TRIAL BY TELEVISION

When Antoine Vayer called Chris Froome's Tour performance "miraculous", his suspicions were based on calculated performance figures. Is this pseudo-science or can monitoring riders' power output play a part in defeating doping?

Writer: Daniel Friebe
Photography: Tim De Waele *

By the first rest-day of the 2013 Tour, surrounded by botoxed former beauties and their sugar daddies on the garden terrace of a plush Vendée hotel, Dave Brailsford could no longer hide the exasperation in his voice.

Just 48 hours earlier, Chris Froome had taken the yellow jersey with a dazzling tour de force at Ax 3 Domaines. It was a win that Brailsford rightly wanted to celebrate, the culmination of months of planning and hard graft. Yet much of what he'd seen in the press the following morning had left him incensed. In *Le Monde*, the former Festina coach, Antoine Vayer, said that by his calculations, Froome's average power output on the final climb had been 446 watts, just two fewer than Lance Armstrong on the same ascent in the 2003 Tour.

Vayer lent further context: "Below 410 watts there's no guarantee but it's human, between 410 and 430 is suspicious, up to 450 is miraculous and beyond that is mutant." None of this was intended as a compliment.

Days later, at the entrance to the start village in Avranches, we recognised the twin tufts of greying hair either side of Vayer's bald dome, his bulging eyes and conspiratorial smile. "Don't repeat this," he whispered, "but tonight I'm meeting Brailsford."

If the ensuing discussion was supposed to broker peace and some level of understanding, it didn't have the desired effect: Vayer's articles implying that Froome's performances were in line with a turbo-charged, EPO-fuelled Armstrong continued, gaining increasing traction at the Tour.

Vayer's reputation as a firebrand would have made him relatively easy to dismiss – that's if other analysts weren't arriving at similar figures with their different formulae. When the practice of estimating power outputs on the basis of a rider's time on a certain climb and his weight was in its infancy, the simplicity of the methods made a mockery of the results. Wind speed and direction were either assumed or ignored, as was air pressure, the influence of drafting, the rolling resistance of the road surface and a host of other factors. Rider weights and even where climbs started and finished were also subject to guesswork. This was before one even considered, for example, whether the climb was the last of five in a 220km, third-week mountain stage, or came after 120 flat kilometres in the first few days of a race.

▲ Froome and Team Sky's strict adherence to all checks and controls did nothing to silence their 2013 Tour doubters

By 2013, though, some of the individuals doing the sums were experts with refined techniques for gathering and processing performance data. One of them, the South African physiologist, Ross Tucker, saw Froome rampage to victory at Ax 3 Domaine and mused that he was "either one exceptional individual or, well, we know the rest, we have seen this movie too often in the sport."

By now, understandably, Sky and Brailsford were beginning to feel embattled. Under pressure to release Froome's power files, Brailsford replied that there were "very few people who can properly interpret and understand that data" but plenty who were happy to indulge in "pseudo-science".

That term – "pseudo-science" – added a Molotov cocktail to an already angry bonfire. The flames have continued to be fanned by more prodigious exploits from Froome. Barely a race now passes without someone estimating the winner's wattage, his power-to-weight ratio in watts per kilo or his VAM – a measure of ➲


LCL
LCL
LCL

5%

The typical difference in maximum sustainable power between a 20-minute and 60-minute effort. Around

25W

for a good pro

◀ Chris Froome's performances at the 2013 Tour were subject to extensive scrutiny by 'experts'

▶ Brailsford has called the practice of power calculation via TV using assumed data 'pseudo-science'

climbing rate per hour invented by the notorious Michele Ferrari. And barely a number can be quoted without a pronouncement, implicit or explicit, on whether the rider is doping or not.

Meanwhile, team coaches such as Sky's Tim Kerrison, FDJ's Fred Grappe and Julien Pinot, Cannondale's Sebastian Weber and BMC's Marco Pinotti – all of whom we spoke to while researching this article – continue to view the new science of performance estimation with a mix of puzzlement, curiosity and weariness. With Weber it's mainly the latter: "I can remember seeing some crazy figures for Contador on Verbier in the 2009 Tour that were being quoted in German papers, so I contacted the guys who'd worked them out. It turned out they'd just asked 15 people to calculate and had taken the average. But these 15 people had numbers between 380 and 490 watts!

"There are so many other examples. When I was working with Katusha last year, we'd changed the tyres we were using and guys were coming back from the Tour de San Luis and Oman saying the new ones felt slow, especially in hot weather. So we tested the tyres on a velodrome to measure the rolling resistance and, in cold conditions in the velodrome, one tyre was 17 watts faster than the other one. That's one tyre. So two tyres would be over 30 watts! And when it's warmer it's worse and when it's a bad road it's worse. So how can these calculations account for just an easy thing like rolling resistance, which is key because the speed is not that important on a climb?"

"It turned out they'd just asked 15 people to calculate and had taken the average. But these 15 people had numbers between 380 and 490 watts!"

Sebastian Weber

For every Weber, though, there is a Pinotti, the recently retired rider who is now a coach at BMC. The Italian says that his generally dim view of the armchair analysts changed when, last winter, a prominent Twitter-based number cruncher from Finland contacted him to request some of Pinotti's old power files for personal research. The account goes by the Twitter handle @ammattipyöräily - 'professional cycling'. Its owner told Pinotti that he or she would compare Pinotti's raw, accurate power data from the 2010 Giro with the estimates generated by his own timings and algorithms. Pinotti was shocked at the result: "Their accuracy was incredible," Pinotti says. "They were never more than about five watts out."

Previously, Pinotti says, the limitations of remote analysis had been all too clear. He sensed – and could see from his own SRM files – that several factors were substantially under- or overestimated by those watching races on TV with a stopwatch and calculator. There was one variable that Antoine Vayer's method ignored and which, in Pinotti's view, caused errors "As far as I could see, he was always overestimating the watts because he underestimated the effect of drafting, especially in the climb's first few kilometres where the pace is highest and the group biggest."

Nowadays, Pinotti is more inclined to endorse the work of the best couch statisticians – especially the mysterious @ammattipyöräily (he or she doesn't identify him or herself in email exchanges with *Procycling*), who is known to collect raw data from professional racers and calculate other riders' wattages by extrapolation. Pinotti worked with Sebastian Weber at T-Mobile and HTC-Columbia but is sceptical about Weber's claims that a change of tyres could save a rider up to 30 watts – and the logical implications for armchair analysis. "In my experience road surface makes the biggest difference but the Finnish guy, for example, builds that into his calculations," Pinotti says.

In the Italian's view – as in Fred Grappe's and Ross Tucker's – knowing or accurately estimating wattages can give us a useful overview of performance trends and clues about doping patterns. But 'pattern' is the key word; apparently unlike Vayer, this trio isn't generally interested in pointing fingers at individual riders or exceptional, isolated performances. There are two main reasons for this: one, an acceptance that inaccuracies can occur, both in the raw data generated by power meters and the amateur sleuths' maths, and, two, because they find the notion of a performance 'ceiling' problematic both intellectually and philosophically. In other words, they all acknowledge that there will always be outliers, freaks of nature who exceed what we previously believed was possible. When Antoine Vayer writes in *Le Monde* that, based on his stats from the 2013 Tour, Chris Froome "must be the best rouleur-grimpeur of all time", a scepticism bordering on sarcasm almost drips from the page. This, though, is also a hypothesis that most objective judges are willing to entertain. Fred Grappe nuances that Froome wouldn't even need to be the very best athlete of his generation: "To do what he does, you'd need a VO_2max in the high 80s, getting towards 90 or even above – which is very rare but you see in three or four riders in most generations. You'd also need to be among the most efficient riders, in other words be able to get the best use of that cardiovascular engine. We've seen past performances that have been consistent with a VO_2max of 100, 102, and that's been because of illegal methods. You have to bear in mind that blood doping can boost VO_2max by seven or eight millilitres, then there can also be a small psychological boost that comes with doping. So you've perhaps had riders with a natural VO_2max that's been very high boosting it by up to 10 millilitres and into that 'unexplored' zone above 95. That's been down to doping – but at some point you may just have to stand back and say, 'This is a physical specimen that has never existed before.'"

Froome and his credibility, though, are really separate issues. A more important question is whether it's ever okay to put a limit on what can be achieved – to say that, as per the Olympic ideal, sporting competition is about going ever stronger, higher, faster, but only up to a pre-agreed threshold. This idea jars with Julien Pinot of FDJ. Marco Pinotti feels the same, although concedes that limits must actually exist, there's just no way of knowing where. "No one will ever win the 100 metres in five seconds. That's beyond the limit, and you can say the same in cycling: to my mind 6.8, 6.9W/kg [at threshold] is beyond the limit. But where do you draw that line?"

And, maybe more pertinently, why would you draw a line – as in, for what purpose? Three-time Tour champion Greg LeMond has long advocated the use of physiological monitoring to replace or augment an anti-doping system currently based on direct or indirect detection of banned drugs. The theory is that, after a certain period of time or number of tests, personalised performance limits could be established. These would act as trip-wires: a rider exceeding his 'natural' capacities would at the very least have to explain himself.

◀ Sky's Kerrison says that doping distracted the sport from refining training practices that are now helping riders

▶ Alberto Contador's searing climb to Verbier in 2009 provoked wildly varied calculated power data

Fred Grappe has also lobbied vocally for a performance passport to sit alongside the biological equivalent. Grappe already employs something similar at FDJ – a profiling system called the Record Power Profile (RPP). He stresses that this is primarily a training optimisation tool, although it would also theoretically reveal suspicious spikes in performance. Constant monitoring of his riders' training and race data is indeed one reason why Grappe is "very close to 100 per cent confident" in FDJ being completely clean and "willing to stake my future on it. Say that if anyone's caught doping, I should lose my job." His fellow FDJ coach, Julien Pinot, won't go quite that far: "Someone can still do something at home to gain a couple of per cent. Because doping's a mental thing, too. A guy could dope just a tiny bit but that combined with the psychological effect can be the difference between winning and losing."

Pinot makes an important point: in the current age of (alleged) micro-dosing and highly refined detection methods, the prospective cheat aims for small but decisive performance gains. And while two or three per cent more oxygen to the muscles might be the difference between winning and losing, it probably won't cause a glaring increase in power output.

Herein, says Sebastian Weber, lies one major problem with power as an indicator of doping: even if the data is accurate and environmental factors discounted, fluctuations will be too small to allow firm judgements. Either that or, in most cases, they could be legitimately explained away. As Weber says, with six weeks of training a rider can conceivably increase his VO_2max by 10ml and his power output by several per cent. Year-on-year comparisons might give us strong clues as to whether a similar jump owed to hard work or chemical enhancement but, without any accompanying chemical evidence, will never be legally acceptable as proof. Moreover, as Weber points out, the challenge of collecting years of accurate data – conceivably starting in a rider's junior days – is simply too daunting for any authority to contemplate, and certainly cash-strapped national cycling federations, national anti-doping agencies or the UCI. Calibration of the capture devices is another concern. Five manufacturers – SRM, Garmin, Powertap, Stages and Quarq – dominate the power meter market. All of their devices have a degree of variance between each unit and require regular calibration by the user. "It's a minefield," says Weber. "I mean, even when you take the same power meter… they calibrate an SRM one way in Italy and another way in America."

The sports scientists at Sky have also sometimes been left scratching their heads. Bradley Wiggins was much scrutinised for his use of non-round chainrings in his 2012 annus mirabilis – and ridiculed by Vayer for Sky's claims about their advantages. What Sky's Head of Athletic Performance, Tim Kerrison, and his colleagues couldn't fathom, though, was why Wiggins's exceptional wattages weren't producing even more speed. It took them a while to find the answer: the shape of the chainrings was simply causing the SRM to over-read his power output.

These caveats aside, what almost everyone, even Weber, acknowledges is that good surveillance of accurate performance data could help to direct internal and external dope-testing. Weber has first-hand experience of precisely this. In the spring of 2012, while working for Katusha, he noticed gross inconsistencies between what the team's sprinter, Denis Galimzyanov, was capable of in training and his impressive displays in races. "At all of the training camps he was the worst rider," Weber recalls. "That spring, 2012, we had a meeting at our service course in Lake Garda, and I said, 'This guy has a red light on'. He had big pressure from Russia, he wasn't connected to anybody on the team, he didn't have the necessary power – he was putting out 300W on a climb and then cracking – and his jump of performance was suspicious in the SRM data. And two weeks later he tested positive for EPO."

Weber says that rather than highlighting the need for performance surveillance, the example of Galimzyanov proves that normal testing is usually adequate. "If someone does something really stupid, dopes massively, you're going to see it in the power data but you'll also see it in the blood and urine tests. The only exception would be someone using an undetectable product but then you can never make a case unless you find out what it is, anyway."

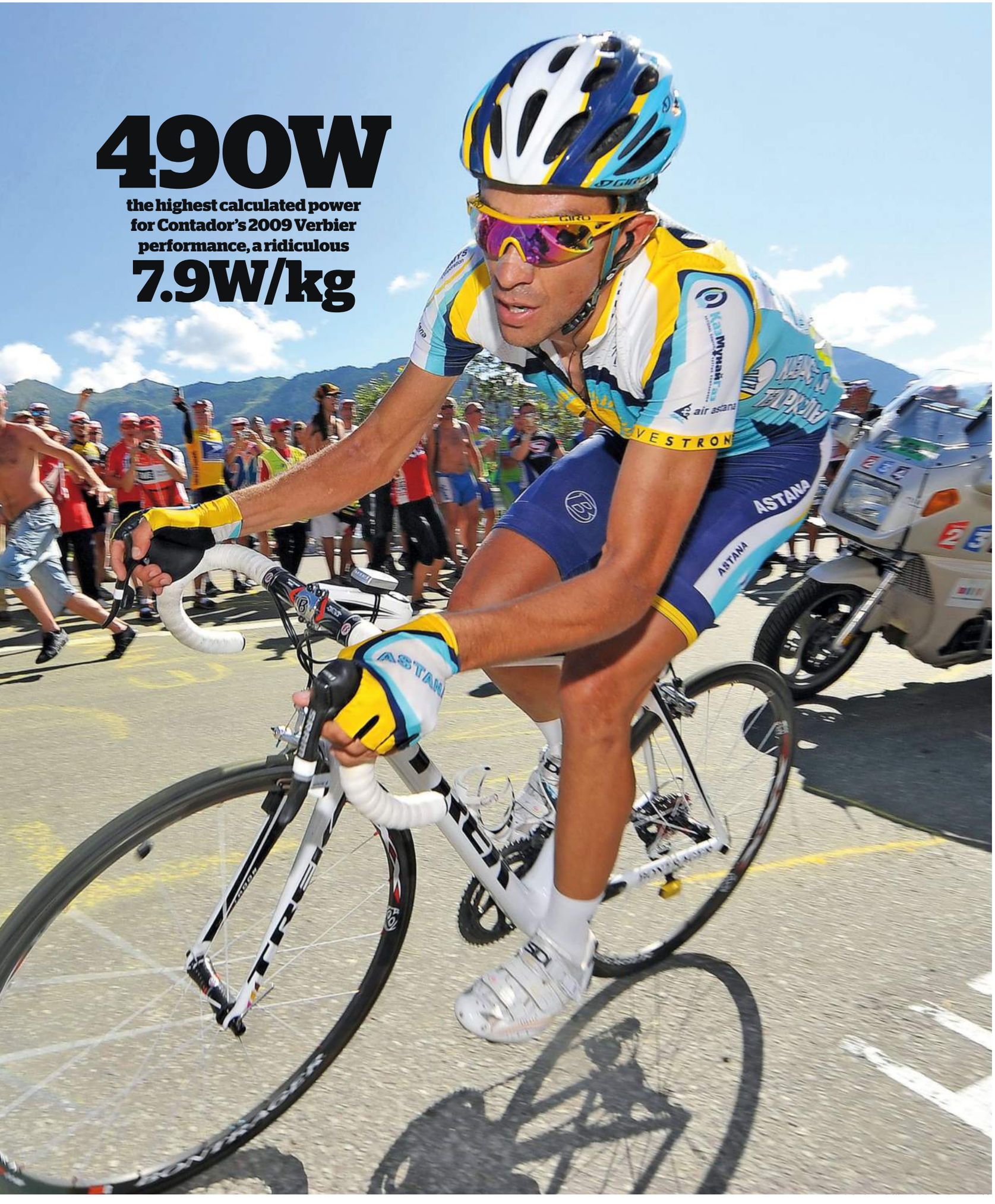
That eventuality, though – the appearance of a new, undetectable doping method with massive performance gains – is surely exactly why authorities in professional cycling (and also teams) should be at least keeping an eye on performance data. Perhaps it wouldn't be immediately clear what was causing climbing times to plunge or power outputs to soar but this realisation would at least be a call to arms, the spur to gather intelligence. As Pinotti notes: "If you look at the evolution of the world record in the 10,000 metres, it's pretty steady except for two big dips, one that coincided with the advent of pacemakers and the other, we know now, with the arrival of haematic doping. It was pretty much the same in cycling, if you just looked at climbing times on L'Alpe d'Huez. I think looking at the performance data of the group can definitely help to identify macro trends."

One such macro-trend, of course, ought to be the current generation inching towards their EPO-enhanced predecessors' high watermark. The question is how quickly the two graphs might intersect. Ross Tucker pondered precisely this in a lengthy post on his 'The Science of Sport' website in June 2013. Tucker's final estimates, allowing for factors including improved technology, training methodology and increased financial incentives, was that it would probably take between 20 and 30 years (Tucker admitted erring towards 30) for a clean rider to challenge, say, Marco Pantani's record of 37:35 on Alpe d'Huez. Even allowing for environmental conditions on the day, with this in mind, it is understandable that alarm bells would start ringing when *L'Équipe* reports that Chris Froome's record on the Col de la Madone is 36 seconds faster than Lance Armstrong's.

Decontextualised dot-joining like this, though, is the source of great frustration to Froome ▶

> **"A guy could dope just a tiny bit but that combined with the psychological effect can be the difference between winning and losing"**
>
> *Julien Pinot*

380W
the lowest calculated power for Alberto Contador's eyebrow-raising performance at Verbier in the 2009 Tour de France, a non-alarming
6.1W/kg


490W
the highest calculated power for Contador's 2009 Verbier performance, a ridiculous
7.9W/kg
air astana
ASTANA
TREK

◀ Conventional testing was enough to reveal that Galimzyanov was both a massive doper and a bit rubbish

and his coaches. It's easy to sympathise: Sky, more than any other team, entered professional cycling with the stated brief of pushing the boundaries of training, technology and – as a consequence – human performance.

Dave Brailsford has long argued that widespread, heavy-duty doping stunted the sport's development and that his team have merely begun to explore areas and resources that for years were ignored. Marco Pinotti at BMC also fully subscribes to the view that EPO slowed cycling's natural evolution: "Ten years ago, almost no one was training at altitude because EPO did the same thing, only much more powerfully. Now if you look at the top 10 from any given grand tour, you'll almost invariably see that they trained at altitude."

The idea that today's riders couldn't possibly rival Armstrong's supersonic climbing speeds irritates Tim Kerrison because of parallels with his time coaching Australian swimmers. As Kerrison explains: "In 2008 and 2009 swimming was fundamentally changed through the introduction of a new generation of non-textile swimsuits, which greatly reduced the (hydrodynamic) drag of the swimmer, resulting in approximately two per cent improvements in times. Almost every world record tumbled in 2009 and when the suits were banned at the end of 2009, there were many claims that it would be decades before these records were broken again. Based on historical rates of progression, it should have taken that long but already one third of these records have been surpassed. If anything, the brief introduction of the high-tech swimsuits highlighted what was possible. The event focused swimmers and their coaches on the factors that can have the biggest impact on performance and on finding legitimate ways to improve.

"The athletes who thrived in that era were not necessarily the most gifted but those who were prepared to do whatever it took to perform, by whatever means"

Tim Kerrison

"The biggest error, in my opinion, is to assume that there is a known limit of human capability," Kerrison adds. "It would be an interesting exercise to engage experts from a range of scientific disciplines to discuss what the limits of cycling performance might be but one thing I am very sure of is that we are not yet close to reaching these limits. As a young coach and sport scientist, I was taught that setting limits on performance would only result in limited athletes. The last thing I want my riders thinking is that something they are aiming to achieve is not physically possible.

"We must also consider that in the doping era of cycling, we didn't necessarily see what the most talented riders were capable of in a well-supported environment," he goes on. "The focus of supporting and developing riders was through doping, so the athletes who thrived were not necessarily the most gifted but those who were prepared to do whatever it took to perform, by whatever means. Specialised medical (doping) support was prioritised over legitimate coaching and performance support structures. The sport suffered as a consequence."

Kerrison could also have mentioned cycling's rapidly growing gene pool and demographic diversity – the fact that until just a few years ago professional cyclists were almost exclusively working-class white Caucasians from western Europe. The American journalist David Epstein also provided evidence for a fascinating theory in his ground-breaking book *The Sports Gene*, what he calls 'The Big Bang of Body Types'. Epstein postulates then demonstrates that, over the decades, professional sports have sifted out the body types best suited to each discipline, causing the general level to improve at a rate that a constellation of other factors can't explain. As Epstein puts it: "Just as the galaxies are hurtling apart, so are the body parts required for success in a given sport speeding away from one another toward their respective highly specialised and lonely corners of the athletic physique universe."

However we measure them, then – with stopwatches, videos or power meters – we will discover that professional cyclists are getting faster and more powerful at a greater rate than we perhaps could have anticipated.

Meanwhile, the war of words between the pseudo-scientists, the armchair experts, the living-room detectives – whatever we decide to call them – and their detractors will become ever louder. Louder and more contentious.

Jonathan Vaughters, manager of the Garmin-Sharp team, explains the sometimes inexact science of power calculation

"The first thing you're going to establish is your ascension speed, how long it takes to gain a certain altitude. From that you can then calculate your Newton-metres, which is the force it takes to lift one kilogram, one metre vertically, in one second. One Newton metre equals one watt, so on that basis you can work out how many watts it would take to do a certain climb if you were riding in a vacuum.

So far so scientific but then you get into the subjective stuff: what is the rolling resistance, air resistance, tailwind, headwind, temperature, the effect of

drafting? This is all very difficult to do. For example, look at the difference in times on the Ventoux when it was resurfaced in 2004. I mean, you can say that we were all on EPO – which we were – but the times just nosedived by about a minute on average. And that was to do with the tarmac and the rolling resistance. So, sure, there's a massive difference but how do you quantify that?

Then you consider that a heavy rider is going to have a greater rolling resistance than a lighter rider, that the riders have their tyres inflated to different pressures...It all starts to get a bit nebulous but that's just the rolling resistance.

Then you think about air resistance, which is affected by temperature, wind direction, wind speed, rider position, air pressure... These may only account for a small margin of the final calculation but if you get a few of them wrong and it causes quite big mistakes. To get some of these variables nailed on, we'd have to go to a wind tunnel, get a guy in his climbing position with his jersey flapping, food in one pocket, a water bottle on his bike, and climbing at 21kph... That's just one guy on one climb. We've never done that and probably never will.

So there's a lot of guesswork involved but quite often the estimates are pretty close to the raw data we get from the SRM. The guy who consistently seems to be the most accurate is @ammattipyöräily on Twitter. He seems to take the steepest portions of climbs, where everyone is spread out and the drafting is minimal, and he evens it out from there. His numbers tend to correlate with what we see on the power meters. You have to be careful because some climbs can, for instance, have two kilometres of flat where the body recovers and can produce anaerobic force again when the slope ramps up. Extrapolating from that is going to be useless."