Intelligent Agent Based Modelling of Real-World Peloton Behaviour in Road Cycling

by Matthew Thoumine, BEng
14274893

Supervised by Colin Johnson

Submitted to The University of Nottingham
September 2021
in partial fulfilment of the conditions for the award of the degree of
Master of Science in Computer Science

I hereby declare that this dissertation is all my own work, except as indicated in the text

Abstract

This paper presents an agent-based model of competitive road cycling, with the ultimate aim of such a model being used to improve the safety of real-world athletes in professional road racing. With this model having being developed from scratch for the purposes of this project, the aim in this paper is for the model to be able to replicate the results produced by previous research published within this field. Agents will model the behaviour of real-world athletes in a peloton by forming groups through flocking. In addition agents will have diverse and changing goals dependent on: their desire to co-operate with other riders; a representation of their physical strength; their current energy expenditure; and the race distance remaining. One of these goals is to draft fellow rider agents. Drafting is the act of riding behind another rider in the pocket of low air pressure produced by the leader, doing so reduces the aerodynamic drag on the following rider which in turn reduces energy expenditure. This is key to cycling tactics. Different methods of calculating the drag reduction have been implemented and compared. The results produced by the model are comparable to those within previous research however some modifications are required. These modifications includes changes to how the model treats non-active agents. Furthermore optimization of the model is also necessary given the length of time taken to completed simulated test runs even when done on a reduced scale.

Acknowledgements

I would like to start by thanking my supervisor Colin Johnson for supporting me in my desire to do a project on something left-field. Without your calm guidance I'm sure this thesis would resemble the ramblings of mad-man rather the reasonable piece of work it has become.

I'd like to thank my parents for keeping me trapped in my room to make sure I was working hard enough. I'd also like to thank my sister for letting me out of said room for food, water, and exercise once a day. Jokes aside, without all of your support I would either have failed to finish this or gone stir crazy.

Finally I would like to thank my laptop, you're getting on a bit but the fact you didn't catch fire while running the simulations is enough.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Although unsurprisingly debated, the birth of road cycling is thought to be a 1,200 meter race held on the 31st May 1868 at the Parc de Saint-Cloud, Paris. This race is recorded as having been won by an Englishman James Moore, who rode on a wooden bicycle with solid rubber tyres (Wikipedia contributors (2021)). However much has changed within the sport in the 153 years since then. That being said if you were to look up the results of every subsequent race after James Moore's you'd likely read a list of names and come to the conclusion that road cycling, like running, is an individual sport. In doing so, you would be very much mistaken. Road cycling is in-fact a team sport.

The top echelon of road cycling is the UCI World Tour, this year (2021) there are 29 road events at World Tour level. These will be divided between one day races, as the name suggests these are events that take place over a since day, and stage races, which take place over multiple days. The most prestigious one day races are known as "the Classics" and within these you have "the Monuments" which are the five "best of the best". The most prestigious stage races are the 3 annual Grand Tours of road cycling: the Giro d'Italia; the Veulta Espana; and the largest cycling race of all the Tour de France. Each of these will have 23 teams of 8 riders each competing in 3 weeks of racing across the roads of their host nation. This means a professional road racer on a World Tour Grand Tour stage will share the road with approximately 183 other riders. The rider to cross the line first each day will have their name etched into the history books, the other 183 will not. And here in lies the conundrum of modern road cycling, an individual is given the glory but a team is required to get them there.

Ask many about crashes in modern road racing and there is likely to be a discussion about "respect within the peloton". Although the "respect" they speak of is not a quantifiable construct their point can be observed. In order to do so one must first understand the composition of a road cycling team. At its most basic a team will consist of one leader and the rest domestiques (or helpers). The job of the leader is to win, while the job of the domestique is to enable the leader to do so no matter the cost to personal results. Historically a team's leader would likely be one of the most, if not the most, experienced rider on the team. As a result of this, they were likely the most well trained and therefore also the most physically capable. This system for years built a natural hierarchy within road cycling. A new professional (neo-pro) would turn up and sacrifice themselves for the team. If they prove their worth then they may be given the chance to win something, prove they could do that and they might be given the opportunity to lead a team. Naturally the longer a professional stayed around the better they became, the more results they gathered, the more respect they got within their team and within the group they raced with. However if a new professional decided to try and win "before their time" they might end-up with no renewal of their contract for the following year. What has all this got to do with crashing? Well although modern cycling still has leaders and domestiques the previous system of deciding which riders were which no longer exists. Now not only can a seasoned modern professional make a name for themselves as a super-domestique (someone so good at the domestique role they transend it) but also modern training methods are available to much younger athletes meaning the team leader role is not just the rider who has been around the longest. This means the hierarchy discussed earlier no longer exists in the same manner, road cycling is more open than even. As a result more teams and riders believe they are capable of winning. Unsurprisingly this increases the pressure to do so. Which in turn leads to more riders taking more risks under higher pressure in order to come out on top.

The solution to this problem? Some will tell you riders need to take fewer risks. I'd argue this position is naive, after all the famous Ayrton Senna quote "[if] you no longer go for a gap that exists, you are no longer a racing driver" is surely applicable to the athletes discussed here. Due to this fact, there is a second camp that argue the race routes themselves are no longer safe for modern professional racing. Whether it be treating the symptom rather than finding a cure, working towards safer road racing is important because any errors in judgement or mistake by organisers can lead to life changing injuries or even fatalities for athletes within a fast travelling peloton. This year alone has seen horror crashes in the first couple of days of two out of the three major Grand Tours. One being, favourite for Giro d'Italia success, Mikel Landa crashing out of the race after impacting with a piece of road furniture are well as the road marshall whose duty it was to make it obviously clear to riders. Another being Tour De France favourite Primoz Roglic after he was brought down in a crash on the first very first day of racing. Bear in mind these are only 2 of the most high profile cases where athletes have been forced to go home after hitting the asphalt, many more are out there just without the same level of media coverage.

1.1 Aims and Objectives

This project therefore will aim to investigate the possibility of simulating cycling events through use of artificially intelligent systems with the overall purpose of improving athlete safety. An accurate model of cyclist behaviour and peloton movement would allow event organisers to run simulations of race routes beforehand. This could let organisers identify areas of increased risk or danger based off of empirical evidence. Using this evidence they could then apply resources accordingly. However getting to this stage is still a long way off and that is one of the reasons this problem is interesting scientifically. Applying intelligent agents to human athletes in a competitive setting such as this is still an under-researched field.

This thesis will contain the following:

- *Section 2 : Research* - A brief review of previous published works on this and the surrounding topics
- *Section 3: Methodology* - An overview of the principles used to develop the model as well as a description of how said principles were implemented
- *Section 4: Results* - An outline of the tests that the model has been put through and the reasons for doing so
- *Section 5: Discussion* - An evaluation of how success implementation of the model has been. Identifying areas of positive outcomes or shortfalls.
- *Section 6: Conclusion* - A final summary of the model and what has been learnt through the development process
- *Section 7: Further Work* - An explanation of some where development of the model could go from its' current form

Chapter 2

Research

As discussed this project plans to utilise an intelligent agent-based model to produce an accurate simulation of a road racing peloton with the eventual aim of being able to predict areas of increased rider danger due to peloton movement on planned racecourses. This involves combining research from the fields of aerodynamics and artificial intelligence to first create a system that will represent real-world physics on riders in the form of agents. These agents will then be given the ability to make tactical decisions. This "thought" process must be as closely matched to human athletes as possible to ensure the best levels of accuracy. Finally, these agents will need to be placed into an environment that represents those faced by real-world athletes. All leading to the opportunity: first to identify what factors create a dangerous situation; and secondly look for patterns in racecourse design that can be avoided to reduce these factors.

2.1 Agent Based Simulations

The use of an agent-based model to represent the intricacies of bicycle racing were first investigated in "Cooperation in Bike Racing- When to Work Together and When to Go It Alone" (Hoenigman et al. (2011)). This research identified that the athletes within bicycle races display behaviour easily replicated by agents in a model. This is due to the fact that agent-based models are a preferred method to "describe and investigate systems of individuals (the "agents") whose global (group) dynamics emerge from simple interactions between the individuals". The study used this to try and predict optimal race strategy calls, cooperative (sharing the workload) vs defective (pursuing self-interest), for riders to make for both themselves and their teammates based on levels of physical exertion. The model and results from this study are useful as they demonstrate how accurate rider agent decision making is when compared that of real-world athletes, even if the goal of the research to solely predict race results is not fully aligned with the objectives of this project. This fact was also pinpointed by Erick Martins Ratamero who used the basis of the model from the previous study to build out research focused “not on the actual results of the race, but rather on simulating the behaviour of a peloton during a race, under different circumstances” (Ratamero (2015)).

Essentially all modern flocking algorithms build upon the work of Craig W. Reynolds. In July 1987 he released the paper Flocks, Herds, and Schools: A Distributed Behavioural Model that pointed at the three criteria required to build a simulated flock with Boids. These behavioural criteria are: “1. Collision Avoidance: avoid collisions with nearby flockmates. 2. Velocity Matching: attempt to match velocity with nearby flockmates. 3. Flock Centring: attempt to stay close to nearby flockmates” (Reynolds (1987)). Erick Martins Ratamero’s work is no different, making use of the Netlogo Flocking Model (U. (n.d.)). However, Ratamero felt it was necessary to adapt this model for an accurate representation of bike racing. One of these adaptations is changing velocity matching (in this case referred to as alignment) to be purely intra-peloton lateral travel, as all competitors travel in the same direction normal alignment matching was deemed “out of place”(Ratamero (2013)).

Furthermore, Erick Martins Ratamero adapted the separating and cohesive forces to be within a vision radius to better replicate human athletes (Ratamero (2013)). This final point is interesting as years later J. Belden released the paper “How vision governs the collective behaviour of dense cycling pelotons” which investigated the shape and organisation of peloton structures. The paper links the riding position of athletes within the peloton to their angle of visual input (Belden et al. (2019)). The research not only points to riders reacting better to certain peloton directional shifts than others because of this link but also to the existence of a change in positioning by riders sub-consciously based on this link. In other words, as athletes get closer to the end of a race their sensory focus narrows (possibly due to higher energetic output) and as a result the group becomes elongated. This is a point of interest as it somewhat contradicts the “common knowledge” within cycling that is the bunch narrows as speed increases due to conscious factors relating to aerodynamics, drafting, and conscious attempts at reducing energy expenditure. The distinction here between conscious (tactical) and subconscious (natural) decision making will likely be important to get an accurate simulation of peloton behaviour in dangerous and high paced areas of a route.

Interesting although Erick Martins Ratamero noted that his previously mention study was flawed due to the lack of intelligent agents making the best strategic decisions, there does not appear to be evidence this avenue was ever pursued. That being said, he was involved in another study "A deceleration model for bicycle peloton dynamics and group sorting" which investigated how pelotons sort and split based on rider energy expenditure on different gradients (Trenchard et al. (2015)). Through this study is developed the "peloton-convergence-ratio" (PCR) that can be calculated from a riders Maximum Sustainable Power (MSO) along with a value for the workload of the front rider and a value for the amount of work saved by the rider in questions position. The result is that when PCR is greater than 1 a rider is physically incapable of holding on (maintaining a "draftable" distance) to the wheel in front and therefore the rider is dropped (the gap to the rider in front becomes uncrossable). This study however again still falls short of displaying any tactical nuance, but it was built on by Hugh Trenchard in "The peloton superorganism and protocooperative behaviour" (Trenchard (2015)). This second study compares peloton drafting behaviour to that of other "biological systems involving energy saving mechanisms". It hypothesises that there is a threshold at which athletes change between one of two states, that of intra-group cooperation and another of inter-group competition. The former state is found to occur most often at low speed where cyclists have a higher tendency to share the non-drafting positions amongst themselves however as speeds and therefore energy expenditure rise each athlete has a threshold at which they no longer have this tendency. This threshold named the "protocooperative behaviour threshold" (PBT) is found to be equivalent to the coefficient of drafting. Nevertheless, this still falls short of accurately representing a peloton.

As touched on previously over the course of a road race athletes make unconscious and conscious decisions based on observations of the athletes around them and their understanding of fundamental race tactics built through years of racing. As a result of each of these individual decisions a peloton of cyclists can be seen to demonstrate emergent behaviour. It is true that this emergent behaviour needs to be replicated to determine which areas of a race route are dangerous and which are not, but it does not necessarily show the full story. Many of these studies rely on the physical limitations of athletes to generate peloton behaviour with less powerful riders falling to the back and more powerful ones remaining near the front. The fact this must naturally occur is undeniable, but cycling is a team sport. Often weaker riders will sacrifice themselves "on the front" riding in non-drafting positions in order to serve a “team-leader” this behaviour could be replicated with intelligent agents. This introduction of tactics into a model is key in allowing race organisers to accurately predict peloton behaviour as it can cause premature or delayed splitting of groups that can have an impact on how many riders enter an area of the course at the same time and what speed they’re going.

2.2 Simulating Environments

As discussed, many of the tactics behind road racing are built upon the concept of drafting. Aerodynamic drag is the main resistance that a fast-travelling cyclist is required to overcome. Obviously to do this requires the cyclist to do work to “punch a hole” in the air in front of themselves. This act generates a turbulent wake in the airflow following said rider. If a second rider positions themselves in the turbulent wake, then they can travel at the same speed as the rider in-front but for considerably less effort due to the change in aerodynamic forces. The amount of effort saved by the rider behind has been up for debate over the years, but the agent based research discussed previously makes use of results from “Energy Expenditure during bicycling” which was published in 1990 (McCole et al. (1990)). This series of experiments involved athletes riding down a 2km stretch of road while their exhaled gases were monitored. In addition, they rely heavily on a formula for calculating drag correction factor (this controls the amount the drag forced is reduced by sitting in the draft) from ”The mathematics of breaking away and chasing in cycling”(Olds (1998)). This formula was calculated using graphical data from an older paper (Kyle (1979)) but in doing so chose to focus specifically on the ”racing position” of cyclists rather than the ”upright position” or the average. The importance of this will be discussed in more detail later. At the time, these studies were ground-breaking however with improvements in technology (both to carrying out studies of this nature but also with monitoring within the sport itself) it might be prudent to use more recent research. For example, the dataset held within “Aerodynamic drag in cycling team time trials” (Blocken, Toparlar, van Druenen & Andrianne (2018)) involving a leading cycling aerodynamicist Bert Blocken. Throughout this research Blocken makes use of wind-tunnel and computational fluid dynamics testing that was not used during the earlier study. Furthermore, the Blocken study offers a wider range of data than that available previously. This additional data not only includes details on how the impact of the draft changes based on the distance between leading and following rider but also how the drafting effect changes once more than 5 riders are included in the paceline. Common knowledge would suggest that the best place to ride to expend the least amount of energy is at the back of a line of riders however Blocken discovered that once there are more than 5 riders present this is no longer the case. In fact, the wake created by the first rider will start to fill in/tail off after the 5th rider meaning a 6th rider will have to "reopen" the hole in the wind and as a result have to work harder than the rider in-front. This effect changes once again when a 7th rider is added and so on as more riders are added. Finally, this study also comments on the effect of drafting within both a tightly packed and slightly more spaced-out peloton, it found that riders within the group can have an average drag at 21.1% and 21.9% respectively when compared to that of an isolated rider. These figures are extremely important for the model to accurately represent the impact of real-world drafting which is a major step required to then implement any form of tactical decision making by agents within the environment.

Of course modelling the environment the agents will operate in goes further than simply aerodynamic forces but to also include the effects of gravitational forces, changes in momentum, and rolling resistance. All of these are neatly outlined in the paper "Validation of a Mathematical Model for Road Cycling Power"(Martin et al. (1998)). Unlike the aerodynamic forces where much is still up for debate, these formulae are overall agreed on by the scientific community. The same formulae such as those outlined by Martin et al. (1998) can be seen in multiple more recent research papers in the field including "Bicycle Aerodynamics: History, state-of-the-art and future perspectives" (Malizia & Blocken (2020)) and "Aerodynamics analysis of uphill drafting in cycling" (van Druenen & Blocken (2021)).

Chapter 3

Methodology

It is important to mention at this stage that there is no access to the previous models used in research therefore the only suitable approach is to design and implement a new one from scratch. This places some limitations on the scope of this project regarding the amount of new developments that are possible. The new model will first need to demonstrate that it can match the results of previous ones before being developed further into some of the areas discussed within the previous section. The model development process can of course be done in a multitude of different ways using a wide range of different software. In this case the model is built in python purely due to it's level of familiarity, therefore reducing the adaption time required through use of a new software solution.

The model makes use of the numpy library to perform vector mathematics to determine the location, velocity, and acceleration of an agent at any one time. The acceleration of an agent is dependent on the four main sections of the model: dynamic parameters; energetic parameters; physiological parameters; and finally decision making.

Dynamic Parameters are those that produce the classic flocking behaviours outlined in Boids Algorithm. They are the fundamental underpinnings of any model involving flocks and in this case can be compared to the subconscious behaviour of riders within a peloton. Each rider will aim to avoid crashing into the riders around themselves (separation), while aiming to stay within the group (coherence), and all this while simultaneously having an eye on the road around them and a desire to get to the finish (alignment/path following).

Energetic Parameters are those that introduce our rider agents to the environment they will operate within, forcing physics upon them. This involves the introduction of resisting forces such as aerodynamic drag, rolling resistance, and the work required to

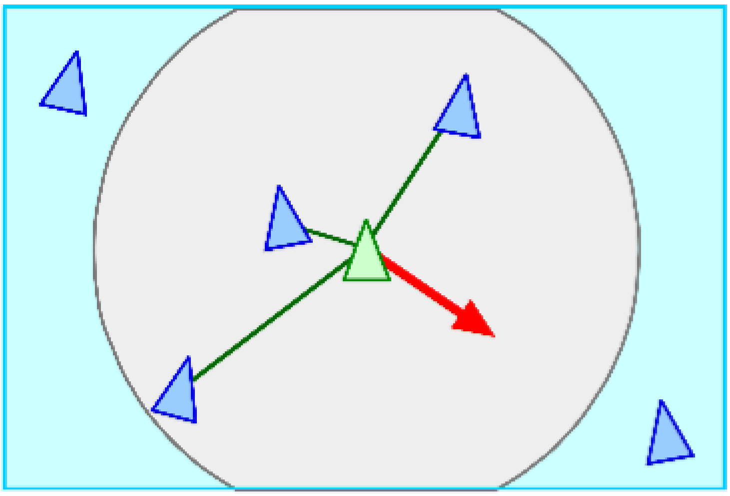


Figure 3.1: Separation Forces, Diagram from Reynolds (2002)

change their own potential and kinetic energies. Of course involved within this are certain physical characteristics such as the mass or frontal area of the rider agent.

Physiological Parameters are where the dynamic and energetic parameters start to co-inside. Dynamic Parameters are what are pushing the rider agent to make a change and the Energetic Parameters are those resisting that change. The Physiological Parameters are where these are combined and confined to the limits of a human's physiological ability.

To finish off there is Decision Making which is as it sounds, this is what causes an agent to make a certain tactical decision. Up until this point much rider/agent behaviour is the same. None want to crash or get dropped, all want to finish and experience physics, and all are limited by their own physiological abilities (be that to a more or lesser degree). Based on these physiological abilities an agent must decide how they are going to behave in order to give themselves the best chance of being victorious.

Dynamic Parameters, Energetic Parameters, and Decision Making all apply a range of forces to the agent. When summed and limited to the Physiological Parameters of the agent these can be applied to produce an acceleration to the agent within the environment.

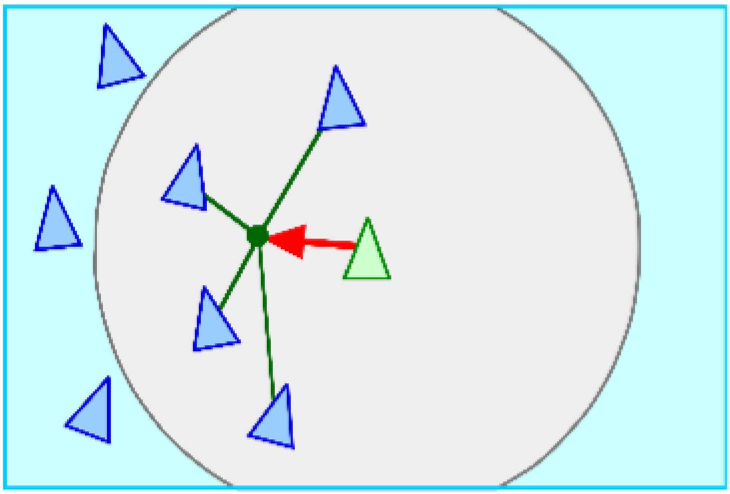


Figure 3.2: Cohesion Forces, Diagram from Reynolds (2002)

3.1 Dynamic Parameters

3.1.1 Separation

To ensure that rider agents do not collide with one another they must experience a force that moves them away from agents within close proximity, see Figure 3.1. For this series of experiments the desired separation of agents will be 2m from center-to-center, in-line with the works of Ratamero (2013). The system will therefore check the positions of all agents relative to the agent in question. For those where the vector between agents is less than the desired separation, this vector is then normalized and scaled with relation to the distance. Therefore a closer agent will result in a larger proportion of the overall force. All these vectors are summed and divided by the number of close agents, before being normalized and given a magnitude equal to the largest possible force (as defined as a rider characteristic). Doing this ensures the fastest possible reaction to being too close to another agent. The steering force applied to the rider is then the difference between this resulting reaction force and the agents current velocity.

3.1.2 Coherence

The act of cohering to a group of agents is very similar to that of separating however there are a few key difference. Again for this series of experiments the desired coherence distance is as that of Ratamero (2013) at 20m. The average location of all riders within a 20m

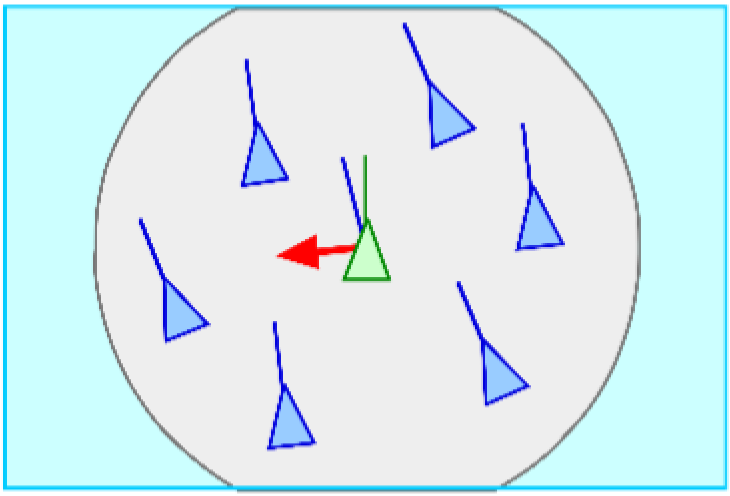


Figure 3.3: Alignment Forces, Diagram from Reynolds (2002)

radius of the rider agent in question is then calculated. Taking this average location vector and subtracting the agent's own location vector gives the desired vector for the agent to move along to cohere towards the center of the group. Again this vector is then normalized and given the same magnitude as the largest possible force in order to create the fastest reaction. Similarly to before the steering force vector is then the difference between this resulting reaction vector and the agent's current velocity. This behaviour can be seen in Figure 3.2.

3.1.3 Alignment

In an ordinary flocking model there would then be an aligning force present, see Figure 3.3. This force would ensure that all members within the flock have a similar heading. As identified by Ratamero (2013) this force is not required for a model of this nature. This is due to the fact that all agents within the model follow the course direction. In a cycling race all competitors have a desire to finish, it is unheard of for a rider to turn around mid-race even if something disastrous were to occur to a teammate behind. Therefore this seeking of the finish has been modelled along with Ratamero (2013)'s mention of a "bias to the center" of the course through use a path following algorithm which is discussed in more detail later.

3.1.4 Removal of visual cone

So far much of this has been reflective of Ratamero (2013)'s work, however one element of their work did not make it into the model, that of a visual cone. Ratamero (2013) suggests the use of a 140° (presumably 70° either side of the rider heading) visual cone to mimic that of the human eye. This would fall inline with the work of Belden et al. (2019), as discussed earlier, whose work could be taken to suggest that an adaptive weighted approach may be a worthwhile investigation. With agents being more perceptive to movements within 2°- 10° then 10°- 30° then 30°- 60° and with agents being least perceptive between 60°- 90°. And with these limits changing based on level of rider exertion.

With all this being said however the introduction of a limiting visual cone appeared to have a negative impact on this model. Figures in the Ratamero (2013) work seems to suggest that for much of this work the peloton stays together in a single bunch, this is not true of what was seen during development of this model. Furthermore the introduction of a visual cone appeared to worsen this further with rider agents in-front often being ”unaware” of the group behind and as a result continuing regardless. This behaviour is not synonymous with professional cycling where riders will often ”sit-up” and rejoin a group behind if they know it is coming. Yes, humans have a visual cone but human professional riders can be seen to look round behind themselves even at seemingly the most tense or extreme moments such as in a bunch sprint. Potentially assuming agents and therefore athletes have little to no idea of what is going on behind them is an incorrect assumption to make. Therefore this model will go forward without the restriction of a visual cone of this nature. However agents in this model will seek preferable drafts that change based on positioning behind fellow riders, this behavioural element will force agents into the classic diamond peloton shape regardless.

3.1.5 Path Following

As mentioned in the alignment section, Ratamero (2013) suggests that the model should contain a bias for agents to head towards the center of the road. This is justified by the fact that real world riders head towards the center to ”avoid danger on the fringes of the road”. Ratamero (2013) notes this is only true in still wind conditions. However this behaviour could also be considered true for headwinds and tailwinds. The most important distinction to make is that this is no-longer true when crosswinds are present, during these weather

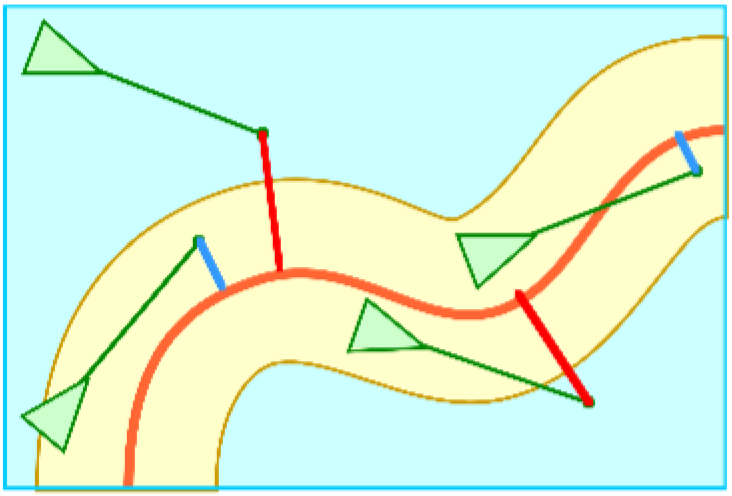


Figure 3.4: Path Following Forces, Diagram from Reynolds (2002)

conditions the peloton changes shape entirely into echelons. Another important point of note is that this behaviour is also only true on a straight road. When riders and therefore the peloton enters a corner there is a tendency to change road positioning to allow the most speed to be carried through the corner. A tight right-hand corner will see riders drift leftward prior, turn rightward into the bend apex, and then drift outward to the left-hand side again in a classic "outside-inside-outside" cornering technique.

How other research for this kind has implemented positional bias of this kind within their models is unclear, however this project will make use of the well known Path Following algorithm created by Craig Reynolds as part of his flocking research (Reynolds (2002)). This algorithm works by comparing the predicted position of an agent, if they are to continue at their current velocity, with the corresponding normal point on the course. If the distance between the predicted location and the normal point on the path is greater than the path radius then the agent is heading off of the path and therefore a force needs to be applied to correct the agent's heading, see Figure 3.4. This differs from the work of Ratamero (2013) as it will keep agents on the road and nothing more, allowing for agents to move freely across the road when attacking which is similar to that of real world riders who try to "break the tow" or reduce the opportunity for others behind to get a draft by attacking away from others and onto the opposite side of the road.

Implementing path following in this method means that there needs to be a defined route. To do this the model has made use of the python library .gpxpy (Krajina (2020)) to use modified GPS Exchange Format (.gpx) files. This is an open file format that is used in a lot of modern real-world tracking solutions. The files contain a series of way-points where each one is described by 3 values, these are latitude, longitude, and elevation. As the agent moves along the route they will increment between the way-points to remain on the current section of route. The reason for using this method is that now real-world routes can be uploaded into the model as well as test environments. Unfortunately the model in its current form cannot give an accurate representation of cornering. The advanced predictive behaviour discussed earlier along with the requirement of accurate calculations of cornering arc alongside tyre properties such as slip angle find it outside of the scope achievable by this project. However path following implemented in this manner does allows for the possibility of this expansion into real-world routes to be made down the line.

Finally, a seemingly obvious but easily overlooked point to mention is the issues that can arise when applying this method to a three dimensional space. Although the route is defined in three dimensions the rider agent is only ever able to function with two degrees of freedom, this of course being within the x and y axes. The z axis (elevation) of a rider is defined by the course and therefore must be restricted, effectively enforcing a plane on which agents are active. If this is overlooked and the z axis is restrained under the path following algorithm it will allow agents freedom within a cylinder around the defined route which in the real-world translates to allowing riders to travel above and below ground to a maximum height/depth of the path radius. This, for obvious reasons, would lead to large inaccuracies in the model.

3.2 Energetic Parameters

A model of the real-world requires the presence of real-world forces, see Figure 3.5. Those present in the model are outlined below and most are initially from the works of Martin et al. (1998) unless otherwise specified.

3.2.1 Aerodynamic Drag

The force that opposes motion of an object through a fluid, in this case aerodynamic drag produced by a rider moving through air. This force is arguably the most important for modeling modern road cycling. Not only are vast resources spent by teams with the aim of helping their athletes cut through the air more efficiently but race tactics are often

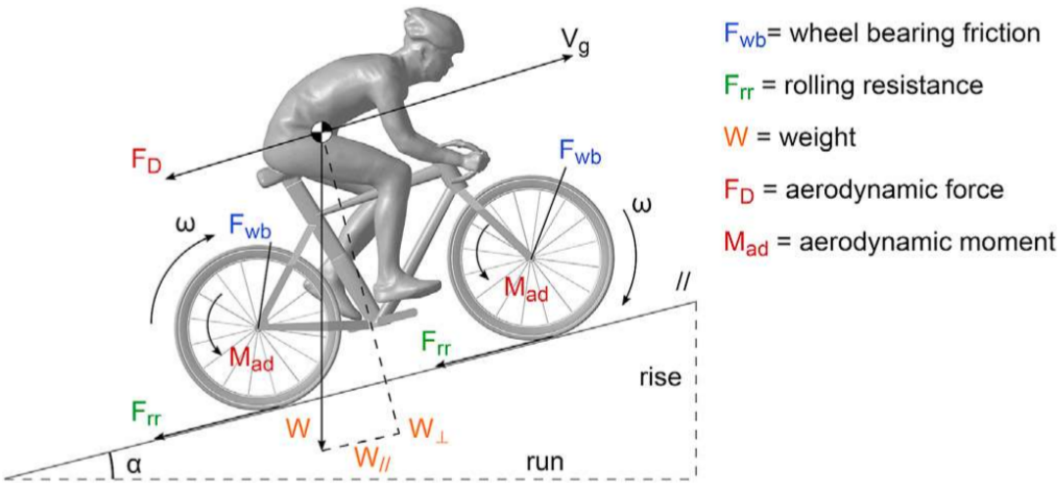


Figure 3.5: Forces Acting Upon a Cyclist, Diagram From Malizia & Blocken (2020)

determined by how much an athlete must "ride in the wind". This concept relates to whether a rider can follow another, making use of the pocket of low pressure behind the leading rider to reduce their own workload. The formula for drag a solo rider is as follows.

$$F_D = 1/2 \rho C_D A V_a^2$$

Where F_D is the Drag Force, rho is air density (in this case assumed a constant $1.225 kg/m^3$), C_D the Coefficient of Drag (calculated from Blocken's work and assumed constant at 0.69 (Blocken, van Druenen, Toparlar, Malizia, Mannion, Andrianne, Marchal, Maas & Diepens (2018) and approximately agreed with by Crouch et al. (2017))), A which is the frontal area of a rider (again taken from Blocken, van Druenen, Toparlar, Malizia, Mannion, Andrianne, Marchal, Maas & Diepens (2018) as 0.423), and finally V_a which is the velocity of the air moving past the rider.

As is commonly known, power is equal to the product of force and velocity. This therefore means that for a solo rider to overcome their own aerodynamic drag they must produce a power equal to $P_A D$ while travelling at a ground velocity of V_G. Giving the following equation.

$$P_{AD} = 1/2 \rho C_D A V_a^2 V_G$$

During this simulation it is assumed that the air is still. Therefore the velocity of the air moving past the rider (V_a) is considered to be equal to the velocity of the rider with respect to the ground (V_g). Meaning the previous equation can be simplified (where V is the rider's velocity).

$$P_{AD} = 1/2\rho C_D A V^3$$

Previous agent based studies discussed make use of the following simplified equation.

$$P_{air} = kCF_{draft}V^3$$

Where CF_{draft} is defined as the drafting coefficient or draft factor. This is the percentage decrease in the drag force, and therefore related decrease in power output, experienced by a following rider due to drafting of a lead rider.

This therefore produces:

$$F_D = 1/2\rho C_D A V_a^2 CF_{draft}$$

$$P_{AD} = 1/2\rho C_D A V^3 CF_{draft}$$

Draft Factor

Within the works of Hoenigman et al. (2011), Ratamero (2013), and then Ratamero (2015), CF_{draft} is calculated through the quadratic formula discovered by Olds (1998) who in turn generated it from the graphs of Kyle (1979). The resulting quadratic formula is outlined below.

$$CF_{draft} = 0.0452d_w^2 - 0.0104d_w + 0.62$$

Where d_w is wheel-to-wheel distance between the proceeding and following rider. This is only considered to be true for $d_w <= 3m$ after-which CF_{draft} is assumed to be 1. This would mean that there is no longer a benefit to be had from drafting more than 3m behind another rider. This assumption has since been found to be incorrect, as can be seen in more modern research. Anecdotally drafting effects of a single lead rider have been measurable up to 20m behind(Dirksen (2017)). Although corroboration of this specific claim could not be found, effects behind a motorcycle can be measured by to 50m away (Blocken et al. (2020) which would suggest claims of 20m are founded. Furthermore use of this formula throughout the entirety of a model would suggest that all riders receive the

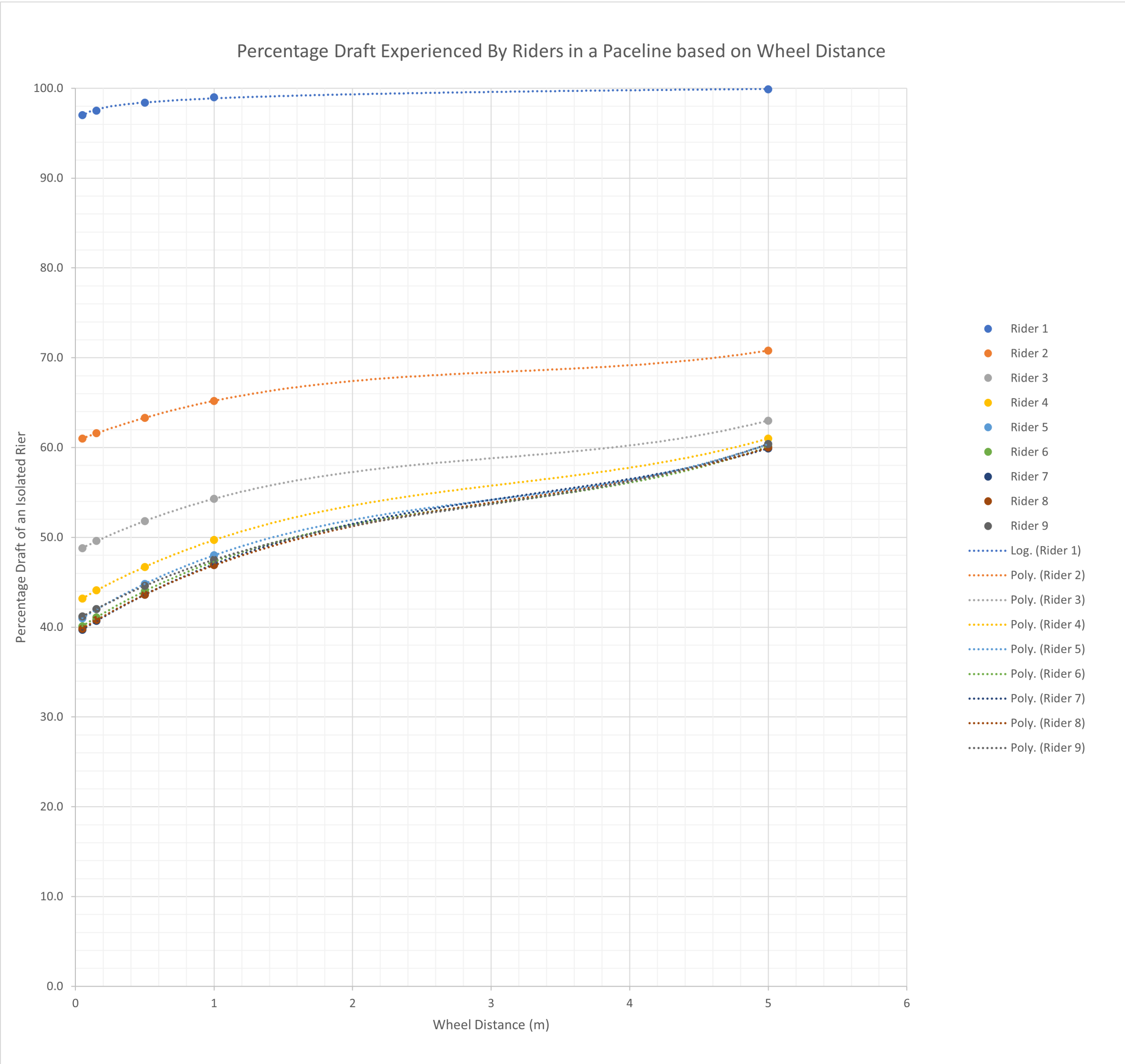


Figure 3.6: Percentage Draft Experienced by Riders in a Paceline, Produced Using Raw Data from Blocken, Toparlar, van Druenen & Andrianne (2018)

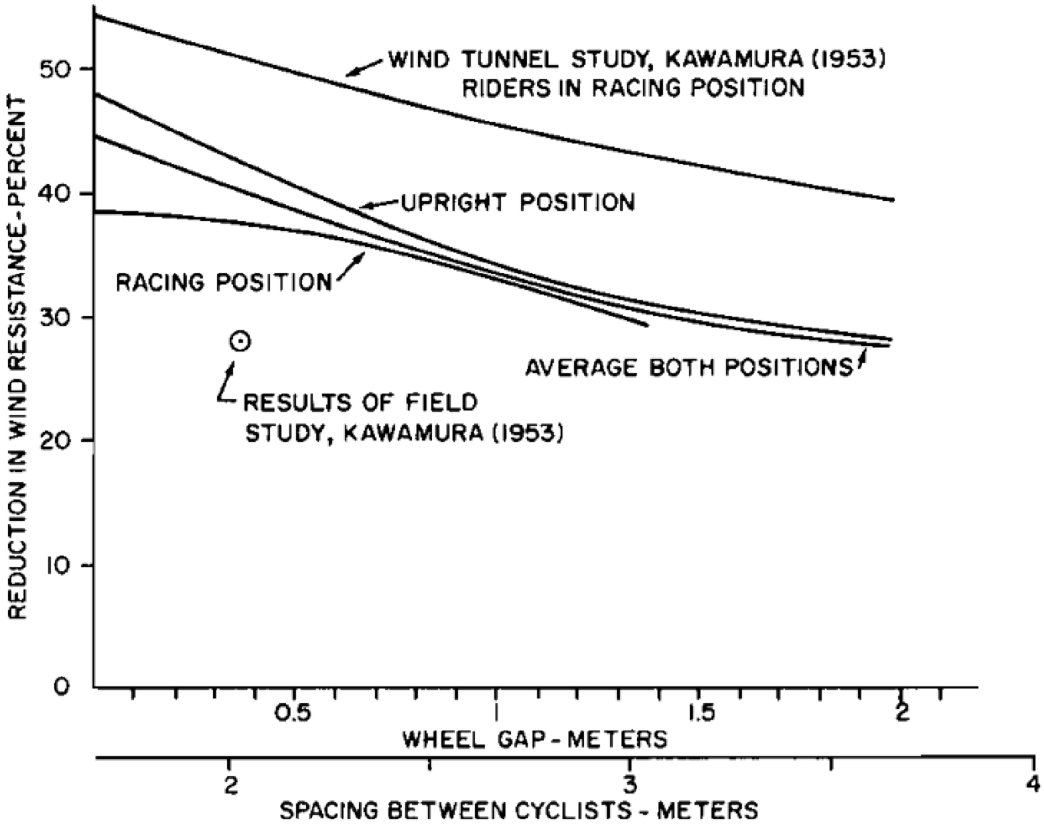


Figure 3.7: Reduction in Wind Resistance with Wheel Gap, Graph From Kyle (1979)

	Position of Rider within line								
Dw (m)	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
0.05	97.0	61.0	48.8	43.2	41.0	40.1	39.7	39.8	41.2
0.15	97.5	61.6	49.6	44.1	42.0	41.1	40.7	40.8	42.0
0.5	98.4	63.3	51.8	46.7	44.8	44.0	43.6	43.6	44.6
1	99.0	65.2	54.3	49.7	48.0	47.3	47.0	46.9	47.5
5	99.9	70.8	63.0	61.0	60.4	60.1	59.9	60.0	60.4

Table 3.1: Percentage Draft of An Individual Rider for Riders within a Paceline

Rider Position	**Trendline Formula**
1	$y = 0.6535ln(x) + 98.88$
2	$y = 0.1738x^3 - 1.6583x^2 + 5.9617x + 60.721$
3	$y = 0.2326x^3 - 2.131x^2 + 7.7499x + 48.445$
4	$y = 0.2394x^3 - 2.2546x^2 + 8.9329x + 42.78$
5	$y = 0.2922x^3 - 2.6225x^2 + 9.779x + 40.547$
6	$y = 0.2906x^3 - 2.6378x^2 + 10.019x + 39.626$
7	$y = 0.2409x^3 - 2.3525x^2 + 9.8722x + 39.236$
8	$y = 0.2492x^3 - 2.3475x^2 + 9.6365x + 39.357$
9	$y = 0.2556x^3 - 2.2372x^2 + 8.7255x + 40.758$

Table 3.2: Trendline Formulas For Riders in a Paceline, For Graph see Figure 3.6

same percentage benefit from the rider in-front independent of rider position within the bunch.

Outlined in Table 3.1 are the draft effects on riders in a line from the work of Blocken, Toparlar, van Druenen & Andrianne (2018). Applying these graphically gives Figure 3.6 which can be seen to show a cubic relationship between the wheel spacing and the percentage of the drag force felt by the following riders. All of these trendline formulae are outlined in Table 3.2. It can be seen that each following rider within the paceline experiences a different version of this relationship depending on their position. Interestingly the leading rider is also seen to benefit from the following riders with this displayed as a logarithmic relationship.

The cubic relationship here will allow riders to experience a draft to approximately 8m which although not the full 20m is still an improvement over the 3m used by the previous models. When this new data is compared to the works of Kyle (1979) there are similarities present. The new trendlines match those displayed by the "Upright Position" or "Average Both Positions" trendlines in Figure 3.7 which is different to the "Racing Position" trendline in the same figure as focused on by Olds. Tests will be carried out using both versions for Draft-Factor:

$$Olds\ CF_{draft} = 0.0452d_w^2 - 0.0104d_w + 0.62$$

$$Blocken\ Based\ CF_{draft} = 0.01 \times (0.1738d_w^3 - 1.6583d_w^2 + 5.9617d_w + 60.721)$$

Angle Correction

During the course of the simulated runs rider agents are going to be receiving a draft off of riders that are not positioned directly ahead of them. To correct for the reduction in the draft due to being off to one side an angle correction factor will be added to the draft-factor. This angle factor will be calculated as:

$$0.01 \times N$$

Where N is the number of the degrees that the rider behind is off of the rider in front. This means if a following rider is drafting in a position with a wheel distance that gives a draft-factor of 0.6 but is offset by 40° to one side they will experience no benefit to the draft. Very little research exists in this sector but this formula does link with the angled results given in Blocken, van Druenen, Toparlar, Malizia, Mannion, Andrianne, Marchal, Maas & Diepens (2018).

3.2.2 Rolling Resistance

Rolling Resistance, also known as rolling friction or rolling drag, is the resisting force present when an object rolls along a surface. In the case of cycling this relates to the tyres rolling along the road. This can be calculated by making use of the formula below.

$$F_{RR} = COS[TAN^{-1}(G_R)]C_{RR}m_T g$$

Where F_{RR} is the rolling resistance, G_R is the Gradient of the surface (rise/run), C_{RR} is the Coefficient of Rolling Resistance (this is dependant upon road surface, tyre material, tyre pressure along with a host of other things) for the purposes of this simulation it is assumed constant at 0.0053, m_T is the total mass of the system (in this case bicycle and rider where the bicycle is $6.8kg$ as per the UCI weight limit and the rider weight is

$70kg$), finally g the standard acceleration due to gravity (as $9.81m/s^2$). Once again the product of this force and the velocity gives the power the athlete is required to produce to overcome this force.

$$P_{RR} = VCOS[TAN^{-1}(G_R)]C_{RR}m_Tg$$

It can be noted that for road grades up to 10%, $COS[TAN^{-1}(G_R)]$ is approximately 1. This has been used by previous studies to simplify this equation, however World Tour races often tackle grades above 10% on deciding stages and therefore it seems prudent to not simply this equation any further.

It is worth noting that during harsh cornering this rolling resistance value is likely to change due to changes in contact patch as well as other changes in tyre behaviour. As discussed earlier for the purposes of this model these have been ignored. Due to the relatively small amount of sideways movement on a straight road this has been deemed an acceptable compromise.

3.2.3 Wheel Bearing Friction

For the purposes of this simulation the friction produced by the wheel bearings have been deemed negligible when compared to the larger forces at play. As a result they have so far been ignored from the model. It is worth noting however that not only can they be calculated using the following formula but also it is common to see teams and athletes replacing standard bearings with ceramic ones in order to reduce this value further, the effectiveness of this is often up for debate.

$$P_{WB} = V(91 + 8.7V)10^{-3}$$

3.2.4 Changes in Potential Energy

This model takes place in three dimensions, due to the changes in elevations the changes in gravitational potential energy need to be taken into account. Gravitational Potential Energy is of course the change in energy held by an object due to change in elevation with relation to the Earth's centre. When this elevation changes, i.e. by riding up or down hill, work is done against or by gravity. This can be calculated in the following formula.

$$W_{PE} = Dm_TgSIN[TAN^{-1}(G_R)]$$

Where W_{PE} is the work done, D is the distance traversed, m_T is the total mass of the bike and rider (for this mass of the rider has been assumed to be 70kg and the mass of the bike has been given as the UCI minimum of 6.8kg making $m_T = 76.8kg$), and G_R is the road gradient. Therefore the power required to produce said work is given by the following equation.

$$P_{PE} = V m_T g SIN[TAN^{-1}(G_R)]$$

Similarly to rolling resistance this equation can be simplified further for roads under 10%. Again this simplification has been ignored due to the nature of road cycling races often crossing over this 10% mark at divisive points within a race.

3.2.5 Changes in Kinetic Energy

As is well known the kinetic energy of a body is related to the mass of a body and its velocity by the following equation.

$$KE = 1/2mV^2$$

Therefore when any change in kinetic energy occurs work must be done on the body. Which can be calculated through the following. Note that the mass is the total mass of both rider and bike m_T.

$$W_{KE} = \Delta KE = 1/2m_T(V_{final}^2 - V_{initial}^2)$$

Power related to kinetic energy (P_{KE}) is the rate of change of kinetic energy.

$$P_{KE} = \Delta KE/\Delta t = 1/2m_T(V_{final}^2 - V_{initial}^2)/(t_{initial} - t_{final})$$

For this model the $V_{final}^2 - V_{initial}^2$ can be equated through the magnitude of the velocity vectors. The time interval over which this change in kinetic energy occurs will be 1 *second* which relates to one iteration of the model.

Note, there is additional kinetic energy stored in the rotating wheels of the bicycle. For the purposes of this model this has been deemed to be negligible. An interesting note if this were to be included in future would be the impact of additional rotating mass with respect to aerodynamic gains from heavier/deeper carbon wheels when compared to lighter/shallower ones.

3.2.6 Total Power Requirements

To get the net power required to overcome all of the previously discussed forces we sum all of these preceding equations to give:

$$P_{NET} = P_{AT} + P_{RR} + P_{WB} + P_{PE} + P_{KE}$$

3.2.7 Frictional Loss within the Drive Train

When riding a bicycle due to the nature of the drive train, a chain drive will multiple gears. There is a loss in the system. The power produced by the rider is therefore not equal to the power delivered to the rear wheel to propel the vehicle forward. The following equation can be used to display this relationship.

$$P_{TOTAL} = P_{NET}/E_c$$

P_{NET} is the power required at the rear wheel to propel the athlete, the E_c is the chain (or drive-train) efficiency.

For the purposes of this model the drive train and bearings are considered to be 100% efficient therefore:

$$P_{TOTAL} = P_{NET}/E_c = P_{NET}/1 = P_{NET}$$

3.3 Physiological Parameters

Up to this point the model is only focused on the energy expenditure of each rider, from here it is necessary to define how a rider reacts to having to spend this amount of energy. For this section we will focus solely on biological limitations of riders through some well known sporting concepts.

Exercise itself can be divided into two categories, aerobic and anaerobic which relate to "with air" and "without air". When a body is unable to get enough oxygen during exercise it switches to anaerobic respiration a by-product of which is lactate/lactic acid and is what causes the well known burning sensation bought on by high intensity exercise. This switch between aerobic and anaerobic happens gradually, however at a point the body is unable to clear enough of the lactate by-product and therefore the concentration of it within the athlete's blood begins to rise rapidly. The point of this change is known as the lactate threshold.

As touched on the switch between aerobic and anaerobic occurs when the body is unable to get enough oxygen therefore an athlete's level of oxygen uptake is important. This characteristic is measurable and referred to as VO_2max, the maximum rate at which a body can take in a volume of oxygen (given in l/min).

3.3.1 Time to Exhaustion

So far we have just discussed defined characteristics of riders (and therefore agents), this still doesn't relate to an athlete's ability to work at that level of extrusion. This will be done in a similar method to both Hoenigman et al. (2011) and Ratamero (2013) by making use of the following equation from Olds.

$$ln(T_{lim}) = -6.351ln(fV0_{2_{max}}) + 2.478$$

This equation calculates the time to exhaustion (T_{lim} in $mins$) of a rider when they are working at a fraction (f) of their maximum oxygen uptake ($V0_2max$). Hoenigman et al. (2011) stipulates that this fraction of maximum oxygen uptake ($fV0_2max$) can be replaced by the ratio of the riders current workload and the riders maximum capable workload for 10 minutes. Giving the following

$$fV0_{2_{max}} = P_{TOTAL}/Max_{10}$$

Where P_{TOTAL} is the average of the power output of the rider over a given minute, calculated through the methods outlined previously. With Max_{10} being the Maximum Power output by a rider for 10 minutes.

This substitution results in:

$$T_{lim} = e^{-6.351ln(P_{TOTAL}/Max_{10})+2.478}$$

Enabling the Time to Exhaustion of a rider to be calculated based on their power output for the last minute and their Maximum Power output over a 10 minute period.

3.3.2 Energy Level

To mimic the energy reserves of a real-world athlete each athlete will start the simulation with an amount of energy which will be reduced with each iteration of the run. To calculate the energy reserve at the start of the test run we will consider a lone rider. A lone rider travelling over a test course of 25km at 45km/h will take just over 30 minutes or exactly 2000 seconds. An average rider of $W/kg = 7.1 or Max_{10} = 500$ would use $1MJ$ of energy. However when riders are operating below recovery power $0.5 \times Max_{10}$ they are deemed to not be using up energy reserves as they can eat enough to maintain energy levels. Therefore the amount of energy riders will need at the start of the simulation outlined is $500kJ$. When riders reach $100kJ$ remaining they are deemed exhausted and their workload is limited.

3.4 Decision making

The final section of the model is to have each agent display tactical decision making. This is implemented in-line with the Hoenigman et al. (2011)'s approach with slight modifications made. Each rider has a co-operation characteristic that defines their probability of choosing a co-operative or defective approach. This will be between 0, never cooperates, and 1 will always cooperate. Within these two defined approaches there is an active and non-active role that are chosen based on the rider agent's current level of energy expenditure. There are some important distinctions to made between what the result of a decision will be as well as how long a rider agent will have to stick with that decision for.

3.4.1 Attacking - Defective and Active

An attack is triggered in two stages, the first is for the rider to decide to defect and the second is if the rider's group's current level of excursion is less than 80% of the individuals Max_{10}. To produce an attack of a realistic nature, the agent's seek target will be made the next point in the route as well as adjust the agent's power to above 90% of the riders Max_{10}, thereby increasing the agent's velocity. This attack will last 3 minutes at which point the rider must stop attacking and pick a co-operative option. Replicating how a real-world rider must leave breaks between attacks.

3.4.2 Pulling a Turn - Co-operative and Active

Co-operative riders who are active replicate a real-world rider's behaviour by "pulling a turn". This role will see riders move toward the front of the peloton by seeking the front rider of their group, once at the front the seek target will change and the rider will "pull a turn". This is when a rider rides on the front of a group creating a draft for those behind. An agent will pursue this behaviour for 5 minutes. Once the rider has "pulled a turn", this can either be successfully making it to the very front of the bunch and offering a draft or spending 5 minutes trying to get to the front, their behaviour is changed to allow the rider to rest.

3.4.3 Seeking a Draft - Non-active

Both defective and co-operative riders can become non-active, when agents are in this category their purpose is finding the best draft in the bunch. This is done by checking the draft-factor forward, backward, left, and right of the riders current position and then applying a force vector in the direction of an improved draft-factor should one exist. Although both active and non-active agents express the same behaviour the key difference is the power expenditure at which both are initiated, a defective rider will always hunt the draft when struggling at 80% or above of their Max_{10} whereas a co-operative rider will often try their best to "pull a turn" even if it means damaging their own chances of being able to complete the simulated race run at all let alone position well. Finally it it worth mentioning that there are no limits on the number of times a rider can sit in the group without becoming active.

3.4.4 Finishing

A final point to all of this is that this behavioural dynamic changes as riders approach the finish. Similarly to Hoenigman et al. (2011) there is an override with the last 5km of a race where agents no-longer respond to the standard decision making and aim to finish with as little energy as possible. This has the effect of causing agents to "ride" as fast as they are physically able to, replicating a real-world rush for the line.

Chapter 4

Results

The model itself will allow for multiple variations of tests to be carried out depending on the parameters altered. The most common parameters to change would include changes to rider properties which sometimes impact behaviour as well as changes to the route being tested over. Then going a step further elements of the model itself can be changed to identify where improvements can be to simulate as close to the real world as possible.

Before any changes to the model itself are made it is worth comparing it to those of previous research, to do this the model outputs two distinct categories of results. The first category mimics elements of the results of Hoenigman by focusing on the how a singular rider with certain pre-determined characteristics performs in relation to the overall group of riders. Whereas the second category monitors factors of multiple riders in real-time to determine how riders interact with one another throughout the course of the simulated run. These two categories of testing will be run twice. The primary testing will be with the intention of comparing the model to those produced in the past. The secondary round of testing will involve the changing of the draft-factor formula with a focus on seeing how these results compare to those of the primary testing with a focus on whether this new Blocken based cubic formula is an improvement over the Old's quadratic.

Although focused on different aspects all of the experimental categories will require some common ground. All will be run over the same test courses. Generated by hand as a test environment, each course is a mapped distance of 25km (that is 25km from start-to-finish on the same plane, disregarding any changes in elevation) with different course gradients of 0% and 10%. Note, this obviously means that there will be a slight variation in overall distance travelled by agents when comparing between each course. In addition to this all of the agents within the simulated runs will have similarly defined characteristics of Max_{10} in watts-per-kilo and co-operation tendency. Each will be normally distributed within the ranges of $6.3 <= Max_{10} <= 8.3W/kg$ and $0.2 <= C_b <= 0.8$ respectively with means of $\mu = 7.1W/kg$ and $\mu = 0.48$, along with a standard deviation of $\sigma = 0.4W/kg$ and $\sigma = 0.2$. At this stage it is worth noting that although these parameters are defined, within this there is space for varience/randomisation meaning the likelihood of any run being exactly the same as a previous run is low. Furthermore, when athlete agents are generated they are placed within a 50m by 5m box at the start of the course with a random velocity, between $0 - 5m/s$ in the x direction and $-1 - 1m/s$ in the y direction, in the direction of the course. This is to simulate the start of a bunch race within road cycling where racers gather in a defined area at the start but in no particular order. It is worth mentioning that in some road cycling races that this is then followed by a "neutral zone" where riders spend time behind a the commionsaire's (official's) car, usually working their way out of a town or city and onto larger more open roads before the race is truly started. This is currently out of scope of the project but it would seem pertinent to model this in the future given that although the peloton is not moving at pace this is usually when riders are the most densely packed and the roads are narrowest meaning accidents at this stage can be just as costly.

4.1 Old's Draft-Factor Tests

4.1.1 Hoenigman Style Tests

To replicate the work of Hoenigman the finishing position of a rider with defined characteristics was recorded. The recorded rider was modified for varying parameters of between $6.3 <= Max_{10} <= 8.3W/kg$ with increments of $0.4W/kg$ and $0.2 <= C_b <= 0.8$ with increments of 0.2. This gives 24 variations covering a range of possible skill levels and cooperation tendencies. Each variation is run 10 times giving a total number of test runs of 240 per course. It is worth mentioning at this stage that Hoenigman et al. (2011) ran significantly more tests, including cooperation tendencies at extremes of 0 (never cooperate) and 1 (always cooperate), as well as completing 50 runs per variation which gives a total number of runs of 1800. In addition Hoenigman tested over a course length of 160km with a field of 150 riders. Unfortunately recording results from simulations of this nature using this model is simply not feasible within current constraints on time and computing power. Current 240 run bulk tests of the nature described can take up to 24 hours to complete, depending on the gradient of the course being used.

0% Gradient Course

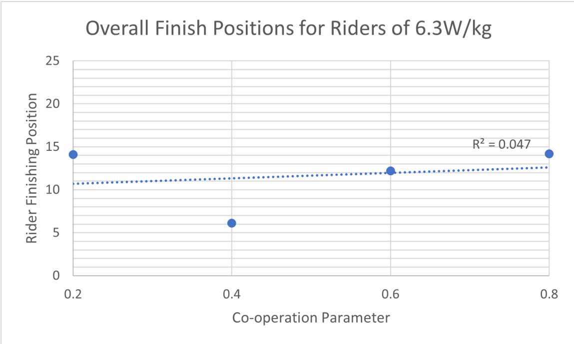
This section contains the graphical results of the average finishing positions of the target rider when defined at different levels of skill and cooperation while being tested over a 25km course at a 0% gradient. All of these results can be seen in Figure 4.1.

The best overall average result when compared across all runs for all possible combinations is for the strongest rider at $8.3W/kg$ with an average placing of 10.6, this in itself is unsurprising (Figure 4.1f). However it is interesting that further down the line, weak riders of only $6.7W/kg$ (Figure 4.1b) and $6.3W/kg$ (Figure 4.1a) appear to outperform stronger riders of $7.9W/kg$ (Figure 4.1e) and $7.1W/kg$ (Figure 4.1c). Although this is only by approximately a single place this could be a sign that riders closer to the mean rider strength miss out. Perhaps their strength level places them in an awkward limbo point for strategy decisions where neither approach enforced is preferable. It is possible that this observation could be reinforced by inspecting the average individual best result across combinations. Here the weakest rider produces a best placing of 6.1 closely followed by the strongest rider at 6.3. Whereas again the worst in this case is the rider with the mean strength. That being said it is worth noting that for both the strongest and weakest rider these results are outside the general trend, see Figure 4.1f and 4.1a, which may suggest that these are rather anomalous results interest.

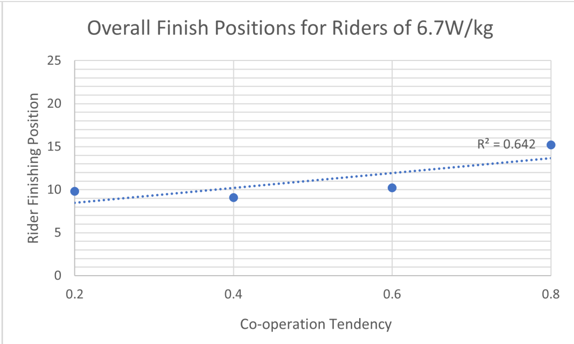
10% Gradient Course

This section contains the graphical results of the average finishing positions of the target rider when defined at different levels of skill and cooperation while being tested over a 25km course at a 10% gradient. See all results in Figure **??**.

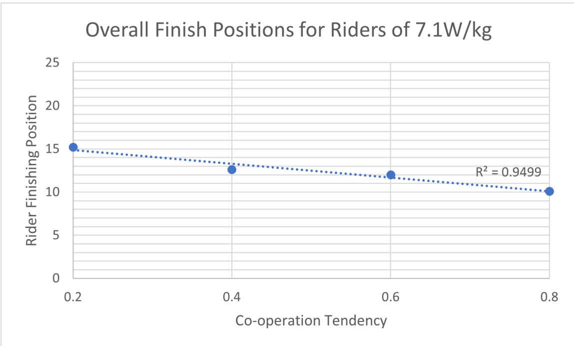
The best performing rider on average over all test runs is the one with a Max_{10} of $7.1W/kg$ (Figure 4.3c). This is different to that of the results of a flat course and the expected result of the strongest rider performing the best. In addition to this the agent with a Max_{10} of $7.1W/kg$ also gets the best individual result out of all skill and cooperation variations. This is obviously the opposite of the results for the flat course where riders close to mean strength were seen to perform the worst. A possible explanation of this change is the power demands over a course with an increased gradient. Obviously on a gradient



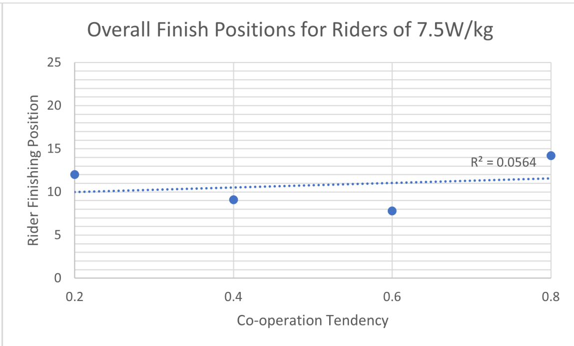
(a) Rider of $Max_{10} = 6.3W/kg$



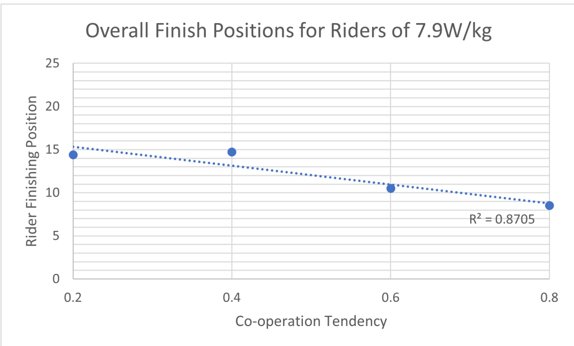
(b) Rider of $Max_{10} = 6.7W/kg$



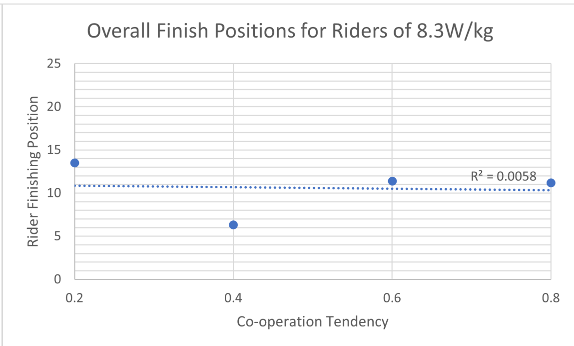
(c) Rider of $Max_{10} = 7.1W/kg$



(d) Rider of $Max_{10} = 7.5W/kg$



(e) Rider of $Max_{10} = 7.9W/kg$



(f) Rider of $Max_{10} = 8.3W/kg$

Figure 4.1: Mean Finishing Position for Riders of Multiple Skill Level and Co-operation Strategy over a 25km Course of Gradient 0%

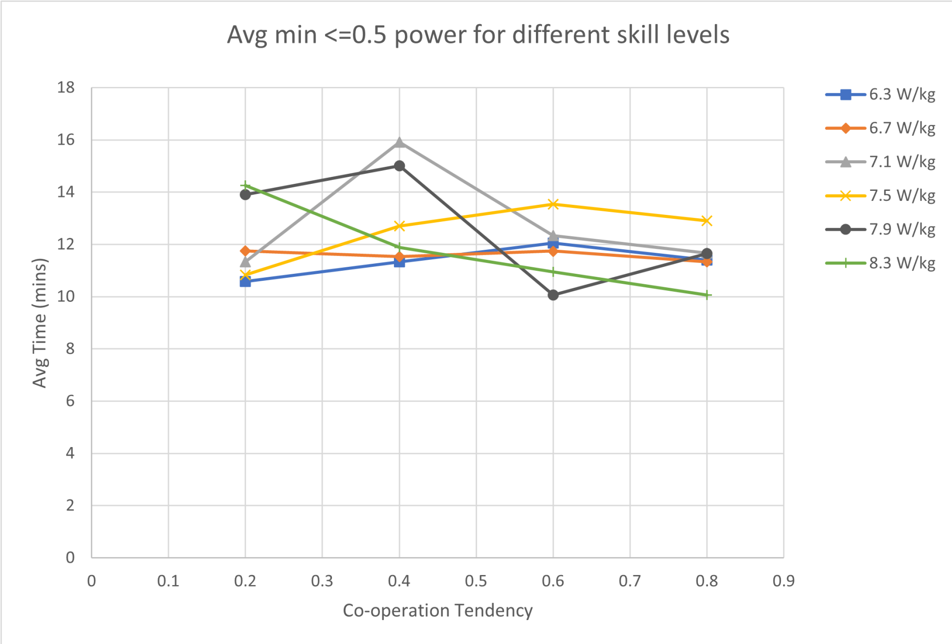
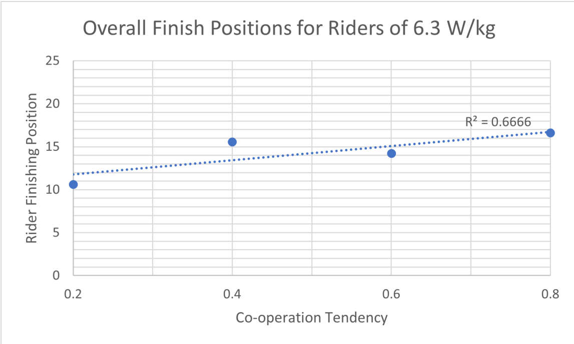


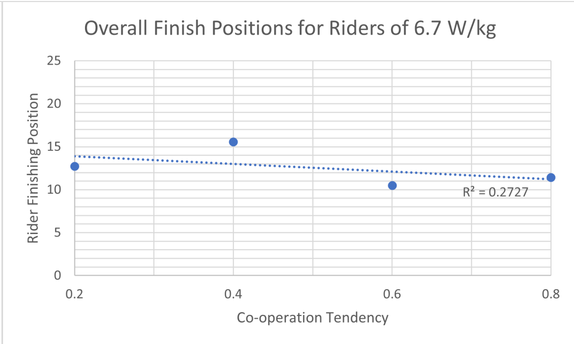
Figure 4.2: Minutes Spent at 50% or Above of Max_{10} For All Rider Strength Levels over a 25km Course of Gradient 0%

riders either must increase their power output to travel at the same velocity or choose to travel at a lower velocity for the same power output. The later, reduction in velocity, causes a reduction in the advantage given by drafting the riders in-front. This would explain why weaker riders are seen to perform so much worse over this route when compared to the flat route. It could also go onto explain why the mean strength riders perform so much better. Rather than being stuck in a situation where they flip-flop between strategies the increased power demands cause the mean strength riders to take the same approach as the weak riders, sitting in the draft and saving energy before attacking near the end. This strategy now fails for the weaker riders because they lack the power necessary to stay with the bunch even in the draft due to the reduced drafting benefit.

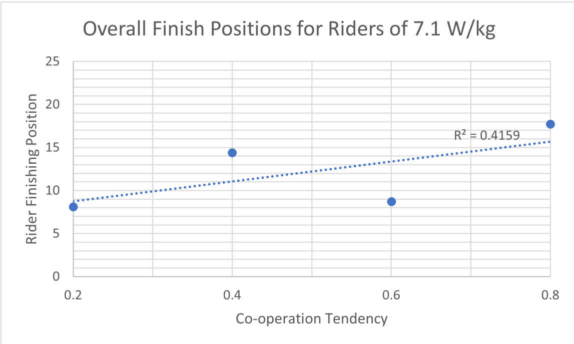
When comparing cooperation tendencies this is less clear cut than on a flat course. On the flat course it can be seen that: the 2 weakest riders both favour a cooperative strategy; the 2 strongest riders both favour a defective strategy; and the 2 middle riders sit the opposite sides of this than expected. On the 10% course extreme weakest and strongest rider maintain this but riders between $6.7W/kg - 7.9W/kg$ all switch preferred strategy. The reason for this behaviour is somewhat unclear.



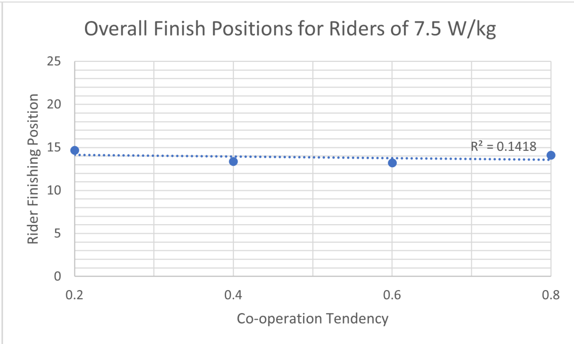
(a) Rider of $Max_{10} = 6.3W/kg$



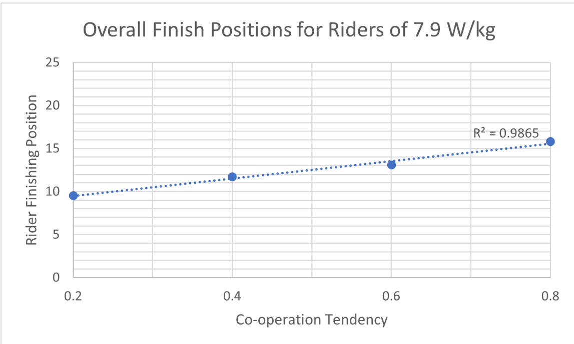
(b) Rider of $Max_{10} = 6.7W/kg$



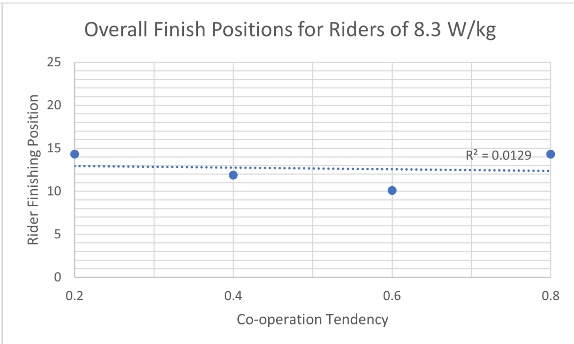
(c) Rider of $Max_{10} = 7.1W/kg$



(d) Rider of $Max_{10} = 7.5W/kg$



(e) Rider of $Max_{10} = 7.9W/kg$



(f) Rider of $Max_{10} = 8.3W/kg$

Figure 4.3: Mean Finishing Position for Riders of Multiple Skill Level and Co-operation Strategy over a 25km Course of Gradient 10%

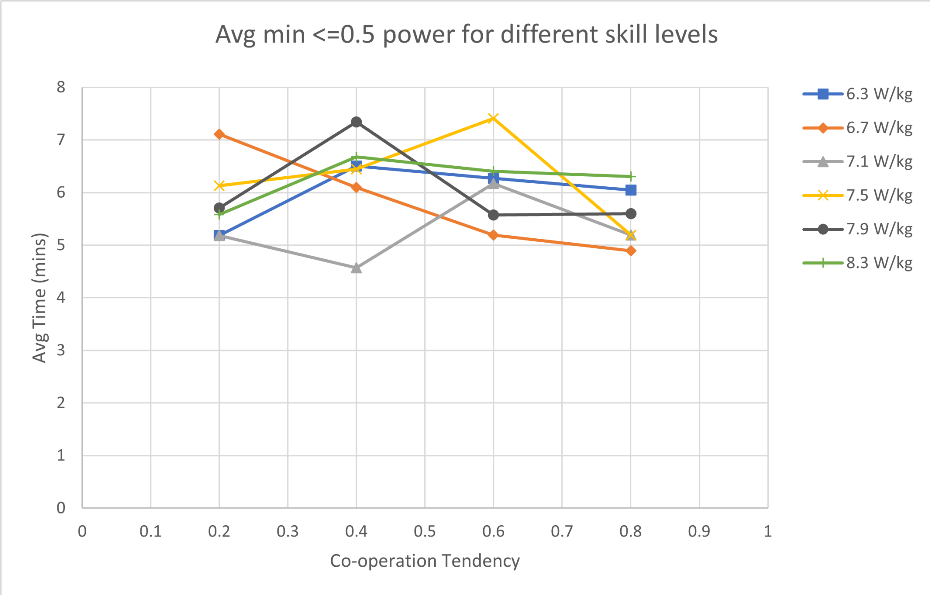


Figure 4.4: Minutes Spent at 50% or Above of Max_{10} For All Rider Strength Levels over a 25km Course of Gradient 10%

Comparison to published results

When comparing these results to those of Hoenigman we can see that both models produce similar results for finishing position of riders. With riders of $6.3W/kg$ and $7.9W/kg$ where Hoenigman's results are available the same trend in co-operation tendency can be observed. It is unfortunate that comparisons cannot be made with rider agents of other strengths due to lack of results given by Hoenigman in their paper. Furthermore how Hoenigman's results are impacted by course choice is not discussed.

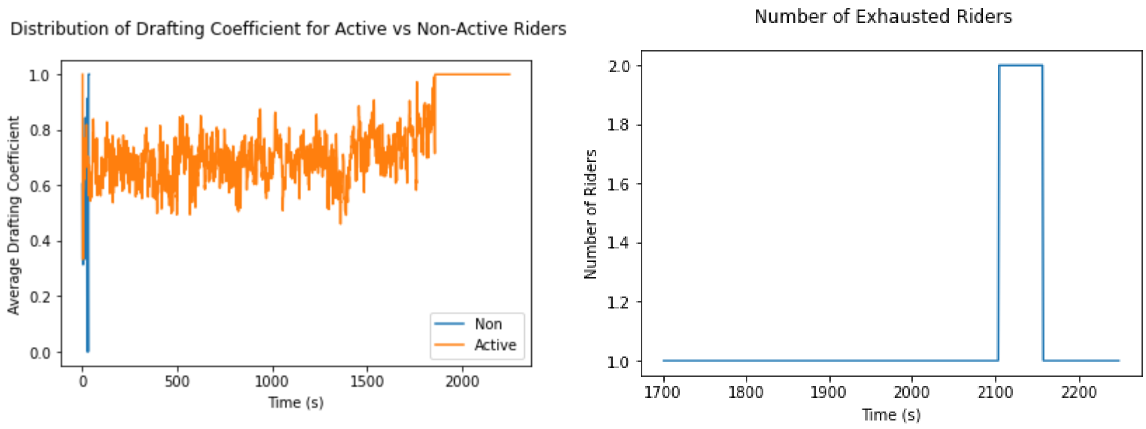
Comparisons between the time spent at less than 50% of Max_{10} power level of different riders is slightly less favourable (Figure 4.2 and Figure 4.4). Hoenigman demonstrates some relatively clear trends across three riders of different strengths ($6.3W/kg$, $7.1W/kg$, and $8.3W/kg$). With the weakest rider working considerably harder as their level of co-operation increase. While the middle and strongest rider work harder to a peak at 0.6 cooperation tendency before starting to tail off as cooperation tendency continues to increase. This behaviour is visible in some specific circumstances, $6.3W/kg$ and $7.5W/kg$ at 0% (Figure 4.2) as well as $6.3W/kg$, $7.5W/kg$, and $8.3W/kg$ at 10% (Figure 4.4), this is not the case across the board which may indicate a significant difference between the models.

4.1.2 Ratamero Style Tests

This second category of test focuses on the behaviour of the riders throughout the simulation rather than on solely end results. To do this properties have been recorded in real time through the simulated run. Each rider is divided into one of three statuses: active, non-active, and exhausted. A rider is defined as active if they are either attacking or pulling a turn. A rider is non-active if they have chosen to sit in the draft. A rider is exhausted if their energy remaining is below 100kJ. The field is again made up of 25 riders defined with the same characteristics as discussed for the previous test. The measured rider is given characteristics equal to that of the mean rider to ensure it doesn't negatively impact the run. Again the runs are completed over a distance of 25km. The number of repeat tests is significantly lower than that for the previous category of tests at only 3 repeat runs. This is because the real-time recorded data is an expression of the behaviour of a simulated run and therefore does not lend itself to being averaged over multiple runs in the same manner as the previous category. Although there is a reduction in number of repeats which would suggest that the length of route or number of riders could be increased this is unfortunately not possible. As data is collected for every rider every second the simulation runs there is a large quantity of data to process post each simulated run. Increasing length or number of riders would significantly up this quantity of data and increase processing time. To try and reduce this processing time the data has been plotted straight to graphs rather than being exported as raw data before being graphed.

0% Gradient Course

Although all three test runs outlined here are different there are a few key points that can be drawn from each test. To start with there are very few non-active riders. This can be seen through the lack of a blue line for much of runs 1 and 3 (Figure 4.5 and Figure 4.7), even in run 2 (Figure 4.6) much of the race simulation occurs without a non-active rider hence the line connecting approximately 50 seconds straight to 1500 seconds which is a point of no non-active riders being present. With the lack of any non-active riders it is therefore understandable that the number of exhausted riders is so low across the 3 runs, with the number of exhausted riders only ever maxing out at 4 riders. A point of note, as can be seen in run 2 and the exhaustion graphs, is that non-active or exhausted riders change state to active as the end of the course approaches this replicates how even the tiredest riders try



(a) Mean Draft Factor of Riders Over Time (b) Number of Exhausted Riders Over Time

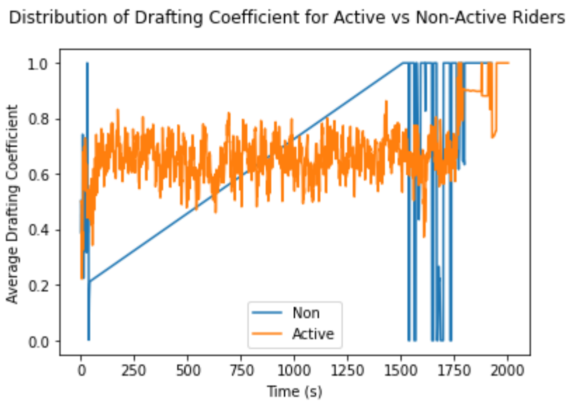
Figure 4.5: Time Based Results for Riders on a 25km Course of Gradient 0%, Run 1

their hardest to get to the line as quickly as possible.

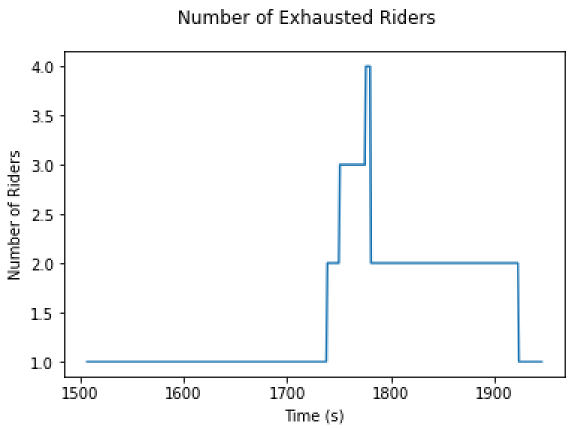
It is worth commenting on the fluctuations of the average drafting factor of all active riders. These rapid fluctuations are expected as riders change positions within the bunch. The most interesting thing however is the trend for this value to increase over the course of a race suggesting that just like in a real race the opportunities to get a good drafting position reduce as the bunches get smaller and smaller towards the finish. Finally it can be seen that the average tops out at 1, this is likely due to their being a few or perhaps even a single rider left on the course. None of these riders have the opportunity to draft off of one another so come into the finish separately experiencing full aerodynamic drag.

10% Gradient Course

It is worth noting again that all three of these results are different due to variation in the simulated runs however as with previous results there are trends across the three that allow a comment to be made about the model. Starting with non-active and exhausted riders. Over the harder course we can see an increase in the level of non-active riders as well as an increase the the amount of riders to be exhausted when compared to the results from the flat course. This appears to happen in stages. You get the start where a few riders are classed as non-active, as with the flat course, but then there is a second wave of non-active riders around the 2000 second mark in all three of the results. This could be linked to how

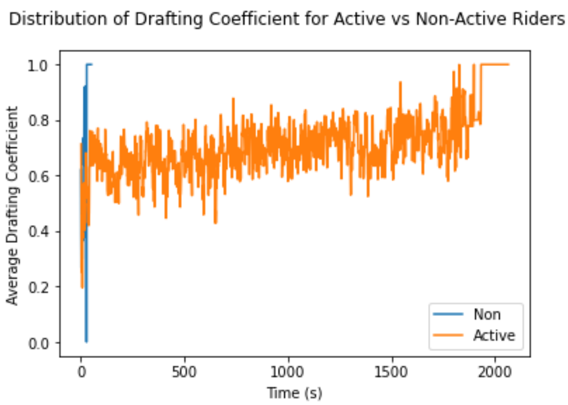


(a) Mean Draft Factor of Riders Over Time

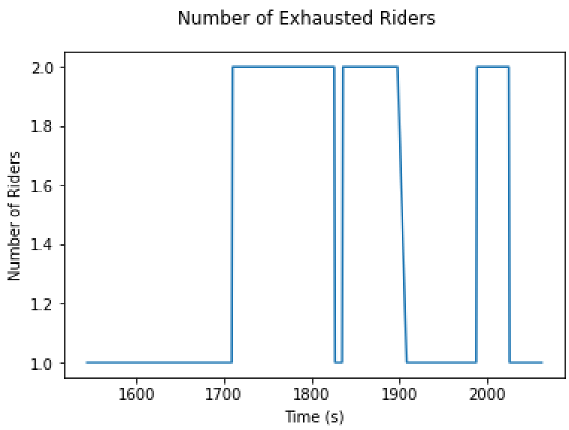


(b) Number of Exhausted Riders Over Time

Figure 4.6: Time Based Results for Riders on a 25km Course of Gradient 0%, Run 2

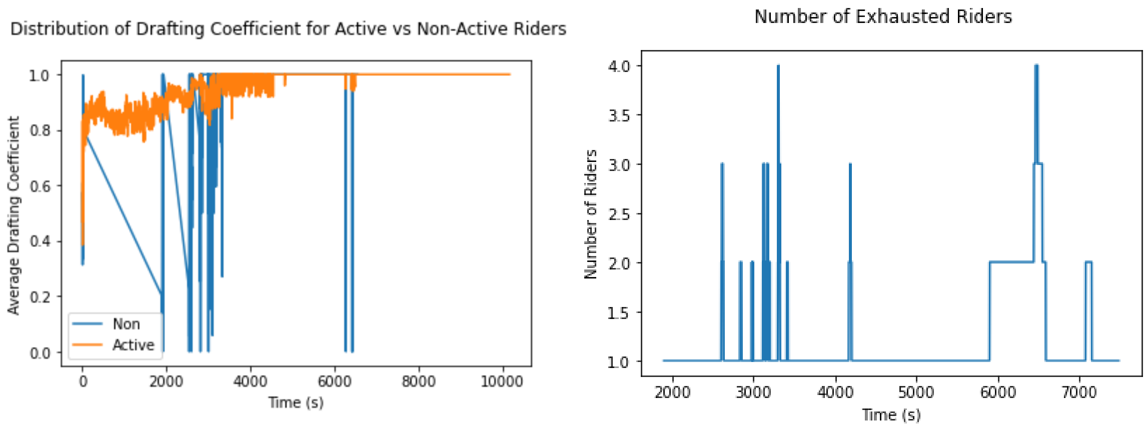


(a) Mean Draft Factor of Riders Over Time



(b) Number of Exhausted Riders Over Time

Figure 4.7: Time Based Results for Riders on a 25km Course of Gradient 0%, Run 3



(a) Mean Draft Factor of Riders Over Time (b) Number of Exhausted Riders Over Time

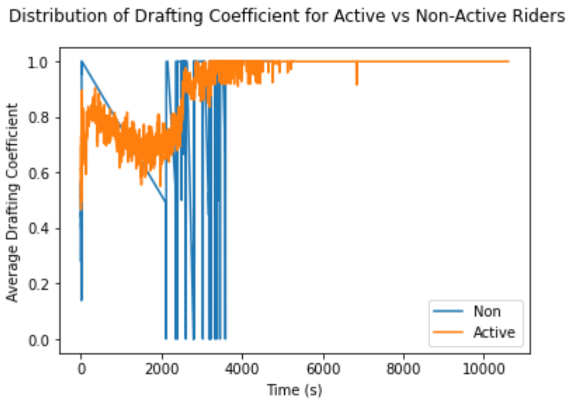
Figure 4.8: Time Based Results for Riders on a 25km Course of Gradient 10%, Run 1

at the start of a road race you get a period where lots of riders attack followed by a relaxing and regathering of the riders behind once a selection has been made.

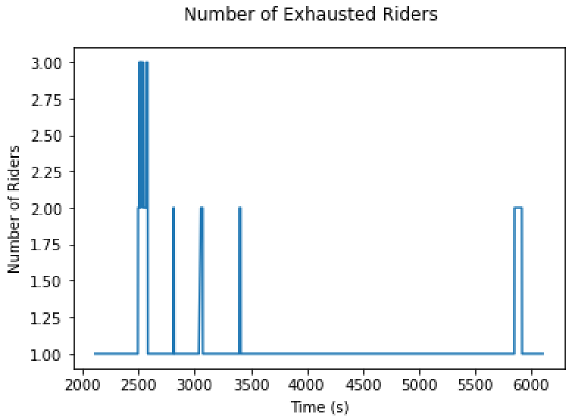
As with the previous results there is a trend for the average drafting factor to increase through the race as the effective draft available to riders decreases with bunch size. This appears more pronounced due the compacting of the graphs due to the increased overall time to complete the course by all riders but is likely to still be a fairly similar trend. An interesting point is the dip against this trend in run 1(Figure 4.8) and even more pronounced in run 2 (Figure 4.9). This backs up the theory that riders have attacked and then regrouped with the percentage aerodynamic drag the riders experience increasing (as attacks happen) before decreasing (as regrouping occurs) which in then followed by increased inactivity (as riders sit in and try to recover). As with the flat course results a single or small collection of separated riders come to the finish well after everyone else. This is even more evident on the more difficult course as would be the case in the real world.

Comparison to Published Works

With a direct comparison between these results and those produced by Ratamero (2013) the active riders can be seen to show much the same behaviour. The drag factor is comparable across both data sets which should be the case. The only slight difference is that these results have a tendency to raise this value towards the finish. As discussed this is with relation to agent pace and group size change towards the finish. Ratamero (2013) did not

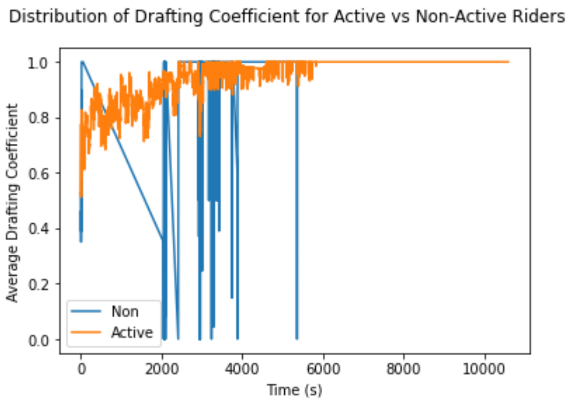


(a) Mean Draft Factor of Riders Over Time

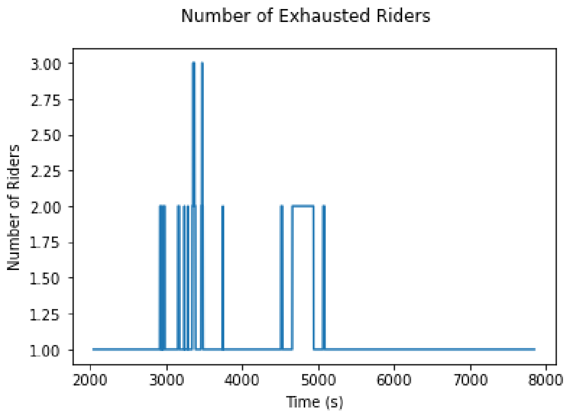


(b) Number of Exhausted Riders Over Time

Figure 4.9: Time Based Results for Riders on a 25km Course of Gradient 10%, Run 2



(a) Mean Draft Factor of Riders Over Time



(b) Number of Exhausted Riders Over Time

Figure 4.10: Time Based Results for Riders on a 25km Course of Gradient 10%, Run 3

model this behaviour of increasing pace toward the finish and therefore this difference is expected.

In contrast to that discussed above the non-active riders seem to show very different behaviour between the two studies. This drastic difference suggests a possible disparity in what the definition of an non-active rider is within the model. Ratamero (2013)'s model forces this status change upon the agents to create the rotational convection behaviour seen within a racing peloton. However this model allows rider agents to change their own status based on perceived parameters with the aim of riders producing rotational convection behaviour on their own. Obviously these two status changes occur under less similar conditions than expected. Unfortunately due to this dissimilarity it is difficult to get an accurate description of the draft factor for non-active riders.

Exhausted riders feature within both models with exhaustion more likely over harder sections of course. In both models these riders can change state between active and exhausted depending on level of excursion. Ratamero (2013) seems to suggest the rate of exhaustion should relate closely to the gradient however it is not indicated over what gradients this model was tested. This relation is difficult to see in the results from this model, possibly due to the small sample size meaning smaller variations appear larger when graphed. It might also be possible a too shallow gradient has been used to test over however it is difficult to confirm this. What can be said for certain is there is a difference in how exhausted riders behave in each model. Ratamero (2013)'s model places exhausted riders within reach of the main peloton allowing them to rejoin once recovered. This model does not do this but rather replicates the real-world in that once a rider has been dropped from the group they are in they have very little, to no, opportunity to re-join it without working considerably harder (which is often not possible physically).

4.2 Blocken Based Draft-Factor

As discussed earlier, previous research in this topic has made use of the Old's formula for draft factor which is used to calculate drafting effects. Since Old's formula was published further work has been done into the aerodynamics of cycling. Therefore a repeat of the previously outlined tests has been carried out, this time making use of the Blocken based cubic formula for draft factor presented within the methodology section.

4.2.1 Hoenigman Style Tests

It has been chosen not to carry out these tests under the new Draft-Factor model. This is mainly because tests of this nature take a long time to run in their current format. When combined with the fact that changes in these results should be predictable once the change of draft-factor is understood through Ratamero style tests, the best option with current restraints has been deemed the removal of these tests entirely.

If Ratamero style tests demonstrate that the average draft-factor has been reduced we would expect to see better results for those combinations of power level and co-operation tendency that make increased use of drafting, such as weaker riders. It could therefore be assumed that the opposite would be true of defecting riders. It would then be reasonable to assume that this situation would be reversed should average draft-factor be shown to increase, indicating a reduction of the effect of drafting. These assumptions are made while acknowledging that they are only that, assumptions. To be concrete on these further testing is necessary but under current restraints not feasible.

4.2.2 Ratamero Style Tests

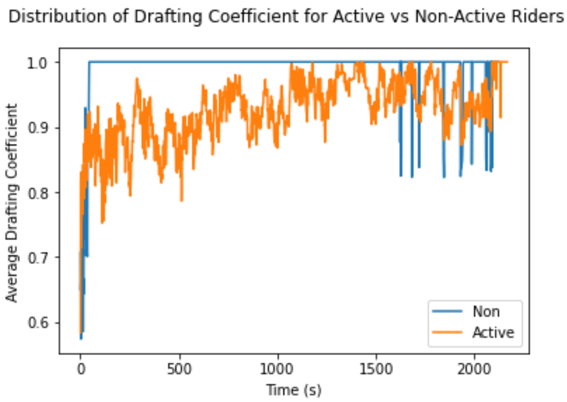
This style of test is focused on the real-time behaviour of rider agents throughout the simulation. The tests were carried out to the same specification as used previously for these kinds of tests. The only difference present was that the draft-factor formula was adjusted to the be Blocken based cubic formula.

0% Gradient Tests

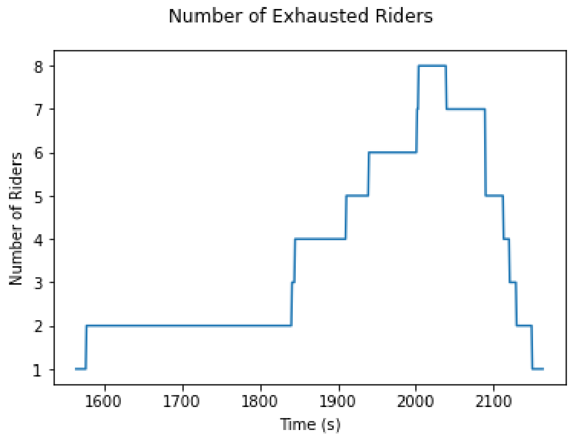
Shown in Figure 4.11, Figure 4.12, and Figure 4.13 are the test results of three runs of the simulation over a 25km course with a gradient of 0% where the draft-factor formula has be modified to that of the Blocken based cubic.

10% Gradient Tests

Shown in Figure 4.14, Figure 4.15, and Figure 4.16 are the test results of three runs of the simulation over a 25km course with a gradient of 10% where the draft-factor formula has be modified to that of the Blocken based cubic.

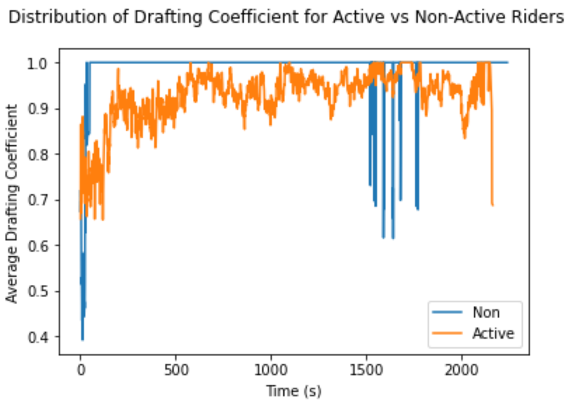


(a) Mean Draft Factor of Riders Over Time

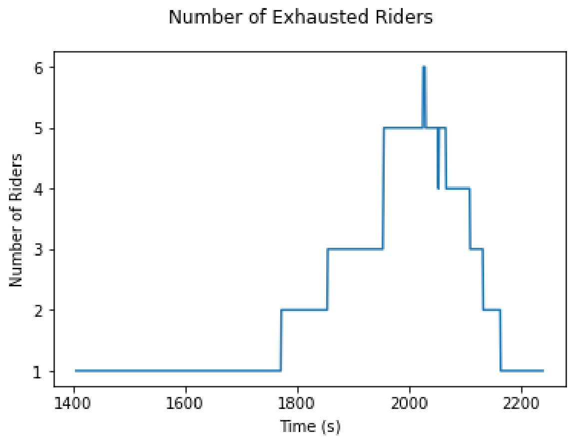


(b) Number of Exhausted Riders Over Time

Figure 4.11: Time Based Results for Riders on a 25km Course of Gradient 0% Using Blocken Based Formula, Run 1

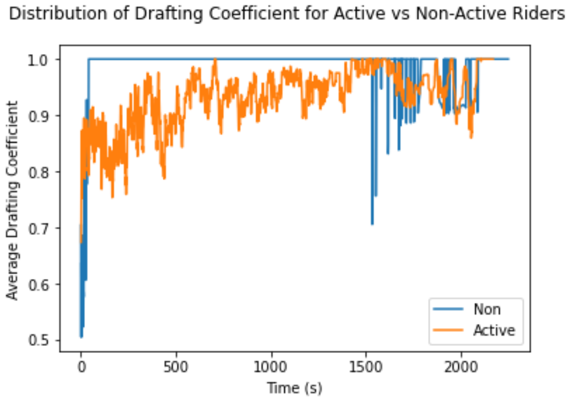


(a) Mean Draft Factor of Riders Over Time

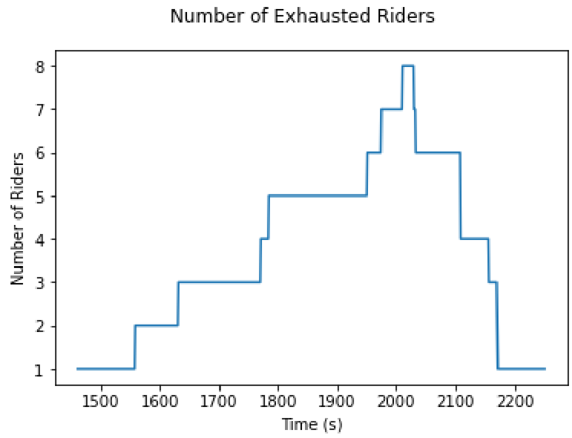


(b) Number of Exhausted Riders Over Time

Figure 4.12: Time Based Results for Riders on a 25km Course of Gradient 0% Using Blocken Based Formula, Run 2

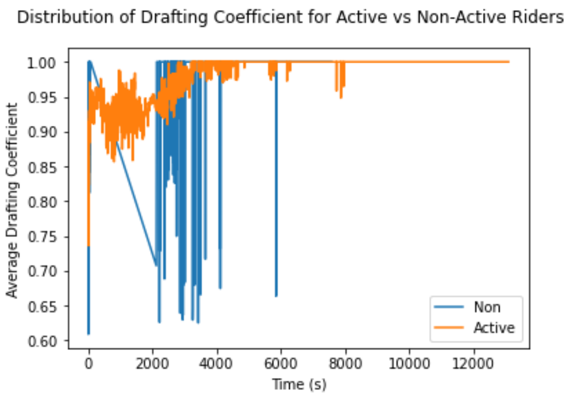


(a) Mean Draft Factor of Riders Over Time

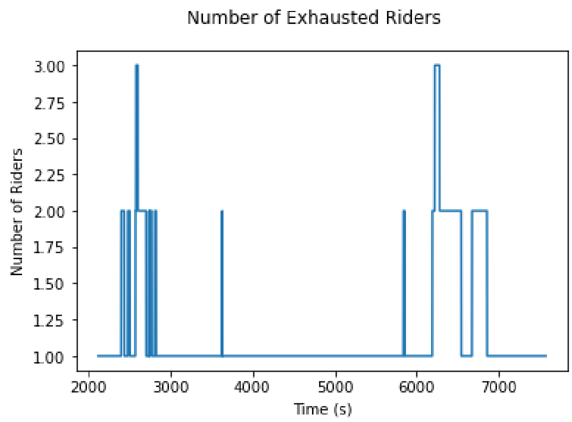


(b) Number of Exhausted Riders Over Time

Figure 4.13: Time Based Results for Riders on a 25km Course of Gradient 0% Using Blocken Based Formula, Run 3

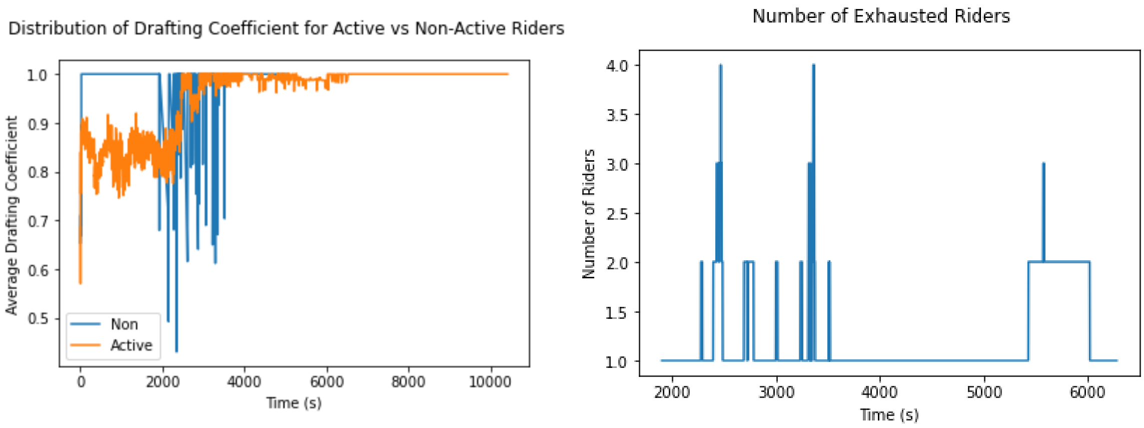


(a) Mean Draft Factor of Riders Over Time



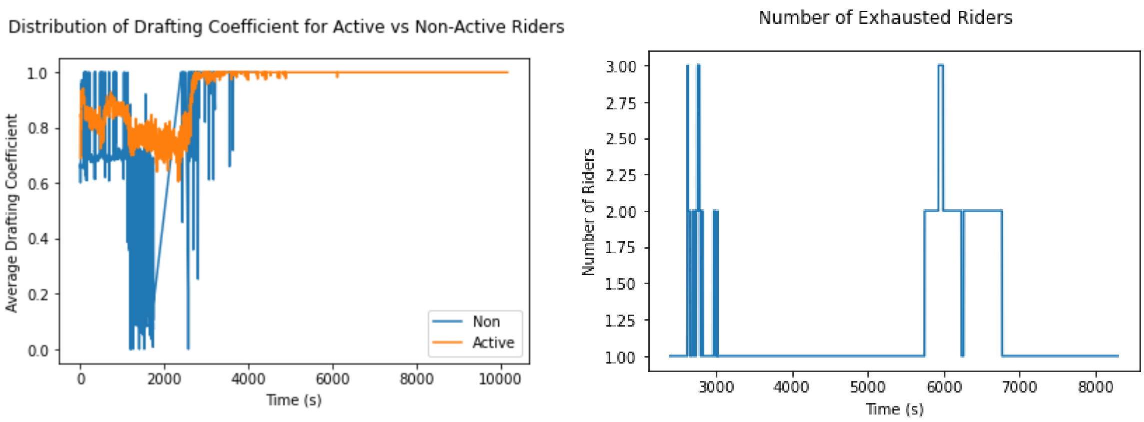
(b) Number of Exhausted Riders Over Time

Figure 4.14: Time Based Results for Riders on a 25km Course of Gradient 10% Using Blocken Based Formula, Run 1



(a) Mean Draft Factor of Riders Over Time (b) Number of Exhausted Riders Over Time

Figure 4.15: Time Based Results for Riders on a 25km Course of Gradient 10% Using Blocken Based Formula, Run 2



(a) Mean Draft Factor of Riders Over Time (b) Number of Exhausted Riders Over Time

Figure 4.16: Time Based Results for Riders on a 25km Course of Gradient 0% Using Blocken Based Formula, Run 3

Comparison To Previous Test Results

When compared to the results gathered during the primary testing phase, making use of the Old's formula, there are some immediately obvious differences between the data gathered. The first of which is that the average draft-factor of riders can be seen to be higher with the Blocken formula than those seen with the Old's formula. This would suggest that when the model uses the Blocken based formula rider agents experience less of a reduction in drag due to drafting. The numbers of exhausted riders would also suggest this to be the case given that more rider agents are exhausted on the Blocken based flat runs than similar Old based ones. As previously touched on the effectiveness of drafting is reduced overall on higher gradient runs hence this is not necessarily displayed on the 10% runs.

Upon initial inspection this seems erroneous given that the Blocken based cubic offers a reduction in aerodynamic resisting forces to a greater distance behind the leading rider. However, this indicates a problem with the simplification of the real-world for modelling. As shown when forming the Blocken based formula, every rider within the paceline has a different version of the formula depending on position. To simplify for the model and to allow for an easy swap with the Old formula only the second rider formula was used. When using this to work out draft factor within the group only the section of the formula relating to a small wheel distance is being used so the benefits of the additional 5m compared to the Old's formula are pretty much negated. To resolve this the model would need to calculate rider position within the bunch and adjust the version of the Blocken based formula used to get a more accurate result. This would likely show the significant decrease in average draft-factor expected.

Chapter 5

Discussion

The series of tests outlined previously demonstrate that the model produced fits within the landscape of currently published works. All tests point to proving that the model in its' current form correctly handles the behaviour of "active" riders. That being riders who are a racing in a pro-active manner, be that pulling a turn on the front of a larger group to offer riders behind shelter or by attacking off the front of a group of other riders. This statement can be backed up by comparing between the results produced based on variations of co-operation tendency using this model with those results published by Hoenigman (Hoenigman et al. (2011)). Both show matching trends for stronger riders to defect and weaker riders to co-operate. Furthermore testing into the drafting of active agents within the model demonstrates an similar ability to receive a benefit from drafting as that shown by Erick Martins Ratamero (Ratamero (2013)). The key differences between the published works and this model is in the handling of non-active agents, those that have chosen to seek a draft and save energy. It would appear that either the definition of a non-active rider or the conditions that triggers an agents behavioural change within the model do not match with published work. The former could be fixed through reclassifying the definition of a non-active rider when results are collected. This would have little impact on the actual behaviour of agents within the model. The later would need for the agent behavioural triggers within the model to be adjusted. These could be adjusted in multiple ways including reducing the opportunities for an agent to become active by increasing the power limits on this behaviour. Another method would be to force time limits on inactivity in a similar method to that of activity. Both attacking and pulling a turn must be carried out for a certain period of time, this is not the case with seeking a draft. This choice was made during implementation to ensure that riders could attack or pull a turn as soon as was necessary and be more responsive to tactical changes. To give a real-world example where this would be the case: a rider is pulling a turn, they finish and no other rider chooses to pull a turn, the rider then attacks from the group as no others are co-operating. It is possible that this choice was wrong and that agents should go through a larger forced "rest" period of inactivity after attacking or pulling a turn.

The tests carried out using the new formula for drag factor turned up unexpected results. The fact these showed an increase in average drag factor rather than a decrease suggests that the model's current form is lacking when handling the aerodynamic forces experienced by riders within the peloton. It is likely this comes from two areas, the first are the assumptions made on the aerodynamic forces experienced by a rider following at different angles behind another. These are based on the limited evidence available, to improve on said evidence would require in-depth Computational Fluid Dynamics and wind tunnel testing. The second is how the drag factor of a rider within the pack is calculated. This currently relies on using the drag factor of the riders ahead. As the Blocken based test results show this is somewhat insufficient. However implementing a method better than this comes with added complexity. The next stage would to be apply varying Blocken based formulae dependent upon position. Fixing these two elements in the model would suggest improvements up to pelotons that are 9 rows deep but wouldn't allow riders any further as no formula exists for the 10th rider. This may improve accuracy in some areas but apply unrealistic restrictions in others. Whether this would be an improvement overall is debatable. Seemingly the only method to get truly accurate aerodynamic data for all riders in the race at a time would be to run computational fluid dynamics at every iteration. Doing so would obviously add a high degree of complexity to the model and require even more compute power.

Aside from the accuracy of the model there are further considerations that need to be taken into account to validate its usefulness. Even though the model is shown to work it is limited by the time it takes to run a set of simulations, even at the shortened test length and with fewer rider agents. To start with the simulation is running on relatively dated hardware and is likely limited in both computing power as well as write speed for collecting results post simulation run. Secondly, although an attempt has been made to optimise the model it still has some downfalls in efficiency. "Premature optimization is the root of all evil" - Sir Tony Hoare. Given the timescale of this project it is difficult to produce an accurately working model and then still have time to optimise effectively in post. That being said the model was built step-by-step, adding functionality with each element. This often occurred with the introduction of a new function. Frequently said function would involve looping extensively, this is unfortunately the nature of the many flocking programs. However, if the program was to be rebuilt at this stage it would be done with the aim of reducing these where possible. It may have been possible to do this in the initial build if more consideration was spent on this in the planning phase. That being said, although the new approach may have led to a more elegant solution it could just have easily led to a non-working solution. Obviously a working solution is preferable however messy. Some other flocking methods are available that reduce the computing power necessary, these regularly involve grouping agents to reduce the number of calculations required per iteration of the simulation. This might be a somewhat effective method of optimization but it could also reduce the amount of freedom offered to individual rider agents within the model. At this point it starts to become a decision based on how accurate the model needs to be verses the level of model simplicity required for compute time.

Against the initial aims of the project the produced model arguably falls short. The model just replicates the results of works already published and does not bring anything revolutionary to the table. That being said many of the initial aims for the project were upon reflection naive. Surprisingly if something hasn't been done before it might be because it's not particularly easy. Many elements of the starting concept model had to stripped out and left in the interest of completion, these are outlined later on within *Section 7: Further Work*. This model in it's current state is not going to be helping race organisers improve rider safety in the professional peloton anytime soon. With all this being said the model could also feasibly be considered a success. It manages to use intelligent agents to replicate real-world rider behaviour within a peloton over a range of courses while fitting in with the latest research on the topic.

Chapter 6

Conclusion

To conclude this project set out with the aim to produce a model able to replicate the behaviour of athlete's within a professional road cycling race through use of intelligent agents. The hope was that such a model could be used to improve the safety of events organised for real-world athletes. Although the model will not be used by race organisers anytime soon it has been built up from scratch and manages to recreate some of the behaviours produced by two current models that are the foundations of research papers on the topic (Hoenigman et al. (2011) and Ratamero (2013)). The model produces active rider behaviour similar to that displayed within these papers with riders experiencing varying levels of drag reduction due to peloton positioning while making tactical decisions based on defined rider characteristics in order to influence overall race results. Unfortunately the model falls short in terms of non-active rider behaviour, however as discussed this is likely a problem with the model that could be rectified with minimal modification. In addition, improvements could be made to the handling of aerodynamics within the pack of riders, although progressing further would lead to additional complexity that may hinder the running of the model. Furthermore the model can be seen to work over a range of courses with varying gradients with base elements in place to allow testing over real-world race routes in the future provided an amount of further development. That being said it is advised that the current model be rebuilt and optimised before too many additional elements are added. This optimisation is vital as current simulation runs take extended periods of time, even over shorter courses and with fewer riders than previous studies. These sets of runs can take up to approximately 24 hours to complete. Admittedly this is on dated hardware but the scale of performance drop off as the size of the course or number of riders increases is dramatic and therefore even if switched over to more modern hardware this advice still remains. If the project was to be done again then very few areas would be completed differently. As touched on the development and implementation of different features within the model may have been achieved more elegantly if the model was produced as a whole rather than a segment at a time. However this could have also led to a model that didn't have any functional elements at all. In short producing this model a second time around is likely to lead to be better solution but to build the model for the first time from scratch the correct approach was taken.

Chapter 7

Further Work

There are many options for further work with this model but the most important one to start with is optimisation. As mentioned multiple times in the later half of this report the next stage would be to improve the model in its current form. This would involve investigating elements where the model is seen to differ from those in previously published works, mainly how the model deals with non-active rider agents. Upon resolution of these issues the performance of the model needs to be drastically improved. The first method of doing so would be to optimise the current code by reducing the number of looping elements to as few as possible and offering break out options within the remaining loops thereby reducing the number of iterations run. Once the model has been optimised sufficiently it is then worth investigating available hardware options for running the model. This could either be in the form of purchasing improved hardware or a more cost effective approach may be to use a cloud computer service such as those offered by Amazon Web Services (AWS). Using a cloud approach would allow easy access to improved hardware as well as offering the option to run simulations in parallel. Being able to run in parallel would massively reduce the time required to run each round of testing. A reduction in run time such as this then makes it feasible to do further testing on the current model such as using longer routes, more riders, and more repeat runs. Doing so would reinforce the effectiveness of the model in its' current form which would be important before adding any additional features.

A plethora of features could be added to this model to enable it to better replicate the real-world, starting with the implementation of aerodynamic forces. As touched on the switch over to a Blocken based formula model, although seemingly unsuccessful in these tests, is a worthwhile development. The results in this report are for a simplistic approaching in implementing Blocken's research, a more complex approach making use of a wider array of Blocken's data (such as a differing formula for each rider based on position) may give more positive results. Included in this would be to add the "up-draft" effect which is discussed in some of Blocken's research. This is the reduction in drag a leading rider experiences by having riders following behind. Furthermore, the current approach for correcting for the angle at which the following rider is behind the leader is mostly based on assumption rather than collected scientific data. Obviously large improvements could be made to this if credible scientific evidence is found on the topic and the model modified in line with it. All of the above would have knock on effects for how the draft for riders within the peloton is calculated and could therefore have dramatic improvements on the overall accuracy of the model.

Moving on from how the drag of a rider agent is calculated but staying on the topic of drafting, the algorithm that allows rider agents to hunt for a draft could be improved. The current implementation is very simplistic and clunky. It compares the current draft the rider experiences to that of if they were slightly forward, backward, left, or right from their current position. This leaves riders with a very limited understanding of the bunch around them with very limited "vision" of the positions available within the bunch. To start with a rider agent needs to have a better understanding of the available drafts in a bunch as a whole, rather than those just around them. A professional road racer will be able to tell roughly where the best drafts are based on pack density, rider size, etc. They can then head that way and once in position refine their pick by moving closer to the wheel in front, swapping lead rider, etc. Implementing a two stage approach such as this could possibly work. Further to understanding where the most desirable draft is, it may also be worth agents understanding the energy costs that would be incurred by trying to get to the required position. There is no point a rider getting to the best draft in the group if they have wasted all of their energy doing so. A staged draft choosing system with a weighting by energy cost may be a good future method of implementation.

As touched on above it is important a rider agent is "aware" of the same things as real world athletes. Although during development a visual cone was seen to be detrimental to this model it is probably still worth further investigation. A more elegant weighted system that allows riders to be "aware" of all around them while with a preference to areas most easily visible to human athletes, i.e. in front, seems like it would be the best approach. Furthermore, the impact of this weighting system changing with level of physical excursion could allow for testing of the theory that this is what changes peloton shape in real-world racing.

What a rider agent is able "to see" has an additional impact on cornering. Although not included in this model it is a logical step to add cornering to a model of this nature as many of the changes to peloton shape occur through tight corners. The group compresses as it slows into the corner and then stretches as it accelerates out, this is sometimes referred to as the accordion effect. To implement cornering correctly will require development to much of the model. This would have to start with the additional physical forces experienced during cornering, including understanding the relationships between: the speed of the rider; the angle the bicycle can be lent over to; the changes in weight distribution and resulting changes in tyre properties both front and rear; as well as the centripetal force experienced by the rider. All this and likely more is required purely to model a cyclist riding through a corner on a defined line. For agents to make their way through a corner in a model such as this they will need to be able to predict the line they intent to take, changing speed and heading accordingly to match. It is likely every rider will be forced to take a slightly different line due to their position within the bunch. Prediction of the corner for each rider will therefore be critical and this links in with the visual cone of the rider and what they are "aware" of. Furthermore different riders likely have different cornering abilities and therefore it may be worth adding an element of error to the line choice or corner prediction depending on rider skill.

With the implementation of cornering the model could be used to test over race routes imported through GPX files as intended. However it's important to note that this is still not entirely accurate to the real-world. GPX files give no indication of road width which can be constantly changing in a road race, the sudden narrowing of the road can cause massive problems for a racing peloton and as a results these situations need to be modelled. It is also worth bearing in mind that GPX files have no information on road furniture such as bollards, crossings, or speed bunch which are all dangerous to a racing peloton. In addition a GPX file will only give the route travelled, therefore only one side of a roundabout. Often in races the peloton will split either side of a roundabout with the riders trying to squeeze onto the faster side. All of these elements are dangerous to cyclists but there's no existing easy method to model them.

Even more developments could be made to the environment the rider agents exist within. Although the aspect of road furniture has been touched on the fact the road surface itself can change hasn't. Some of the most prestigious road races in the world are carried out over cobblestones which are obviously a completely different prospect to smooth asphalt. Changes in surface could introduce changes in rolling resistance, level of grip through corners, it could even cause riders deemed "uncomfortable or uncertain" on the differing terrain to make poorer decisions or leave larger gaps to the riders around them. Other environmental changes could include the weather. As with surface changes, rain will decrease the amount of grip available but could also reduce a riders vision. Extremes of temperature are also going to change how riders perform, those used to heat will perform worse in the cold and visa versa. Finally and arguably the most important element of the weather in cycling, the wind. Wind direction plays a massive role in road cycling. There's no point attacking into a headwind, you'll tire quickly and get caught. A tailwind is better, it lessens the impact of the draft for a rider sitting behind you. The best, or worst, however is a crosswind. These are the most divisive in road cycling, causing the peloton to split into many smaller groups all rolling rotations within themselves. Crosswinds can make or break a rider's race. Being able to do well in crosswinds is often a skill of its own in road cycling. It seems obvious therefore that a model of road cycling would need these wind effects developed into it to be truly accurate.

To finish we'll look at representing the real-world riders themselves. Up until this point rider parameters ranging from power output, weight, cooperativeness, or even possibly cornering skill, have all been decided by normal distributions. In modern sport vast quantities of data are pulled from athletes, this data is often used by third parties. Notable examples of this include sports based computer game developers. This real-world athlete data is then used to create a model of said athlete within the game, often in-game athlete statistics will update depending on how the athlete is currently performing. Some athlete data of this nature is current available for professional road cyclists online and a model could make use of this to make a simulation accurate even down to the statistics of every athlete taking part.

After all of this expansion into the possibilities of what developments could be made to this model it feels worthwhile to ground this back into the current reality. Every addition to the model will undoubtedly come with an increase in complexity which will impact on the performance of the model. The most important part of this model is that it must be usable. To improve the safety of professional road cyclists the model can't afford to take months, weeks, or even days to run for each route. To have an actual impact the model needs to delicately balance efficiency with accuracy. Too accurate, at the cost of efficiency, and it won't be used. Too efficient, at the cost of accuracy, and athletes get hurt.

References

Belden, J., Mansoor, M. M., Hellum, A., Rahman, S. R., Meyer, A., Pease, C., Pacheco, J., Koziol, S. & Truscott, T. T. (2019), 'How vision governs the collective behaviour of dense cycling pelotons', *Journal of the Royal Society Interface* **16**, 1–9.

Blocken, B., Malizia, F., Druenen, T. & Gillmeier, S. (2020), 'Aerodynamic benefits for a cyclist by drafting behind a motorcycle', *Sports Engineering* **23**.

Blocken, B., Toparlar, Y., van Druenen, T. & Andrianne, T. (2018), 'Aerodynamic drag in cycling team time trials', *Journal of Wind Engineering I& Industrial Aerodynamics* **182**, 128–145.

Blocken, B., van Druenen, T., Toparlar, Y., Malizia, F., Mannion, P., Andrianne, T., Marchal, T., Maas, G.-J. & Diepens, J. (2018), 'Aerodynamic drag in cycling pelotons: New insights by cfd simulation and wind tunnel testing', *Journal of Wind Engineering I& Industrial Aerodynamics* **179**, 319–337.

Crouch, T. N., Burton, D., LaBry, Z. A. & Blair, K. B. (2017), 'Riding against the wind: a review of competition cycling aerodynamics', *Sports Engineering* **20**, 81–110.

Dirksen, F. (2017), 'Drafting how far is close enough', `https://www.swissside.com/blogs/news/the-deal-with-drafting?locale=en`. [Online; last accessed 8-September-2021].

Hoenigman, R., Bradley, E. & Lim, A. (2011), 'Cooperation in bike racing - when to work together and when to go it alone', *Complexity* **17**, 39–44.

Krajina, T. (2020), 'gpxpy – gpx file parser', `https://pypi.org/project/gpxpy/`. [Online; last accessed 8-September-2021].

Kyle, C. R. (1979), 'Reduction of wind resistance and power output of racing cyclists and runners travelling in groups', *Ergonomics* **22**(4), 387–397.

Malizia, F. & Blocken, B. (2020), 'Bicycle aerodynamics: History, state-of-the-art and future perspectives', *Journal of Wind Engineering and Industrial Aerodynamics* **200**, 104134.
URL: *https://www.sciencedirect.com/science/article/pii/S0167610520300441*

Martin, J., Milliken, D., Cobb, J., McFadden, K. & Coggan, A. (1998), 'Validation of a mathematical model for road cycling power', *Journal of Applied Biomechanics* **14**, 276–291.

McCole, S. D., Claney, K., Conte, J.-C., Anderson, R. & Hagberg, J. M. (1990), 'Energy expenditure during bicycling', *Journal of Applied Physiology* **68**(2), 748–753.

Olds, T. (1998), 'The mathematics of breaking away and chasing in cycling', *European journal of applied physiology and occupational physiology* **77**, 492–7.

Ratamero, E. (2015), Modelling peloton dynamics in competitive cycling: A quantitative approach, Vol. 464, pp. 42–56.

Ratamero, E. M. (2013), 'Moped: An agent-based model for peloton dynamics in competitive cycling', *icSPORTS 2013 - Proceedings of the International Congress on Sports Science Research and Technology Support* pp. 134–143.

Reynolds, C. (1987), 'Flocks, herds, and schools: A distributed behavioral model', *ACM SIGGRAPH Computer Graphics* **21**, 25–34.

Reynolds, C. (2002), 'Steering behaviors for autonomous characters'.

Trenchard, H. (2015), 'The peloton superorganism and protocooperative behavior', *Applied Mathematics and Computation* **Volume 270**, Pages 179–192.

Trenchard, H., Ratamero, E. M., Richardson, A. & Perc, M. (2015), 'A deceleration model for peloton dynamics and group sorting', *Applied Mathematics and Computation* **251**, Pages 24–34.

U., W. (n.d.), 'Netlogo'.
URL: *http://ccl.northwestern.edu/netlogo*

van Druenen, T. & Blocken, B. (2021), 'Aerodynamic analysis of uphill drafting in cycling', *Sports Engineering* **24**, 10.

Wikipedia contributors (2021), 'Cycle sport — Wikipedia, the free encyclopedia', `https://en.wikipedia.org/w/index.php?title=Cycle_sport&oldid=1042394345`. [Online; last accessed 8-September-2021].

Appendix A

Supplementary Materials I

Code and Results are available at the following Github Repository:
https://github.com/matt-thoumine/COMP4031_Project.git

Appendix B

Supplementary Materials II

B.1 Preliminary UG & PGT Research Ethics Checklist

The University of Nottingham

School of Computer Science

Preliminary UG & PGT Research Ethics Checklist

(Last updated 2018-10-23)

*This form is required for **all UG and Taught MSc dissertation projects** to confirm whether the proposed work requires ethical review.*

SECTION I. Applicant Details

1. Name	Matthew John Thoumine
2. Status	**Postgraduate Taught Student**
3. Email address	eaymjt@nottingham.ac.uk

SECTION II. Module Details

4. Module name/number or MA/MSc/MPhil course and department	COMP4003
5. Supervisor's name	Colin Johnson
6. Supervisor's email address	Colin.Johnson@nottingham.ac.uk

SECTION III. Project Details

Project title	Intelligent Agent-Based Modelling Using Swarm Intelligence to Simulate Peloton Movement
Proposed start date	**01/08/2021**

	Please answer each question by ticking the appropriate box:	**Yes**	**No**
1.	Does the study involve human participants?		x
2.	Does the study involve personal data, and/or include the processing of data in the public domain in such a way as to invade the privacy of the individuals concerned? (see CS Intranet/Workspace for guidance on what constitutes personal data.)	x	
3.	Does the study involve the use of animals or biological materials?		x

Complete whichever of the following applies:

If the answer to all questions in Section III is **NO** and therefore ethical review is **not** required: ☐

I confirm that (i) the information provided is an accurate description of the proposed project,
(ii) I understand that I do **not** need to submit a full ethics application and
(iii) if my project changes such that ethical review is required then I will submit a revised form.

Signature of Applicant: ………………………………………………………… Date:……………………………

OR

If the answer to at least one question in Section III is **YES** and therefore ethical review **is** required: ☒

I confirm that (i) the information provided is an accurate description of the proposed project,
(ii) I will submit a full ethics application including a **UG & PGT Research Ethics Checklist** (available from the CS Intranet/Workspace) and
(iii) I will not start any work requiring ethical approval until the **full** ethics application has been approved.

Signature of Applicant: …MJT…………………………………… Date:…27/06/2021……

Note: if this you are submitting this form electronically to Moodle or from your University email then type your full name

Please follow the instructions provided by the module convenor to submit this form.

B.2 Full UG & PGT Research Ethics Checklist

School of Computer Science
Research Ethics Checklist

for Taught Module Activities
(UG, PGT and PGR students)

- This checklist must be completed for every research project that involves human participants, use of personal data and/or biological material, be*fore* potential participants are approached to take part in any research.
- Any significant change in the design or implementation of the research should be notified to cs-ethicsadmin@cs.nott.ac.uk and may require a new application for ethics approval.
- It is the applicant's responsibility to follow the University of Nottingham Code of Research Conduct and Research Ethics and any relevant academic or professional guidelines in the conduct of the study. **This includes providing appropriate information sheets, consent forms and recruitment materials, and ensuring confidentiality in the storage and use of data.**
- Completion of this form confirms that you have read and understood the guidelines on the CS Intranet regarding:
  - what is defined as *personal data;*
  - what is required for *valid consent;*
  - the key requirements of the Data Protection Act (2018), which includes GDPR
- The supervisor is responsible for exercising appropriate professional judgement when completing this form.
- **Sections I to V should be completed by the student undertaking the study, in discussion with their supervisor.**
- The **student** is responsible for submitting the completed form according to the instructions provided by the module convenor (normally on Moodle).

SECTION I: Applicant Details	
1. Applicant's name	Matthew Thoumine
2. UoN Email address	eaymjt@nottingham.ac.uk
3. Status	**PGT Student**
4. Student ID	**14274893**
5. Degree name	**Computer Science MSc**
5. Module name/number or MA/MSc/MPhil course and department	**COMP4003: Research Project in Computer Science (MSc CS)**
6. Supervisor's name	Colin Johnson
7. Supervisor's email address	**Colin.Johnson@nottingham.ac.uk**

SECTION II: Project Details	
1. Project title	**Intelligent Agent-Based Modelling to Simulate Peloton Motion**
2. Proposed start date and latest end date of study	**Start: 06/07/2021** **End: 30/09/2021**
3. Date and version of this submission	**Date: 06/07/2021** **Version: 1**
4. Type of submission?	**First submission**
5. Application ID (if known[1])	

6. Description of Project, including aims/objectives and procedures. *Please include any information which may affect the consideration of the ethics involved, e.g. how participants will be recruited and rewarded, data to be collected/used (see also II.7), location of study, unusual circumstances, age range of participants:*

The aim of the project is to use intelligent agents to model individual rider and then therefore peloton movement within a road bicycle race. The simulation could then be applied to real world race routes with real world rider data to predict which sections of the course present the greatest risk to rider safety due to peloton movement.

Agents would be based off the data of professional cyclists in this year's (2021) Tour De France. These athletes would not be recruited or rewarded, and all would be aged between 22 and 41 years old. Data used for the agents in the simulation would all be already publicly available and collected from www.procyclingstats.com.

Simulation testing routes would again be from this year's Tour De France route. Race organisers do not supply a GPS route file to the public and the races "roadbook" (that includes all the information about the race) only includes inaccurate maps of routes. Therefore, the best method for getting accurate GPS files of the route is to download gpx route files from a riders Strava profile from www.strava.com. These files only include values for latitude, longitude, and elevation of the route ridden by an athlete with no data connecting directly back to the athlete.

[1] Normally each ethics application will be allocated an ID by the University *after* its initial submission

7. Will data from the project potentially support an academic publication? *(Not just a dissertation or assessment.)*	**No**

8. Will personal data (including photos, video or audio) or biological materials be collected, recorded or used?

Yes

If Yes, please give details below.
*Where several types of data are involved be explicit about which statements apply to which types of data (e.g. "**video**: …; **audio**: …"). Include details of any potential risks to subjects (e.g. from re-identification) and how these are mitigated. See the guidance on research data notes.*

What data *(or materials) will be collected or used*	Participants Names, Weight, and Height
*What if any **constraints** apply to use of this data (or materials)*	Data collected from www.procyclingstats.com if not constrained from use in research projects. However the retrieval of data is restricted from use of "any robot, spider, or other device" aswell as "any automated scripting tool or software". (See https://www.procyclingstats.com/info/general-terms-and-conditions) Procyclingstats has been reached out to for clarification, but a reply has not been received. The best method would therefore be to collect the required data from the site manually.
***How** will this data (or materials) be:*	
collected or obtained	Agent data would be collected from www.procyclingstats.com Route data would be collected from www.strava.com
processed before analysis	*e.g. anonymised or de-identified (if relevant)*
stored and secured	*(a) during the research* – On the University OneDrive *(b) after the research – (Depends, might say not be stored, or stored on moodle as part of submission)*
formatted	*Agent Data* - CSV file (excel file) *Route Data* - GPX file
organised	*e.g. standard structure in guidance notes –* - Raw Data (University or Public) - Working Data (University or Public) - Analysis and Results (University or Public)
analysed	Data will be used as input data for a simulation. Results of the simulation are focused on the peloton as a whole and therefore be given as statistics of the population.
reported in publications, including reports and dissertations	Results will focus on the population/peloton. Any comment about an individual athlete will be in relation to how an agent based on said athlete behaves rather than the individual themselves.
***How** and **when** (if ever) will this data (or materials) be:*	
reused	No
archived, indexed, published or otherwise made available to others	*include details of any (further) anonymization, how it will be made available and to whom*
deleted or destroyed	Data will be deleted December 2026

If human subjects are involved then at what point(s) can they ***withdraw*** *and what will happen in each case? (if no human subjects are involved enter "Not Applicable")*
What will happen to this data if/when you ***leave*** *the University?* It will be retained by supervisor for period of 5 years
9. Will personal data or commercially sensitive (i.e. "restricted") data be collected or stored? **Yes** *If Yes, please give details below for the University data asset inventory.*

Title *of data asset*	*Student project 14274893 data set*
What *personal/sensitive information (fields) does it contain?*	Name, Weight, Height
Data ***owner***	*Colin Johnson*
Data ***stewards*** *(and responsibilities)*	*Matthew Thoumine*
Data ***users***	
Data ***location***	*i.e. sufficient to find the data asset*

SECTION III: Research Ethics Checklist (Part 1)	
Please answer all questions:	**Yes/No**
1. Does the study involve participants who are unable to give informed consent (e.g., children, people with learning disabilities or dementia[2], prisoners, your own students)?	No
2. Will the study involve participants who are particularly vulnerable[3]?	No
3. Will it be necessary for participants to take part in the study without their knowledge and consent at the time (e.g., covert observation of people in non-public places)?	No
4. Will it be necessary for participants to be kept in ignorance, misled or deceived at any point in the study (e.g., if revealing the full aims of the project during the consent process would undermine the research)?	No
5. Will the study involve the discussion of sensitive topics (e.g., sexual activity, drug use)?	No
6. Will participants be asked to discuss anything or partake in any activity that they may find embarrassing or traumatic?	No
7. Is it likely that the study will cause offence to participants for reasons of ethnicity, religion, gender, sexual orientation or culture?	No
8. Are drugs, placebos or other substances (e.g., food substances, vitamins) to be administered to the study participants or will the study involve invasive, intrusive or potentially harmful procedures of any kind?	No
9. Will body fluids or biological material samples be obtained from participants? (e.g., blood, tissue etc)	No
10. Is pain or more than mild discomfort likely to result from the study?	No
11. Could the study induce psychological stress or anxiety or cause harm or negative consequences beyond the risks encountered in normal life?	No
12. Will the study involve prolonged or repetitive testing for each participant?	No
13. Will financial inducement (other than reasonable expenses and compensation for time) be offered to participants?	No
14. Is this medical or clinical research on human participants?[4]	No
15. Will the study involve the recruitment of patients, staff, tissue sample, records or other data through the NHS or involve NHS sites and other property?[5]	No
16. Will the study involve the use of animals?[6]	No

[2] If participants are adults who lack the mental capacity to give informed consent then you must obtain approval from an “appropriate body” approved by the Secretary of State (instead of this committee).

[3] “who is or may be in need of community care services by reason of mental or other disability, age or illness; and who is or may be unable to take care of him or herself, or unable to protect him or herself against significant harm or exploitation” (Department of Health (2000): *No Secrets: guidance on protecting vulnerable adults in care)*

[4] If Yes then you must obtain approval from the Faculty of Medicine and Health Sciences REC (instead of this committee).

[5] If Yes then you must obtain NHS REC and R&D approvals from the relevant Trusts (instead of this committee).

[6] For work with animals always seek advice from the University’s Animal Welfare and Ethical Review Body (AWERB). If the animal(s) are vertebrates or cephalopods then you must obtain approval from AWERB (instead of this committee).

SECTION III: Research Ethics Checklist (Part 2)	
Please answer all questions:	**Yes/No/NA**
1. For research conducted in public, non-governmental and private organisations and institutions (such as schools, charities, companies and offices), will approval be gained in advance from the appropriate authorities?	Yes
2. If the research uses human participants, personal data or the use of biological material, will explicit consent be gained?	No
3. Will participants be informed of their right to withdraw from the study at any time, without giving explanation?	No
4. If data is being collected, will this data be anonymised before publication or sharing?	No
5. Will participants be assured of the confidentiality of any data?	No
6. Will all data be stored in accordance with the Data Protection Act?	Yes
7. Will participants be informed about who will have access to the data?	No
8. If quotations from participants will be used, will participants be asked for consent?	Yes
9. If audio-visual media (voice recording, video, photographs etc) will be used, will participants be asked for consent?	Yes
10. If digital media (e.g. computer records, http traffic, location logs etc) will be used, will participants be asked for consent?	Yes
11. If the research involves contact with children, will appropriate safeguards be in place (e.g. supervision, DBS checks if required)?	Yes
12. If research data itself is to be published, shared or reused (e.g. alongside a publication or in an archive) will participants be asked for consent?	No

- If you have answered ‘No’ to all questions in SECTION III Part 1 and ‘Yes’ to all relevant questions in SECTION III Part 2 the project is deemed to involve **minimal risk** - go to the signature page.

- If you have answered ‘Yes’ to any of the questions in Part 1 or ‘No’ to any of the questions in Part 2 the project is deemed to involve **more than minimal risk**. Please explain in SECTION IV why this is necessary and how you plan to deal with the ethical issues raised.

SECTION IV: If the project involves more than minimal risk, please explain why this is necessary and how you plan to deal with the ethical issues raised
The project is deemed to have more than minimal risk because participants would not be contacted at any stage and that their data would not be anonymised. In this project participants will not be contacted, this is because all participants are professional athletes, the data being collected about each one is already publicly available (name, height, weight). Furthermore, there is likely to be up 184 records of athlete data (1 record per athlete all combined into an overall dataset) to collect. Contacting each athlete individually to gain consent is not feasible given the timescale of the project. The dataset will not be anonymised because it allows for comparisons to be made between the agents who represent the athletes in the simulations and real-world athletes. If an athlete's name is used during the report as an identifier with relation to the simulation, it will be with the focus on the agent representing said athlete rather than the athlete themselves. This is purely as a method of grounding the research in the real-world. As discussed, GPS route data (in the form of a GPX file) will need to be downloaded from an athletes Strava profile. This data is publicly available for anyone and downloading a route is such a fashion is an encouraged method of usage within the Strava eco-system. The resulting GPX route file would contain no data linking back to athlete who's profile the route was downloaded from. It is necessary to use this data as it will further test the accuracy of the model produced as part of the research. As well as allowing the results gathered to be directly compared to events that occurred as part of this year's Tour De France.

RESEARCH ETHICS CHECKLIST – SIGNATURE PAGE

SECTION V: Applicant Declaration		
Please confirm each of the following statements:		**Yes/No**
The project is deemed to involve **minimal risk** as defined in SECTION III		**No**
I confirm that I have read the University of Nottingham Code of Research Conduct and Research Ethics		Yes
I confirm that I have read the guidance documents listed on page 1		Yes
I confirm that the information provided in this application is correct		Yes
I confirm that I have included all of the associated documents* with this application		Yes
Signature of applicant**	Matthew John Thoumine	
Date		

** All applications for projects involving human participants (or their tissue) must be accompanied by an information sheet, privacy notice, consent form and recruitment materials (e.g. posters, flyers, text for emails) where relevant.*

*** if this you are submitting this form electronically to Moodle or from your University email then type your full name*

- The **student** is responsible for submitting the completed form (and associated documents*) once they have agreed it with their supervisor, according to the instructions provided by the module convenor (normally on Moodle).

- The **supervisor** is responsible for checking and confirming that the application is complete and appropriate, and either referring it to the CS Research Ethics Committee, or approving it themselves (see approval policy below). This should be recorded according to the instructions provided by the module convenor (normally on Moodle).

SECTION VI: Supervisor Approval Policy	
By approving an application, the supervisor is confirming that:	
The participant information sheet or leaflet is appropriate for this research project	**Yes or NA**
The procedures for recruiting participants and obtaining informed consent are appropriate	**Yes or NA**
The collection and handling of data is appropriate and in accordance with the Data Protection Act	**Yes or NA**
The supervisor must specify which of the following applies:	
The supervisor has received training in research ethics and the project involves minimal risk (as assessed in section III) and therefore this project **DOES NOT REQUIRE** consideration by the Research Ethics Committee	
(or if applicable) This project conforms to a standard model associated with this module/course that has already been approved the Research Ethics Committee and therefore **DOES NOT REQUIRE** further consideration by the Research Ethics Committee *Approved model application (ID or title/date):*	
(Otherwise) This project **DOES REQUIRE** consideration by the Research Ethics Committee	