. In the snapshot at 80 s, the fill level of the middle tank dropped below the exit of the pipe of the left tank, and thus the flow from the left tank was at maximum capacity M_U.

With this understanding, we used predicted flow from the left tank as $\dot{V}_{\text{O}_2}$ predictions of hydraulic$_{\text{weig}}$ and compared them with collected data in Weigend et al. (2022b).

7.2 Comparison with data from participants

We designed our work Weigend et al. (2022b) to investigate and confirm the limitations of hydraulic$_{\text{weig}}$ should it be considered for predicting rates of aerobic and anaerobic metabolism. To verify hydraulic$_{\text{weig}}$ limitations, we compared $\dot{V}_{\text{O}_2}$ predictions of hydraulic$_{\text{weig}}$ with the collected $\dot{V}_{\text{O}_2}$ data of exercising participants. Then, we related the observations to assumptions and constraints of the M-M model.

CP and W' are required to apply $\text{hydraulic}_\text{weig}$. These parameters are estimated from TTE performance tests during which participants exercise at constant high intensities until volitional exhaustion. We scrutinised $\text{hydraulic}_\text{weig}$ $\dot{V}_{O_2}$ predictions on collected data from these TTE tests.

7.2.1 Data collection

Data collection was approved by the Human Research Ethics Committee at Western Sydney University (HREC Approval Number: H13975). Five recreationally active participants (4 males and 1 female, age 32 years $\pm$ 7.8 years, weight 73.6 kg $\pm$ 5.81 kg, $\dot{V}_{O_2\text{max}} = 3.59$ L/min $\pm$ 0.62 L/min) gave informed and written consent to participate and to have their anonymised data published. All participants were familiar with maximal exercise efforts. Exercise tests were conducted on an SRM - High Performance Ergometer (Jülich, Germany) in hyperbolic operation mode, which adjusts power dynamically to cadence changes to maintain a constant power output. Breath-by-breath $\dot{V}_{O_2}$ data were collected using the Quark CPET system by COSMED (Rome, Italy). The equipment was calibrated prior to each trial. Each of the 5 participants completed 6 exhaustive exercise trials. To ensure that participants were fully rested, they were asked to avoid strenuous exercise 24 h prior to the tests. Tests were scheduled more than 24 h apart, at roughly the same time of day. The tested population of five participants is small because the investigated $\dot{V}_{O_2}$ predictions of the $\text{hydraulic}_\text{weig}$ model showed a large error for 100% of the tests. This error is large enough to conclude that our model should not be used to obtain accurate $\dot{V}_{O_2}$ predictions without collecting further evidence.

7.2.1.1 Ramp test

All participants had to perform an initial ramp test to obtain the appropriate power settings for subsequent TTE tests. After a 3-min warm-up at 50 W, the power increased by 30 W per min for males and by 20 W per min for the female. Once power rose above 110 W, participants were instructed to maintain a self-chosen cadence between 80 RPM and 100 RPM. The point of volitional exhaustion was defined as the first time point when the cadence dropped by more than 10 % below the intended cadence for more than 3 s. This definition is similar to procedures by Caen et al. (2019, 2021); Sreedhara et al. (2020). The highest 10-s moving average power output achieved was defined as the peak power output (P_peak).

7.2.1.2 TTE tests

After the ramp test, each participant completed 5 constant power TTE trials at distinct powers in random order. The powers were set to 100 %, 92.5 %, 85 %, 80 %

and 77.5 % or 75 % of P$_{\text{peak}}$ to obtain a range of TTEs between 2 min and 12 min. Participants were blinded to their exercise power. Again, each test started with a 3-min warm-up at 50 W before power was set to the randomly chosen percentage of P$_{\text{peak}}$. During exercise, participants were asked to cycle at their self-chosen cadence from the ramp test. The point of volitional fatigue was defined as the first time point when cadence dropped by more than 10 % below the intended cadence for more than 3 s. Throughout all tests, breath-by-breath $\dot{V}_{\text{O}_2}$ data were collected. For each test, the highest achieved 30-s moving average of measured $\dot{V}_{\text{O}_2}$ was considered as the peak oxygen uptake ($\dot{V}_{\text{O}_2\text{peak}}$) of that test. TTEs, as well as the power-meter data of the SRM ergometer, were also recorded for later analysis.

7.2.2 Data analysis

Hydraulic$_{\text{weig}}$ and its evolutionary fitting process require CP and W' to make predictions. These parameters were obtained by fitting the critical power model to conducted TTE tests of a participant.

7.2.2.1 Fitting the models

To obtain CP and W' that best fitted a participant, as the Section 2.2.1 describes, three forms of the critical power model were fitted to conducted TTE tests. These forms were Equation (2.5) as the nonlinear power-time relationship, Equation (2.6) as the linear power-1/time relationship, and Equation (2.3) as the linear work-time relationship.

The goodness-of-fit of each of these equations was determined from the SEE associated with fitted CP and W'. The goodness-of-fit of a model was considered sufficient if SEE associated with CP was < 5 % of CP and the SEE associated with W' was < 10 % of W' (Jones et al., 2019; Caen et al., 2021). The best individual fit for a participant was selected by whatever model resulted in the smallest sum of the SEE associated with CP as % of CP, plus the SEE associated with W' as % of W' (Black et al., 2015; Jones et al., 2019; Caen et al., 2021).

Then, hydraulic$_{\text{weig}}$ was fitted to CP and W' derived from the best individual fit using the evolutionary algorithm outlined in Chapter 5. Considering the notation in Figure 4.1, fitting hydraulic$_{\text{weig}}$ meant finding a set of parameters for $[LF, LS, M_U, M_{LS}, M_{LF}, \theta, \gamma, \phi]$ that made the hydraulic model resemble expected exercise responses according to the critical power model for energy expenditure, and according to the published recovery ratios for energy recovery (Caen et al., 2019; Weigend et al., 2021). To determine these parameters for our participants, we used

the automatised evolutionary computation procedure of our `threecomphyd`[1] Python package, which was published in Weigend et al. (2021).

7.2.2.2 Oxygen uptake predictions

To assess the quality of predicted $\dot{V}_{O_2}$ kinetics of fitted hydraulic$_\text{weig}$ models, the predictions were compared with the collected breath-by-breath $\dot{V}_{O_2}$ data of the TTE tests. Specifically, we opened and closed the tap of hydraulic$_\text{weig}$ according to collected SRM power-meter data and recorded flow from the left tank as predicted $\dot{V}_{O_2}$.

   Our objective measure to assess the quality of $\dot{V}_{O_2}$ predictions was the difference between the time at which the simulated flow from the left tank (U) was predicted to reach its peak, and the time at which the observed breath-by-breath $\dot{V}_{O_2}$ data reached $\dot{V}_{O_2\text{peak}}$. As an additional visual comparison, we plotted normalised predicted flow from U together with normalised actual $\dot{V}_{O_2}$ dynamics. Predicted $\dot{V}_{O_2}$ dynamics (flow from U) were normalised with the maximal flow M_U. The 30-s averaged real $\dot{V}_{O_2}$ uptake measurements were normalised with the observed $\dot{V}_{O_2\text{peak}}$ of that test.

7.2.3 Results

This section presents the by Weigend et al. (2022b) reported model fittings and predictions with little comment. Prediction shortcomings and causes are discussed in detail in Section 7.3.

7.2.3.1 Ramp test results and model fittings

The average P_peak of the ramp tests of all participants was 327 W $\pm$ 52 W. The shortest TTE was excluded from the estimation of CP and W' for one participant because it was too short (113 s). For all participants, the linear power-1/time model resulted in the best individual fit and resulted in an averaged CP of 223 W $\pm$ 40 W and W' of 148 912 J $\pm$ 2 869 J. Individual critical power model fitting results and associated SEEs are summarised in Table 7.1.

   Using the notation of Figure 4.1, the fitted hydraulic$_\text{weig}$ models had an average LF of 14 330 J $\pm$ 2 463 J, LS of 38 575 J $\pm$ 6 605 J, M_U of 222 W $\pm$ 40 W, M_{LS} of 90 W $\pm$ 15 W, M_{LF} of 15 W $\pm$ 3 W, θ of 0.7 $\pm$ 0.05, γ of 0.02 $\pm$ 0.01, and ϕ of 0.26 $\pm$ 0.04. Individual results for these parameters are summarised in Table 7.2.

Table 7.1: An overview of critical power model fitting results for all participants. SEE% denotes the standard error associated with the parameter as a percentage of the parameter, e.g, SEE% of CP is the SEE associated with CP divided by CP.

participant	best fit model	CP		W'	
		W	SEE%	J	SEE%
1	linear power-1/time	211	3.1	13 240	8.7
2	linear power-1/time	292	2.4	19 143	8.0
3	linear power-1/time	238	0.7	10 820	3.6
4	linear power-1/time	199	1.8	16 790	6.3
5	linear power-1/time	174	1.5	14 469	3.8
avg±std		223 ± 40	1.9 ± 0.8	$14\,892 \pm 2\,869$	6.1 ± 2.1

Table 7.2: An overview of parameters of fitted hydraulic$_{\text{weig}}$ models.

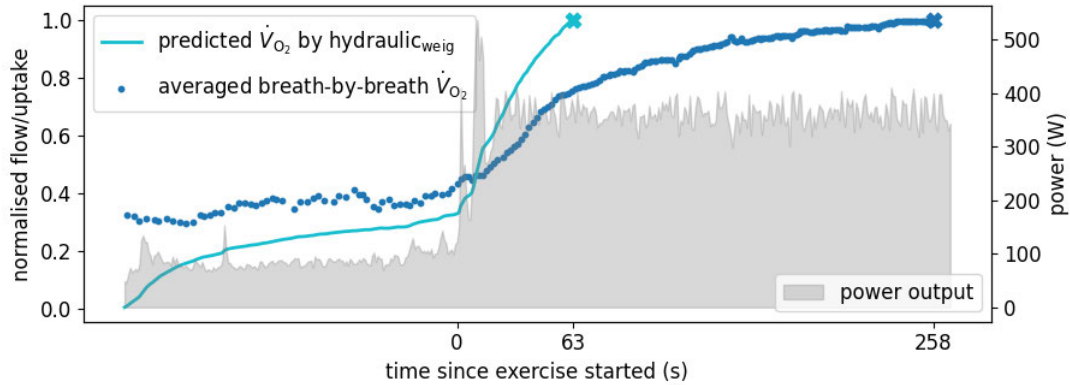| participant	LF(J)	LS(J)	M_U(W)	M_{LS}(W)	M_{LF}(W)	θ	γ	ϕ
1	12 562	36 679	210	82	14	0.71	0.02	0.27
2	18 245	50 537	291	117	20	0.7	0.02	0.27
3	11 914	38 269	238	71	11	0.79	0.02	0.2
4	16 196	37 141	198	95	15	0.68	0.01	0.27
5	12 733	30 250	173	87	14	0.64	0.02	0.31
avg ± std	$14\,330 \pm 2\,463$	$38\,575 \pm 6\,605$	222 ± 40	90 ± 15	15 ± 3	0.7 ± 0.05	0.02 ± 0.01	0.26 ± 0.04

Figure 7.2: Measured power-meter output and $\dot{V}_{O_2}$ of Participant 2 during warm-up and a TTE test at 364 W. Measurements before second 0 are part of the 3-min warm-up at 50 W. The plotted line displays predicted $\dot{V}_{O_2}$ uptake (flow from U) of a hydraulic$_{\text{weig}}$ model fitted to CP and W' of Participant 2. The averaged collected breath-by-breath $\dot{V}_{O_2}$ data is depicted as dots. The x-symbols mark the time points when hydraulic$_{\text{weig}}$ predicted $\dot{V}_{O_2\text{peak}}$ (63 s) and when collected averaged breath-by-breath reached $\dot{V}_{O_2\text{peak}}$ (258 s). The prediction error was 195 s.

7.2.3.2   Oxygen uptake predictions

Of all 25 constant power tests, the tests of Participant 4 at 288 W and Participant 5 at 255 W were excluded from the $\dot{V}_{O_2}$ analysis. In these two cases, unrealistic drops in $\dot{V}_{O_2}$ indicated that the breathing mask was not fixed tightly enough and leaked air when participants lowered their head too far. The unrealistic drops are clearly recognisable and depicted in Figures A.4 and A.5 in the Appendix.

The defined objective measure to assess the quality of $\dot{V}_{O_2}$ predictions was the time difference between predicted and observed $\dot{V}_{O_2peak}$. Furthermore, normalised flow from Ae and measured breath-by-breath data were plotted for visual comparison. An example is depicted in Figure 7.2. The power output measured by the SRM power-meter is plotted as the grey area in the background. It is overlaid with predicted and observed $\dot{V}_{O_2}$ kinetics. The time at which hydraulic$_{\text{weig}}$ predicted $\dot{V}_{O_2peak}$ was 63 s after the commencement of exercise. The time at which the actual $\dot{V}_{O_2peak}$ was observed was 258 s after the commencement of exercise. Therefore, the prediction error was 195 s. In addition, it is observable that the predicted $\dot{V}_{O_2}$ started at 0 and increased slowly during the warm-up.

The example Figure 7.2 is representative for all tests. As summarised in Table 7.3, on average, $\dot{V}_{O_2peak}$ was observed after $282\,\text{s} \pm 134\,\text{s}$ of exercise, while hydraulic$_{\text{weig}}$ predicted $\dot{V}_{O_2peak}$ after $65\,\text{s} \pm 24\,\text{s}$ of exercise in the respective test. In all the tests investigated, hydraulic$_{\text{weig}}$ predicted a much faster rise in $\dot{V}_{O_2}$ and a too early $\dot{V}_{O_2peak}$ with an average prediction error of $216\,\text{s} \pm 113\,\text{s}$. Table 7.3 summarises our results. The best hydraulic$_{\text{weig}}$ prediction was 67 s too early. The worst prediction was 461 s too early. The prediction error decreased as the exercise power increased.

7.3   Discussion

Previously, the theoretical hydraulic performance models of Morton (2006) and Sundström (2016) promised predictions for metabolic responses during exercise; for example, for lactic, alictic, and aerobic energy sources. But their models were not suitable for real-world applications because the required parameters to apply these models, such as, precise lactic energy capacities in joules, were impossible to obtain from individual athletes (Morton, 2006).

Hydraulic$_{\text{weig}}$ was designed to predict energy expenditure and recovery during intermittent exercise, but because its tanks and pipes resemble the M-M model so closely, it is tempting to interpret hydraulic$_{\text{weig}}$ in a metabolic context. With our study Weigend et al. (2022b), we confirmed, on data collected from 23 performance

---

[1]`https://github.com/faweigend/three_comp_hyd`

Table 7.3: The summary of hydraulic$_{\text{weig}}$ prediction errors for $\dot{V}_{\text{O}_2\text{peak}}$. The example from Figure 7.2 (Participant 2 with 364 W) is in row 8. The column 'observed' details the seconds it took from the onset of exercise to reach $\dot{V}_{\text{O}_2\text{peak}}$ (blue x-symbol at 258 s in Figure 7.2). The column 'predicted' details the seconds hydraulic$_{\text{weig}}$ predicted it would take (azure x-symbol at 63 s in Figure 7.2). The prediction error in the last column is the difference between these two times.

| | | time until $\dot{V}_{\text{O}_2\text{peak}}$ (s) | | |
participant	power (W)	observed	predicted	prediction error (s)
1	243	481	72	409
1	259	230	53	177
1	275	196	45	151
1	299	156	41	115
1	324	120	34	86
2	332	484	88	396
2	343	297	78	219
2	364	258	63	195
2	396	169	54	115
2	428	137	44	93
3	252	479	113	366
3	265	355	85	270
3	280	230	72	158
3	307	154	53	101
3	330	109	42	67
4	224	581	120	461
4	230	445	94	351
4	245	330	85	245
4	265	266	67	199
5	214	309	68	241
5	221	292	55	237
5	234	249	44	205
5	276	152	32	120
avg±std		282±134	65±24	216±113

tests of 5 participants, that hydraulic$_{\text{weig}}$ is not suitable for $\dot{V}_{\text{O}_2}$ predictions. The following discussion provides further information on the causes of poor metabolic predictions and highlights differences to the assumptions of the M-M model.

7.3.1 Predictions of the oxygen uptake slow component

From the onset of high-intensity exercise, hydraulic$_{\text{weig}}$ consistently predicted $\dot{V}_{\text{O}_2}$ to rise too quickly. As summarised in Table 7.3, $\dot{V}_{\text{O}_2}$ was predicted to reach its peak after an average of 65 s $\pm$ 24 s while observed kinetics were slower and took 282 s $\pm$ 134 s. The average difference between $\dot{V}_{\text{O}_2\text{peak}}$ predictions and observations was 216 s $\pm$ 113 s. Considering that TTE tests lasted between 2 min to 12 min, prediction errors of more than 3 min on average made clear that hydraulic$_{\text{weig}}$ could not predict realistic $\dot{V}_{\text{O}_2}$ kinetics.

Further, it is observable in Table 7.3 that the prediction error increases with decreasing power. Thus, the longer the exercise, the larger the error in predicted $\dot{V}_{\text{O}_2}$. These results are in contrast to remarks by Morton (2006) for $\dot{V}_{\text{O}_2}$ prediction capabilities of his M-M model. We elaborate the reasons for such poor predictions in the following section.

7.3.1.1 M-M model

Tank positions and sizes of the M-M model are determined by the values θ, γ and ϕ. As summarised in Section 2.4.2, Morton (1990) developed several constraints on these parameters to determine a realistic arrangement of tanks. From his developed constraints, Morton (1990) argued that the only realistic configuration of the three component hydraulic model is the one depicted in Figure 2.9, which corresponds to configuration type M in Section 4.4.

In his review, Morton (2006) highlighted how the M-M model predicts the $\dot{V}_{\text{O}_2}$ slow component phenomenon. Barstow and Mole (1991) empirically showed that $\dot{V}_{\text{O}_2}$ uptake quickly reaches a steady state at a constant exercise intensity below the moderate-heavy boundary (LAT). However, at exercise above LAT, an initial rapid increase in $\dot{V}_{\text{O}_2}$ uptake is followed by a slower continuous rise. This slower rise is called the $\dot{V}_{\text{O}_2}$ slow component.

As observable in Figure 7.3, the slow component is well captured by a hydraulic model that is configured according to type M and satisfies the constraints on the M-M model by Morton (1990). Depicted on the right in Figure 7.3 are the predicted $\dot{V}_{\text{O}_2}$ dynamics (flow from U) during constant high-intensity exercise. With all tanks filled at the beginning, the dynamics play out as follows: during the warm-up, the tap is not opened wide and thus liquid level in the middle tank drops slowly and flow from U increases slowly. Then exercise starts after one min, the fill level in
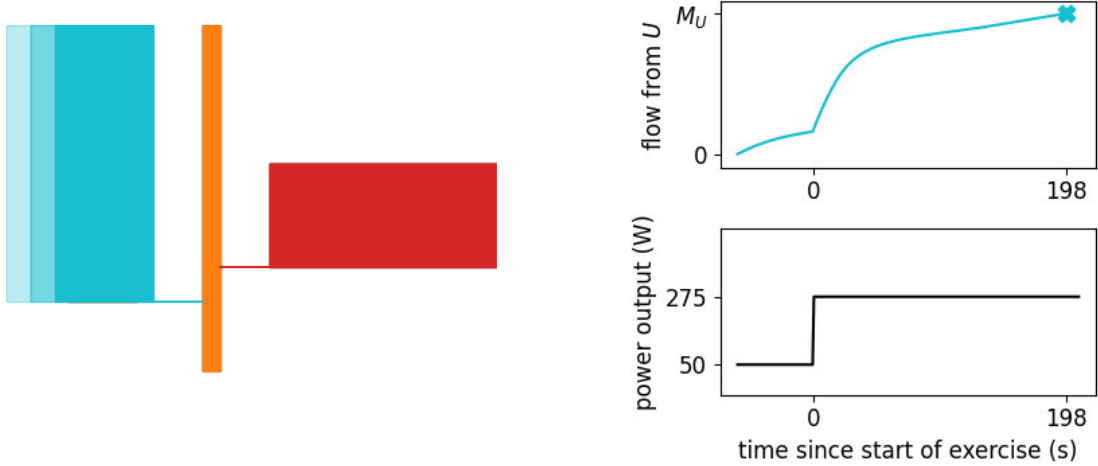
Figure 7.3: **Left:** A hydraulic$_{\text{weig}}$ model configured in a way that follows the physiological constraints on the M-M model discussed in Section 7.3.1.1. **Right:** The top plot displays flow from the left tank (U) into the middle tank (LF). The displayed model simulates constant intensity exercise at 275 W with a warm-up at 50 W until maximal flow through U is reached (x-symbol). We interpreted this flow as $\dot{V}_{O_2}$ uptake predictions. The $\dot{V}_{O_2}$ slow component is observable in predicted $\dot{V}_{O_2}$ dynamics.

the middle tank drops quickly and therefore flow from U increases quickly. The exercise intensity is high enough so that the dropping fill level of the middle tank reaches the top of the right tank, and the right tank also starts to contribute. This additional flow makes the fill level of the middle tank drop slower and therefore, flow from U increases slower from this point. Morton (2006) state that the way in which the M-M model simulates the $\dot{V}_{O_2}$ slow component can be compared with the mathematical formulation by Barstow and Mole (1991).

7.3.1.2 Hydraulic$_{\text{weig}}$ model

Because of the evolutionary fitting procedure that adjusts hydraulic$_{\text{weig}}$ model parameters freely, it is not guaranteed that hydraulic$_{\text{weig}}$ conforms to the constraints on the M-M model. Fittings can be of all of the 20 configuration types in Section 4.4. As reported in Section 7.2.3, the average parameter γ of the fitted hydraulic$_{\text{weig}}$ was 0.017 ± 0.005 and the average ϕ was 0.263 ± 0.038. These values indicate that none of the hydraulic$_{\text{weig}}$ models adhered the constraint of Morton (1990) that $\gamma > \phi$. This had a direct effect on $\dot{V}_{O_2}$ predictions of hydraulic$_{\text{weig}}$.

As an example, on the left in Figure 7.4 is the system of tanks fitted to Participant 1. Here, the parameters are $\gamma = 0.02$ and $\phi = 0.27$, so $\gamma < \phi$. In addition, the top of the right tank is much lower than the tank of the model depicted in Figure 7.3. The second constraint of Morton (1990) suggests that the top of the
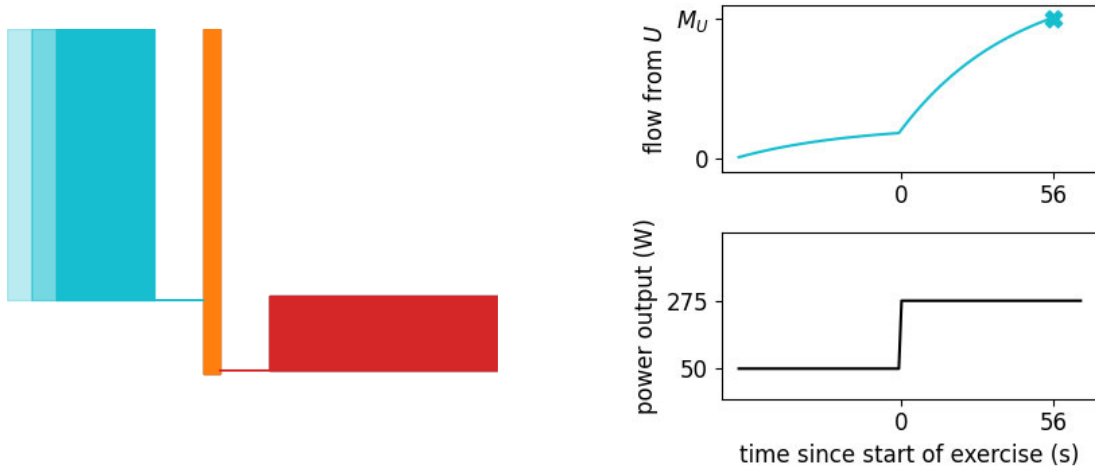
Figure 7.4: **Left:** A three component hydraulic model fitted to CP 211 W and W' 13 240 J of Participant 1. Values of all fitted parameters are in Table 7.2 in the Appendix. **Right:** Flow from U when the left model simulates a warm-up at 50 W and then a constant high-intensity exercise at 275 W until maximal flow through U is reached (x-symbol).

right tank has to be at approximately 40 % of the height of the left tank, which is also not satisfied with the fitted model in Figure 7.4.

As a result, $\dot{V}_{O_2}$ predictions of the model fitted to Participant 1 did not resemble the $\dot{V}_{O_2}$ slow component. The onset of flow from the right tank had almost no effect on the exponential increase of flow from the left tank. This explains the prediction errors in Table 7.3 and confirms on the example of $\dot{V}_{O_2}$ predictions that the hydraulic$_{weig}$ model allows for unrealistic predictions for metabolic responses during exercise.

7.3.2 Warm-up oxygen uptake predictions

Another example of unrealistic $\dot{V}_{O_2}$ predictions occurs at the beginning of exercise during the warm-up. Because the middle tank of hydraulic$_{weig}$ is filled initially, it will always predict the complete absence of $\dot{V}_{O_2}$ at the first time step of a simulation. As observable in Figures 7.2 to 7.4, this caused unrealistic $\dot{V}_{O_2}$ predictions at the beginning of exercise tests. Predicted $\dot{V}_{O_2}$ started at 0 and increased slowly throughout the 3 min of warm-up at 50 W. The size of the middle tank was fixed for the M-M and hydraulic$_{weig}$ models. Therefore, if the flow from U is interpreted as $\dot{V}_{O_2}$, these models are guaranteed to predict the absence of $\dot{V}_{O_2}$ at the beginning of exercise and when the athlete is fully recovered.

Lidar et al. (2021) acknowledged this issue and introduced a new parameter to reduce the size of the middle tank. Using the notation in Figure 4.1, Lidar et al. (2021) made the top of LF adjustable and introduced a parameter ψ, which defines

the distance between the top of LF and the top of U. The closer ψ to $1 - \phi$, the larger the possible flow from U, even if LF is filled. This additional parameter is not present in the M-M or hydraulic$_{\text{weig}}$ models and is another argument for why hydraulic$_{\text{weig}}$ cannot make realistic predictions for $\dot{V}_{\text{O}_2}$.

7.4   Summary

The findings presented in this chapter confirm that the M-M model and hydraulic$_{\text{weig}}$ have to be seen as distinct models for distinct purposes. The M-M model was designed to predict metabolic processes during exercise. Hydraulic$_{\text{weig}}$ failed to make realistic metabolic predictions by deviating considerably from $\dot{V}_{\text{O}_2}$ observations of all investigated 23 performance tests of 5 participants (Weigend et al., 2022b). Furthermore, the fitted hydraulic$_{\text{weig}}$ model configurations did not adhere to the physiological constraints of Morton (1990) for the M-M model.

This chapter addressed the third research question from Section 3.2 and concluded that hydraulic$_{\text{weig}}$ cannot predict metabolic responses. Instead, as shown in the previous Chapter 6, hydraulic$_{\text{weig}}$ was optimised such that it addresses shortcomings of established work-balance models for energy recovery predictions during intermittent exercise. In conclusion, the findings of this chapter and the previous Chapter 6 suggest that hydraulic models cannot be optimised to predict energy recovery and metabolic processes at the same time. Currently, hydraulic$_{\text{weig}}$ can only be recommended for predicting energy recovery during intermittent exercise.

Many opportunities remain for future work to improve the hydraulic$_{\text{weig}}$ or the M-M model. These are outlined, together with the final conclusion of this work, in the next chapter.

Chapter 8

Conclusions

This work developed and investigated the hydraulic$_{\text{weig}}$ model, which represents a new approach for how hydraulic performance models can be used. The code and collected data are publicly available in the repositories `pypermod`[1] and `threecomphyd`[2].

The established work-balance models predict energy recovery during intermittent exercise from CP and W' of an athlete (Skiba and Clarke, 2021). Physiologically, CP is defined as the threshold between heavy- and severe-intensity exercise (Jones et al., 2019). The capacity W' limits the time an athlete can exercise at a severe intensity above CP. In recent literature, work-balance models have been found to overly simplify recovery dynamics (Caen et al., 2019; Skiba and Clarke, 2021).

We designed hydraulic$_{\text{weig}}$ to predict energy recovery during intermittent exercise and to address the shortcomings of established work-balance models. Chapter 5 defined a pathway to obtain hydraulic$_{\text{weig}}$ predictions from CP and W'. Chapter 6 retrospectively compared energy recovery predictions of work-balance and hydraulic$_{\text{weig}}$ models on published data from five studies. The prediction capabilities of hydraulic$_{\text{weig}}$ outperformed established work-balance models on all metrics and rendered it a strong direction for future research in performance modelling. Chapter 7 highlighted the differences of hydraulic$_{\text{weig}}$ to the original M-M model and confirmed that it cannot be used to predict metabolic responses during exercise.

In the following sections, the opportunities and challenges for future research are outlined. Then, this work concludes with the significance and implications of the presented findings.

8.1   Future Work

In the following section, opportunities for future research directions are divided into three categories and related to the findings of the previous chapters.

---

[1] https://github.com/faweigend/pypermod
[2] https://github.com/faweigend/three_comp_hyd

8.1.1 More comparable studies

Because $W'_{\text{bal-ode}}$ models and hydraulic$_{\text{weig}}$ are intended to be applied to a wide range of athletes and conditions, it is important to scrutinise model predictions on various data sets that cover a range participants and test conditions. The various formats of observations extracted from previous studies in Chapter 6 highlight the need for more and more comparable studies on energy recovery dynamics.

In Chapter 6 we combined extracted recovery observations, which came with associated uncertainties that could not be considered in MAE, RMSE and AIC$_{\text{c}}$ scores. As an example, the standard deviations of observed recovery ratios by Caen et al. (2021) depicted in Figure 6.2 were greater than reported standard deviations by Ferguson et al. (2010) depicted in Figure 6.4.

We could not incorporate these standard deviations into goodness-of-fit metrics because of how different recovery ratios were reported. Caen et al. (2021) and Ferguson et al. (2010) reported averaged observed recovery ratios with standard deviations, for the Bartram data set we had to use W'_{bart} predictions as observations for comparison. In Weigend et al. (2021), we derived our values from Caen et al. (2019) without standard deviations, and for the Chidnok data set we fitted constant values for $\mathcal{T}_t$ for $W'_{\text{bal-ode}}$ models to their reported times to exhaustion to obtain comparable recovery ratios in percent.

Larger data sets are vital for more educated investigations of recovery models and their improvement in future work. We see the combination of data sets in Chapter 6 as a step towards this direction. In order to improve and compare models more holistically, it is important that more comparable studies are conducted in the future and combined into a larger test bed for performance models.

8.1.2 Fitting procedure

Chapter 6 observed the improved prediction quality of hydraulic$_{\text{weig}}$, but this came at the cost of a time-demanding fitting process. The evolutionary fitting procedure outlined in Chapter 5 requires CP and W' of an athlete as inputs and then obtains fitted hydraulic$_{\text{weig}}$ configurations using an evolutionary computation approach. Different CP and W' values require a new hydraulic$_{\text{weig}}$ configuration to be fitted. Additionally, as reported in Section 5.2.5, fittings require computation times of 5 hours or more on 7 cores of an Intel$^{\circledR}$ Xeon$^{\circledR}$ CPU E5-2650 v4 @ 2.20GHz each.

On the contrary, obtaining $\mathcal{T}_t$ for work-balance models can be solved in milliseconds and they can be applied to any CP and W' combination without fitting $\mathcal{T}_t$ anew. Therefore, the application of hydraulic$_{\text{weig}}$ is a time consuming task in comparison with the application of work-balance models. To improve the feasibility of

application, it will be worthwhile to optimise the hydraulic model fitting process in future work.

8.1.3 Hydraulic$_{\text{weig}}$ predictions

Although Chapter 6 showed that hydraulic$_{\text{weig}}$ achieved overall promising accuracy in energy recovery predictions and could address shortcomings of work-balance models, several prediction dynamics require further investigation.

The comparison on the $D_{\text{CP}}0$ case of the Bartram data set (see Figure 6.1) revealed that hydraulic$_{\text{weig}}$ predicts a slight recovery during exercise at CP. No liquid from U flowed back into the system but the ongoing flow from LS to LF still caused the fill level of LF to rise during the recovery bout. Recovery while exercising at CP intensity is a controversial assumption that is made to an even stronger extent by the original work-balance model of Skiba et al. (2012). Such dynamics have to be considered when the models are used for predictions and are an important direction for future investigation.

Additionally, the analysis of fitting results in Section 5.3 reported that hydraulic$_{\text{weig}}$ consistently underpredicted TTEs at high power outputs by up to 20 s. These shorter predictions are part of the in Section 5.2.2 outlined necessary trade-off between energy expenditure and recovery prediction accuracy. Adjustments to hydraulic$_{\text{weig}}$ may allow a more favourable trade-off. For example, as discussed in Section 7.3.2, Lidar et al. (2021) introduced an additional parameter ψ to their version of the M-M model, which, when introduced to the hydraulic$_{\text{weig}}$ model, could improve short TTE predictions in future work.

8.2 Significance and Implications

The current focus and strength of hydraulic$_{\text{weig}}$ are energy recovery predictions during intermittent exercise, where promising new investigation opportunities emerge. Platforms such as Strava[3] or Golden Cheetah[4] provide constantly growing databases of real-world intermittent exercise training and competition data. With affordable power metres, smartwatches, and online cycling apps such as Zwift[5], the interest in intermittent exercise performance models and the amount of available data will grow.

Therefore, the future holds exciting possibilities for new iterations of work-balance or hydraulic models. The earlier we begin to develop pathways to investigate

---

[3] https://www.strava.com
[4] https://www.goldencheetah.org
[5] https://www.zwift.com

these models on such real-world data, the better. Hydraulic$_\text{weig}$ is one model iteration of this pursuit. It is applicable to limited real-world data and promises to enable athletes to predict their performance capabilities more accurately, such that they can make more educated decisions about pacing strategies and training goals.

8.3 Roundup

Depending on focus and setup, different available performance models become more or less suitable for the analysis. Automated solutions like Automated Machine Learning (AutoML) can help users to select models, which make the most accurate predictions for their data (He et al., 2021). Next to the comparison of prediction accuracy for energy recovery during intermittent exercise, we also scrutinise the advantages and limitations of hydraulic$_\text{weig}$ regarding the physiological interpretations of its parameters.

As previously reported by Skiba and Clarke (2021); Morton (2006), and also as discussed in this work, there is no one-size-fits-all performance model. Figure 8.1 depicts a simplified spectrum, where simple and applicable performance models are on the left and complex and theoretical models are on the right, one could say that the critical power model is on the far left, because it has only two parameters and requires just a few performance tests to be applied. Then, the hydraulic models of Morton (2006); Lidar et al. (2021) and Sundström (2016) would be on the right because they have eight or more parameters and are cumbersome to apply to athletes due to their required in-depth knowledge about metabolic and bioenergetic capacities.

On this spectrum, the hydraulic$_\text{weig}$ model bridges the gap and remains somewhere in the middle. It can be applied to CP and W' and outperforms $W'_\text{bal-ode}$ models in intermittent exercise predictions. However, it should not be used for pre-
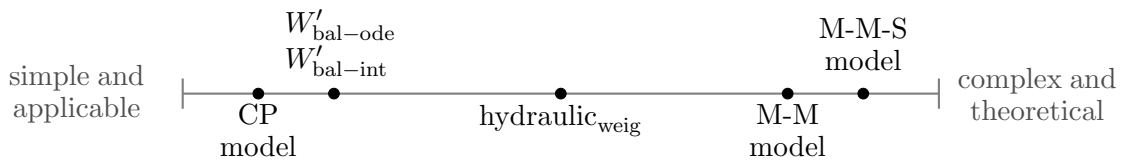

Figure 8.1: A simplified spectrum where simple and applicable models are on the left and complex and theoretical models are on the right. The critical power model (CP model) is on the far left. Work-balance models by Skiba and Clarke (2021) include energy recovery predictions and therefore are slightly further towards the right ($W'_\text{bal-ode}$ and $W'_\text{bal-int}$). The hydraulic models by Morton (2006) and Sundström (2016) are on the far right (M-M model and M-M-S model) because they predict metabolic processes. Our hydraulic$_\text{weig}$ model fits in the middle.

dictions of $\dot{V}_{O_2}$ or alactic or lactic energy sources as the original hydraulic models were intended to be used.

We strive to progress performance model development and to help users make informed decisions for their analysis. As summarised in this chapter, there are many directions for creating new models and improving existing ones. To support this pursuit, we embedded all our data, the code of compared models and hydraulic model advances in the open-source python packages `pypermod`[6] and `threecomphyd`[7].

Appendix A

Supplementary Information

A.1 Label update

When the hydraulic$_{\mathrm{weig}}$ model was first proposed in Weigend et al. (2021), we chose to name the left infinitely big tank the aerobic energy source (Ae) and the two limited tanks the anaerobic fast (AnF) and anaerobic slow energy source (AnS). These names have a bioenergetic connotation and imply that the flow from the aerobic source still represents $\dot{V}_{\mathrm{O}_2}$ uptake as it did in the M-M model and the model of Margaria (1976). The results of Chapter 7 show that hydraulic$_{\mathrm{weig}}$ does not predict realistic $\dot{V}_{\mathrm{O}_2}$ dynamics and should not be used for $\dot{V}_{\mathrm{O}_2}$ predictions.

Therefore, it was sensible to change the labels of hydraulic$_{\mathrm{weig}}$ to more abstract names. The new labels are those depicted in Figure 4.1. Instead of Ae, we re-labelled the left tank as an unlimited energy source (U). Instead of AnF, the middle tank is now a limited fast energy source (LF), and instead of AnS, the right tank is a limited slow energy source (LS). These changes do not apply to any other hydraulic models and only affect hydraulic$_{\mathrm{weig}}$. They are an important step to protect users from misinterpreting fitted hydraulic$_{\mathrm{weig}}$ models and their prediction results.

A.2 Convergence plots

This section presents additional convergence plots of an example run of our evolutionary fitting algorithm from Chapter 5. With reference to the hyperparemters discussed in Chapter 5, we set the parameters to 30 generations, 40 cycles, 64 population size and 7 islands. For each of the 40 cycles Figure A.1 shows the minimal distance to (0,0), i.e., the most accurate solution. We can observe that the accuracy of the best solution improved rapidly over the first 5 cycles and then levelled off. The last five cycles (34-39) do not show any improvement and it is reasonable to state that the algorithm converged to an optimum.
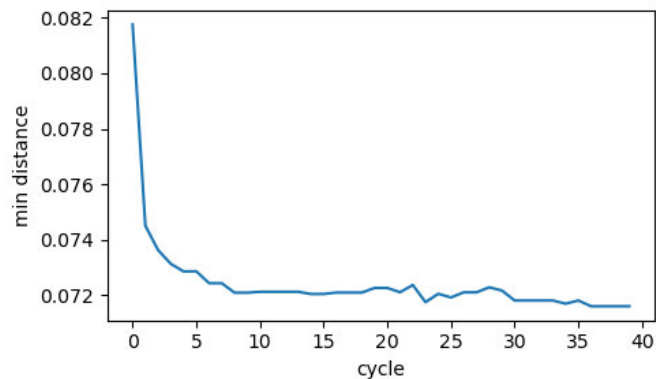
Figure A.1: The accuracy of the evolving best solutions during an example fitting with our evolutionary algorithm from Chapter 5. The minimal distance is our established accuracy measure from Section 5.2.2. It improves quickly during the first cycles and then levels off as the algorithm approaches an optimum.

The evolving Pareto fronts are depicted in Figures A.2 and A.3. The plots are in direct reference to Figure 5.1 and we refer to Section 5.2.2 for more explanations on the error measurements along the two axis. Figure A.2 shows the Pareto fronts of all 7 islands and their most accurate solutions when the fitting algorithm completed the first cycle and the final results after 40 cycles are shown in Figure A.3. We observe that all islands converged to a very similar optimal solution.
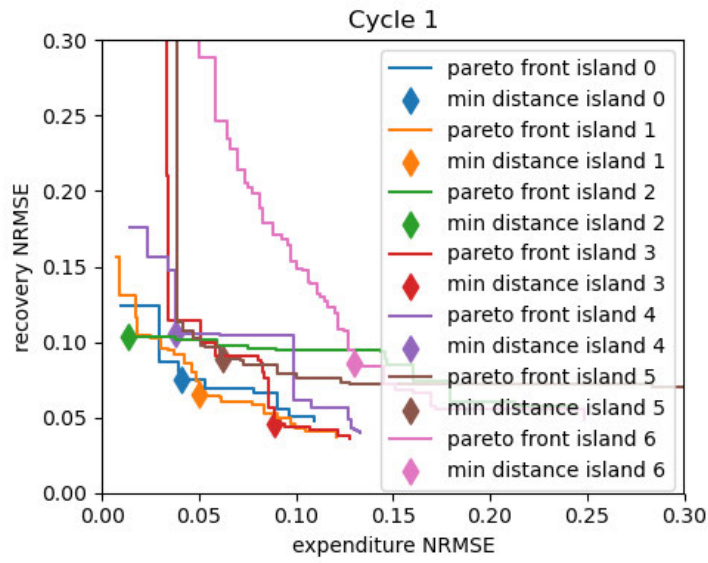
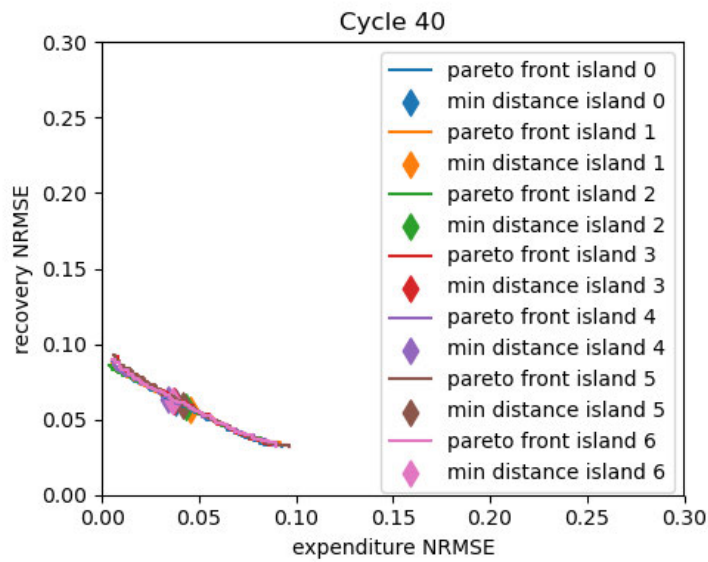Figure A.2: Pareto fronts of the evolutionary fitting after the first cycle.

Figure A.3: Pareto fronts of the evolutionary fitting after after 40 cycles. All islands converged to a very similar optimum.
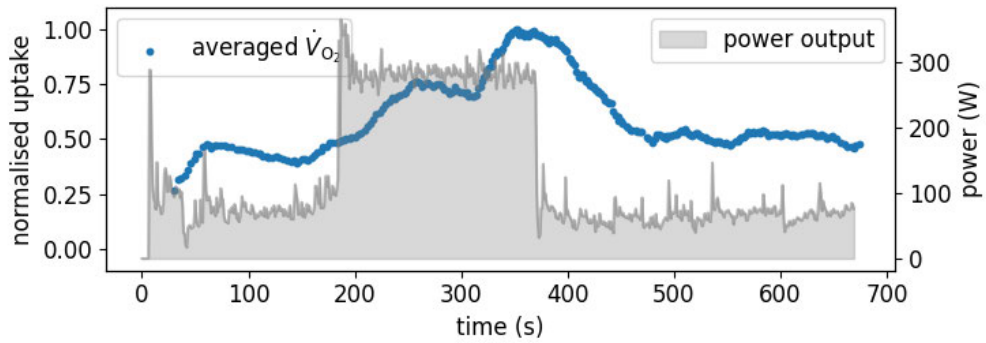
A.3 Oxygen uptake study

Figure A.4: Measured power output (grey area) and $\dot{V}_{O_2}$ uptake (dots) of Participant 4 during a constant intensity exercise trial at 288 W. $\dot{V}_{O_2}$ drops occurred because the face mask leaked air when the athlete lowered their head too far.
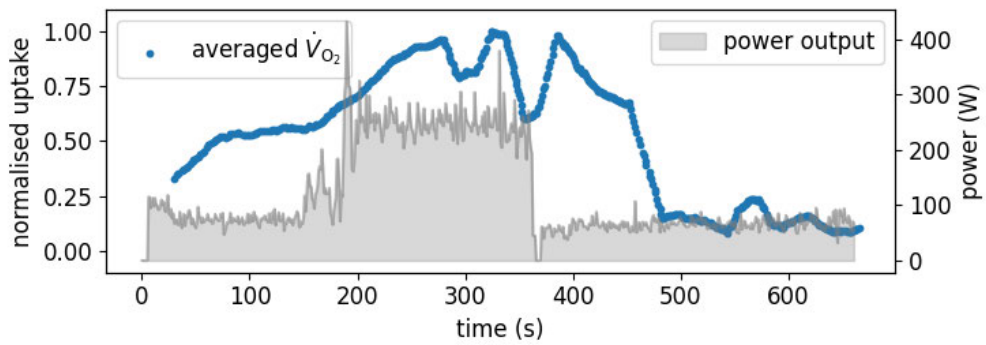

Figure A.5: Measured power output (grey area) and $\dot{V}_{O_2}$ uptake (dots) of Participant 5 during a constant intensity exercise trial at 255 W. $\dot{V}_{O_2}$ drops occurred because the face mask leaked air when the athlete lowered their head too far.

Bibliography

Barstow, T. J. and Mole, P. A. (1991). Linear and nonlinear characteristics of oxygen uptake kinetics during heavy exercise. *Journal of Applied Physiology*, 71(6):2099–2106.

Bartram, J. C., Thewlis, D., Martin, D. T., and Norton, K. I. (2018). Accuracy of W' recovery kinetics in high performance cyclists - modeling intermittent work capacity. *International Journal of Sports Physiology and Performance*, 13(6):724–728.

Beaver, W. L., Wasserman, K., and Whipp, B. J. (1986). A new method for detecting anaerobic threshold by gas exchange. *Journal of Applied Physiology*, 60(6):2020–2027.

Behncke, H. (1993). A mathematical model for the force and energetics in competitive running. *Journal of Mathematical Biology*, 31(8):853–878.

Behncke, H. (1997). Optimization models for the force and energy in competitive running. *Journal of Mathematical Biology*, 35(4):375–390.

Binder, R. K., Wonisch, M., Corra, U., Cohen-Solal, A., Vanhees, L., Saner, H., and Schmid, J.-P. (2008). Methodological approach to the first and second lactate threshold in incremental cardiopulmonary exercise testing. *European Journal of Cardiovascular Prevention & Rehabilitation*, 15(6):726–734.

Biscani, F. and Izzo, D. (2020). A parallel global multiobjective framework for optimization: pagmo. *Journal of Open Source Software*, 5(53):2338.

Black, M. I., Jones, A. M., Bailey, S. J., and Vanhatalo, A. (2015). Self-pacing increases critical power and improves performance during severe-intensity exercise. *Applied Physiology, Nutrition, and Metabolism*, 40(7):662–670.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2):261–304.

Bäck, T. and Schwefel, H.-P. (1993). An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, 1(1):1–23.

Caen, K., Bourgois, G., Dauwe, C., Blancquaert, L., Vermeire, K., Lievens, E., Van Dorpe, J., Derave, W., Bourgois, J. G., Pringels, L., and Boone, J. (2021). W' recovery kinetics after exhaustion: a two-phase exponential process influenced by aerobic fitness. *Medicine & Science in Sports & Exercise*, 53(9):1911–1921.

Caen, K., Bourgois, J. G., Bourgois, G., Van Der Stede, T., Vermeire, K., and Boone, J. (2019). The reconstitution of W' depends on both work and recovery characteristics. *Medicine & Science in Sports & Exercise*, 51(8):1745–1751.

Chai, T. and Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – arguments against avoiding rmse in the literature. *Geoscientific Model Development*, 7(3):1247–1250.

Chidnok, W., Dimenna, F. J., Bailey, S. J., Vanhatalo, A., Morton, R. H., Wilkerson, D. P., and Jones, A. M. (2012). Exercise tolerance in intermittent cycling: application of the critical power concept. *Medicine & Science in Sports & Exercise*, 44(5):966–976.

Chorley, A. and Lamb, K. L. (2020). The application of critical power, the work capacity above critical power (W'), and its reconstitution: a narrative review of current evidence and implications for cycling training prescription. *Sports*, 8(9):123.

Clarke, D. C. and Skiba, P. F. (2013). Rationale and resources for teaching the mathematical modeling of athletic training and performance. *Advances in Physiology Education*, 37(2):134–152.

de Jong, J., Fokkink, R., Olsder, G. J., and Schwab, A. (2017). The individual time trial as an optimal control problem. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*, 231(3):200–206.

Dekerle, J., Brickley, G., Hammond, A. J. P., Pringle, J. S. M., and Carter, H. (2006). Validity of the two-parameter model in estimating the anaerobic work capacity. *European Journal of Applied Physiology*, 96(3):257–264.

Drake, J., Finke, A., and Ferguson, R. (2022). Modelling human endurance: Power laws vs critical power. preprint, Physiology.

Efron, B. and Tibshirani, R. (1993). *An introduction to the bootstrap*. Number 57 in Monographs on statistics and applied probability. Chapman & Hall, New York.

Eiben, A. and Smith, J. (2015). *Introduction to evolutionary computing*. Natural Computing Series. Springer Berlin Heidelberg, Berlin, Heidelberg.

Ferguson, C., Rossiter, H. B., Whipp, B. J., Cathcart, A. J., Murgatroyd, S. R., and Ward, S. A. (2010). Effect of recovery duration from prior exhaustive exercise on the parameters of the power-duration relationship. *Journal of Applied Physiology*, 108(4):866–874.

Gastin, P. B., Costill, D. L., Lawson, D. L., Krzeminski, K., and McConell, G. K. (1995). Accumulated oxygen deficit during supramaximal all-out and constant intensity exercise. *The American College of Sports Medicine*, 27(2):255–263.

Good, P. I. (2000). *Permutation tests: a practical guide to resampling methods for testing hypotheses.* Springer, New York. OCLC: 681912126.

Gorostiaga, E. M., Garcia-Tabar, I., and Sánchez-Medina, L. (2022). Critical power: Over 95 years of" evidence" and" evolution". *Scandinavian journal of medicine & science in sports*, 32(6):1069–1071.

He, X., Zhao, K., and Chu, X. (2021). AutoML: A survey of the state-of-the-art. *Knowledge-Based Systems*, 212:106622.

Hill, A. V. (1925). The physiological basis of athletic records. *Nature*, 116:544–548.

Hill, D. W. (1993). The critical power concept: a review. *Sports Medicine*, 16(4):237–254.

Hoogkamer, W., Snyder, K. L., and Arellano, C. J. (2018). Modeling the benefits of cooperative drafting: is there an optimal strategy to facilitate a sub-2-hour marathon performance? *Sports Medicine*, 48(12):2859–2867.

Housh, D. J., Housh, T. J., and Bauge, S. M. (1989). The accuracy of the critical power test for predicting time to exhaustion during cycle ergometry. *Ergonomics*, 32(8):997–1004.

Hunter, S. K., Duchateau, J., and Enoka, R. M. (2004). Muscle Fatigue and the Mechanisms of Task Failure:. *Exercise and Sport Sciences Reviews*, 32(2):44–49.

Jones, A. M., Burnley, M., Black, M. I., Poole, D. C., and Vanhatalo, A. (2019). The maximal metabolic steady state: redefining the 'gold standard'. *Physiological Reports*, 7(10):e14098.

Jones, A. M. and Vanhatalo, A. (2017). The 'critical power' concept: applications to sports performance with a focus on intermittent high-intensity exercise. *Sports Medicine*, 47(S1):65–78.

Jones, A. M., Vanhatalo, A., Burnley, M., Morton, R. H., and Poole, D. C. (2010). Critical power: implications for determination of VO2max and exercise tolerance:. *Medicine & Science in Sports & Exercise*, 42(10):1876–1890.

Karatzaferi, C., de Haan, A., Ferguson, R., van Mechelen, W., and Sargeant, A. (2001). Phosphocreatine and ATP content in human single muscle fibres before and after maximum dynamic exercise. *Pflügers Archiv*, 442(3):467–474.

Lidar, J., Andersson, E. P., and Sundström, D. (2021). Validity and reliability of hydraulic-analogy bioenergetic models in sprint roller skiing. *Frontiers in Physiology*, 12:726414.

Margaria, R. (1976). *Biomechanics and energetics of muscular exercise.* Oxford University Press, Oxford University Press, Walton Street, Oxford, OX2 6DP.

Mattioni Maturana, F., Fontana, F. Y., Pogliaghi, S., Passfield, L., and Murias, J. M. (2018). Critical power: how different protocols and models affect its determination. *Journal of Science and Medicine in Sport*, 21(7):742–747.

McCrary, J. M., Ackermann, B. J., and Halaki, M. (2018). EMG amplitude, fatigue threshold, and time to task failure: A meta-analysis. *Journal of Science and Medicine in Sport*, 21(7):736–741.

Monod, H. and Scherrer, J. (1965). The work capacity of a synergetic muscular group. *Ergonomics*, 8(3):329–338.

Moritani, T., Nagata, A., Devries, H. A., and Muro, M. (1981). Critical power as a measure of physical work capacity and anaerobic threshold. *Ergonomics*, 24(5):339–350.

Morton, R. H. (1986a). On a model of human bioenergetics II: maximal power and endurance. *European Journal of Applied Physiology and Occupational Physiology*, 55(4):413–418.

Morton, R. H. (1986b). A three component model of human bioenergetics. *Journal of Mathematical Biology*, 24(4):451–466.

Morton, R. H. (1990). Modelling human power and endurance. *Journal of Mathematical Biology*, 28(1):49–64.

Morton, R. H. (1996). A 3-parameter critical power model. *Ergonomics*, 39(4):611–619.

Morton, R. H. (2006). The critical power and related whole-body bioenergetic models. *European Journal of Applied Physiology*, 96(4):339–354.

Morton, R. H. and Billat, L. V. (2004). The critical power model for intermittent exercise. *European Journal of Applied Physiology*, 91(2-3):303–307.

Muniz-Pumares, D., Karsten, B., Triska, C., and Glaister, M. (2019). Methodological approaches and related challenges associated with the determination of critical power and curvature constant. *Journal of Strength and Conditioning Research*, 33(2):584–596.

Noordhof, D. A., Skiba, P. F., and de Koning, J. J. (2013). Determining anaerobic capacity in sporting activities. *International Journal of Sports Physiology and Performance*, 8(5):475–482.

Peronnet, F. and Thibault, G. (1989). Mathematical analysis of running performance and world running records. *Journal of Applied Physiology*, 67(1):453–465.

Poole, D. C., Burnley, M., Vanhatalo, A., Rossiter, H. B., and Jones, A. M. (2016). Critical power: an important fatigue threshold in exercise physiology. *Medicine & Science in Sports & Exercise*, 48(11):2320–2334.

Qingfu Zhang and Hui Li (2007). MOEA/D: a multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 11(6):712–731.

Riegel, P. S. (1977). Time predicting. *Runner's World Magazine*, 12(8).

Riegel, P. S. (1981). Athletic Records and Human Endurance: A time-vs.-distance equation describing world-record performances may be used to compare the relative endurance capabilities of various groups of people. *American Scientist*, 69(3):285–290.

SciPy 1.0 Contributors, Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, i., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., and van Mulbregt, P. (2020). Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature Methods*, 17(3):261–272.

Skiba, P. F., Chidnok, W., Vanhatalo, A., and Jones, A. M. (2012). Modeling the expenditure and reconstitution of work capacity above critical power. *Medicine & Science in Sports & Exercise*, 44(8):1526–1532.

Skiba, P. F. and Clarke, D. C. (2021). The W' balance model: mathematical and methodological considerations. *International Journal of Sports Physiology and Performance*, 16(11):1561–1572.

Skiba, P. F., Clarke, D. C., Vanhatalo, A., and Jones, A. M. (2014). Validation of a novel intermittent W' model for cycling using field data. *International Journal of Sports Physiology and Performance*, 9(6):900–904.

Skiba, P. F., Fulford, J., Clarke, D. C., Vanhatalo, A., and Jones, A. M. (2015). Intramuscular determinants of the ability to recover work capacity above critical power. *European Journal of Applied Physiology*, 115(4):703–713.

Sreedhara, V. S. M., Ashtiani, F., Mocko, G. M., Vahidi, A., and Hutchison, R. E. (2020). Modeling the recovery of W' in the moderate to heavy exercise intensity domain. *Medicine & Science in Sports & Exercise*, Publish Ahead of Print.

Sreedhara, V. S. M., Mocko, G. M., and Hutchison, R. E. (2019). A survey of mathematical models of human performance using power and energy. *Sports Medicine - Open*, 5(1):54.

Sugiura, N. (1978). Further analysts of the data by akaike's information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, 7(1):13–26.

Sundström, D. (2016). On a bioenergetic four-compartment model for human exercise. *Sports Engineering*, 19(4):251–263.

Sundström, D., Bäckström, M., Carlsson, P., and Tinnsten, M. (2015). A four compartment model on human exercise bioenergetics. *Procedia Engineering*, 112:4–9.

Sundström, D., Carlsson, P., and Tinnsten, M. (2014). Comparing bioenergetic models for the optimisation of pacing strategy in road cycling. *Sports Engineering*, 17(4):207–215.

Vanhatalo, A., Doust, J. H., and Burnley, M. (2007). Determination of critical power using a 3-min all-out cycling test. *Medicine & Science in Sports & Exercise*, 39(3):548–555.

Vanhatalo, A., Jones, A. M., and Burnley, M. (2011). Application of critical power in sport. *International Journal of Sports Physiology and Performance*, 6(1):128–136.

Vinetti, G., Taboni, A., Bruseghini, P., Camelio, S., D'Elia, M., Fagoni, N., Moia, C., and Ferretti, G. (2019). Experimental validation of the 3-parameter critical power model in cycling. *European Journal of Applied Physiology*, 119(4):941–949.

Ward-Smith, A. (1985). A mathematical theory of running, based on the first law of thermodynamics, and its application to the performance of world-class athletes. *Journal of Biomechanics*, 18(5):337–349.

Watt, M. J., Heigenhauser, G. J. F., Dyck, D. J., and Spriet, L. L. (2002). Intramuscular triacylglycerol, glycogen and acetyl group metabolism during 4 h of moderate exercise in man. *The Journal of Physiology*, 541(3):969–978.

Weigend, F. C., Clarke, D. C., Obst, O., and Siegler, J. (2022a). A hydraulic model outperforms work-balance models for predicting recovery kinetics from intermittent exercise. *Annals of Operations Research*.

Weigend, F. C., Gray, E., Obst, O., and Siegler, J. (2022b). Benefits and limitations of a new hydraulic performance model. arXiv:2207.14295 [q-bio].

Weigend, F. C., Siegler, J., and Obst, O. (2021). A new pathway to approximate energy expenditure and recovery of an athlete. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 325–326. arXiv:2104.07903 [cs].

Whaley, M. H., Brubaker, P. H., Otto, R. M., and Armstrong, L. E., editors (2006). *ACSM's guidelines for exercise testing and prescription*. Lippincott Williams & Wilkins, Philadelphia, Pa, 7th ed edition. OCLC: ocm56415071.

Whipp, B. J., Huntsman, D. J., Storer, T. W., Lamarra, N., and Wasserman, K. (1982). A constant which determines the duration of tolerance to high-intensity work. *Federation proceedings*, 41(5):1591–1591.

Winter, E. M., Abt, G., Brookes, F. C., Challis, J. H., Fowler, N. E., Knudson, D. V., Knuttgen, H. G., Kraemer, W. J., Lane, A. M., Mechelen, W. v., Morton, R. H., Newton, R. U., Williams, C., and Yeadon, M. R. (2016). Misuse of "power" and other mechanical terms in sport and exercise science research. *Journal of Strength and Conditioning Research*, 30(1):292–300.