Relationship Between Common Measures of Training Stress and Maximum Mean Power During Road Cycling Races

Hamish Ferguson

A thesis submitted to Auckland University of Technology in fulfilment of the requirements for the degree of Master of Philosophy

2014 School of Sport and Recreation

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

ATTESTATION OF AUTHORSHIP

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signed Hamish Ferguson Date 9 July 2015

ACKNOWLEDGEMENTS

A huge thanks to Will Hopkins, for his amazing expertise in the area of sport science, and statistics. For having the patience to guide me into new territories of sports performance analysis. Big thanks to him and Audrey for a comfy bed on my visits to Auckland and I wish him all the best in his retirement ☺

Carl Paton for being my second supervisor and his input based on years of research in cycling and sports science.

To my good friend Geoff Chase whom I helped to relocate in New Zealand from San Francisco. It's been awesome to see you rise from *just* being a PhD to now a Distinguished Professor at University of Canterbury. Your academic input on this manuscript was priceless and highlighted the need for a local supervisor for this project. Our many chats and the odd single malt whisky have paid off!

Many thanks to Andy Coggan for continually pushing the boundaries of how we can measure, and what we can learn about cycling and human performance. I am sure your *hobby* will provide me with many areas for future research.

To my lovely Mum, we celebrated 15 years of your passing and still the pain is there. I wish I could have shared this adventure with you. To my sister Sarah, I love you and all you have achieved. Your pride in what I have accomplished keeps me going.

To all the riders I coach, you are my motivation to keep learning. All this knowledge is meaningless if I have no one to apply it to! The very conclusions of this study spell out how I must do more than just collect data, I must communicate more and draw out all the necessary information to shape your preparation and racing plans to achieve your very best!

Finally the Canterbury, New Zealand and International cycling community a huge thanks to those who were participants in the research. It will have an impact on performance, small, but I feel the conclusions give us some clear directions on how we incorporate power meter data into the overall picture of cycling performance. Thanks to the community for constantly challenging and testing me on a daily basis!

PRESENTATIONS

Conference presentation.

Ferguson, H.F., Paton, C.D. & Hopkins, W.G. (2014). Measures of training stress in cyclists do not usefully predict maximum mean power in competitions. Science behind the Tour de France Conference. Leeds, UK. (Author contribution percentages: HF, 60%; WH, 30%; CP 10%).

PREFACE

The topic of racing and training with an on-board cycle ergometer, commonly referred to as a power is introduced in Chapter 1. Chapter 2 presents a literature review focusing on power meter function, accuracy and validity. Additionally, the use of a power meter to assess event demands and monitor acute training effect and chronic training adaptations. Chapter 3 presents the study design, and the statistical approach used to address the research question. Results, interpretations and discussion of findings are presented in Chapters 4-5. Finally Chapter 6 presents the overall conclusions and directions for future research.

ABSTRACT

In preparation for endurance cycle races, cyclists carry out a large volume of training to attain the necessary fitness to perform. These loads must be managed wisely to be optimally prepared for race day. In the early 1970s Banister and colleagues introduced empirical models that describe the relationship between training load and performance ability. Banister suggested that Performance = Fitness – Fatigue and proceeded to introduce mathematical sophistication to this underlying premise by incorporating decay constants for both fitness and fatigue. Banister suggested that training load rapidly influenced fatigue but only slowly influenced fitness. However, with recovery fitness was well maintained while fatigue quickly dissipated. More recently, commercially available software packages have made it easier for coaches and cyclists to engage in these concepts.

The TrainingPeaks™ software incorporates a performance manager, which is based on an impulse-response model for managing the training loads of cyclists based on data recorded by on board cycle ergometers called power meters. The aim of this study was to determine how well the performance manager model predicts the performance ability of competitive cyclists in road time trials, individual road races and multi-day events.

Nationally and Internationally competitive cyclists (20M, 5F) submitted power meter files for a six- to eight-month period. Measures of fitness, fatigue and freshness were derived in the performance manager from the day before each competition. Maximum mean powers (MMPs) for 5-s, 60-s, 5-min and 20-min durations were recorded from each

race. Mixed modelling was used to estimate the linear relationship between changes in fitness, fatigue and freshness, and changes in the MMPs during competition.

Expressed as coefficients of variation (CV), within-cyclist variation in MMP from competition to competition ranged from 15% (5-s MMP) to 4.1% (20-min MMP). These CVs were too large for the MMPs to track the usual changes in performance that cyclists would show between competitions. When the bottom half of each cyclist's MMPs were discarded, only 5- and 20-min MMPs in time trials had CVs that could track reasonable changes in performance (~2.5%). However, the mixed models showed effects of fitness, fatigue and freshness on MMPs that were either unclear or too weak to be useful.

This study casts doubt on the use of fitness, fatigue and freshness measures to assess training load and the use of MMPs to assess performance in road cycling. Different models of measuring training loads should be investigated. Contextual information about each competition ride might reduce the error in MMPs by allowing filtering or adjusting for poor performances, but other measures of performance from competitions may be needed to determine whether fitness, fatigue and freshness are worth monitoring.

CHAPTER 1: INTRODUCTION

Cycling is a highly competitive sport, with opportunities for Olympic representation or employment in a professional cycling team. Large numbers participate in events all round the World and the best-performed athletes gravitate towards hotbeds of cycling in Europe, Asia, America and Australia to compete in international events. Performance at this level involves a huge investment in time and money to attain the peak performances that can be very lucrative for individuals, teams and nations, in terms of financial rewards and future opportunities.

Elite cyclists perform training on the bicycle and to a lesser extent other activities like resistance exercise, running or cross-country skiing. Large volumes of time are spent in preparation for events, even if those races are of a very short duration. The majority of time is spent preparing for competition on the bicycle. From the early 1980s bicycle-based ergometers were developed to measure forces generated by the rider. Such measurements maintain training specificity and provide a more true measure of sports specific forces. In 1989 Uli Schoberer launched the SRM brand of mobile bicycle ergometer, which became to first successful commercial model to be produced. Competitive cyclists then started using mobile ergometers, called power meters, to record power output during training and competition rides.

The most common power meters in use are crank based power meters as they offer the convenience of being able to be used in both training and competition with the most ease. Rear hub power meters require two wheels, with one for competition and the other for training. This requirement adds to the costs and reduces the convenience of use. The issue with current pedal based power meters are that they require a considerable amount of skill and precision to set up, especially if a rider uses a power meter on multiple bicycles.

Most well performed nations employ coaches and sport scientists to monitor the performance of their riders. The goal is to ensure the considerable investment made is tracking towards an improvement in their overall performance. The process involves collecting data on the demands of international competition, measuring the rider to determine where they sit in relationship to the demands, and monitoring the training process. The goal is to maximise gains and minimise excess fatigue, burnout and injury that hinder these improvements. Power meters offer a continuous measurement that is considered one of the more objective ways to perform assessment and monitoring.

Cyclists train and compete with a power meter, and at the end of each ride can upload the data from the receiver to various forms of analytical and storage software. There are numerous commercial and open source software programmes available to cyclists and coaches to view charts and summary data from their rides. For this study the popular Training Peaks WKO+ Version 3.0 software (Peaksware LLC, Denver, CO) was used to analyse the data.

There are four main typical uses of power meter data in the training and racing process of the rider. A rider will use power meter data as a means of testing their current level of performance. They will perform a variety of tests using watts as a dependent variable.

These tests can be duration based and, depending on the riders preferred event, the durations of 4 seconds to 60 minutes are common. They can also be distance-based, where a track cyclist may perform a test over a 200-m to 4000-m distance using duration and mean power as dependent variables. All of these provide training and racing metrics in terms of power, which is readily measured and can be obtained objectively and consistently.

Riders and coaches also use the data from testing to try and determine training and racing intensities. A common metric in power meter training is the functional threshold (Allen and Coggan 2010). This value is the maximal power that one can sustain for a 60-min period when well rested and motivated. It can be best determined by performing a 60-min maximal effort. A functional threshold can also be determined using a variety of estimates based on tests of shorter duration, or using various critical power calculations. Based on the functional threshold, riders and coaches prescribe zones of power meter training that target different physiological adaptions depending on the frequency, intensity, time and type of training performed. The goal is to identify weaknesses, from testing, that are important to the riders chosen event, and then prescribe power-specific training to address them.

Power meter data is used to assess riders physiology in cycling competitions. In timed events they can analyse the power produced over the ride and also the distribution of power within the ride. Figure 1 is an example of a chart from an individual time trial event. For mass start events, such as a road race, the coach and rider will look at the power produced at key moments of the race, such as hill climbs, headwind sections or contextual information to determine how well the power was delivered to tactical effect, and how well the rider conserved energy within the bunch. Figure 2 is a chart from a road race.

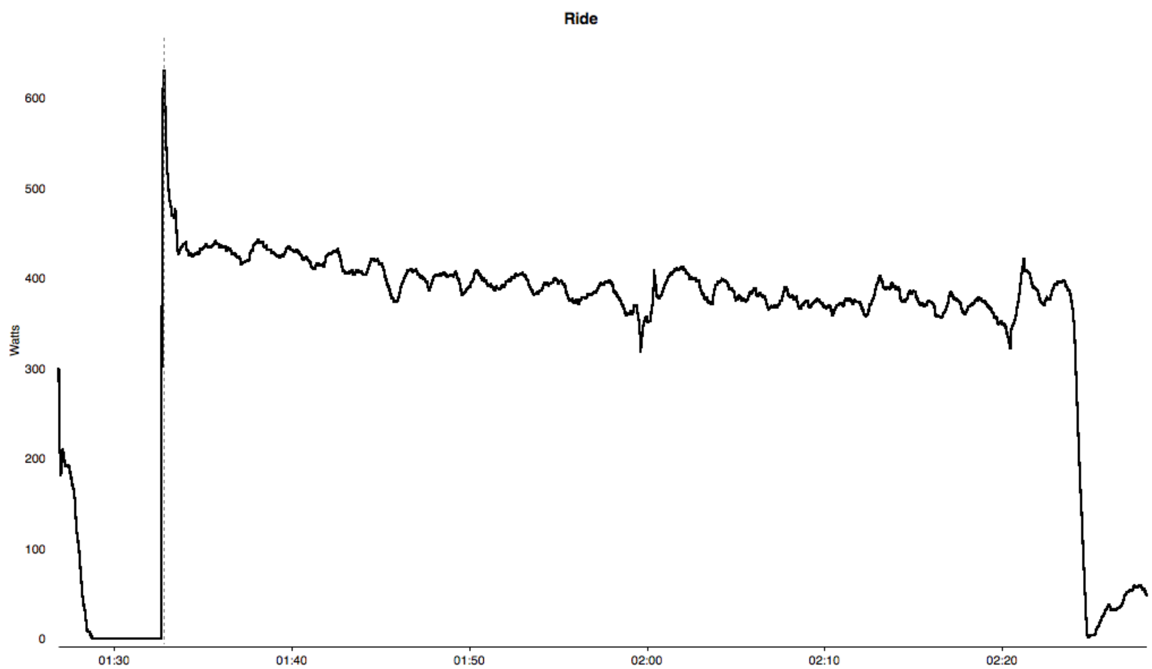


FIGURE 1. Power meter file from a 40-km individual time trial.

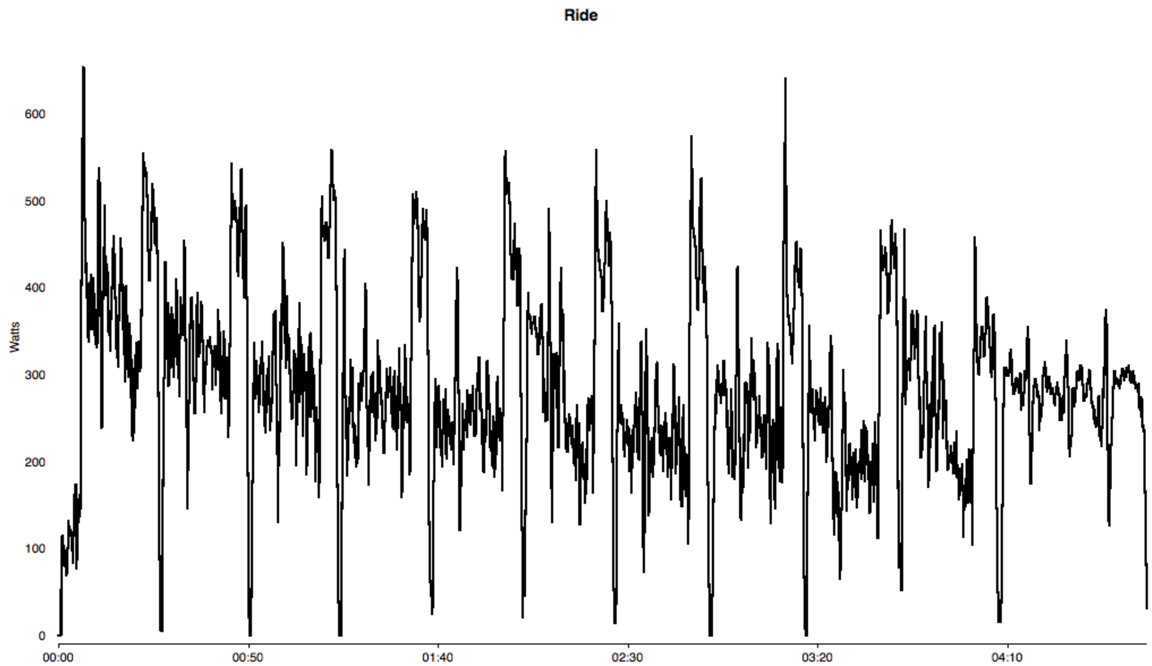
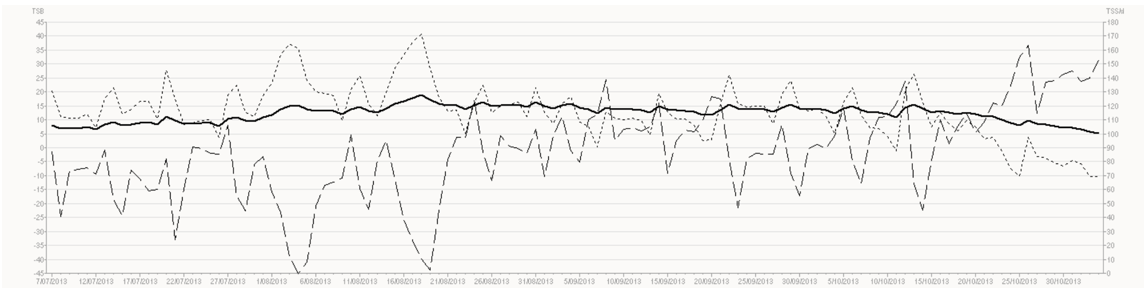


FIGURE 2. Power meter file from a 180-km circuit road race.

As a result of all these applications, it is now more commonplace to look at power meter data over longer periods than just one race, or from test to test. For each ride a training stress score can be determined. This number is determined by using time and intensity. A normalized power, using a rolling 30-s average to effectively smooth the data, can also be determined for the ride or any part of it. The intensity relative to a specified functional threshold is squared and multiplied by the duration of the ride to determine the training stress score (Allen and Coggan 2010). In WKO+ the training stress score is used as the basis of the performance management chart (Allen and Coggan 2010). The performance manager uses an impulse–response algorithm to present three lines on a chart: in particular a chronic training load line that is an indication of fitness; an acute training load that is an indication of fatigue; and the difference between the acute and chronic training loads, which creates the third line called the training stress balance or freshness.

Figure 3 is an example of the performance management chart that illustrates a performance chart for a rider over a period of months.



– Fitness ···· Fatigue -- Freshness

FIGURE 3. Performance management chart

Coaches use the performance management chart like the one in Figure 3 to determine if a rider has sufficient fitness to perform at the level to which they aspire. They are also used to ensure cyclists do not accumulate high loads of fatigue in training and racing, and thus arrive at goal events with sufficient freshness to achieve their best performance. Hence, power-measurements have become a core component in planning and managing fitness and training. This study will attempt to quantify if the numbers generated have any utility in the cycling performance process.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This literature review covers the validity and reliability of cycle based power meters, and the ways they are calibrated. Power meters are used to assess the demands of cycling and the characteristics of different types of cycle races. They are also used to assess riders in the field and to assess a rider's performance in cycle races. From these data, training is planned and conducted by the rider. From effort to effort, power meter data is used to assess the acute training responses, and, much longer term, the chronic adaptations to riding. For context this review also deals with measures of long-term monitoring of training in other sports.

2.2 Cycle power meters

Since the release of commercially available power meters their use has grown beyond small groups involved in the World Tour or national cycling teams. They are a common accessory on many racing bikes at all levels, and even for recreational cycling use. They still represent an expensive option compared to using a speedometer, heart rate monitor or global positioning system measure. This makes the assessment of a power meter crucial to measure the validity and the reliability of the unit. Equally important is the ongoing calibration to ensure the data recorded can be compared against previous measures and for comparison with data from other riders or even other sports.

The validity of power meters has been investigated on most commercially available options. Recently, more meters have come onto the market and are awaiting testing. Most power meters have been tested against a calibrated ergometer or using a dynamic testing rig. As the first commercially available power meter, the SRM brand, where strain gauges are housed in the crank spider that the chainrings attach to, is the most commonly tested model. More recently, the SRM ergometer has become the standard for power meters, and has been used measure the validity and reliability of other laboratory and cycle based ergometers. SRM power meters have been compared with the Kingcycle (Balmer et al. 2000a; Smith 2008), Monark 814e (Balmer et al. 2004), Axiom ergometer (Bertucci et al. 2005b), PowerTap (Bertucci et al. 2005a; Gardner et al. 2004; Paton and Hopkins 2006a) and the Velotron ergometer (Abbiss et al. 2009).

Results were varied in terms of agreement and the conditions used in testing. There was disagreement between the SRM and Kingcycle (Smith 2008), and for the Monark 814e (Balmer et al. 2004) and Velotron (Abbiss et al. 2009). The agreement between the two types of power meter depended on the type of test performed. There was agreement between both the PowerTap and Axiom ergometers and the SRM (Bertucci et al. 2005a). Peak power was measured using a Monark ergometer, a Kingcycle ergometer and SRM cycle based ergometer and compared with performance in a 40 kilometre road time trial (Smith 2008). Peak power output on the Kingcycle was 3.6% higher than the SRM. All three forms of measuring peak power output were strongly correlated with performance in the road time trial.

Another early model of power meter was the Ergomo, where the strain gauge was placed in the bottom bracket (axle on lower part of bicycle that left and right crank arms attach to). It was unique in that it measured power from the left side crank only and at the receiver the power recorded was doubled. Recently the Stages power meter became available and contains a strain gauge only in the left side crank. Rotor has released a new model that also features a strain gauge in the left hand crank only. Neither Stages nor Rotor have been evaluated at present. The Ergomo was evaluated by Duc (2007) and Kirkland (2008). Duc found the Ergomo less valid and reliable compared to an SRM and PowerTap. Kirkland found acceptable accuracy for the Ergomo but compared to SRM and Monark it was not as reliable.

Polar offered an early model based on the tension of the bicycle chain. This meter was assessed by Millet et al. (2003) and Hurst (2006). Millet found the Polar S710 to be valid and reliable for endurance cycling while Hurst incorporated a wider range of riding intensity in the testing and found large differences between the Polar S710 and an SRM. Polar have recently released a new Pedal based power meter with strain gauges in the pedal axles of both pedals. The Polar Keo was assessed by Sparks (2014) who found poor validity and reliably compared to SRM. Garmin have also released the Vector model power with strain gauges in the pedal but this model has not been assessed in the literature, nor have other crank spider based models produced by Power2Max, InfoCrank, Quarq and Pioneer.

The initial calibration of SRM power meters was performed using a dynamic calibration rig (Jones and Passfield 1998) or laboratory ergometer (Martin et al. 1998). Wooles et al. (2005) developed a newer method of calibration involving hanging a known mass from the crank to apply force that can be measured on the receiver and used to check if the calibration is correct. Currently only the SRM and Quarq brands allow users the option of recalibration of their power meters.

From the available research we can conclude that the SRM brand power meter offers a valid and reliable means of recording watts from cycling. The PowerTap brand offers similar confidence. Other models of bicycle based power meters and some static models do not offer the same promise. Several newer models have yet to be tested scientifically. Thus it is important to also understand the potential deficiencies when analysing other studies or comparing across studies using different power meters.

2.3 Use of a power meter to test various cycling models

Models are used to test various equations and can be used to estimate the effects of changing different parameters of the model. Numerous equations have been developed when using a power meter. The utility of these various models is that they allow the mathematical estimation of various changes in position on the bike, gearing, different equipment, rolling resistance riding on different surfaces, and on the bike sources of friction using different types of bearings, ceramic verses standard, as examples. This use of models offers both a cost and time saving with estimation using valid equations.

Martin et al. (1998) performed a calibration test on the SRM power meter and then conducted a series of road trials comparing the data against a model they developed. They found a strong relationship (R^2 =.97) between the model and the SRM. They then used the model to estimate the effects of various changes. Lukes et al. (2006) developed a model for riding on a banked velodrome that used SRM data as an input that effectively modelled 4000-m individual pursuit times. The SRM was used to validate this model (Lukes et al. 2012) finding less than 2% error.

Nine models of cycling power, excluding the Martin et al. (1998) model, were compared with SRM measurement (González-Haro et al. 2007). Estimates from Candau et al. (1999) and Di Prampero et al. (1979) provided the best models compared to SRM data. Underwood and Jermy (2010) developed a model for the individual pursuit and found the model accurate within 3% of data from the SRM. A model of outdoor cycling has also been used to develop indoor training systems that recreate the riding loads of riding outside (Dahmen et al. 2011).

Power meters are used to test models of aerodynamics and to test models to determine frontal area measures. Both allow for comparisons of different aerodynamic options for racing without spending time and money testing. Martin et al. (2006) tested sprint cyclists accelerating to model changes in aerodynamics. There were no substantial differences between the model and measures of frontal area taken from wind tunnel testing. A PowerTap meter was used to develop a model that measured frontal area and rolling resistance (Lim et al. 2011a). The PowerTap detected differences between riding on the tops of handlebars, drops of the handlebars, and riding with tyres inflated to either 60 psi or 120 psi. Bertucci et al. (2013) used an off-road version of the PowerTap to assess aerodynamic drag while seated and the difference in rolling resistance between knobbled and smooth tyres. Again the power meter was able to measure differences in position, tyre type and tyre pressure.

Several models of work performed while cycling has been compared with data recorded with a power meter. Power meter data has also been used to model aerodynamics. Both types of models have been used with success. These models can save money and time by reducing the need for expensive testing in the laboratory or testing of rider aerodynamics in a wind tunnel. They thus present a means of extending power meter data to better coach a rider.

2.4 Using a power meter to assess the cyclist

Assessment is the first step in the goal setting process a coach and athlete will carry out to plan for a goal race. Prior to the advent of power meters all testing work performed in watts was conducted in the laboratory. Testing in the field was based on times over specified courses. Power meters are now allowing the measurement of performance in competition offering a higher level of validity, both over an event and for specific critical or contextually significant points within an event.

Hawley and Noakes (1992) looked to estimate the relationship between peak aerobic wattage (W_{peak}) and $V\mathrm{O2}_{max}$ in a ramp test and then performance in a 20 kilometre time trial. Data from 54 male and 46 female subjects showed that W_{peak} explained 94% of the variation of $V\mathrm{O2}_{max.}$ A further 19 subjects then completed a 20km time trial and W_{peak} explained 82% of the variation. Balmer et al. (2000b) completed a similar study comparing maximal aerobic power to performance in a 16.1-km time trial. They found a very strong relationship (r = 0.99) between maximal aerobic power and time trial power output, but not finish time (r = 0.46). In contrast to previous studies by Hawley et al. and Balmer et al., a study comparing Wpeak and power output in both a 20-min and 90-min time trial had mixed results (Bentley et al. 2001), showing a relationship between a ramped power test and average watts in a 16.1-km time trial but not the completion time. This would suggest that other influences like gradient, road surface, rider position among other things have an influence on completion times for individual events.

Power meter data was used to determine the smallest worthwhile changes of cyclists power over a season between training phases (Paton and Hopkins 2005). Large improvements were seen over a season in both an incremental test and 4-km power test. However, the amount of variation within the base phase of training would make it difficult to determine if training strategies were having a worthwhile effect. Thus, Paton and Hopkins suggest that any experiments be used in the pre-competition and competition phases where the variation in performance is much smaller. These results indicate that short-term power results may not effectively track improvement well enough to guide training.

Based on findings that laboratory tests were a valid means of predicting cycling time trial performance Quod, et al. (2010) looked to test a method that would predict performance in mass start road cycling where tactics, bunch dynamics and power variations influence the outcome. They measured the power profile in the laboratory (5, 15, 30, 60, 240 and

600-s) and then power meter data from 10 cyclists recorded from competition. The study found no difference in the laboratory power profile and peak watts recorded from competition.

Pinot and Grappe (2010) proposed a variation for cycling based on recording peak powers over durations from 7 min to 2 h. They labelled the plot of results the *Power Profile*. Power outputs of five cyclists of different abilities ranging from second-category French cyclists to World Tour level were measured over a seven-month period. They determined 12 durations of peak powers expressed in watts per kilogram. The 12 durations were 1, 5, 30 s, 5, 20, 30, 45 min, 1, 2, 3, 4, and 5 h. The data was able to illustrate the changes in performance over the measurement period and that the more exercise an athlete performed the greater the change in power output over the season. The study highlighted the use of power output measurement to track changes in performance over a season and as a method to test various forms of training.

In a subsequent paper the *Power Profile* was renamed the *Record Power Profile* which was used to assess if power output in 17 cyclists over a 10-month period was able determine any differences between rider category (professional or elite), and the type of rider (sprinters, climbers and flat specialists) (Pinot and Grappe 2011b). They found that power output helped to create a signature that did differentiate between professional and elite cyclists and also found that the power profile indicated differences between hill climbers, sprinters and flat specialists. Hence, the results were encouraging for elite level riders in terms of using power to monitor and perhaps guide training.

Based on a model developed by Peronnet and Thibault (1989) that measured running performance based on times over several durations, Pinot and Grappe (2011a) used a variation based on power outputs of durations from 5 minutes to 2 hours to see if power output was comparable to the Peronnet and Thibault model. A total of 20 cyclists (elite and professional) supplied power meter data from training and competition over a 7-month racing season. They found that a variation of the model using power meter data was comparable with the Peronnet and Thibault model used to model world-record performances in running. This study provides information for the coach and rider that they can use to set realistic power targets for event specific durations to train towards.

These studies looking at maximum mean power values have relevance to this study as it examines the use of maximal mean power values as a measure of performance in competition. Max mean power values from competition have been shown to relate to power values from laboratory testing, which has higher reliability. Thus, their use to assess rider performance in events and training appears justified.

2.5 Using a power meter to measure demands of racing

The second step in the performance goal setting process is determining the demands of competition that cyclists would be expected to meet. In the past all, coaches and riders had to go on was distance and estimated time measures for certain events. Downloadable heart rate monitors have also added clarity to this picture. Several studies have now added measurement with a power meter to help riders and coaches understand what is required to perform in events, and thus, to help them plan training to an objective measure of that demand based on the riders current abilities.

SRM power meters were used to measure the demands of women's world cup single day cycling events (Ebert et al. 2005). Women in the Australian cycling team performed a graded exercise test to determine their lactate threshold, anaerobic threshold, and graded exercise test peak power. Zones determined from the testing were compared with power files from top 20 placing's to illustrate time spent in each zone, time spent for different absolute power levels, time spent in relative power levels, and differences between hilly and flat races. Elite women in single day races spent most of their race time riding between 100-300 watts, 2-5 watts per kilogram, and in flat races had lower maximum mean power (MMP) for 180 to 300 s compared to hilly courses. Maximal mean power is the average power for a set duration. In flat races riders spent more time above 7.5 watts per kilogram. This provides valuable information on what a women needs to compete at this level and tactical insight on how women's single day road races are typically raced in terms of power that is easily measured and can be used to guide training efforts.

Ebert et al. (2006) repeated their 2005 study with male cyclists collecting power files over 207 events in a six year period from 31 subjects. Their findings included large periods at very low power outputs mixed with durations of very high power output above a maximum aerobic power. This study gives insight on the way men's road races are raced where climbs are key moments in the race and flat periods are ridden at lower power, while waiting for uphill opportunities.

In another study, 6 professional male cyclists performed a graded exercise test to determine ranges, then competed in a six day cycling event using an SRM power meter (Vogt et al. 2006). Heart rate was also measured during the event. During the mass start road stages the participants spent most of their time riding close to the lactate threshold. Heart rate, when compared to power output, underestimated the time spent in the zone between lactate threshold and the anaerobic threshold. This study highlights the advantages of using a power meter over heart rate monitoring to understand the demands of competition, and indicates that power is perhaps a more objective measure than heart rate even though the former measures external work and the latter cardiovascular stress.

Both heart rate and power were measured from 15 professional cyclists competing in the Tour de France to document the power demands of the event (Vogt et al. 2007b). Comparisons were made between flat, semi-mountainous, and mountain stages. The findings illustrated that different types of stages required a different power profile. Vogt et al. (2007a) presented a case study of a male professional cyclist competing in the Giro d'Italia to show the variation of power between hilly and flat stages. Flat stages showed a higher variation in power, while the hill stages were characterised by periods of sustained power. These studies present a challenge to riders and coaches preparing for a three week stage race where different terrains and types of event, individual time trial, team time trial and mass start, provide different demands in terms of power that require different methods of preparation.

Eight male and 10 female cyclists used a PowerTap while competing in a three stage race that included a 4-km time trial, four road races and a short circuit race (Lim et al. 2011b). The men and women competed separately and the men completed their races over the same distance in faster times and at higher absolute power levels. All participants completed testing to determine their lactate threshold and, while there were differences in finishing times and absolute wattage, the power as a percentage relative to the measured lactate threshold between male and females was similar.

Power meters can be used in off road cycling events to measure the demands of various events. Macdermid and Stannard (2012) measured the physiological and mechanical work performed by seven participants on a cross-country race using a PowerTap. They found a high level of variability in the power recorded, with periods of low power and, on the incline sections short periods of very high power. This result highlights an issue with mountain-bike races running for around two hours, but which are classified as an endurance event. However, power-meter data shows that such events involve periods of anaerobic exercise, in either case; the power meter data presents a more precise measure.

Power meter data have allowed sport scientists and coaches to better understand the demands of cycling events. Comparison of these demands with cyclist assessment can be used to plan training to prepare to meet these demands based on the gap between a rider's current abilities and performance levels required to compete in goal events.

2.6 Using a power meter to monitor daily training

When the gap between a cyclist's current abilities and the demands of a particular event are known, the coach and rider can develop a specific, power-based training plan. This plan will consist of daily sessions that aim to build a riders ability to compete in the goal event. Daily sessions are monitored to ensure a rider is training with the goal event in mind. A power meter helps to ensure that training sessions are specific and the training loads are appropriate for the level of the rider.

Daily training loads were quantified as training stress score's (TSS) using power meter data (Garvican et al. 2010). The change in training loads measured in accumulated TSS scores that decay over a 42-d period are used to determine a chronic training load (CTL). CTL values were compared with changes in haemoglobin mass (Hb_{mass}) in female cyclists. The study also compared Hb_{mass} with changes in MMPs. Garvican highlighted issues with the failure of CTL to account for sleep, nutrition or travel on training loads. Hb_{mass} varied by 3.3% in female cyclists over a season and both changes in CTL and MMPs played a part in this variation. This study illustrates that, while measures of daily training load measured by a power meter can help with the training process, there are other data that contribute to the process and must be managed.

Swart et al. (2009) and Robinson et al. (2011) both performed studies that compared two groups training at a similar workload, basing their intensity either from a heart rate monitor or a power meter. Neither study found a significant difference in performance testing between methods and both studies concluded that a cheaper heart rate monitor was the better method of determining training intensity in training. However, these studies ignore the many other uses of a power meter outside of establishing training intensity. It also ignores heart rate training issues, that include cardiac drift, heart rate lag and that heart rate can not measure intensity beyond the maximal oxygen uptake (Jeukendrup and Diemen 1998).

Nimmerichter et al. (2011) monitored power and heart rate output from 11 subjects over an 11-month period to compare training related variables with performance measures. For low intensity training, both power output and heart rate provided useful information to monitor the training process. However, heart rate had limited application for higher intensities compared to the power meter and Section 2.5 has already illustrated how periods of high intensity characterise the tactically critical periods of most road cycling events.

These studies highlight the usefulness of monitoring power. However, it can also be argued that, for low intensity training, heart rate is still a valid measure of training intensity and determining training loads. Once the cyclist starts riding at higher intensities the power meter offers an advantage of tracking performance. This result would be especially relevant for sprint cycling, track endurance cycling and the parts of road cycling events determined by periods of high intensity.

2.7 Using a power meter to track training loads over time

It is easy to determine the effectiveness of individual training sessions based on comparing the session against previous data to see progress. Adding goal event data or estimated power for events ensures the sessions are headed in the right direction. The next step in the performance development process is using individual session data to build a picture of the long-term effects of training over time. Again, to determine if training is having the desired effect and also to ensure that training loads are appropriate. Training loads that are too low will not lead to progress and loads that are too high will also cause a stagnation or regression in performance.

Bannister and colleagues initial described performance as a function of a training impulse that incorporated both a fitness and a fatigue component (Banister et al. 1975; Calvert et al. 1976).

A training-impulse score (TRIMP) based on heart rate data recorded in 7 professional cyclists competing in one of the three Grand Tours (France, Italy or Spain) to try and model energy expenditure (Foster et al. 2005). The study found that the energy expenditure between different races was very similar as were the TRIMP scores. TRIMP is a heart rate based stress score and is illustrated in Section 2.6 for low intensity events heart rate data can be an effective method of tracking training loads.

Seven young cyclists were measured before and after a 14-wk period of training. From the ride data, an objective training load was determined using a TRIMP model and combined with subjective data to model training loads, monotony, strain and fitness-fatigue (Delattre et al. 2006). They found differences between how hard the riders thought they were training and how hard they were actually training. The rider's velocity at both maximum oxygen uptake and ventilatory threshold increased, and it was concluded that both subjective and objective data led to a more complete picture of the long term training process than just physiological measures alone. The training stress scores from this model were based on heart rate data. It still remains to be seen if training scores based on power data will model the actual effects of training and racing better than TRIMPs.

Jobson et al (2009) reviewed research on the analysis and utilisation of training data for cyclists. In previous years, cyclists have used either duration or the volume of kilometres covered over a certain period. Jobson points out that either method fails to take account of intensity. Heart rate and time have been used to generate a TRIMP score to try and take into account of both the training load and the training stress. With the increasing use of power meters, cyclists are now able to measure work output when training and racing in the field.

The Jobson review (2009) highlights an issue of measuring average power due to the variability inherent in measuring watts. TrainingPeaks™ software uses a normalized power multiplied by time to determine a training stress score. The calculation of training stress scores is complicated by the need to have a regularly updated functional threshold to determine an intensity factor for each ride to allow the calculation of the training stress score. This necessitates regular testing to maintain an accurate threshold to ensure an accurate training stress score is determined.

A case study followed the 1500-m running performance of an Olympic athlete (McGregor et al. 2009). The aim of the study was to assess whether a performance manager model would relate to outcomes in competition over a seven-year period. Training logs were used to generate a Training Stress Score (TSS) for each training sessions. The TSS was then used in a performance manager model to determine acute and chronic training loads. Performances from 800-m and 1500-m running races were used to generate Mercier scores. A Mercier score is based on the time relative to international standards of times for all the common distance in track running races. The Mercier scores were correlated with TSS, acute and chronic training loads. The predictions of the simplified performance manager model correlated with the Mercier scores from competition. It is the first published study to use training metrics, such as training stress scores, acute training loads, chronic training loads, training stress balance and performance managers discussed in Allen and Coggan (2010). A challenge for the current study is finding measures of performance in the sport of cycling where more factors determine the outcomes than running, where times regardless of tactics, are a better determinant of outcome.

Wallace, Slattery and Coutts (2014) performed a similar study to McGregor using seven trained runners who completed 15 weeks of training. Training dose was determined three ways: ratings of perceived exertion, TRIMP based on heart rate, and a running training-stress score. A model of performance (weekly 1500-m running time trial), fitness (submaximal and resting heart rate) and fatigue (heart-rate variability and Profile of Mood States) was developed based on each of the training dose calculations and the relationship between the models was performed. Modelled performance correlated with actual performance in all methods of determining training dose: $r = 0.60 \pm 0.10$ for session RPE, 0.65 ± 0.13 for TRIMP, and 0.70 ± 0.11 for running training-stress score. The correlations between modelled fitness and fatigue were moderate to moderate-to-large. These results provide more evidence for a training stress score over TRIMP and perceived exertion however a small number of subjects over a shorter period of time mean results may not be repeatable.

Most coaches adopt a form of varying training loads and training types over a season. One form of planning a training year is periodization (Smith 2003). Although different groups use different terminology for different phases, a central theme is a base phase aimed at building general fitness, a build phase aiming at specific fitness and a competition phase aimed at performing at the cyclists best. The relevance to this study, based over a 6-8 month period, is that cyclists will be racing and training through different phases as they prepare for these major events.

2.8 Justification for use of analysis and research methods

The first step of this study is to ensure that the performance measures used are reliable. With good reliability we can be sure that any changes in performance measures reflect the changes in independent variables. The coefficient of variation (CV) is used to determine if the performance measures are reliable (Hopkins 2004).

To determine the smallest worthwhile change in performance required to be competitive in various cycling events Paton and Hopkins (2006b) reported the variation between races. Higher between cyclist variation for average cyclists was reported compared to elite riders, and depending on the type of event were found. The lowest within cyclist variation was seen in road races, ~0.5%, and the highest in mountain biking events, ~2.5%. Smallest worthwhile changes of ~0.5% (kilometre time trial), ~0.6% (road time trial) and ~1.2% in mountain bike races. Mass start roads where riders compete in a group precluded the estimation of a worthwhile difference, as tactics and other factors predominated.

Pinot and Grappe (2011b) found a between-subject CV of 6.1% and 13.1% for the thirteen MMPs they recorded. Higher variations were seen in their data for the durations of 1-s, 5-s and 30-s (11.7% - 13.1%) than durations of 60-s to 240-min (6.1% - 8.8%). The aim of the study was measure MMPs for each rider over a season and compare riders of similar ability, Professional and Elite, classifying them in relation to their specialities in cycling: sprinters, climbers and flat cyclists. In Quod et al. (2010) CV was not determined for MMPs tested in the laboratory and from field based data. Standard deviations from the shorter durations between-subject MMPs were larger (74-125W) than longer (25-42W). Our study is the first to report CV within-subjects for cycle races. No means, SD or CV were reported for Pinot and Grappe's 2010 and 2011a papers.

2.9 Research Question

The previous research in the area of using a power meter to monitor the competition preparation process has been conducted using case studies or small groups. The McGregor et al. (2009) case study was based over several years. However, other researchers have used study durations of 2-3 months. Many of the studies have used international-level elite male road cyclists competing at the very highest levels in the sport.

CHAPTER 3: METHODS

3.1 Introduction

This is an observational study and analysis of participants submitting cycling power meter data over a period of 6-8 months.

3.2 Subjects

Advertisements were placed in several New Zealand social media pages seeking participants in the study (Appendix 1). Participants were invited to share their power meter files for a period of six months or greater. Exclusion factors were significant periods of time with no files submitted within the time frame and data sets that included no race files.

Participants needed to have either an SRM or Quarq power meter to be involved in the study. Both of these brands can be checked for calibration, and, if out of an acceptable range, can be re-calibrated. A calibration would have needed to be performed within 3 months of submitting their power meter files.

Thus, participants needed to be proficient users of a power meter. In particular, participants had to know how to zero the power meter before each ride. At a basic level a power meter operates like a set of scales where, if the scales do not read zero with no weight, then any subsequent measurement will be inaccurate. The same applies to power meters, which should read zero with no force being applied.

Participants also needed to know how to upload the power meter file to a computer to be transferred to the researcher to carry out the analysis. This process recreates the way an athlete would record their training and racing rides, and then upload the files so a coach or sport scientist can view and analyse them. Individual files were downloaded to the researcher via the commercial TrainingPeaks^{TM} website and were uploaded in bulk (6-9 month period of training and racing). Hence, the study operates in a similar fashion to the coach-athlete interaction seen in the real world.

Based on these conditions Nationally and Internationally competitive cyclists (20M, 5F), age 29 ± 9 y, weight 71 ± 7 (mean ± SD), were able to provide recordings from a six-month period. All participants were informed of the procedures, risks, and benefits of the study (Appendix 2), completed an informed voluntary consent form (Appendix 3), and completed a pre-exercise questionnaire (Appendix 4). The AUT University Ethics Committee (AUTEC) granted ethical consent for this study and the use of the (anonymised) data before participants were recruited for involvement in this study (Appendix 5).

3.3 Study protocol

The participants submitted power meter files for a period of 6-8 months. In this time, all training and racing files were downloaded to a computer and then sent via email or uploaded to an online training site from which they could be downloaded. An account was created for each participant in TrainingPeaks WKO+ 3.0^{TM} software, and each file was added to each subject's account. This data management and transfer method is common practice among cyclists who use power meters to supply data so the coach or sport scientist can view the file. TrainingPeaks WKO+^{TM} allows analysis of the individual ride and uses the collation of data over time to try and determine levels of fitness, fatigue, and freshness for competition.

Within the WKO+^{TM} software a measure of power in each second was recorded. From this record over time a variety of measures were calculated. First, the power was normalised using a 30 second rolling average to reflect the higher physiological cost of riding at higher intensities. The normalised power was then compared with the functional threshold of the rider to determine an intensity factor (IF) for the ride. Functional threshold was determined for each subject based on 95% of their 20min power (Allen and Coggan 2010). Based on IF squared times duration a training stress score (TSS) was determined for the ride. Functional threshold for each rider was determined, using a critical power model, based on mean maximal powers for 3-min, 8-min and 20-min (Allen and Coggan 2010).

The measurement of TSS over time comes from the WKO+ 3.0^{TM} software using a performance manager based on the training-impulse model to provide measures of fitness, fatigue and freshness. Chronic training load (CTL) is a measure of fitness based on the accumulation and decay of TSS/day over a time period, usually with a time constant 42 days. Acute training load (ATL) is a measure of fatigue based on the accumulation and decay of TSS/day over a time period, usually 7 days. Training stress balance (TSB) is a measure of freshness based on a simple equation of CTL – ATL.

Participant data was analysed with the performance manager in WKO+^{TM} software to derive measures of fitness, fatigue and freshness. Specifically they used power meter files from training, one-day mass start road races, one-day individual time trials and mass start road races from a multi day events. For each race file TrainingPeaks WKO+^{TM} was used to determine the 5 second, 60 second, 5 minute and 20 minute maximum mean power.

3.4 Analysis

Maximum mean powers (MMPs) for 5-s, 60-s, 5-min and 20-min durations were determined from the data files using the TrainingPeaks WKO+^{TM} software for each race and sorted into individual time trials, single day road races and multiday road races. Only data from participants who supplied two or more files from the same kind of race were retained for analysis.

A mixed-model procedure (Proc Mixed) was used in Statistical Analysis System (Version 9.4; SAS Institute, Cary, NC) for subsequent analyses. MMPs were log-transformed before analysis to remove skew, and effects and errors from the analyses were back-transformed to percent units. The within-subject variability of each MMP for each of the three kinds of race was first determined with a simple reliability model consisting of individual subject mean values (specified with random effect for subject identity) and the residual error, representing the within-subject variability and expressed as a coefficient of variation (CV). Because the time-trial distances ranged from 3 km to 90 km, an adjusted CV was also determined for the time trials by including the log of the distance as a fixed-effect predictor to adjust for the effect of distance.

Initial analyses showed that the CV of the MMPs was too large to allow for any possibility of useful relationships between MMP and the fitness, fatigue and freshness measures. Therefore analyses were done using only the top half of each rider's MMP values. This choice eliminates races where the participant either did not race maximally for tactical or training reasons, or the course did not facilitate high power outputs because of downhill sections and tailwind sections. It can also eliminate road races where participants were part of a large group of riders sheltered from the wind. This approach thus uses the residuals in the reliability analyses to select those races providing the highest MMP values would be more or most representative of maximal performance in an event.

For these races, the fitness, fatigue and freshness scores from the previous day were extracted and each score was used as a predictor in separate analyses for each of the MMP values. The mixed model was the same as for the reliability analyses, with the addition of each of the training measures as a fixed effect. The magnitude of the effect was evaluated for two within-cyclist SDs of the training measure (Hopkins et al. 2009), as follows:

1. Each cyclist's mean training measure was first standardized to zero
2. The resulting measure was then standardized to an overall SD of 0.5

When included in the model, the coefficient for this measure is the effect of two within-subject SD of the measure on the MMP. For the time-trial events the effect of doubling the distance of a time trial on each MMP was evaluated.

Uncertainty in the estimates of CV and of effects on MMPs is presented as 90% confidence limits, in ×/÷ form for CV and in ± form for effects. Magnitude-based inference was used to make conclusions about true effects (Hopkins et al. 2009). The smallest worthwhile change was set to 1.0%. If the confidence interval for the true effect overlapped substantial positive and negative values, the effect was deemed unclear. Otherwise, the effect was deemed clear and was qualified with the probability that the effect was substantial (possibly, 25-75%; likely, 75-95%; very likely, 95-99.5%; most likely, >99.5%).

CHAPTER 4: RESULTS

4.1 Introduction

From the 25 participants, five provided multiple files of time trial event data, 25 provided multiple road race data files and 16 provided multiple multi-day data files. No subjects were excluded from the study for either long periods of no data or for supplying a data set with no race files. A total of 6232 data files were submitted for racing and training. Of these files 502 were for race days comprising of 52 time trial files, 249 road race files and 201 files for multi-day races.

4.2 CV for max mean power in racing

Table 1 shows the coefficient of variation (CV) for all races and with subsequent adjustments. CV for MMP_{5s} from racing ranged between 13 and 15% for all events. Because of the level of variation, the analysis was repeated using only the top half of the MMP's for each subject and for each event. The top quarter or top third of MMPs were not used, because this would have left too few data points for subsequent analyses (even using the data from half of the races, many of the effects turned out to be unclear).

A separate analysis of the time trial data was performed based on the distance of the time trial events. These are presented in Table 1. For a time trial, adjusting for the distance improved CV for the MMP_{5s} and MMP_{60s} by negligible amounts, but CV for the other MMPs were improved. Using the top half of the power values reduced all CV by up to a factor of 2. As the duration of MMPs increased, the CV for the top half of races dropped for each type of race, with the exception of multi-day events between MMP_{5m} and

MMP_{20m}.

TABLE 1. Within-cyclist race-to-race variability in maximum mean power (MMP) expressed as coefficients of variation (CV, %) for all data without adjustment, for all time trials adjusted for distance, and for the top half of each cyclist's MMPs adjusted for distance (time trials) or unadjusted (road races). The top-half CV is the smallest of the three values of residuals in the analyses for the effects of fitness, fatigue and freshness.

	Time-trials			One-day road-races		Multi-day road races	
MMP	All data, unadjusted	All data, adjusted	Top half, adjusted	All data, unadjusted	Top half, unadjusted	All data, unadjusted	Top half, unadjusted
MMP_{5s}	15.0	14.9	12.5	14.9	6.7	13.0	5.5
MMP_{60s}	10.4	9.8	6.5	10.8	7.7	9.7	5.8
MMP_{5m}	5.4	3.6	2.5	7.0	4.8	7.4	4.2
MMP_{20m}	4.2	3.4	2.4	7.2	4.1	10.4	4.4

Uncertainty (90% confidence limits): time trials all data, ×/÷1.23; time trials top half, ×/÷1.40; road races all data, ×/÷1.10; road races top half, ×/÷1.14.

Means, between-cyclist SD and within-cyclist SD for fitness, fatigue and freshness and MMPs are described in Tables 2-4. The within-cyclist SD for fitness, fatigue and freshness was 11-26 of the performance manager measures (TSS/d or TSB) for each cyclist, for each type of event.

The main observation from the summary data presented in Tables 2-4 is the within-cyclist SD for time-trial MMPs were smaller, as the duration of MMP became longer, implying approximately four times greater consistency in MMP performance between the shortest and longest durations. For road races the range in within-cyclist SD between MMP_{5s} and MMP_{20m} was only a factor of 2, and for multi-day road races there was little difference between the MMPs.

TABLE 2. Descriptive statistics of time trials corresponding to the top half of time trial MMPs. Fitness, Fatigue and Freshness measures are from the day before competition. Within cyclists SD are the mean of the individual cyclists' SD.

	Mean	Between-cyclist SD	Within-cyclist SD
Time trial distance	24 km	18 km	14 km
Time trial duration	33 min	24 min	20 min
Fitness	98	17	11
Fatigue	106	20	22
Freshness	-9	17	18
MMP_{5s}	889 W	30%	12%
MMP_{60s}	500 W	24%	7%
MMP_{5m}	397 W	19%	4%
MMP_{20m}	358 W	20%	3%

All data except those providing 5 cyclists entering 4.2 ± 1.6 time trials (mean ± SD; range 3-7). Only four of those cyclists contributed to the MMP_{20m} data 4.3 ± 1.9 time trials (mean ± SD; range 3-7).

TABLE 3. Descriptive statistics of single day races corresponding to the top half of road race MMPs. Fitness, fatigue and freshness measures are from the day before competition. Within cyclists SD are the mean of individual cyclists' SD.

	Mean	Between-cyclist SD	Within-cyclist SD
Fitness	80	22	11
Fatigue	86	26	16
Freshness	-9	15	16
MMP_{5s}	927 W	27%	6%
MMP_{60s}	485 W	23%	7%
MMP_{5m}	349 W	21%	4%
MMP_{20m}	300 W	21%	3%

Data provided by 25 cyclists for entering 4.2 ± 1.6 road races (mean ± SD; range 1-18).

TABLE 4. Descriptive statistics of multi day road races corresponding to the top half of road race MMPs. Fitness, fatigue and freshness measures are from the day before competition. Within cyclists SD are the mean of individual cyclists' SD.

	Mean	Between-cyclist SD	Within-cyclist SD
Fitness	81	22	15
Fatigue	96	23	26
Freshness	-5	11	20
MMP_{5s}	961 W	25%	5%
MMP_{60s}	494 W	19%	6%
MMP_{5m}	357 W	18%	4%
MMP_{20m}	300 W	19%	4%

Data provided by 16 cyclists for entering 6.4 ±4.8 multi day races (mean ±SD; range 1-18).

4.3 Effects of doubling time trial distance on MMPs

Estimation of the effect of doubling the distance of the time trial events is presented in Table 5. Results showed a likely to very-likely substantial change in effect for MMP_{5m} and MMP_{20m}. Effects were unclear for the shorter duration MMPs. For each of the durations, the effects of doubling of time trial distance on MMPs after adjusting for fitness, fatigue and freshness were similar.

TABLE 5. Percent effects of a doubling time-trial distance on maximum mean power in the analysis of the effects of fitness, fatigue and freshness the day before. Analyses are for time trials that provided each cyclist's top-half values of MMP. Data are mean effects and 90% confidence limits.

	Fitness	Fatigue	Freshness
MMP_{5s}	-2.3, ±5.1	-2.3, ±5.2	-1.4, ±4.8
MMP_{60s}	-1.5, ±3.4	-2.5, ±3.2**	-1.8, ±3.0
MMP_{5m}	-3.8, ±1.8***	-4.0, ±1.6**	-3.8, ±1.6***
MMP_{20m}	-3.2, ±2.1***	-3.2, ±1.8***	-3.0, ±1.8***

Magnitude-based inferences evaluated in relation to a smallest substantial change of 1% as follows: *possibly substantial, **likely substantial, ***very likely substantial. All other effects were unclear.

4.4 Effect of fitness, freshness and fatigue on MMPs.

Figure 4 illustrates the scatter of delta fitness, fatigue and freshness for delta MMP_{20m} in multi-day races. It highlights the poor relationships between training and performance. Similar scatterplots were seen for all duration of MMPs and for each type of event.

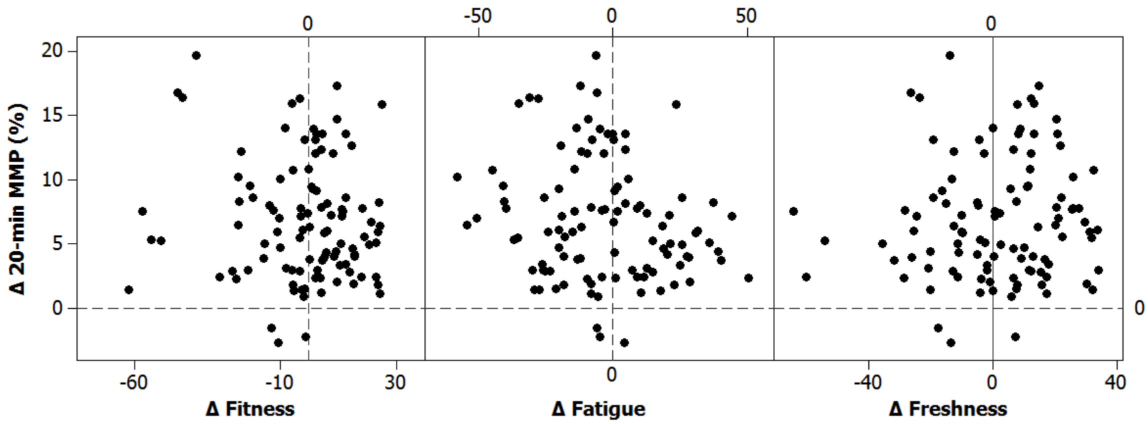


FIGURE 4. MMP_{20m} for multi-day races. Top half of rider data. Delta MMPs are log transformed. Delta fitness, fatigue and freshness raw data expressed as above or below the mean of each participant.

The effects of fitness, fatigue and freshness measures on MMPs for all three types of competition are presented in Table 6. For all types of events the effects of delta fitness, fatigue and freshness on the majority of shorter duration delta MMPs were unclear and when possibly substantial they were in the wrong direction to what we would expect from the impulse-response model. More possibly substantial and likely substantial findings were found for longer duration delta MMPs however these were also in the wrong direction. It was only for multi-day races where the delta fitness and freshness for delta MMP_{20m} where in possibly substantial and in the direction we would expect from the impulse-response model.

TABLE 6. Percent effects of two within-subject SD of fitness, fatigue and freshness the day before a competition on maximum mean power (MMP) in the competition. Analyses are for competitions with each cyclist's top-half values of MMP (after adjustment for distance of time trials). Data are mean effects and 90% confidence limits.

	Fitness	Fatigue	Freshness
Time trials			
MMP_{5s}	4.6, ±11.1	4.6, ±11.2	7.6, ±10.7
MMP_{60s}	-2.0, ±5.6	3.1, ±5.6	-3.7, ±4.9
MMP_{5m}	-0.2, ±2.3	1.3, ±2.1*	-0.3, ±2.1
MMP_{20m}	-0.2, ±2.4	-0.4, ±2.1	0.6, ±2.1
One-day road races			
MMP_{5s}	-0.1, ±2.1	-0.1, ±2.1	-1.1, ±2.1*
MMP_{60s}	-1.6, ±2.4*	-0.5, ±2.4	-0.6, ±2.4
MMP_{5m}	0.4, ±1.5	0.1, ±1.5	-0.3, ±1.5
MMP_{20m}	-0.9, ±1.5*	0.5, ±1.3*	-0.9, ±1.3*
Multi-day road races			
MMP_{5s}	-4.0, ±1.8**	-3.9, ±1.8**	0.5, ±2.0
MMP_{60s}	0.8, ±2.0	-0.4, ±1.9	0.7, ±2.0
MMP_{5m}	-0.3, ±1.4	-0.5, ±1.4*	-0.2, ±1.4
MMP_{20m}	0.5, ±1.3*	-1.8, ±1.4**	0.8, ±1.5*

Magnitude-based inferences evaluated in relation to a smallest substantial change of 1% as follows: *possibly substantial, **likely substantial, ***very likely substantial. All other effects were unclear.

CHAPTER 5. DISCUSSION

The first aim of the study was to assess the reliability of maximum mean power (MMP) data recorded in competition. The shorter the duration of the MMP the higher the variation that was seen. This variation was seen in each of the time trial, road race and multi day race data.

Based on this variation, a reliability analysis was performed to adjust for the distance of the time trial events. Subsequently, only the top half of each rider's MMP values from each different type of competition was used to mitigate each cyclist's variability. Adjusting for time trial distance had a negligible effect on variation. However, using the top half of MMP values did reduce the CV for each type of race, providing a potentially more objective measure.

The second aim of the study was to determine if the fitness, fatigue and freshness measures from TrainingPeaks WKO+^{TM} software could be used to accurately predict performance in road cycling events. The improved CV, using the top half of the MMP values, was used to address the second part of the research question. Strong relationships were found when the distance of the time trial was doubled with MMP_{5m} and MMP_{20m} when any effects of fitness, fatigue and freshness measures were taken into account. When assessing a two standard deviation change in fitness, fatigue and freshness measures on the change in MMP values from all three different types of racing, most of the results were unclear. Many of the possible or likely changes in MMP associated with changes in fitness, fatigue and freshness, were contrary to what the performance manager model may predict. In particular, these counter intuitive results may indicate the dominance of external factors not measured, quantified in this study, or that the model used to assess the data is flawed.

One of the main relationships found in the results of this study was the improvement in CV as the duration of MMPs increased. It was only in the multi-day races, and only for MMP_{20m} that a possibly substantial relationship was seen between MMP values and the fitness, fatigue and freshness measures.

In these cases fitness and freshness were positively associated and fatigue negatively associated for multi-day races as would be predicted in a performance manager model. However, for single day road races the opposite occurred where fitness and freshness were negatively associated and fatigue positively associated with MMP_{20m} from road race competition. This counter intuitive result may indicate that tactics and other external factors may dominate one day road race events in comparison to time trials and longer multi day events.

The improvement in CV as duration increases for all types of events is most likely due to the more aerobic nature of road cycling events. In particular, MMP_{5s} and MMP_{60s} can only contribute to a very small portion of the duration of even the shorter (1- to 2-h) road races or time trial events. Even for short road time trial events of 3-8 kilometres riders are coached to not make very short maximal efforts and to try and pace their rides to maximise their anaerobic contribution to the performance.

When only the top half of each rider's MMP values were used in the reliability analysis a reduction of CV was seen for all events and all durations. The drop in CV would suggest that seasonal strategies and tactics are involved with producing MMP values measured in competition. In addition, some races are used for training and maximal efforts are not the rider’s intention further reducing the value of MMP data for such races. The use of races for training is more likely in the early part of the season where building fitness is a priority over developing race-winning power.

When aiming to win a race it is more likely that a rider will try and conserve energy by not making maximal power efforts unless needed. A conservative racing strategy could increase the possibility that, in any given race file that, where one maximal duration effort for may be seen the other MMPs may be lower. This would affect the CV when collecting MMP data for several durations from each race.

For time trials, adjusted for top half of power, CV for 20-min was 2.4%, which is an acceptable level of variation. As cyclists are coached to adopt an even pace in time trial events from 2- to 180-km it is expected that one would see better reliability from a 20-min effort in these events. With limited carbohydrate stores and the negative consequences of anaerobic energy supply on endurance performance, it is expected that time trials of 40 km or less are performed at or above threshold levels, so pacing is all the more crucial. Hence, we do not see high maximal MMP_{5s} and MMP_{60s} in this event. Similarly, higher MMP_{5m} may only be seen in the shorter distance timed events.

Increasing time-trial distance led to a decrease in MMP values, which is expected and gives some confidence that there is meaningful information from power data by providing a sanity check for the measurements. Pinot and Grappe (2011a) had a comparable result when they doubled the duration of their record power profile, which is essentially similar to MMP values used here.

That fitness, fatigue and freshness measures were all similar, indicating that none had an effect on MMP values after adjusting for time trial distance was unexpected. It would have been expected to see a negative effect of doubling the distance for fitness and freshness on time trial performance however this was confounded by a negative effect for fatigue suggesting increasing fatigue levels led to higher MMP values in time trials. However once the bottom half of riders MMPs from time trials were removed only 5 subjects provided data for MMP_{5s}, MMP_{60s} and MMP_{5m}, and 4 subjects provided MMP_{20m} data. Thus, the results may be biased by reduced sample sizes.

The effect of a two standard deviation change in fitness, fatigue and freshness on a change in MMP values from all three different types of competition was mostly unclear. This counter intuitive outcome is most likely due to the higher level of variability of the MMP recorded in competition. Other reasons for these results would be the interaction between the fitness, fatigue and freshness measures at any one point during a season and also their interaction over a season. These latter issues point towards the *tactical* training over a season as riders target specific races and goal performances.

In particular, over a season the coach and rider must plan periods of the season where they target different aspects of a rider's preparation for competition. Riders engage in periods of training and racing where performance in racing is not the goal. Building high levels of fitness to withstand future racing demands will lead to higher levels of fitness and fatigue measures, and a corresponding lower level of freshness. As the competition goals become a priority the levels of fatigue are reduced with the aim of increasing freshness measures and maintenance of the fitness levels.

This study used competition data from all dates within a six to eight month period where some or all such periods may be encountered for a given rider. Thus, a rider could conceivably be competing while under very high levels of fatigue. Lower level competition MMP values may have been accounted for by only using the top half of race power in competition. Equally possibly, it may also suggest that future analysis only use MMP values from goal events. However, such a reduction in data may limit the number of measures recorded from competition.

The McGregor et al. (2009) case study found a positive relationship between freshness measures and the Mercier Scores they used as a performance measure. Their study of one runner used competition and training data from a seven-year period featuring fewer events than road cycling. In addition in their study, it would be expected that performances, especially in an 800-m or 1500-m running race, would be relatively maximal, and thus measures from competition are less influenced by tactics than road cycling. Also important to make the distinction between running and cycling that riding has no eccentric component, thus, affecting the loading and differences in recovery of efforts between the two sports. The study presented here used a larger number of road cycling competitions, which increases the likelihood of events performed at less than 100% effort.

In the comparison of different methods of quantifying training loads Wallace et al. (2014) found an improvement in 1500-m running performance could be modelled by RPE, TRIMP and a running training stress score. The running training stress score that was estimated correlated the best with running performance. The 1500-m running race may offer a better measure of competition performance than road cycling events. In cycling, a comparable event would be the 4000-m individual pursuit in track cycling, which has less variability than mass start road races.

The first main implication of the results in this study is that to objectively and directly assess performance in road cycling events there is a need for better measures than just maximum mean power. The aim of road cycling events is to win or support a team rider to win and not attain maximal powers for any duration. For individual events MMPs may also be dependent on the course, if strong tailwinds, declines or technical sections limit the delivery of power.

The second main implication is that fitness, freshness and fatigue measures in themselves may not be the best option to predict performance before competition. Other contextual details are needed to advise cyclists about their state of preparedness for competition, so they can plan for each event based on measureable, objective knowledge of their current ability. Pre race information on a riders current ability may include basic physiological measures such as resting heart rate, heart rate variability, weight upon rising on race day (hydration levels) and hours of sleep, qualitative measures like quality of sleep, Profile of Mood States, and ratings of perceived exertion. Finally they can include invasive measures from blood samples such as lactate levels, serum urea and cortisol. Hence, it appears that MMP values are not necessarily sufficient without context.

With the level of competition, and the rewards and opportunities seen in the sport for elite performance, the planning and management of performance is a crucial part of any cyclist's preparation. With the rewards in cycling it highlights the need for simple and more easily obtainable measures, ideally from competition and training itself to allow the cyclist more time between riding and racing to recover both physically and mentally but taking measures from specific training and the racing itself rather than performing laboratory based tests. The opportunity to model performance in the field based on cycling based measures is promising area of research, as it requires very little input beyond riding the bicycle from the cyclist. Hence, the input of power meter data into the performance manager model allows for quantitative analysis of the actual work performed. However, it is only useful if the results generated have a valid and reliable effect on performance.

Although not documented in the literature, there can be a level of error when recording, downloading and transmitting power meter data. It was left to the participants to ensure their power meter was calibrated before starting the study and to have zeroed the meter before each ride. A power meter is also dependent on temperature, so it is conceivable that changes in temperature occurred during longer rides and races that would have changed the zero-offset of the meter.

In addition, different power meter receivers record data from the power meter in different ways and the transmission protocols differ from receiver to receiver. Different models of receiver also use different software and methods to download to a computer and different software packages analyse and manipulate the data in varied ways. All of these technology differences may also add subtle error or noise to the data that could not be considered here.

To address this issue, a level of consistency was maintained by ensuring subjects were competent power-meter users and that they all used a consistent method of downloading t and transmitting their files to the researcher In addition, the researcher only used one software package to determine MMP values from competition and the subsequent fitness, fatigue and freshness scores from racing and training. However, the process has a level of error starting with the power meters themselves displaying an error of ± 2% (Gardner et al. 2004) which is within or near some of the adjusted CVs, such as the MMP_{20m} for time trials that were reported here.

This study used racing and training data from a group of elite and international-level cyclists. For the international cyclists, riding is their full time profession, however for the elite cyclists the majority are employed outside of competing in cycle races. Working full time, part time or being a full time cyclist adds to the contextual data a coach and rider must assess when planning for peak performance in competition. The subjects used in this study may limit the findings when applying the conclusions to World elite-level riders and sub-elite riders who have different conditions to train and race in compared to elite and international level riders.

This study is the first to assess the performance manager, in TrainingPeaks WKO+™, model based on data from a power meter. Further research is needed to improve the reliability of power meter data from competitions to fully assess the fitness, fatigue and freshness measures and their efficacy to predict performance. In addition, more work is needed to understand the performance manager and other impulse-response models within the context of a cycling season, where the types, demands and levels of competition vary along with and the effects of periodisation in their impact on the performance of the rider.

A potential first step in finding better measures of performance would be to start with testing performance in time trial events only with a larger number of subjects compared to this study as timed events were the most consistent and least affected by external factors. A further step would be conducting research using road cycling based only on key events, such as national championship or international championship where competition data is expected to be of a maximal in nature.

In addition studies must account for the types of events where MMP duration chosen reflects the power most likely to influence race outcome. For example, in a track cycling sprint this duration is 15-s, for a 4000-m track pursuit it is 4-min while a mountainous road race might demand 20-60-min power depending on the length of the climbs. Potentially, maximal powers from key events could test the utility of the performance manager model to predict outcomes for specific types of events and also to test if the model has relevance to short duration events in track cycling, where developing high levels of fitness may compromise short duration power output due to adaptations to type I muscle fibres at the cost of developing type IIa and type IIx fibres.

CHAPTER 6. CONCLUSIONS

Key findings of this research are that maximal mean powers recorded from competition have high variability. The variability makes it difficult to use them as a dependant variable in research to determine if any worthwhile change has occurred. Reducing the amount of maximal mean power data by removing the lower half of data did improve the variability of road cycling racing performance data.

Even when only the top half of maximal mean powers for competition were used, the effects of fitness, fatigue and freshness measures from the impulse-response model in TrainingPeaks WKO+ 3.0™ software were unclear and often in the wrong direction to what would be expected.

The implications of this research are that coaches and sport scientists need to find better measures of performance for road cycling events than maximum mean powers. Once better measures of cycling performance are used, it will be possible to test the efficacy of the fitness, fatigue and freshness measures. With the many variables that shape a performance outcome in road cycling, determining a specific outcome measure may prove too great a challenge, and performance models should be developed that incorporate as many of these variables as possible. For the coach and rider it means looking beyond just the power meter data recorded in training and competition for insight into potential for improved performance.

Future research should focus on the determination of outcomes from road cycling events. Until the variability is reduced it will be hard to estimate the effects of various interventions or the utility of various models on cycling performance.

REFERENCES

1. Abbiss CR, Quod MJ, Levin G, Martin DT, Laursen PB (2009) Accuracy of the Velotron ergometer and SRM power meter. International Journal of Sports Medicine 30: 107-112
2. Allen H, Coggan A (2010) Training and racing with a power meter (2nd Ed.). VeloPress, Boulder, CO
3. Balmer J, Bird SR, Davison RCR, Doherty M, Smith PM (2004) Mechanically braked Wingate powers: agreement between SRM, corrected and conventional methods of measurement. Journal of Sports Sciences 22: 661-667
4. Balmer J, Davison R, Coleman D, Bird S (2000a) The validity of power output recorded during exercise performance tests using a Kingcycle air-braked cycle ergometer when compared with an SRM powermeter. International Journal of Sports Medicine 21: 195-199
5. Balmer J, Davison RCR, Bird SR (2000b) Peak power predicts performance power during an outdoor 16.1-km cycling time trial. Medicine & Science in Sports & Exercise 32: 1485-1490
6. Banister E, Calvert T, Savage M, Bach T (1975) A systems model of training for athletic performance. Australian Journal of Sports Medicine 7: 170-176
7. Bentley DJ, McNaughton LR, Thompson D, Vleck VE, Batterham AM (2001) Peak power output, the lactate threshold, and time trial performance in cyclists. Medicine & Science in Sports & Exercise 33: 2077-2081
8. Bertucci W, Duc S, Villerius V, Grappe F (2005a) Validity and reliability of the Axiom Powertrain cycle ergometer when compared with an SRM powermeter. International Journal of Sports Medicine 26: 59-65
9. Bertucci W, Duc S, Villerius V, Pernin J-N, Grappe F (2005b) Validity and reliability of the PowerTap mobile cycling powermeter when compared with the SRM device. International journal of sports medicine 26: 868-873
10. Bertucci WM, Rogier S, Reiser RF (2013) Evaluation of aerodynamic and rolling resistances in mountain-bike field conditions. Journal of Sports Sciences 31: 1606-1613
11. Calvert TW, Banister EW, Savage MV, Bach T (1976) A systems model of the effects of training on physical performance. Systems, Man and Cybernetics 2: 94-102
12. Candau RB, Grappe F, Ménard M, Barbier B, Millet GY, Hoffman MD, Belli AR, Rouillon JD (1999) Simplified deceleration method for assessment of resistive forces in cycling. Medicine & Science in Sports & Exercise 31: 1441-1447
13. Dahmen T, Byshko R, Saupe D, Röder M, Mantler S (2011) Validation of a model and a simulator for road cycling on real tracks. Sports Engineering 14: 95-110

14. Delattre E, Garcin M, Mille-Hamard L, Billat V (2006) Objective and subjective analysis of the training content in young cyclists. Applied Physiology, Nutrition, and Metabolism 31: 118-125
15. Di Prampero P, Cortili G, Mognoni P, Saibene F (1979) Equation of motion of a cyclist. Journal of Applied Physiology 47: 201-206
16. Duc S, Villerius V, Bertucci W, Grappe F (2007) Validity and Reproducibility of the Ergomo® Pro Power Meter Compared With the SRM and Powertap Power Meters. International Journal of Sports Physiology and Performance 2: 270-281
17. Ebert T, Martin DT, McDonald W, Victor J, Plummer J, Withers R (2005) Power output during women's World Cup road cycle racing. European Journal of Applied Physiology 95: 529-536
18. Ebert TR, Martin DT, Stephens B, Withers RT (2006) Power output during a professional men's road-cycling tour. International Journal of Sports Physiology and Performance 1: 324-335
19. Foster C, Hoyos J, Earnest C, Lucia A (2005) Regulation of energy expenditure during prolonged athletic competition. Medicine & Science in Sports & Exercise 37: 670-675
20. Gardner AS, Stephens S, Martin DT, Lawton E, Lee H, Jenkins D (2004) Accuracy of SRM and Power Tap Power Monitoring Systems for Bicycling. Medicine & Science in Sports & Exercise 36: 1252-1258
21. Garvican LA, Martin DT, McDonald W, Gore CJ (2010) Seasonal variation of haemoglobin mass in internationally competitive female road cyclists. European Journal of Applied Physiology 109: 221-231
22. González-Haro C, Ballarini PG, Soria M, Drobnic F, Escanero JF (2007) Comparison of nine theoretical models for estimating the mechanical power output in cycling. British Journal of Sports Medicine 41: 506-509
23. Hawley JA, Noakes TD (1992) Peak power output predicts maximal oxygen uptake and performance time in trained cyclists. European Journal of Applied Physiology and Occupational Physiology 65: 79-83
24. Hopkins WG (2004) How to interpret changes in an athletic performance test. Sportscience 8: 1-7
25. Hopkins WG, Marshall S, Batterham A, Hanin J (2009) Progressive statistics for studies in sports medicine and exercise science. Medicine & Science in Sports & Exercise 41: 3-12
26. Hurst HT, Atkins S (2006) Agreement between polar and SRM mobile ergometer systems during laboratory-based high-intensity, intermittent cycling activity. Journal of Sports Sciences 24: 863-868
27. Jeukendrup A, Diemen AV (1998) Heart rate monitoring during training and competition in cyclists. Journal of Sports Sciences 16: 91-99
28. Jobson SA, Passfield L, Atkinson G, Barton G, Scarf P (2009) The Analysis and Utilization of Cycling Training Data. Sports Medicine 39: 833-844
29. Jones S, Passfield L (1998) The dynamic calibration of bicycle power measuring cranks. The Engineering of Sport: 265-274

30. Kirkland A, Coleman D, Wiles J, Hopker J (2008) Validity and Reliability of the Ergomo® pro Powermeter. International Journal of Sports Medicine 29: 913-916
31. Lim AC, Homestead EP, Edwards AG, Carver TC, Kram R, Byrnes WC (2011a) Measuring changes in aerodynamic/rolling resistances by cycle-mounted power meters. Medicine & Science in Sports & Exercise 43: 853-860
32. Lim AC, Peterman JE, Turner BM, Livingston LR, Byrnes WC (2011b) Comparison of Male and Female Road Cyclists under Identical Stage Race Conditions. Medicine & Science in Sports & Exercise 43: 846-852
33. Lukes R, Carré M, Haake S (2006) Track cycling: an analytical model. The Engineering of Sport 6. Springer, pp. 115-120
34. Lukes R, Hart J, Haake S (2012) An analytical model for track cycling. Journal of Sports Engineering and Technology 226: 143-151
35. Macdermid PW, Stannard S (2012) Mechanical work and physiological responses to simulated cross country mountain bike racing. Journal of Sports Sciences 30: 1491-1501
36. Martin JC, Gardner AS, Barras M, Martin DT (2006) Modeling sprint cycling using field-derived parameters and forward integration. Medicine & Science in Sports & Exercise 38: 592-597
37. Martin JC, Milliken DL, Cobb JE, McFadden KL, Coggan AR (1998) Validation of a mathematical model for road cycling power. Journal of Applied Biomechanics 14: 276-291
38. McGregor SJ, Weese RK, Ratz IK (2009) Performance modeling in an olympic 1500-m finalist: A practical approach. The Journal of Strength & Conditioning Research 23: 2515-2523
39. Millet G, Tronche C, Fuster N, Bentley D, Candau R (2003) Validity and reliability of the Polar® S710 mobile cycling powermeter. International Journal of Sports Medicine 24: 156-161
40. Nimmerichter A, Eston RG, Bachl N, Williams C (2011) Longitudinal monitoring of power output and heart rate profiles in elite cyclists. Journal of Sports Sciences 29: 831-839
41. Paton C, Hopkins W (2005) Seasonal changes in power of competitive cyclists: implications for monitoring performance. Journal of Science and Medicine in Sport 8: 375-381
42. Paton CD, Hopkins WG (2006a) Ergometer error and biological variation in power output in a performance test with three cycle ergometers. International Journal of Sports Medicine 27: 444-447
43. Paton CD, Hopkins WG (2006b) Variation in performance of elite cyclists from race to race. European Journal of Sport Science 6: 25-31
44. Peronnet F, Thibault G (1989) Mathematical analysis of running performance and world running records. Journal of Applied Physiology 67: 453-465
45. Pinot J, Grappe F (2010) The 'Power Profile'for determining the physical capacities of a cyclist. Computer Methods in Biomechanics and Biomedical Engineering 13: 103-104

46. Pinot J, Grappe F (2011a) The Peronnet–Thibault mathematical model applied to the record power profile in cycling. Computer Methods in Biomechanics and Biomedical Engineering 14: 219-220
47. Pinot J, Grappe F (2011b) The Record Power Profile to Assess Performance in Elite Cyclists. International Journal of Sports Medicine 32: 839-844
48. Quod MJ, Martin DT, Martin JC, Laursen PB (2010) The power profile predicts road cycling MMP. International Journal of Sports Medicine 31: 397-401
49. Robinson ME, Plasschaert J, Kisaalita NR (2011) Effects of high intensity training by heart rate or power in recreational cyclists. Journal of Sports Science and Medicine 10: 498-501
50. Smith DJ (2003) A framework for understanding the training process leading to elite performance. Sports Medicine 33: 1103-1126
51. Smith MF (2008) Assessment influence on peak power output and road cycling performance prediction. International Journal of Sports Science and Coaching 3: 211-226
52. Sparks SA, Dove B, Bridge CA, Midgely AW, McNaughton LR (2014) Validity and Reliability of the Look Keo Power Pedal System for Measuring Power Output During Incremental and Repeated Sprint Cycling. International Journal of Sports Physiology and Performance (in press)
53. Swart J, Lamberts RP, Derman W, Lambert MI (2009) Effects of high-intensity training by heart rate or power in well-trained cyclists. The Journal of Strength & Conditioning Research 23: 619-625
54. Underwood L, Jermy M (2010) Mathematical model of track cycling: the individual pursuit. Procedia Engineering 2: 3217-3222
55. Vogt S, Heinrich L, Schumacher YO, Blum A, Roecker K, Dickhuth H, Schmid A (2006) Power output during stage racing in professional road cycling. Medicine & Science in Sports & Exercise 38: 147-151
56. Vogt S, Schumacher YO, Blum A, Roecker K, Dickhuth H-H, Schmid A, Heinrich L (2007a) Cycling power output produced during flat and mountain stages in the Giro d'Italia: a case study. Journal of Sports Sciences 25: 1299-1305
57. Vogt S, Schumacher YO, Roecker K, Dickhuth HH, Schoberer U, Schmid A, Heinrich L (2007b) Power Output during the Tour de France. International Journal of Sports Medicine 28: 756-761
58. Wallace LK, Slattery KM, Coutts AJ (2014) A comparison of methods for quantifying training load: relationships between modelled and actual training responses. European Journal of Applied Physiology 114: 11-20
59. Wooles AL, Robinson AJ, Keen PS (2005) A static method for obtaining a calibration factor for SRM bicycle power cranks. Sports Engineering 8: 137-144

APPENDICES

Appendix 1. Participants Wanted Advertisement

Research Participants Wanted!

12-week study tracking changes in cycling specific fitness in cyclists racing and training with a power meter.

Subjects must own their a SRM or Quarq Power Meter and be prepared to submit all data files from racing and training for the study period.

The research is seeking Ethical approval to ensure the data is only used for the scientific purposes. Parental agreement will be required for those age 16-17 years of age.

For more information contact Hamish Ferguson

mob 027 221 1533 email hamish.ferguson@xtra.co.nz

Appendix 2. Participant Information Sheet

Participant Information Sheet

Date Information Sheet Produced: 1 June 2013

Project Title

Maximum Power Profiles from selected durations of racing or training in competitive cyclists compared with power profiles determined by lab based testing.

An Invitation

Hello I am Hamish Ferguson. I have been a Cycling Coach for 21 years and am carrying out a research project to count towards a Masters in Philosophy at AUT. I will be carrying out a 12 week study measuring power using an on board cycle ergometer (power meter) and comparing this with the data we get from a lab based test. I am seeking 15-20 subjects to include in this study. You will need to own a power meter. I will check the calibration of the power meter for you. You will need to submit the data from each training ride and race during the study period and carry out a lab based power test. You are free to leave the study at any period during the study.

The data gained from the study will only be used for the study purposes and will not be used for any coaching or selection purposes. Data collected from a power meter forms a part of the cycle coaching process and should never be used to compare between riders for coaching or selection purposes without taking into account a wide variety of different variables like body weight, personal aerodynamics, riding equipment, skill and tactical strategies. While Hamish Ferguson is a personal cycling coach the aim of this project is to test a methodology that will benefit cycling as a sport, coaches, riders and sport science.

What is the purpose of this research?

The aim of the study is to test whether a power meter offers a more valid (measures what we want to measure) and reliable (test-retest) method of tracking changes in cycling specific fitness than a lab test. It is expected the study will be published in a Sport Science journal, be presented at Sport Science Conferences and be used to facilitate the coaching process for cycling. It will also serve as the basis for future work in the area of Cycling Performance Analysis.

How was I identified and why am I being invited to participate in this research?

We chose Facebook and the Canterbury Cycling Web pages to seek volunteers for the study to ensure we gained the attention of all potential subjects who own a SRM or Quarq brand power meter. We selected these two models as both can be checked for accuracy and can be re-calibrated if necessary.

What will happen in this research?

Subjects will need to have the calibration of the power meter checked at the start of the study and at 4 week intervals. They will need to participate in a lab based fitness test at the conclusion and end of the study and submit power meter data from every training and racing ride of the study period. The lab based testing will incorporate a 5sec, 60sec, 5min and 20min power test.

This image has been removed by the author of this thesis for copyright reasons

This image has been removed by the author of this thesis for copyright reasons

What are the discomforts and risks?

There is no risk or discomfort to participation in the study beyond the normal completion of cycling based training and carrying out cycling specific fitness tests. No invasive tests will be performed. Data will only be used for study purposes and will not be used for coaching or selection purposes.

What compensation is available for injury or negligence?

In the unlikely event of a physical injury as a result of your participation in this study, rehabilitation and compensation for injury by accident may be available from the Accident Compensation Corporation, providing the incident details satisfy the requirements of the law and the Corporation's regulations.

What are the benefits?

The research project will further the performance analysis process for cyclists determining if they can spend more time training and racing to enhance and measure their fitness rather than spend time away from training and racing performing lab based tests.

How will my privacy be protected?

All subject data will be kept confidential and at the completion of the study will be stored with the main supervisor for the mandatory six year period at which point it will be deleted.

What are the costs of participating in this research?

There is no cost to participation. The time commitment is 2 x 1 hour tests, 4 x 20min power meter calibration check and 5 min's each day to send the power meter data from each ride to myself.

What opportunity do I have to consider this invitation?

We need your acceptance into the study by May 15, 2013.

How do I agree to participate in this research?

All subjects will fill in a consent form and a Par-Q to ensure they are of good health to complete the study.

Will I receive feedback on the results of this research?

Each subject will receive individual feedback about their performance in the study and the lab tests.

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, *Prof Will Hopkins, will.hopkins@aut.ac.nz, phone 021 804 736*.

Concerns regarding the conduct of the research should be notified to the Executive Secretary, AUTEC, Dr Rosemary Godbold, *rosemary.godbold@aut.ac.nz* , 921 9999 ext 6902.

Whom do I contact for further information about this research?

Researcher Contact Details:

Hamish Ferguson

Hamish.ferguson@xtra.co.nz

Project Supervisor Contact Details:

Prof Will Hopkins

will.hopkins@aut.ac.nz

Approved by the Auckland University of Technology Ethics Committee on *1 May 2013*, AUTEC Reference number *Hopkins13122042013*.

Appendix 3. Informed consent form.

Consent Form For use when laboratory or field-testing is involved.	AUT UNIVERSITY TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Project title: Maximum Power Profiles from selected durations of racing or training in competitive cyclists compared with power profiles determined by lab based testing.

Project Supervisor: Prof Will Hopkins

Researcher: Hamish Ferguson

○ I have read and understood the information provided about this research project in the Information Sheet dated 1 June 2013.

○ I have had an opportunity to ask questions and to have them answered.

- ○ I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.
- ○ I am not suffering from heart disease, high blood pressure, any respiratory condition (mild asthma excluded), any illness or injury that impairs my physical performance, or any infection.
- ○ I agree to take part in this research.
- ○ I wish to receive a copy of the report from the research (please tick one): Yes○ No○

Participant's signature: ....

Participant's name: ....

Participant's Contact Details (if appropriate)………………………..

……………………………………………………………………………………..

……………………………………………………………………………………..

……………………………………………………………………………………..

Date: ……………………………………………………………………………..

Approved by the Auckland University of Technology Ethics Committee on 1 May 2013 AUTEC Reference number Hopkins1312_01052013

Note: The Participant should retain a copy of this form.

Appendix 4. Pre Participation Questionnaire (Par-Q)

This image has been removed by the author of this thesis for copyright reasons

Appendix 5. Ethics Approval

AUT UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

A U T E C

S E C R E T A R I A T

1 May 2013

Will Hopkins

Faculty of Health and Environmental Sciences

Dear Will

Re Ethics Application: **13/12 Maximum power profiles from selected durations of racing or training in competitive cyclists compared with power profiles determined by lab based testing.**

Thank you for providing evidence as requested, which satisfies the points raised by the AUT University Ethics Committee (AUTEC).

Your ethics application has been approved for three years until 1 May 2016.

As part of the ethics approval process, you are required to submit the following to AUTEC:

- A brief annual progress report using form EA2, which is available online throughhttp://www.aut.ac.nz/researchethics. When necessary this form may also be used to request an extension of the approval at least one month prior to its expiry on 1 May 2016;
- A brief report on the status of the project using form EA3, which is available online throughhttp://www.aut.ac.nz/researchethics. This report is to be submitted either when the approval expires on 1 May 2016 or on completion of the project.

It is a condition of approval that AUTEC is notified of any adverse events or if the research does not commence. AUTEC approval needs to be sought for any alteration to the research, including any alteration of or addition to any documents that are provided to participants. You are responsible for ensuring that research undertaken under this approval occurs within the parameters outlined in the approved application.

AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to obtain this.

To enable us to provide you with efficient service, please use the application number and study title in all correspondence with us. If you have any enquiries about this application, or anything else, please do contact us at ethics@aut.ac.nz.

All the very best with your research,

Madeline Banda

Acting Executive Secretary

Auckland University of Technology Ethics Committee

Cc: Hamish Ferguson hamish.ferguson@xtra.co.nz