PRACTICAL CONSIDERATIONS FOR ENDURANCE TRAINING WITH POWER

– Written by Simon Jones, United Kingdom

BACKGROUND

Records show that human power measurement was first performed well over 100 years ago in the laboratory. However, it took another 80 years for the power meter to evolve into an accurate measuring tool, useable in the field, then a further 10 years for the data to be understood, stored and analysed in a format that could be easily interpreted.

The need for accurate laboratory measurement of power was required by researchers in order that physiological responses could be benchmarked with confidence. This physiological data was often used to profile athletes for research studies; however, the focus was the physiological responses such as VO_2 max, anaerobic threshold, gross efficiency and alike, rather than the power output per se. Examples of how power output has been traditionally used are given below.

In more recent times, the focus has shifted from the laboratory to the field. Power is the currency of performance and an accurate descriptor of in-field training. Generally speaking, interest has shifted from physiological supply, to power output; the demand side.

Since the year 2000, growth in the number of manufacturers developing power meters has been exponential. The power meter, once only a tool for professionals or research institutions, is now at a price point where a serious amateur can afford to invest and enjoy the benefits of training with more data. Along with this increased availability, the number of coaches and scientists using field-based power measurement has also expanded, each building their own significant data sets, stored on a number of different analytic platforms.

MAKING SURE THE DATA IS ROBUST

Most power meters are incorporated into the drivetrain of the bicycle and measure power by calculating two parameters; force from the deformation of strain gauges and angular velocity. This gives a measure of torque over time which can in turn be converted to power, measured in watts (W).

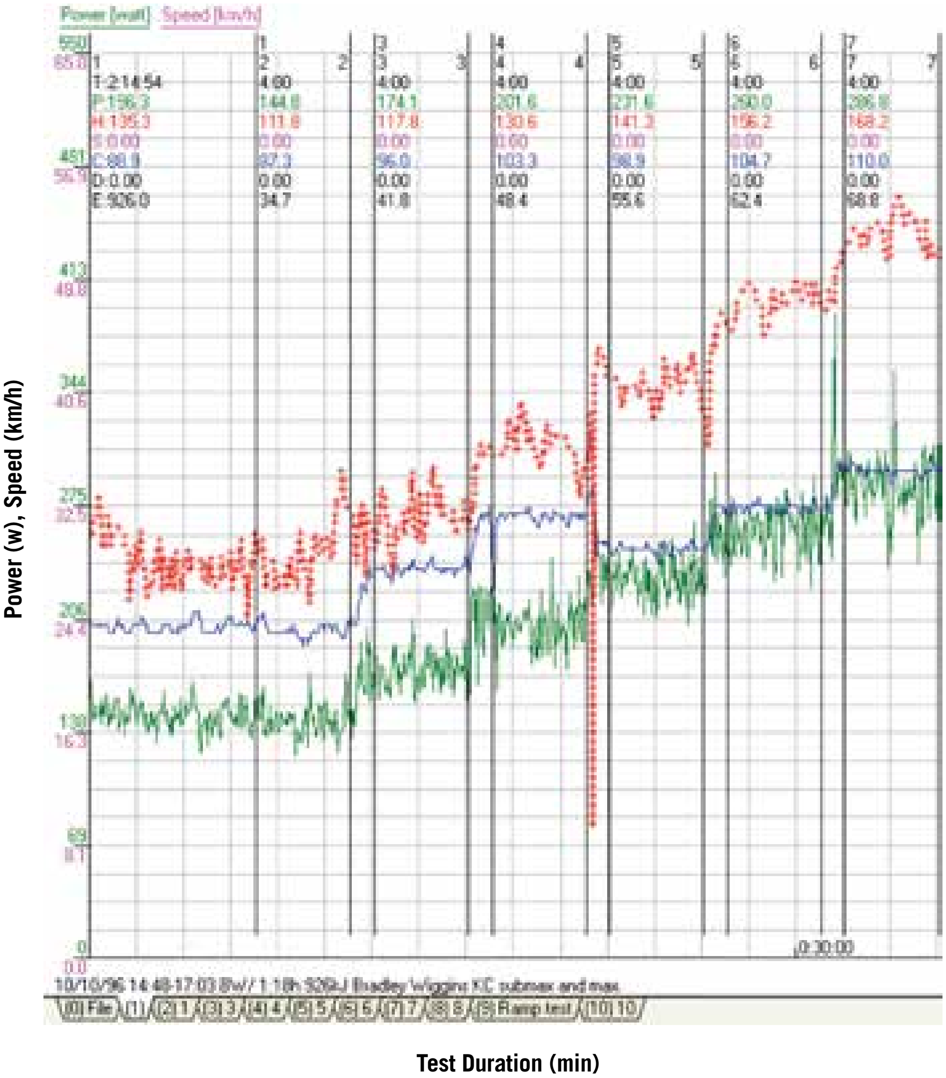


Figure 1: A typical cycling submaximal and ramp test to exhaustion output.

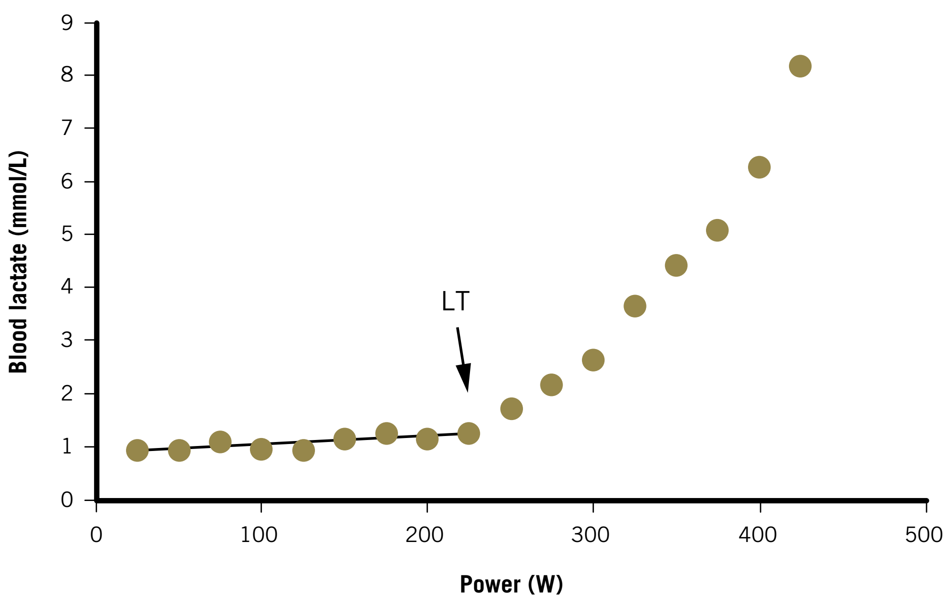


Figure 2: A typical power lactate curve showing breakpoint above baseline. LT=lactate threshold.

Strain gauges are affected by temperature, so most power meters require a daily zero offset. Some newer power meters automatically account for temperature changes, which reduces the need for daily zero offset, making reproducible measurements far simpler to obtain and regular use far more straightforward.

Power meters have enabled a deeper understanding of racing and training demands faced by the rider. The forces acting against a rider vary according to the environment; road surface, altitude, weather and wind, terrain, drafting, as well as body shape and size. If any one of these parameters should change, the energy demands will change as a result. Power enables us to discover more about the environment and through a systematic process helps us to understand more about how to improve performance

It's generally accepted that power meters are valuable tools for understanding cycling requirements and performance. It is worth considering however, that power is merely a reflection of physiological energy turnover. It's the physiology and biomechanics and indeed the 'will power' that enable the mechanical work to be performed. Power meters enable us to see the link between physiology and field-based quantification of mechanical work.

This is an important notion to consider when using power meters during training or for performance analysis. An understanding of the physiology is critical when interpreting power data and will enable the coach or athlete to maximise the use of power when prescribing training or analysing performance.

The following example may help to illustrate this further. Power measured while cycling quantifies the mechanical work done and can be classified as a measure of external training load. It could be argued that to optimise exercise planning and understand the responses, power should be used with 'conventional', perhaps less popular modes of internal training load, such as heart rate and RPE. This is well cited in the scientific literature[1].

EVENT DEMANDS

Getting the training right for an individual requires a good understanding

Figure 3: Crank-based power meter.

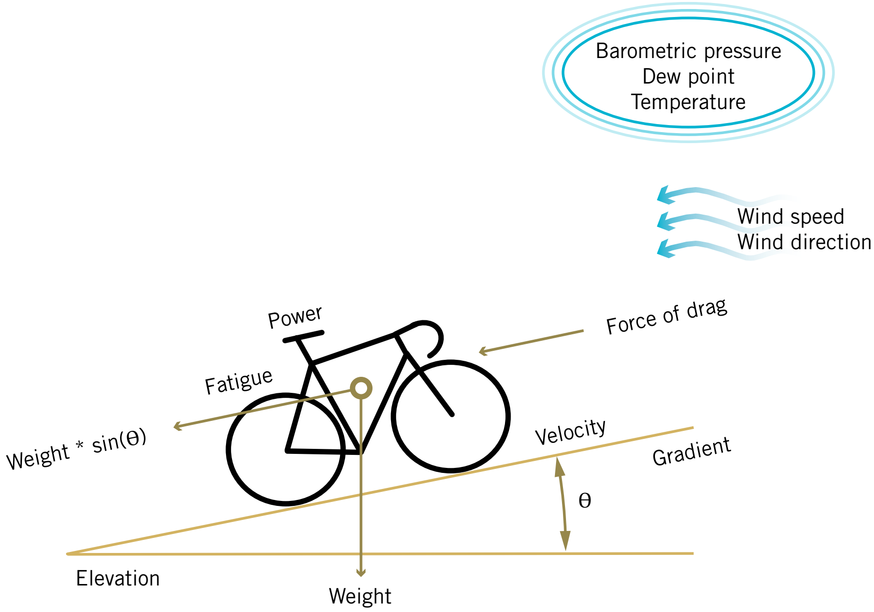


Figure 4: Modelling cycling energy requirements with first principles physics equations.

of the performance requirements of the event. Clearly, the needs for a 4 km team pursuit are vastly different to those of a 3-week stage race. Modelling performance with regards to physics (forces acting on a rider using 1st principles) has been used for years by exercise scientists. This method can help understand what the physical power requirements are for a given task. Once this is understood, specific sessions can be devised and power targets set.

Using the method in Figure 4, it's a fairly straightforward process, which can be done using a spreadsheet, to calculate the power requirements where wind resistance is minimal, such as climbs in excess of 7% gradient. There are several useful web-based resources that enable online estimations of power for a given situation and rider characteristics (www.cyclingpowerlabs.com, www.analyticcycling.com) This may be useful when assessing event demands and deciding what targets to use in training efforts.

QUANTIFICATION OF TRAINING LOAD

Power meters also enable a greater understanding of the frequency and distribution of intensity within a competition or training session. Andy Coggan, probably the most respected architect of power-based training methodology, talks of training and testing being equivalent. This concept is easy to understand with a power meter. Every session is now measureable and quantifiable. Power meters enable, to a large degree, the normalisation of training data; if it's less windy or you make an effort on a different climb, power still describes the external workload and the sessions can be directly compared.

POWER DURATION

The power duration concept is one of the simplest tools available in cycling. This curve is a graphical representation of a rider's personal best efforts, as measured for different durations. These time intervals generally start with short durations, moving to longer ones. The comparison of maximal power output for a specific duration can provide useful insight into the performance capabilities of an individual or comparison of individual strengths and weaknesses. The power duration curve is a very useful when performance comparison is required. For example, pre-season training, race to race, different races periods – a simple overlay of two periods can illustrate what the current performance level is compared to that previously. Power in this context provides the all-important performance benchmark that heart rate, speed and results aren't able to provide.

The power duration in Figure 5 simply quantifies the best duration-power outputs obtained during a given period of time. These lines show comparative power durations for 2 years. The blue line, is the current year and green line the previous year. Comparisons can be quickly identified, enabling the coach or rider to understand in power terms where they lie in relation to a personal best performance. These graphs can also be displayed relative to body mass, which can provide that all-important climbing index. The relevance of this is quite simple and clear to see; a timeframe-to-timeframe comparison, so progress (or lack thereof) can be easily observed and gaps in performance or training efforts identified. Other practical applications of this curve include pacing for a particular effort. This could be a climb or a time trial, to provide the rider with a goal power or benchmark on which to base their race strategy. In other words, it shows the maximal power output that might be expected for an effort of a given duration. If, for example, a rider makes an effort of 5 minutes and produces 350 W, 350 W will be represented on this curve until such time the rider exceeds this value.

A limitation with these power curves is that their use and relevance depends on the rider riding maximally for a given duration. Otherwise, the curve reflects training effort rather than true performance capacity. If these curves are to be used to their full potential, its advised to include a series of maximal efforts at regular intervals to ensure the curve reflects current performance levels and not just training.

CYCLING POWER DEMANDS ARE INTERMITTENT AND RANDOM

Anyone who has used a power meter on the open road will know how intermittent

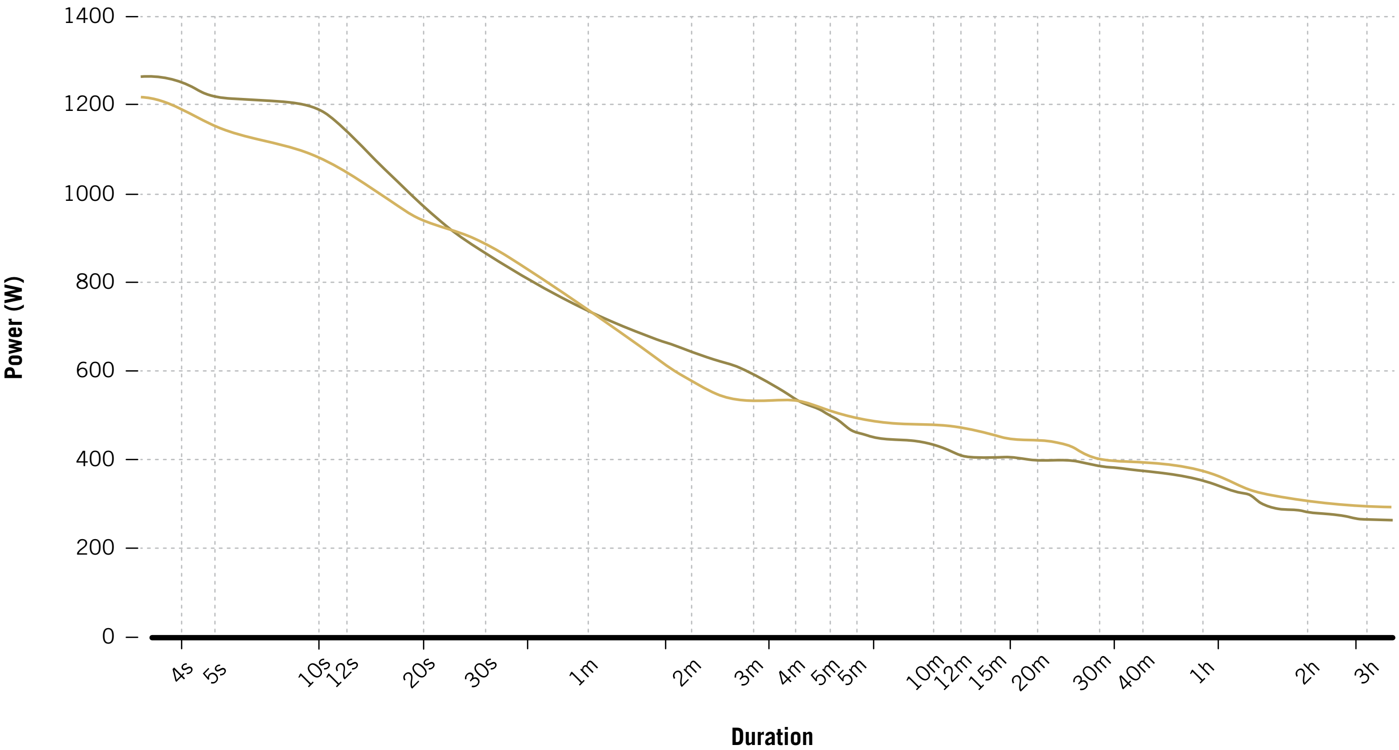


Figure 5: A typical power duration curve overlaying two competitive seasons.

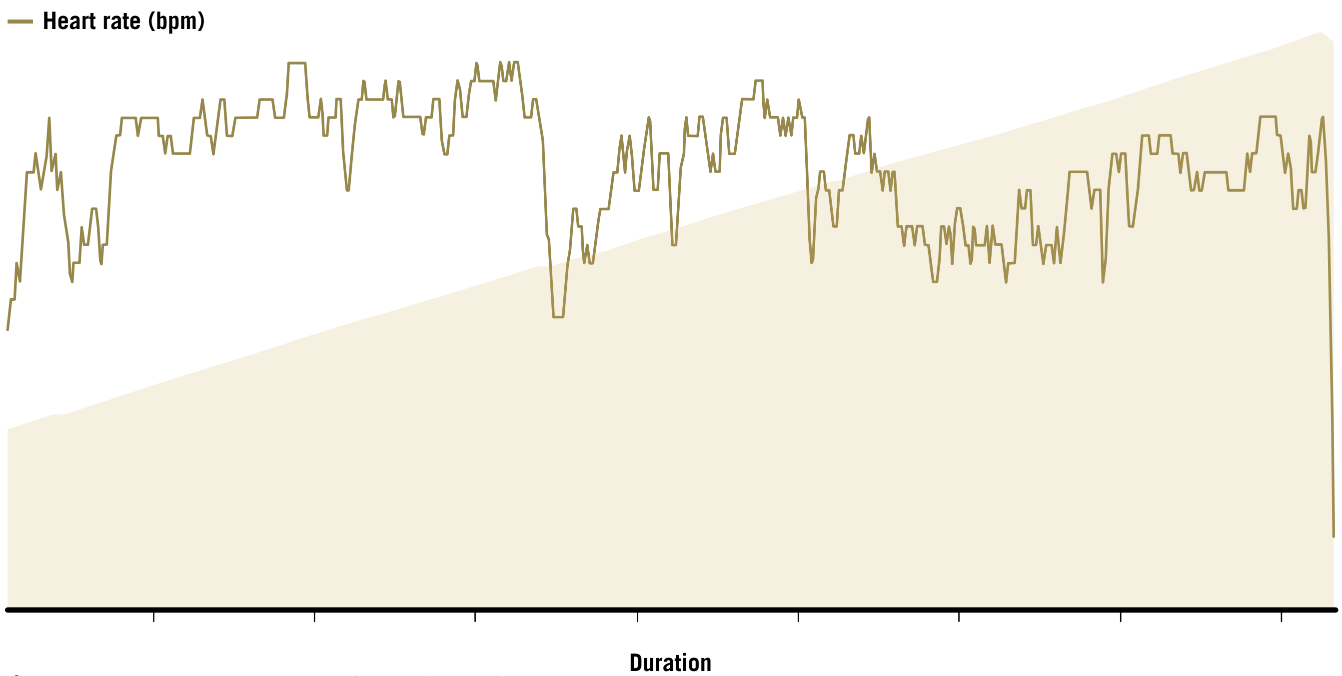


Figure 6: Heart rates responses during uphill cycling.

and variable the demands of cycling are. Perhaps at the most basic level, it can be argued that this information enables coaches to target training sessions with greater understanding and accuracy. There has recently been an evolution in the types of training now prescribed; steady state, constant blocks of intervals, through to intermittent 'spike' intervals with variable cadence and power – these efforts mirror the mechanical load more specifically, while still providing an endurance training effect.

Figure 6 shows a basic heart rate trace from a mountain climb in a major World Tour competition. The heart rate shows the physiological strain and intensity, but it's doesn't tell the whole story. Power meters, however, are able to show exactly what is driving the physical response.

Figure 7 shows the same climb but with both power and heart rate. You can clearly see that at the start of the climb, the intensity of the effort is well above threshold power. More interesting, is the variance and intermittent nature of the effort. It's well documented that stochastic (intermittent and variable) exercise creates a greater level of fatigue. So despite what seems to be a fairly constant intensity when viewing

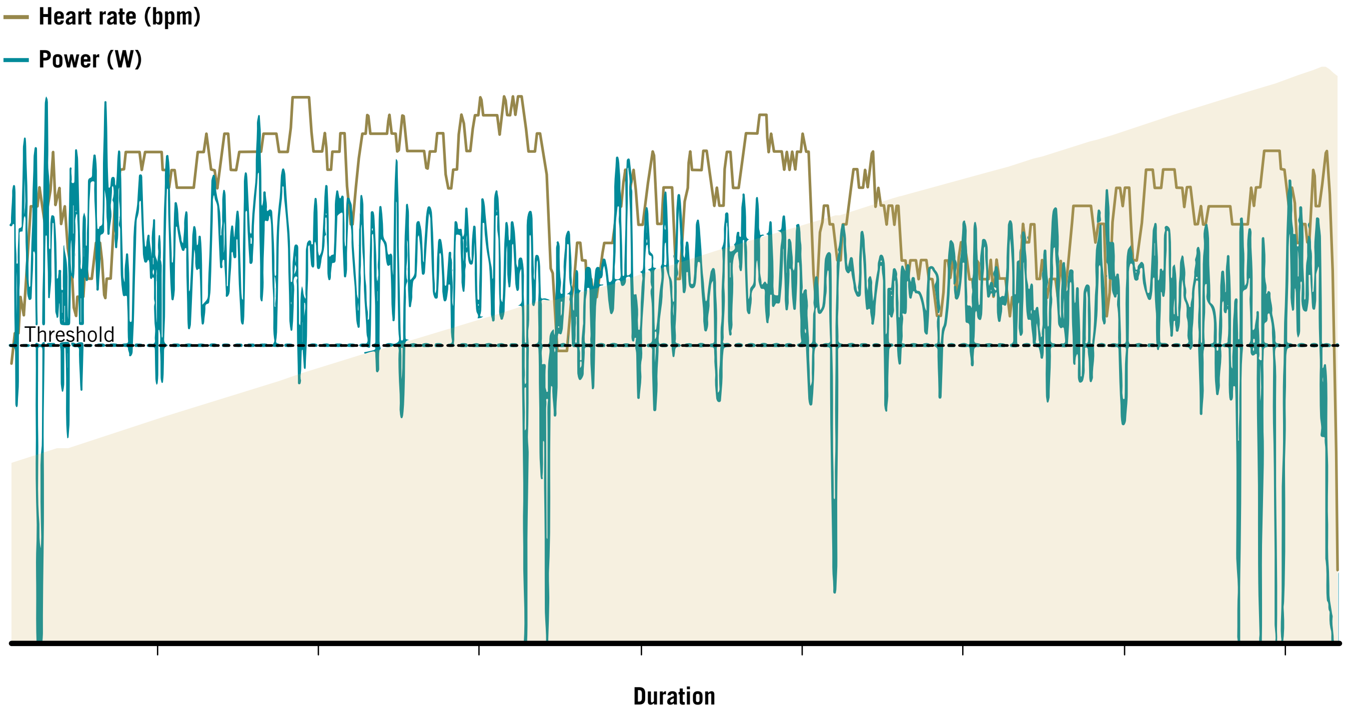


Figure 7: Heart rates and power responses during uphill cycling.

heart rate, power is providing another picture. It's interesting to see on a climb that performance power is so variable and that such variability is likely to be increasing the physiological cost.

This simple graph shows the benefits of using both a physiological and a mechanical measure of physical strain. Using power and heart rate together provides a more complete picture of 'demand and response'. From a practical standpoint, we now know that performance power is variable, it's conceivable that a degree of variability is required in daily training to ensure that training meets the demands of competition.

PACING

There are some limitations of using heart rate as a tool for pacing. Figure 7 shows that heart rate is relatively insensitive to the actual power demands at the start of an effort. The power trace clearly shows that the most intense section is at the beginning. The acceleration in heart rate shortly after the start is driven by the immediacy of the climb and the surge in mechanical demand by the rider. Heart rate will always lag behind the true mechanical demands, which is why power is such a good tool for pacing at the onset of a climb, time trail or training interval.

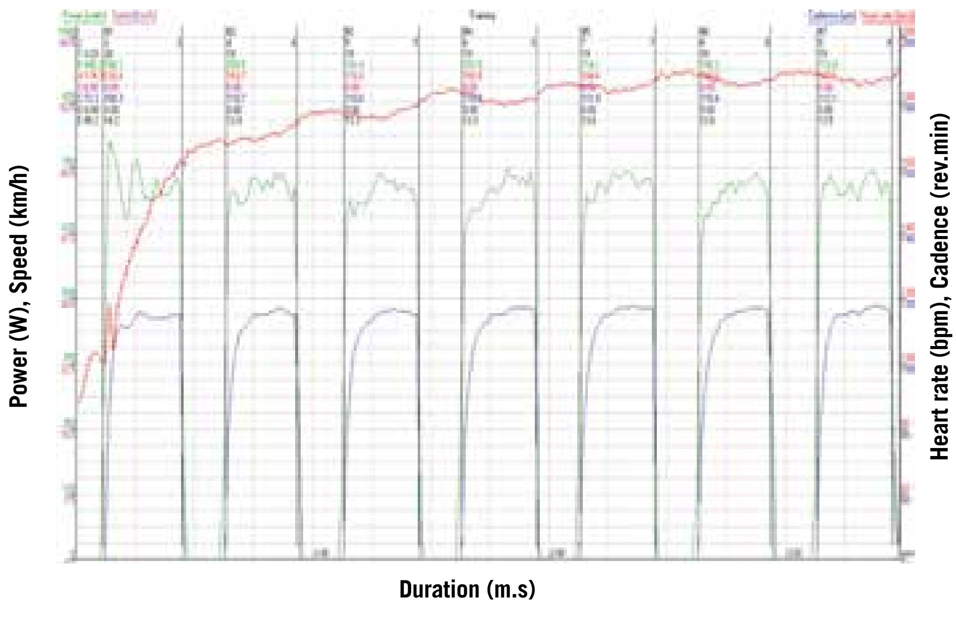


Figure 8: Power (green) and heart rate (red) responses to an anaerobic capacity interval session.

ANAEROBIC INTERVALS

Power meters really come into their own during specific, high-intensity sessions, such as capacity or lactate tolerance intervals. Capacity refers to the maximum amount of work that can be generated with a given duration. Tolerance is more of a physiological term, that refers to the physiological systems that buffer fatigue over a given duration. Both capacity and tolerance can be (indirectly) measured by power.

Figure 8 shows the power and heart rate for a 20 second on, 10 second recovery interval session. Power is a robust tool to use for specific training objectives, ensuring the precise goal is achieved. Note the lag in heart rate, demonstrating that it is not a useful measure of exercise intensity.

TRAINING ANALYSIS SOFTWARE

It's beyond the scope of this article to go into details of the various software packages available to analyse training. The Training Peaks platform and its sister analytics engine WKO was the first to arrive back in 2003. It is possibly the most widely-used package for analysis of training based on power output. However, it is worth bearing in mind the famous and commonly heard phrase, 'all models are wrong, but some are useful'. With this caveat, the algorithms and methodology in Training Peaks has its uses.

The key parameters to this model are Acute Training Load (ATL) and Training Stress Score (TSS). TSS reflects the intensity and duration of the exercise session – 1 hour of riding at the Functional Threshold Power (FTP) is assigned a TSS of 100. ATL represents the current degree of freshness or recovery status. This is an exponentially-weighted average of training load over a period of 5 to 10 days (this period is referred to as a time constant, TC). The more you train consecutively, the greater the training load and ATL. The formula for ATL is;

$$ATL=ATL_y+((TSS-ATL_y))/(TC_a)$$

Where ATL_y = yesterday's ATL, TSS = current Training Stress Score and TC_a = ATL Time Constant. Time constants are measured in days and can be customised to fit individual riders. For example, in the case of a rider who recovers quickly, the constant is shortened. Practically, it may be advisable to leave the time constant set at 6 or 7 days in most instances

Chronic Training Load (CTL) represents fitness using an exponentially-weighted average of training over a 42 day period. It's well known that the adaptations to endurance training are not instantaneous. Performance improvement will require many weeks of training before being evident. Thus, the CTL represents the past 6 weeks' training load. The formula for CTL is;

$$CTL=CTL_y+((TSS-CTL_y))/(TC_c)$$

Where CTL_y = yesterday's CTL, TSS = current Training Stress Score and TC_c = CTL Time Constant (42 days, as explained above).

Training Stress Balance (TSB) is simply the difference between the CTL and ATL and represents form. Practically this can be used as a measure of when a rider is ready to race. The formula for TSB is;

$$TSB = CTL – ATL$$

A negative TSB value represents a high training load, as would occur during a period of intense training. Alternatively, a taper leading up to an event should correspond with an increasing or positive TSB, where ATL is reduced relative to the current CTL.

Once the system has been used for a short while, an obvious issue arises. You can achieve a similar TSS with two rides of different intensities and duration that demand a very different physiological strain. For example, riding 5 to 6 hours at low intensity (Zone 1 to 2) may give a TSS of 250, yet the same or similar TSS can be observed with 2.5 hours of higher exercise intensity. Clearly the metabolic cost, physiological strain and training load (internal load) of these sessions will be very different.

Consider the practical scenario of performing a ride with low carbohydrate availability – a training session that's frequently employed to promote certain desirable training adaptations and reductions in body fat. A 6-hour ride without carbohydrate is much harder than an equivalent session with adequate carbohydrate intake, the impact more severe. However, if the rider manages to produce the same average watts, the TSS will be same for the two sessions. This suggests that the training load is the same. Clearly this is not the case at the metabolic level and measures of perceived exertion will reflect this fact.

Heart rate, specifically the percentage of time spent below or above threshold (similar to what Seiller[2] recommends), along with training volume and RPE provide a simple but effective measure of training stress. Combine this with period comparisons from the power duration curve and this can further enhance learning and training optimisation.

Also of practical value in these various coaching platforms is the training diary: a database to store and capture training data, ride annotations and provide the necessary communication link between rider and coach, irrespective of where they might be in the world. They are also very helpful tools for reviewing rider performance data, whether this is over a short cycle or from season to season.

FUNCTIONAL THRESHOLD POWER

A natural progression from the power duration chart concept is to introduce a common term; functional threshold power

"Heart rate will always lag behind the true mechanical demands, which is why power is such a good tool for pacing at the onset of a climb, time trial or training interval"

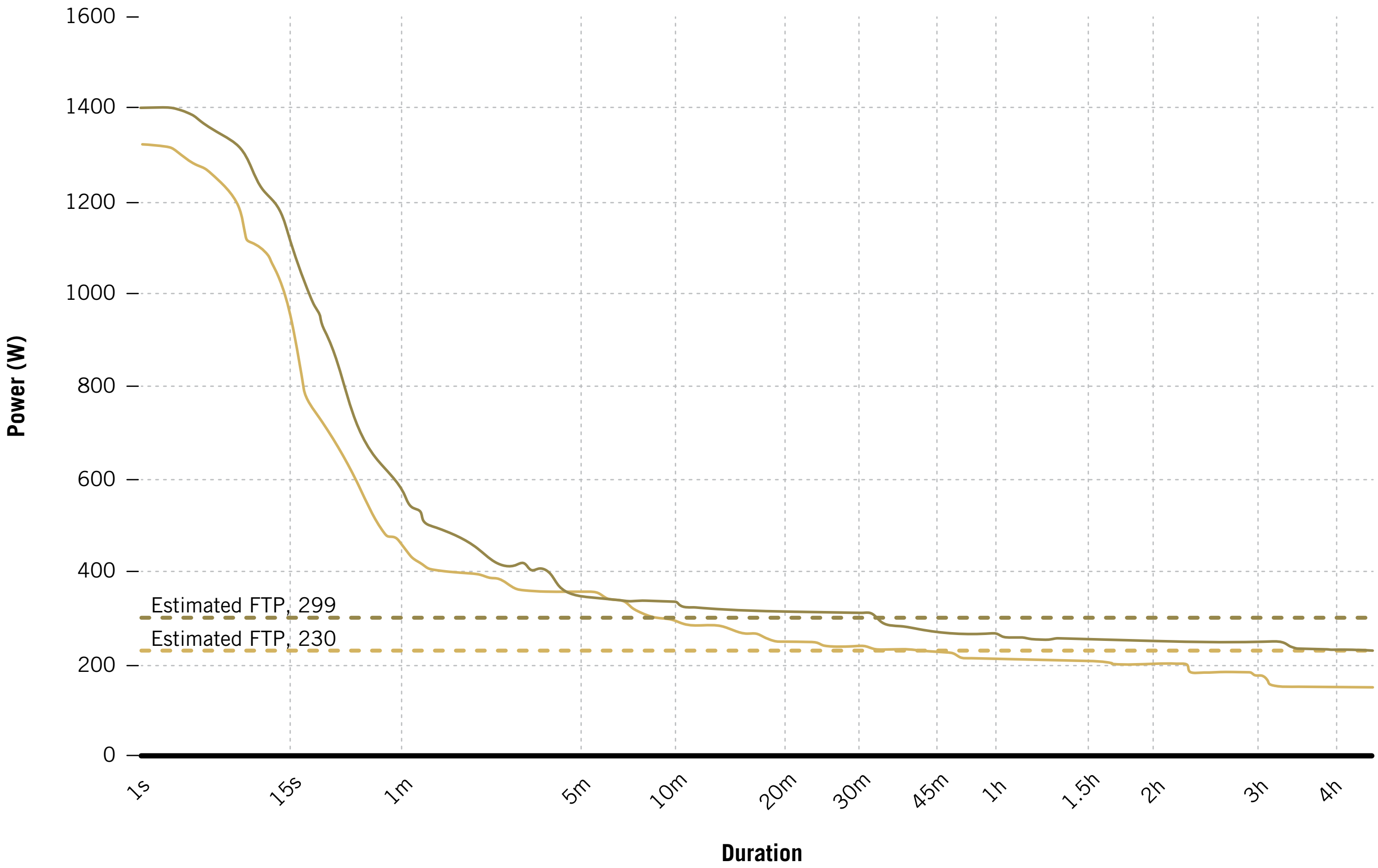


Figure 9: A typical power duration curve overlaying two competitive seasons with estimated FTP – single rider comparison. FTP=functional threshold power.

(FTP). This training methodology links a simple performance test of 20 minutes duration to a measure of threshold or functional threshold (defined as 95% of 20-minute performance power). This subsequently provides the DNA for a performance management process.

Practically, this works fairly well in some circumstances. Twenty minutes is a reasonable duration over which to make regular efforts without requiring excessive mental focus. One aspect to bear in mind is that in reality, a watt measured in one situation or context might not be equivalent to a watt measured in another. For example, a maximal effort produced on a 5 to 7% climb often leads to a different result to that obtained from a similar effort on the flat. This isn't widely researched in the scientific literature, however its well-known among coaches that more power can be produced while climbing a hill compared to cycling on the flat. The same applies to measurements made into a headwind compared to those with an accompanying tailwind Thus, in order to ensure good repeatability, FTP tests should be performed under similar conditions (road, elevation, wind direction).

Figure 9 shows the power duration curve for an amateur cyclist over a 2-year period. The chart also shows the estimated threshold power or FTP, which is estimated from power duration curve and shown by the dotted line.

A large difference of 69 W in threshold power is apparent between time periods. This differece is calculated by comparing the 20-minute power duration performance in the selected period. This method provides a quick and easy illustration of the difference in performance capability without going into a laboratory.

TRAINING PLAN VS. ACTUAL PERFORMANCE

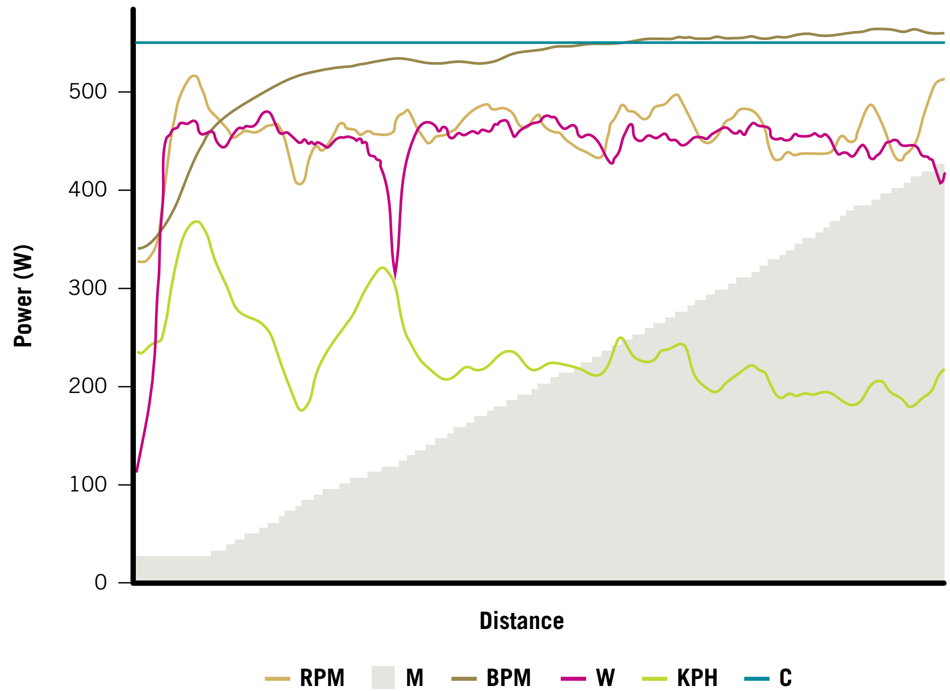
At the very simplest level of coaching, knowing if a session has been executed according to the plan is key to monitoring training goals.

Take a common interval session; four repetitions of 5 minutes above 'threshold', or FTP, where the purpose of the session is to produce fatigue and develop high levels of lactate. Using heart rate alone, it's not easy to determine if the rider is at the correct workload due to cardiac drift (a progressive increase in heart rate with time). This is likely to be related to an increase in core temperature, as these efforts were performed while climbing (high power output, yet low airflow).

In Figure 10, if heart rate (red line) were used as a training target, it would be reasonable to assume that the power (pink line) achieved would decrease, as the rider follows his physiological feedback. With power (pink line), the training objective and real-time feedback to the rider of plan vs actual is displayed. As power is highly sensitive, the feedback is rapid and provides the rider with the opportunity to modify their effort according to this variable. Power measurement enables much more precise training and represents a more objective tool for assessing whether a training session has been performed as intended.

AERODYNAMIC MATTERS

The invention of the mobile power meter has enabled riders and coaches to use indoor velodromes much like they would a wind tunnel. Aerodynamic drag accounts for up to 90% of the power requirements at 50 km/h. It's easy to see why, over the past 20 years, there has been so much interest



	MIN	AVG	MAX	
Power	135	456	555	W
Heart rate	118	175	187	bpm
Cadence	64	92	105	rpm
Speed	16.7	25.2	39.2	kph
Pace	03:36	02:23	01:32	min/ kmh
Elevation	214	263	322	m
Temperature	25	25	25	C

Figure 10: Power (magenta), heart rate (dark gold), speed (green) and gradient (grey) during a 5-minute training effort on the road.

Table 1

Run	Test	CdA [m²]	P_{aero} [W]	Time/km [s]
1	Baseline TT position	0.2290	332.5	73.30
	Average Standard deviation	0.2268 0.0037	329.2659 5.3215	73.0605 0.3957
9	Test-bike TT position: head down	0.2137	310.3	71.63
	Average Standard deviation	0.2133 0.0028	309.7713 4.1249	71.5897 0.3170
13	Baseline TT position: head down	0.2125	308.6	71.50
	Average Standard deviation	0.2149 0.0020	311.9858 2.8468	71.7604 0.2182
17	Baseline TT position: head down; shoulders in	0.1995	289.7	70.01
	Average Standard deviation	0.1996 0.0027	289.8411 3.9162	70.0202 0.3152
21	Test-bike TT position: head down; shoulders in; 10mm rise skis/pads	0.1956	284.0	69.55
	Average Standard deviation	0.1988 0.0022	288.6794 3.1792	69.9268 0.2575

Table 1: Track testing results for varying riding positions and posture. CdA=aerodynamic drag coefficient.

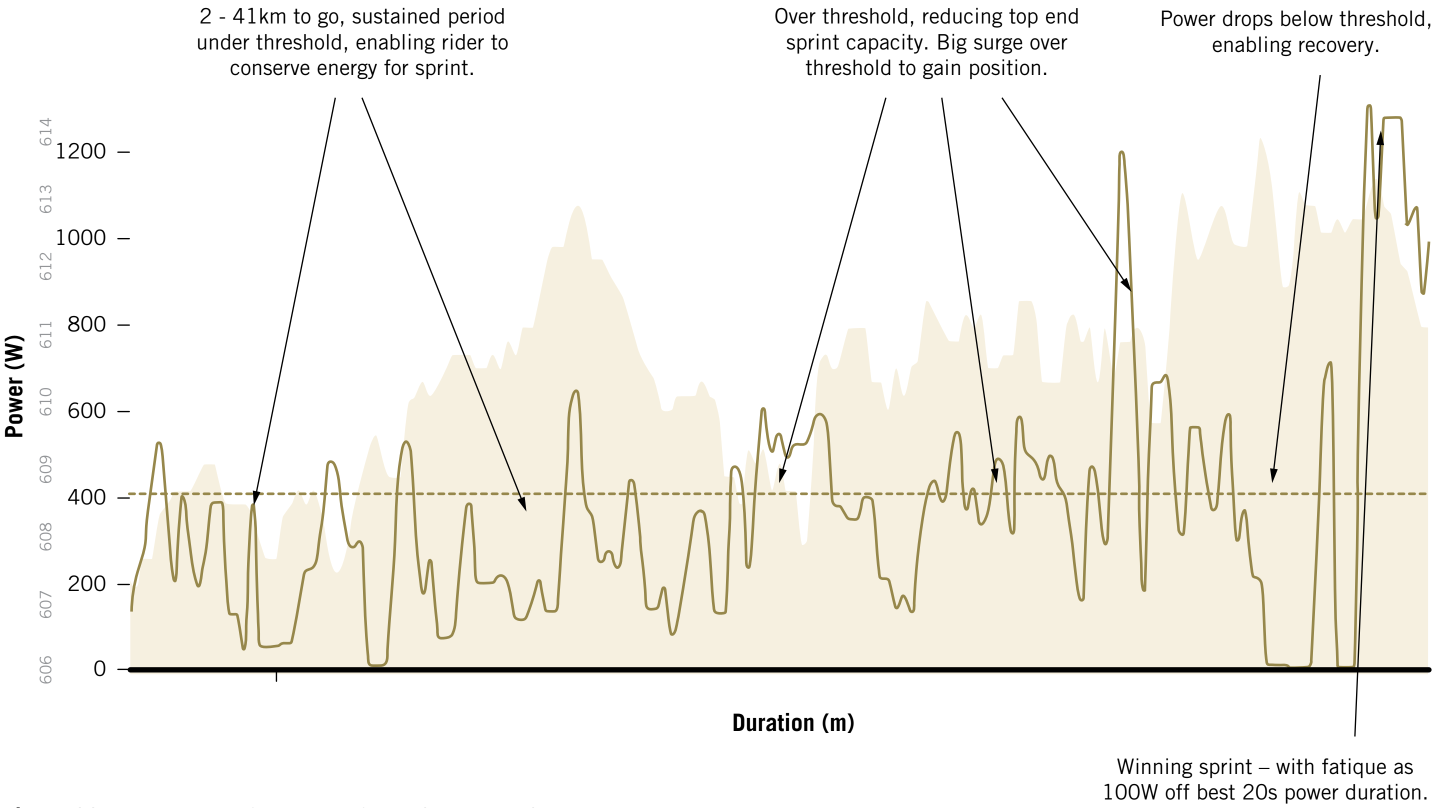


Figure 11: Power trace of a successful sprint stage win.

in aerodynamic bikes, clothing and riding positions. There are clear benefits and drawbacks to both wind tunnels and track testing; however, both approaches are useful and are sensitive methods for measuring the aerodynamic drag directly (wind tunnel) or indirectly, via the power required at a given speed, using first principles physics (power is a function of V^3).

This of course would not be possible without the powermeter. The performance gains in the above example from run 1 to run 21 is something in the region of 40 W or 3.5 seconds/km at 50 km/h. A significant and evidence-based performance enhancement. Power output in these field trials can be used to calculate the coefficient of drag area. Coefficient of drag area is a useful aerodynamic index allowing data to be normalised, which enables comparisons to be made across different testing sessions. In general, a lower coefficient of drag area would support a faster performance.

ANALYSING SPECIFIC TECHNICAL ASPECTS OF PERFORMANCE

Sprint analysis

Power measurement is very useful for understanding the relationship between technical and physiological parameters. Simply overlaying power with video, for the purpose of rider and coach education can have substantially more impact than words or numbers alone.

When sprinting, much of the performance is about the tactical and technical execution, which enables the rider to conserve valuable energy to unleash in the final seconds.

Although skill in cycling is a difficult concept, the power meter can illustrate the outcomes of skill – if the objective is to conserve energy and retain sprint capacity, comparison of sprint performance and the time spent under threshold can be helpful in dissecting a good or excellent performance.

Performance analysis of time trial

Another practical use of power data is to overlay power with video to help explain performance. Power data overlaid to video provides rider and coach with a useful way of reviewing the competition or training session. Although this is a simple process it can be extremely valuable, particularly when a specific and measurable racing or training objective has been set in advance. For example, this could be the use of power to support a pacing strategy in a time trial. Reviewing the power data with video may be more engaging than simply inspecting a graph and is likely to make such feedback more productive.

TRAINING MODELS – IS THERE A METHOD THAT WORKS OPTIMALLY?

Getting the right balance of training and recovery has long been a priority for coaches, riders and researchers. In recent years more knowledge has become available from both scientific training studies and applied research involving the actual training habits of elite athletes. The findings of these recent studies from the modern era contrast sharply with previous models and theories of periodisation, which generally recommended a systematic progression of training with a carefully programmed manipulation of intensity, volume and duration of training.

Research by Seiler[3] has showed that, broadly speaking, endurance athletes achieve high levels of performance by splitting their training 80-20. That is, 80% low-intensity (below threshold or FTP) and 20% at higher intensities (above FTP). In addition, Seiler's work[2] found that many training-related increases in physiological variables came about through increases in training volume, rather than intensity. This isn't to say that intensity isn't important –

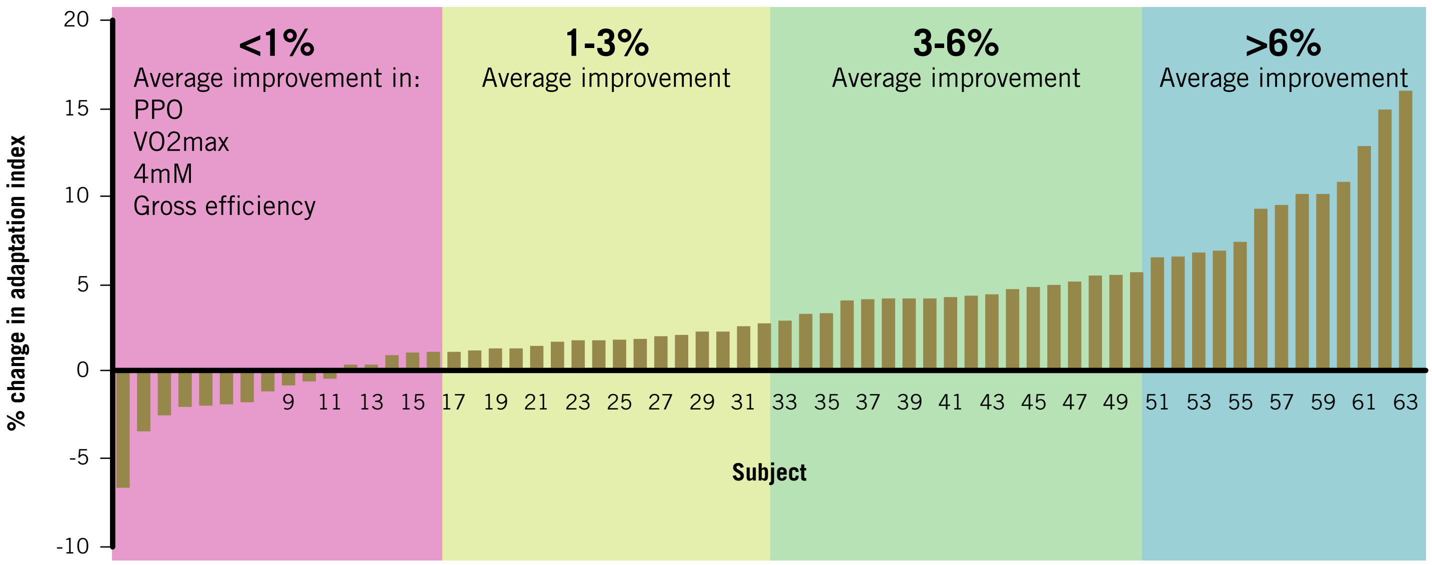


Figure 12: Large variance in subject improvement over 12 weeks of endurance training.

it is critical to performance – however, on average only two high-intensity session per week appear to be required.

In contrast to traditional training periodisation models, with variance in volume and intensity, it seems that many successful athletes maintain a relatively constant weekly training volume throughout the year.

Researchers and coaches have spent many years trying to develop the optimal training programme. As training studies are particularly difficult to implement, good scientific evidence to support a particular training model is not easy to find. However, recent work by Seiller has been very useful in describing the components of an empirically-based successful training schedule and the variability in response observed between individuals, as depicted in Figure 12.

These results clearly show the large variance in training adaptation to an identical training model. While some individuals improved by 10 to 15%, others showed lesser degrees of adaptation and a small proportion actually experienced a reduction in performance.

It's tempting to speculate that training prescription is much less complicated than previously thought. Training should focus on specific physical, technical and tactical event demands. A reasonable suggestion might be to maintain a relatively constant weekly training volume, yet target one or two key sessions each week, simultaneously monitoring internal training load by a preferred tool or system. Power measurement can assist this process as a means of tracking performance gains (or losses). Given the variability in individual responses to any given training schedule, it would appear unwise to rigidly adhere to a programme developed for another athlete.

References

1. *Halson SL. Monitoring training load to understand fatigue in athletes. Sports Med 2014; 44 Suppl 2:S139-S147.*
2. *Seiler KS, Kjerland GO. Quantifying training intensity distribution in elite endurance athletes: is there evidence for an "optimal" distribution? Scand J Med Sci Sports 2006; 16:49-56.*
3. *Seiler S. What is best practice for training intensity and duration distribution in endurance athletes? Int J Sports Physiol Perform 2010; 5:276-291.*
4. *Sylta Ø, Tønnessen E, Hammarström D, Danielsen J, Skovereng K, Ravn T et al. The effect of different high-intensity periodization models on endurance adaptations. Med Sci Sports Exerc 2016 [Epub ahead of print].*

Simon Jones B.Sc.
Head of Innovation
Team Sky
Bath, United Kingdom

Roger Palfreeman M.B.Ch.B. (Hons), B.A. (Hons), Dip.S.E.M., F.F.S.E.M.
Clinical Liaison for Sports Science
Aspetar – Orthopaedic and Sports Medicine Hospital
Doha, Qatar

Former Team Doctor, BMC Racing Team and Team Sky

Contact: jones@teamsky.com