Optimal Power Distribution for a Cycling Individual Time Trial Using Reinforcement Learning

Triston Grayston, Bryce Davidson, Patrick Dann, Oliver Brown
Path 2, with a splash of originality

Abstract

Anna Kiesenhofer, an accomplished Austrian cyclist with a Ph.D. in applied mathematics, demonstrated the efficacy of mathematical modelling in enhancing a cyclist's time trial performance. Her work sparked many others to research the topic. While utilizing differential equations derived from Newton's laws for cycling, some of these models overlooked physiological constraints imposed by the human body. Building on Feng et al. [2022]'s work, our research incorporates critical power (CP) and anaerobic work capacity (AWC) as vital physiological concepts influencing race time. CP signifies the maximum sustainable power output without fatigue, while AWC measures the energy a cyclist can exert beyond their CP. AWC is finite, depleting during efforts exceeding CP and replenishing when cycling below it. This paper consolidates mathematical insights from Feng et al. [2022], Boswell [2012], and Ashtiani et al. [2019]. The resulting model is employed to train a reinforcement learning (RL) agent on a course defined in PyGame. We compare results from the RL model with the optimal solution proposed by Boswell [2012], highlighting the robustness and versatility of reinforcement learning in optimizing cycling performance.

1 Introduction

Cycling pacing in time trials is a crucial aspect of performance. It involves maintaining a consistent speed and power output throughout the course. Traditional thinking suggests an even-pacing strategy, where you aim to ride at the same intensity throughout. However, how does this change if we introduce changes in gradient?

Critical power (CP) in cycling is another key concept. It represents the highest average power you can sustain over a long period of time Poole et al. [2016]. It represents the threshold power between the anaerobic system and the aerobic system. The aerobic system can generate power for long periods of time by replenishing energy through the use of oxygen, while the anaerobic system uses finite energy stored in muscles to produce more power than the aerobic system.

Anaerobic work capacity (AWC), also known as W', represents the size of your anaerobic energy system. The larger this tank, the more power you can release above your CP level Poole et al. [2016]. This is especially important during short, intensive efforts, ie sprinting up a hill. The relationship between pacing, CP, and AWC is intricate. Pacing strategies in time trials often involve working near your CP threshold, while your AWC determines how much power you can release above your

CP level. Therefore, understanding and optimizing these three aspects can significantly improve your cycling performance in time trials.

To calculate an athlete's AWC and CP, Ashtiani et al. [2019] invited test subjects to perform a range of experiments, and found reliable methods to determine these values. The method they decided worked best involved a "3-minute all out trial", where the subjects would pedal as hard as they could for 3 minutes. This method worked well because it simulates real life correctly - they are in the same scenario that they would be in a race.

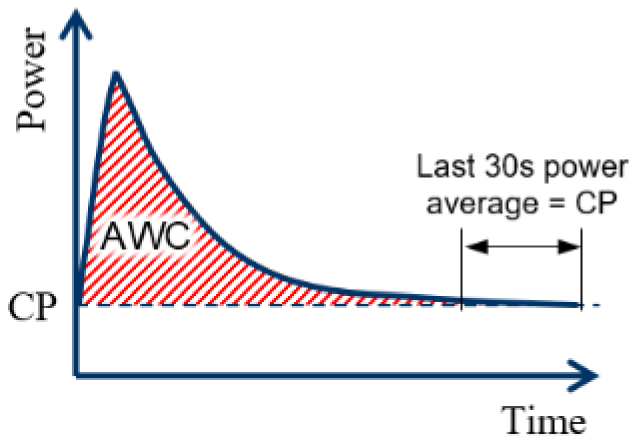

Fig. 2: 3-minute-all-out test protocol. The average power at the last 30 seconds of the tests is considered to be *CP*, and the area between power plot and *CP* is equivalent to *AWC*.

Figure 1: The 3-minute-all-out test.

It is important to note the rider and course we are assuming as well. Currently, we are working with power levels and AWC comparable to a middle-aged man rather than an elite cyclist, assumed to be 70kg, on a course that has a 15% uphill for 1 km, then a 15% downhill for 1km, and continues switching elevation per 1 km until 10 kms have elapsed. In theory, our model will be able to train and find the optimal power for any given rider and course, and we are currently working on a GPX file parser to create environments for our model from real world location data. The ultimate goal of this project is to create an app that will allow you to connect your Strava, choose a saved route as the course, and compute the optimal power distribution in real time. This app would also keep the rider notified if they are putting in too much or too little power (and continuously readjust the optimal power), which assumes they have a power meter connected to their bike. This is a reasonable assumption, as anyone who is serious enough to need optimal power distributions will have a power meter for training.

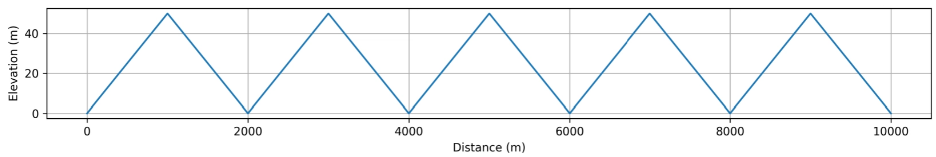

Figure 2: Our testing course.

1.1 Project Overview

In professional athletics, any marginal performance improvement is highly sought after. Sports optimization is a vast field of research. In July 2021, Anna Kiesenhofer competed in the women's Olympic road race, without the aid of a coach or a professional team. Because she was competing without the help of a coach or team, she was not considered a main competitor for qualifying. However, Anna was able to put her Ph.D. in applied mathematics to use and created a model to discover the optimal power distribution over the specific 137 km long Olympic course. By following the power recommendations she found, she won the gold medal for the event. This demonstrated the magnitude of the improvements available by following the optimal power distribution. This sparked a wave of research into optimal power distribution. We build on models created by Feng et al. [2022], Boswell [2012], and Ashtiani et al. [2019], combining the derivations that have been developed and applying RL to solve for the optimal distribution.

1.2 RL Introduction

The type of machine learning we opted to use is a form of reinforcement learning combined with deep neural networks, called a Deep Q-Network (DQN). The mathematics behind this form of ML is enough to build an entirely different explanation project on, so we will briefly introduce the methods it uses without diving too deep.

Reinforcement Learning, in general, is a style of ML that optimizes based on the Reward Hypothesis which is defined by Sutton as "all of what we mean by goals and purposes can be well thought of as maximization of the expected value of the cumulative sum of a received scalar signal (re-ward)." Sutton and Barto [2018] The agent needs to undergo a training process to discern favourable and unfavourable states/actions. Following this training, the agent gains the ability to identify specific moves or states and determine a certain value function for each state.

For our project, we use the distance travelled in the bike race to determine the state reward and the cyclist's power output for the action reward. Instead of associating positive rewards with nearing the finish line, we employ negative rewards throughout the race, decreasing in severity as the agent approaches the end. This approach maintains the incentive for the agent to consistently improve, avoiding a scenario where positive rewards might lead the agent to believe its current

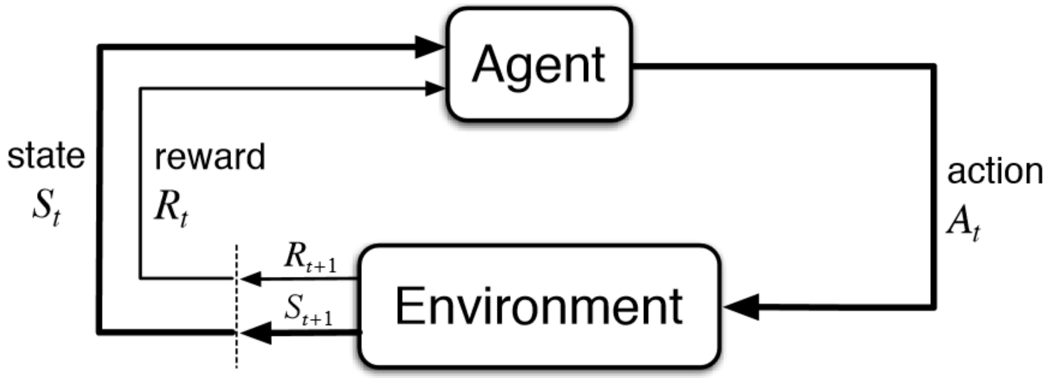

Figure 3: Description of the Reward Hypothesis taken from Sutton and Barto [2018].

position is satisfactory. By providing less negative rewards for progress, the agent is motivated to continually strive for a state where it receives rewards for its improved position.

2 Model Derivation

Our model is derived using physical laws and calculated data shown in Feng et al. [2022] and Boswell [2012]. We have chosen to model our simulation using the same equations, emphasizing the difference in computational complexity required and the speed at which an optimal solution is found.

In the models, we denote position as a function of time with $x(t)$, velocity as a function of time with $v(t)$, the "steepness" or "gradient" of the course to be $\theta(t)$, the mass of the rider as M, and the mass of the bicycle as m.

2.1 Fatigue Equations

To begin, we implement the physiological aspects. We want equations that show W increasing when pedalling below CP and W decreasing when pedalling above CP. Our equations are similar to the ones used in Feng et al. [2022] which are derived in Ashtiani et al. [2019] on experimental data. The equations used are shown below:

$$\frac{dW}{dt} = \begin{cases} -(P_{\text{rider}} - CP) & \text{for } P_{\text{rider}} \geq CP \\ -(0.0879 P_{\text{rider}} + 204.5 - CP) & \text{for } P_{\text{rider}} < CP \end{cases}$$

Here 0.0879 and 204.5 are experimentally determined constants to match realistic human results. These constants were determined by fitting adjusted recovery power vs. actual applied power during interval tests found in Ashtiani et al. [2019]. Where the subject was required to perform repeated bouts of high-intensity intervals. The adjusted recovery power (power level sustained during recovery) was then plotted against the actual applied power (corresponding all-out power level after the recovery period) to generate a linear relationship found in the figure below Ashtiani et al. [2019].

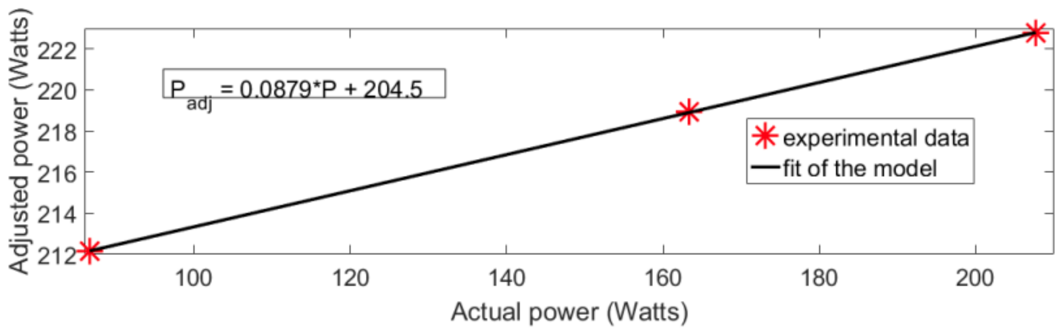

Figure 4: Experimental data found in Ashtiani et al. [2019].

2.2 Physical Equations

Next, we look at the dynamics behind riding a bicycle. We have multiple components to this, and each is broken down before showing the combined equation.

While riding a bike, there are many forces acting against a cyclist. One must overcome gravity, wind resistance, and resistance from rolling. Let's look at each portion individually:

2.2.1 Gravity

The power required to overcome gravity is shown in Boswell [2012] to be:

$$\Delta_{\text{gravity}} = (M + m)g\sin(\theta(x))v$$

Where g is the acceleration due to gravity, and is assumed to be a constant $9.81\,\text{m/s}^2$.

2.2.2 Air Resistance

The power required to overcome wind resistance is shown in Boswell [2012] to be:

$$\Delta_{\text{wind}} = A(v - w\cos(\varphi(t)))^2 v$$

Where A is the frontal area of a rider which we assume to be constant, w is the wind speed which we assume to be constant, and $\varphi(t)$ is the angle that the wind is pushing the rider, which we also assume to be constant and head-on.

2.2.3 Rolling Resistance

The power required to overcome rolling resistance is shown in Boswell [2012] to be:

$$\Delta_{\text{rolling}} = (M + m)gR(x)v$$

Where g is gravity again, and $R(x)$ is a coefficient to take into account the conditions of the ground over each point in the course. We assume $R(x)$ is constant, and equal to 0.00466. This constant is stated in Boswell [2012].

2.2.4 Combined Equation

Putting these all together, we know from physics that the derivative of kinetic energy with respect to velocity is power, and say that the difference between the power applied by the cyclist and the resistance forces gives us the rate of change of kinetic energy. This yields:

$$\frac{d}{dt}\left(\frac{1}{2}(M + m)v^2\right) = P_{\text{rider}} - \Delta_{\text{gravity}} - \Delta_{\text{wind}} - \Delta_{\text{rolling}}$$

And, assuming mass is constant:

$$(M + m)v\frac{dv}{dt} = P_{\text{rider}} - [\sin(\theta(x)) + R(x)](M + m)gv - A(v - w\cos(\varphi))^2 v$$

And this differential equation is assumed to have initial conditions $x(0) = 0$, $v(0) = 0$.

To make sure our model works well with RL, we have discretized position and time much like Ashtiani et al. [2019]. After discretizing, some algebra on the above differential equation gives us a formula for calculating velocity over each interval, which is:

$$V_i = \left[\left(P_{\text{rider}} - \Delta_{\text{gravity}} - \Delta_{\text{wind}} - \Delta_{\text{rolling}} - \frac{1}{2}(M + m)v^2\right)\left(\frac{2}{M + m}\right)\right]^{1/2}$$

3 Assumptions

- Change in AWC is directly proportional to the time spent above or below CP, with a proportionality constant of $(P - CP)$.
- We assume constant rolling resistance throughout the entire race.
- We assume a constant wind speed of 0 throughout the race, making the wind speed term w and the angle of headwind φ both inconsequential. Thus the air resistance is simply a function of v, but we chose to leave these values in the model so that they can be accounted for with ease.
- We average the gradient of the course over every discretized distance, 5 metres.
- We assume the course is straight (no turns are accounted for).

- We assume this is an individual race, that is there will only be one agent at a time.
- We assume the rider is performing at their specified CP/AWC, which is experimentally determined prior to modelling.
- We assume the RL model converges to the optimal strategy.
- We assume the course does not have a high gradient (nothing more than 15%, anything higher is a very steep race!).
- We assume the cyclist is able to perform close to the recommended power output consistently, thus not taking varying motivational levels or mental barriers into account.
- We assume 100% mechanical efficiency for bike powertrain from power on pedals (no power is lost due to minor inefficiencies such as a chain rubbing against the derailleur).
- No falling or crashes.

4 Model Analysis

Our model is still underway; we are continuously tweaking the RL parameters and finding improvements. However, our model has converged to a strategy that works better than randomly choosing a power output at any moment, and although it may not be the optimal solution it does work decently. The strategy found is going at the maximum current power nonstop, which we know is not the correct strategy, as we have chosen certain strategies and shown that when applied to our model, it outperforms the RL solution. By changing the reward function we hope to achieve increasingly accurate results.

4.1 Comparison with Boswell [2012]

In Boswell [2012], their DE model is used to calculate how a hypothetical rider would perform given our course, and a few other courses that are the same structure but with varying gradients (0%, 5%, 10%). The rider performance is compared with a few varying strategies; each strategy is to increase power by a certain percentage when going uphill and decrease by a certain percentage when going downhill. However, the physiological aspects that we implemented were not taken into consideration, so our times were much slower but more realistic. Here is the table given in Boswell [2012], followed by graphs of our model with the 15% gradient course:

Our hypothetical rider took 2257 and 2279 seconds to complete the course, where the first number was using a strategy of 100% of their current available power on the uphill and 80% of their current maximum power on the downhill. The second time was using a strategy of applying 100% of their current maximum power at all times, which ultimately converges to CP and yields an average power of 238 watts. As you can see, in both graphs of ours, the rider very quickly converges to their CP, but in the 100/80 strategy, they gain a little boost of AWC after the downhill portion, allowing them to speed up on the next incline. Because they applied less power on the downhill, the average power was *less* than the 100% strategy, but also faster. This is essentially the whole idea behind optimal pacing strategies; well-timed power output can defeat higher power output. We believe that although this strategy provided results closer to realism than Boswell [2012]'s findings, RL could find an even better strategy, specifically tuned for this course.

Table I. Times (s) to complete 10 km courses by solving equations (5) with constraint (7) using constant ($\bar{P}$) and variable (gradient: +5%, +10%, +15%) pacing strategies for the theoretical cyclists of mass $M - 63$, 70, and 77 kg.

			2 × 5 km			10 × 1 km			20 × 0.5 km		
			Gradient			Gradient			Gradient		
Power applied	Mass (kg)	10 km 0%	5%	10%	15%	5%	10%	15%	5%	10%	15%
Constant $\bar{P}$	$M - 63$	882	1053	1454	1933	1023	1368	1787	988	1266	1619
	$M - 70$	886	1090	1550	2085	1052	1443	1907	1007	1320	1705
	$M - 77$	890	1129	1648	2238	1082	1518	2025	1027	1373	1789
Variable + 5%	$M - 63$	882	1033	1407	1859	1007	1329	1726	976	1237	1571
	$M - 70$	886	1067	1497	2002	1034	1400	1840	995	1288	1655
	$M - 77$	890	1104	1589	2148	1062	1472	1953	1014	1339	1736
Variable + 10%	$M - 63$	882	1015	1364	1791	993	1294	1671	966	1210	1528
	$M - 70$	886	1047	1449	1928	1018	1362	1780	984	1260	1609
	$M - 77$	890	1081	1536	2066	1045	1430	1888	1002	1308	1687
Variable + 15%	$M - 63$	882	999	1325	—	980	1262	1620	957	1186	1488
	$M - 70$	886	1029	1406	—	1004	1327	1725	974	1233	1566
	$M - 77$	890	1060	1489	—	1029	1392	1829	991	1280	1642

Note: Notation and parameter values are given in the text. Entries denoted by "—" correspond to instances where it was impossible to satisfy the mean power constraint in equation (7).

Figure 5: Performance of a rider with varying strategies shown in Boswell [2012].

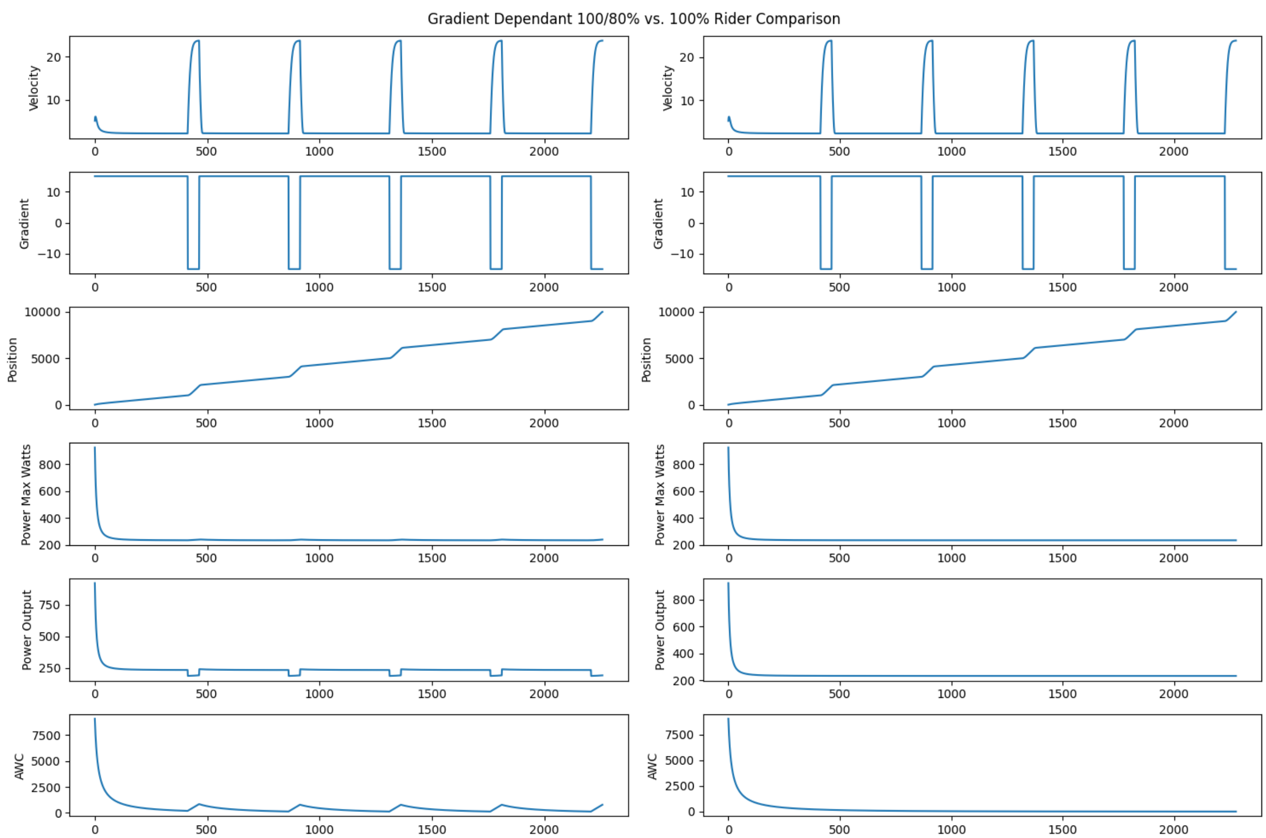

Figure 6: Performance of a rider with varying strategies shown from our model.

It's clear that our rider did not outperform theirs, but that was expected given the unrelenting power assumption of Boswell [2012], and the rider they assumed had a much higher average power (323 watts) than ours (234 watts).

4.2 Comparison with Ashtiani et al. [2019]

In Ashtiani et al. [2019], it is noted that the entire simulation takes less than 1 minute to run on a standard computer. While this time is impressively quick, it is also noted that the method they employ is less generalizable to other courses, meaning they sacrificed applicability to other courses for computational efficiency. As well, they demonstrate the efficacy of their optimal solution with a graph showing the time that it took to complete a 10.3 km course with 2 gradual uphills and a slight downhill in between. Although we cannot directly compare our results to theirs as the courses differ, we can note that the time to complete our course with a rider holding the same mass and CP/AWC values was on the same magnitude as the time to complete their course.

4.3 Strengths

- RL is very general, any course can be added and the solution may be computed quickly.
- It would be easy to apply this theoretical solution to a real-life situation.
- CP and AWC are accounted for.
- Relatively easy to find any particular rider's CP and AWC.

4.4 Weaknesses

- Relatively high number of assumptions (100% mechanical efficiency, bike and person act as one object, riding position isn't accounted for).
- Not trained on many courses (but could be with more time!).
- The constants used aren't general; rolling resistance and some physiological constants should be determined for each scenario.
- Convergence rates of DQN is not the highest when compared to other RL architectures.

5 Conclusion

The approach of applying reinforcement learning can be used as a more general and robust solution to already proposed models. Any cyclist committed enough can utilize this architecture to get a suggested optimal pacing solution, given they know their own CP and AWC values discoverable by simple tests. This has many useful applications in industry, leading to a potential device which illustrates to cyclists the optimal power output based on factors such as course, weight, CP, and AWC. The strategies learned from this architecture can aid riders in maximizing future performance goals.

6 Bibliography

References

Faraz Ashtiani, Vijay Sarthy M Sreedhara, Ardalan Vahidi, Randolph Hutchison, and Gregory Mocko. Experimental modeling of cyclists fatigue and recovery dynamics enabling optimal pacing in a time trial. In *2019 American Control Conference (ACC)*, pages 5083–5088. IEEE, 2019.

Graeme P Boswell. Power variation strategies for cycling time trials: a differential equation model. *Journal of Sports Sciences*, 30(7):651–659, 2012.

Jiafan Feng, Xiaowen Liu, and Shaochun Deng. Optimal pacing strategy modeling of cycling individual time trials. In *Journal of Physics: Conference Series*, volume 2282, page 012003. IOP Publishing, 2022.

David C Poole, Mark Burnley, Anni Vanhatalo, Harry B Rossiter, and Andrew M Jones. Critical power: an important fatigue threshold in exercise physiology. *Medicine and science in sports and exercise*, 48(11):2320, 2016.

Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction.* The MIT Press, second edition, 2018. URL `http://incompleteideas.net/book/the-book-2nd.html`.

7 Appendix

7.1 Python code

To save space on our document, we have decided to add our GitHub repository which shows all of our code for each part. Given that we are planning to continue with this project, it may be important to note that updates will be made, but commit history is available showing where we were at each step. Below is the link: `https://github.com/tristongrayston/BikerGymEnv/tree/master`