Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 147 (2016) 607 – 612

Procedia Engineering

www.elsevier.com/locate/procedia

11th conference of the International Sports Engineering Association, ISEA 2016

Tour de France Modeling: 2015 Results and Comparisons with Elite Cyclist Power Data

Chad Michael Hobson[a], John Eric Goff[a,*]

[a]Lynchburg College, Department of Physics, 1501 Lakeside Dr, Lynchburg, VA 24501, USA

Abstract

For the past dozen years, our research group has been refining a physical model used to predict the winning time for each stage of the Tour de France. Our model is based upon a series of incline planes and incorporates real stage data and cyclist power output, as well as air and rolling resistances. We report on our most recent model modification in which we utilized allometric scaling to adjust our model cyclist's power output based upon varied rider masses for different stage types. We also provide a comparison between our model and published power data for top level cyclists and recent Tour de France winners such as Chris Froome and Vincenzo Nibali. This juxtaposition showcases not only how well our model predicts stage-winning times, but also the extent to which our model matches reality. We finally report on how our model performed in predicting the winning time for each stage in the 2015 Tour de France.

Peer-review under responsibility of the organizing committee of ISEA 2016

Keywords: cycling; modeling; human power output; drag area; Tour de France

1. Introduction

Our research group began modeling the Tour de France in 2003. The original model was based upon a series of incline planes and used discrete power levels for the rider determined by the angle of the incline plane. Our goal was to predict the winning time for each stage of the Tour de France. We were entirely concerned with what the best athlete could do on any given stage. In 2003 and 2004, we were able to predict a majority of stages to under 10% error [1,2]. Since the development of this original model, our goal has remained the same while the model has become more intricate. The first major change to the model came in 2013 when we transitioned from discrete models of drag and power to continuous ones. The 2013 Tour de France saw unprecedented high speeds in the mountain stages in the second half of the race, leading to four stages with greater than 8% error on our predictions [3]. This was somewhat disheartening in that the previous two years of predictions with the discrete model had only one prediction with error over 8% each [4]. The 2014 Tour de France saw the next significant change to our model. In previous years, we used the same cyclist mass for every stage. In 2014 we incorporated the general trend that heavier cyclists tend to win the flat stages, whereas lighter cyclists dominate the mountains. Based on this observation, we averaged winning rider

* Corresponding Author. Tel.:+1-434-544-8856 ; fax: +1-434-544-8646.
E-mail address: goff@lynchburg.edu

1877-7058
Peer-review under responsibility of the organizing committee of ISEA 2016
doi:10.1016/j.proeng.2016.06.250

masses for flat, medium mountain, and mountain stages from previous Tours de France to have a cyclist mass for each type of stage (team time trials and individual time trials were considered to be flat stages). We then scaled cyclist power with mass using the general rule of thumb that power scales with mass to the two thirds power [5]. Our results for the 2014 Tour de France were promising as all of our predictions were under 8% error, including five stages under 1% [6]. Beginning in 2011 we have posted our predictions before each stage on a blog that has attracted worldwide attention [7].

The modifications we have made to our model may raise the question of whether or not we are actually predicting the motion of the wining cyclist or simply generating excellent predictions based on some type of fortuitous averaging. The 2015 Tour de France saw more riders than ever wearing devices to measure power output. Several riders published power data from certain stages or specific climbs throughout the race. This paper will show how our model's power output accurately represents recently published power data from elite cyclists such as Chris Froome and Vincenzo Nibali, winners of the 2015 and 2014 Tours de France, respectively. For the 2015 Tour de France, we made a slight tweak to our power scaling based upon the results from the 2014 race, so this paper also reports on how our predictions compared to the actual winning times.

2. 2015 Model Description

Our model is founded on the idea of treating each stage profile as a series of incline planes and numerically solving Newton's second law equation to determine how long it takes our model rider to traverse the course. Each year the Tour de France website displays stage profiles for every stage in the race. Figure 1 shows the profile for stage 17 of the 2015 Tour de France [8]. Information from these profiles is often not sufficient to create a series of incline planes that accurately represents the stage profile. If we used only the points of distance and elevation for stage 17 given in Figure 1, for example, we would leave out the valley between Col d'Allos and Pra Loup. We thus look elsewhere to obtain a sufficient number of data points for our stage profile [9].

Our model aims to tease out the main physics of cycling: power, drag, and friction. We are less concerned with elements of the race that are not necessarily physical or predictable. These factors include weather, team strategies, crashes, and sharp turns. As previously stated, we used the most recent version of our model featuring allometric scaling based on rider mass for the 2015 race. After seeing this model's performance in 2014, we decided to shift our scaling exponent so that power scales with rider mass to the 0.85 power. For the specifics on why the change was made and how the model works in detail, see our previous work [3,6]. For the first time in the history of our Tour de France research, we took a more active approach to our predictions during the race. Previous years saw us computing all 21 of our predictions prior to the start of the race and leaving the rest up to the riders. For 2015, however, we attempted to adapt to the riding styles and conditions of the race by implementing several scaling factors on our power model based upon previous stages. If our predictions came in quite slow for back-to-back stages, for example, we would examine our prediction for the upcoming stage and possibly scale the power down to accommodate for the

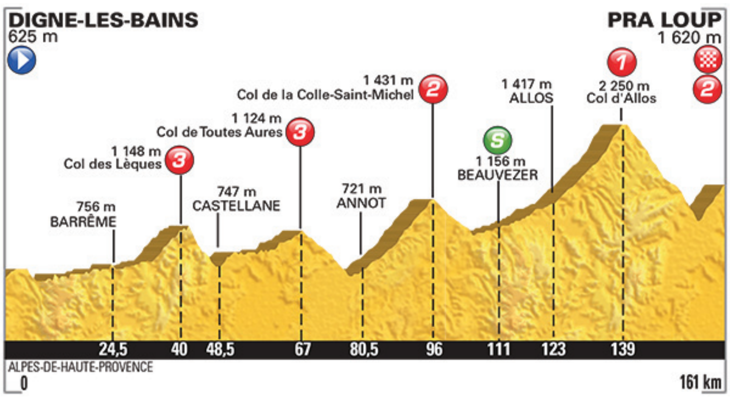

Fig. 1. Profile for stage 17 from the 2015 Tour de France [8].

strategies or fatigue of the cyclists. These changes were obviously done before the start of each stage, yet this was the first time we had manipulated any of our predictions while the race was in progress. We have used scaling factors for time trials in the past, and 2015 was no different. Outside of the time trials we used scaling factors on six stages; these factors ranged from 0.9 to 1.245. We decreased our model power on stages 7, 8, 12, and 20; we increased it on stages 3 and 16. These factors were generally motivated by trends we noticed over the previous stages as well as the changes in distance between stages. If we predicted a fairly short stage very well and the next stage was of a similar profile yet featured an extra 80 km of distance, we would consider decreasing the power slightly to compensate.

3. Model Comparison to Power Data

Our model's goal is to predict the winning time of each stage of the Tour de France, so all of our work essentially boils down to 21 stage times that are compared to what elite cyclists are able to accomplish. These final predictions are not, however, our only concern. We wish to predict the motion of the rider throughout the entire stage, which would then lead to an accurate prediction of the finishing time. For a fast calculation, we could estimate an average speed of an elite cyclist based on similar stages in the past and use it to calculate a prediction. This method may end up being exceptionally accurate, but it does not fulfill our goal. We wish to accurately model the situation at hand and not simply generate the final result. We require expressions for drag and rolling resistance that are consistent with reality, and recent research suggests that this is the case [10]. We also require that our model rider produce power outputs that are comparable to those of elite cyclists. Many top cyclists wore power meters during the 2015 Tour de France and were courteous enough to publish some of the data, thus giving us the opportunity for comparison [11–14].

Table 1 shows a comparison of our model's power to the actual power data of four different cyclists. When calculating our model's power for these climbs, we used the mass of the rider for whom we had power data instead of using our averaged flat, medium mountain, or mountain stage winner masses as described in our previous work [6]. Chris Froome is arguably the best current cyclist in the world, given his impressive victory in the 2015 Tour de France. He validated this claim on stage 10. Froome pulled away on the final climb to the Col de Soudet, and our power for this climb is clearly consistent with that of a top-level cyclist. Froome finished first on stage 10, so our model was correct in matching the power of the best cyclist of that day. Vincenzo Nibali did not place first on stage 15 because the race came down to a sprint finish. Nibali was given the same time as the top finisher for stage 15 because he finished as a part of the lead group, so his measured power is a good indication of the best cyclist for stage 15. Our model's power is again consistent with power outputs to better than 1% error for the winner of the 2014 Tour de France. Our model power was slightly larger than the actual power of Markel Irizar. Though he finished 93rd overall, Irizar was a member of the breakout group on stage 16, so he likely exerted an elite cyclist's amount of power when he came across the Col de Cabre towards the middle of the stage. The main difference between Irizar and a cyclist like Froome is that Froome can maintain the power output by Irizar for an entire stage, whereas Irizar failed to hold his break and dropped to finish the stage in 22nd place. Irizar probably lost his endurance during this first difficult climb of the stage, so our slight overestimation of power makes sense both in that respect as well as in the idea that he may not be able to output the power of an elite cyclist capable of winning stage 16.

Thibaut Pinot is the rider for the last three climbs described in Table 1. At 25 years of age, Pinot is a relatively young rider, but has had outstanding performances in the past two Tours de France. He finished third in the general classification and first in the young rider classification in 2014 [8]. He placed fourth in both the mountain rider and

Table 1. A comparison of actual power and our model's power for six different climbs in the 2015 Tour de France [11–14].

Climb	Rider	Actual Power (W)	Model Power (W)	% Difference
Col de Soudet - Stage 10	Chris Froome	391.5	391.7	0.05
Col de l'Escrinet - Stage 15	Vincenzo Nibali	306	308.4	0.78
Col de Cabre - Stage 16	Markel Irizar	320	327.0	2.19
Col de la Colle Saint-Michel - Stage 17	Thibaut Pinot	293	292.6	-0.14
Col d'Allos - Stage 17	Thibaut Pinot	331	296.2	-10.50
Pra Loup - Stage 17	Thibaut Pinot	362	340.4	-5.97

young rider classifications, as well as 16th in the general classification, in 2015 [8]. For the first climb on stage 17, Pinot stayed relatively close to the other riders, not gaining or losing any large chunks of time from the leader. It therefore seems reasonable our power matches his to better than 1% error. Our power for Pinot's ascent to the Col d'Allos is, however, an underestimation of his actual output by more than 10%. Our model predicted that it would take Pinot about 2′ 10″ longer than it actually did. Pinot was not in the lead going into this climb; instead Simon Geschke was in front by two minutes at the start of the ascent. Geschke went on to win the stage, so he was the rider that we would have liked to predict. Pinot has a mass of 63 kg and Geschke has a mass of 64 kg [8]. We used Pinot's mass for these calculations of power because it is his data that we obtained. The 1 kg increase in mass when examining Geschke correlates to roughly a 4 W increase in power, which would bring our model power closer to the actual data. Pinot was able to gain a minute of time back on Geschke during this climb to the Col d'Allos alone, so he actually out-performed the best cyclist on stage 17 during this climb. If we had power data for Geschke, it is likely that our model would have better matched his output based on the time difference between his and Pinot's climbs. Unfortunately for Pinot, a crash on the descent set him back to 1′ 50″ behind Geschke going into the final ascent to Pra Loup. He placed 4th on the stage and was able to gain back 14″ on Geschke during the last climb. Our model predicted Pinot taking about 22″ longer than he did, thus implying that we only missed the winner's ascent time by 8″. Pinot outdid the leading cyclist on the climb, but the race did not fall his way. Examining the time differences between Geschke and Pinot again implies that had we been able to examine power data for the best cyclist of stage 17, our model would likely have alligned better than it did for Thibaut Pinot's more sporadic performance.

There are obviously a myriad of factors that could impact the amount of power a rider outputs on a given climb. The ascent may be full of hairpin turns or riders could face a strong headwind, trying to push them back down the mountain. We are also obviously unsure of which rider will win each stage, meaning that we cannot incorporate the exact mass of the rider we would like to predict. We instead do our best to use masses that are consistent with previous winners of similar stage types. Even if our power model matches exceptionally well when using the winning rider's mass, our actual prediction may not be so accurate because we could only estimate this mass before hand. With so much unpredictability that is inherent to the Tour de France, having power data that is reasonably accurate for all climbs, and exceptionally precise for some, provides a solid base for our model. Considering that young riders such as Pinot can crank out extraordinary amounts of power, we will have to reconsider the abilities of elite cyclists in future years. Our current model of rider power, however, appears to be accurate enough to provide a good description of each stage.

4. Results

The 2015 Tour de France was relatively short in comparison to recent years; all races dating back to 2002 have been longer. Chris Froome won the 2015 race with an average speed of 39.64 kph, the slowest since 2010. The race featured seven flat stages (F), five medium mountain stages (MM), seven mountain stages (M), one individual time trial (ITT), and one team time trial (TTT). The mountain stages must have been especially brutal to drop speeds so low for such a relatively short race. Table 2 shows the results for our model for the 2015 Tour de France.

The start of the race took place in the Netherlands, followed by two days in Belgium before reaching France. Our first two predictions came in very slow. We had an excellent prediction for stage 3, however, it was marred by crashes. A clean race would have seen our prediction be quite slow, much like the first two stages, but sometimes the unpredictable aspects of the race fall in our favor instead of working against us. We did increase the power on this stage as noted in the model description due to the high speeds cyclists achieved over the first two stages. Stage 1 was unexpected. Rohan Dennis' average speed of 55.45 kph is a Tour de France time trial record. Predicting a record-breaking performance is nearly impossible because our model is based upon previous stages and power information. We must learn when boundaries are being pushed. Riders may have pushed the envelope a tad too much on the first three stages as speeds drastically dropped on the next three. It seemed as though the peleton took these three flat stages rather easy because there was no significant chance for any rider to gain much time in the general classification. Even with an excellent physical model, we can never predict the strategies that teams will employ on a given stage.

Over the next eight stages, we had what is likely the most successful streak of predictions in the history of our research. From stage 7 to stage 14, all of our predictions fell below an absolute percent difference of 1.85%, including 5 stages under 1% difference. These stages were not impacted by serious crashes, rainfall, or headwinds that may

Table 2. Model results for the 2015 Tour de France.

Stage	Actual	Predicted	Difference	% Difference
1-ITT	14′ 56″	16′ 38″	01′ 4″’	11.38
2-F	3h 29′ 03″	3h 45′ 51″	16′ 48″	8.04
3-MM	3h 26′ 54″	3h 28′ 35″	01′ 41″	0.81
4-F	5h 28′ 58″	5h 04′ 45″	-24′ 13″	-7.36
5-F	4h 39′ 00″	4h 19′ 08″	-19′ 52″	-7.12
6-F	4h 53′ 46″	4h 24′ 45″	-29′ 01″	-9.88
7-F	4h 27′ 56″	4h 32′ 22″	04′ 57″	1.85
8-MM	4h 20′ 55″	4h 19′ 45″	-01′ 10″	-0.45
9-TTT	32′ 15″	31′ 50″	-00′ 25″	-1.29
10-M	4h 22′ 07″	4h 24′ 38″	02′ 31″	0.96
11-M	5h 02′ 01″	4h 58′ 09″	-03′ 52″	-1.28
12-M	5h 40′ 14″	5h 38′ 04″	-02′ 10″	-0.64
13-MM	4h 43′ 56″	4h 46′ 07″	02′ 26″	0.86
14-MM	4h 23′ 59″	4h 21′ 59″	-01′ 44″	-0.66
15-F	3h 56′ 35″	4h 12′ 01″	15′ 26″	6.52
16-MM	4h 30′ 10″	4h 47′ 26″	17′ 16″	6.39
17-M	4h 12′ 17″	4h 21′ 10″	08′ 53″	3.52
18-M	5h 03′ 40″	4h 50′ 12″	-13′ 28″	-4.43
19-M	4h 22′ 53″	4h 03′ 33″	-19′ 20″	-7.35
20-M	3h 17′ 21″	3h 15′ 50″	-01′ 31″	-0.77
21-F	2h 49′ 41″	2h 35′ 06″	-14′ 35″	-8.59
Total	84h 46′ 14″	82h 57′ 17″	-1h 48′ 57″	-2.14

cause drastic changes in pace. A few of the stages did have fairly high temperatures for the riders to endure, but the speeds did not appear to be too severely altered. We started off well with another mid-race adjustment. We decreased power slightly due to the slow pace on the previous three stages, and we saw great success. Aside from stages 7 and 9, which were flat and team time trial stages, respectively, all stages in this range were more focused on climbs. Harsh climbs limit the options for strategies for most teams. The race is no longer dictated by the peleton, but instead by the elite climbers of the field. No rider competing for the general classification can afford to take an easy day during these mountainous stages because it is on these stages that riders lose or gain large chunks of time. The race is thus more predictable because we may count on at least one of the elite cyclists to push in an attempt to get the coveted yellow jersey. As shown in the previous section, our power for Chris Froome's final climb on stage 10 was off by less than 0.1%, so it is not surprising that our prediction for the entire stage fell below 1% error.

After stage 14, our predictions started to tail off again. Aside from stage 20, which we predicted to better than 1% error, all of our percent differences fell between a magnitude of 3.52% and 8.59%. The only stage dictated by weather was the final ride to the Champs-Élysées. Riders typically take this stage as a celebratory ride, so we decreased model power to account for this. Heavy rain, however, made speeds even slower than we anticipated. Tour de de France organizers, due to the weather and nature of the final stage, determined that all cyclists would get the same time for the stage 68.5 km before the finish, thus sealing Froome's victory and our largest error on the back end of the race. Riders pushed the pace on stage 15, which broke our streak of predictions under 1.85%. We envisioned a similar pace for the following stage, so we boosted power of stage 16 to compensate. We did not have the same level of success as with our previous in-race modifications, however the adjustment did push our prediction in the right direction. For the three stages prior to the finish, our predictions came in fast. At this point in the race, it is easily imaginable that the riders are exhausted and simply cannot output the same power as on the opening day. Our model does not account for any deterioration of the riders throughout the race, so we are not particularly surprised by this trend. We did, however, attempt to compensate for this trend by dropping power slightly on stage 20. This tweak helped us to achieve our seventh prediction with less than 1% error.

Outside of Rohan Denis' record-breaking performance in the opening time trial, all of our predictions fell below a 10% margin of error. This is not as excellent as the past couple of years, however, we have come a long way from the original model in 2003. Though we did have some relatively large errors, we also predicted seven stages with errors under 1%. This feat alone was enough to get our research group attention form several major media outlets including *The Washington Post* [15] and *CNN International* [16]. The average magnitude of error for medium mountain and mountain stages was 2.34%, but our average magnitude of error for flat stages was 7.05%. Based on the stark difference between our accuracy on flat stages compared to mountainous stages, we conclude that flat stages are more difficult to model because they are more susceptible to team strategies. To counter this, our research group has tried implementing a form of scaling to account for the strategies and fatigue of riders throughout the race. There could be other factors involved in the necessity for these scaling factors, such as the reduction in drag and power when riders are in the slipstream of other cyclists in the peleton [10]. We are still pleased, however, with our predictions for the 2015 Tour de France, especially given that we predicted one third of the stages to better than 1%.

References

[1] Hannas BL and Goff JE. Model of the 2003 Tour de France. Am J Phys 2004;72:575-9.
[2] Hannas BL and Goff JE. Inclined-plane model of the 2004 Tour de France. Euro J Phys 2005;26:251-9.
[3] Ramsey BA and Goff JE. Predicting Tour de France stage-winning times with continous power and drag area models and high speeds in 2013. Proc IMechE, Part P: J Sports Engineering and Technology 2014;228:125-135.
[4] Goff JE. Predicting winning times for stages of the 2011 Tour de France using an inclined-plane model. Proc Eng 2012;34:670-5.
[5] McMahon TA. On size and life. 1st ed. New York: Scientific American Books/W. H. Freeman and Company, 1983.
[6] Hobson CM and Goff JE. Improving Tour de France modeling with allometric scaling. Proc IMechE, Part P: J Sports Engineering and Technology 2015.
[7] http://johnericgoff.blogspot.com/.
[8] http://www.letour.fr/.
[9] Google Maps Find Altitude (homepage on the Internet), http://www.daftlogic.com/sandbox-google-maps-find-altitude.htm
[10] Barry N, Burton D, Sheridan J, Thompson M, and Brown NAT. Aerodynamic drag interactions between cyclists in a team pursuit. Sports Eng 2015;18:93-103.
[11] http://www.teamsky.com/teamsky/home/article/59618#kQDehyhWVovopup5.97.
[12] http://www.srm.de/news/road-cycling/2015-tour-de-france-stage-15/.
[13] http://www.srm.de/news/road-cycling/2015-tour-de-france-stage-16/.
[14] http://www.srm.de/news/road-cycling/tour-de-france-2015-stage-17/.
[15] http://www.washingtonpost.com/news/speaking-of-science/wp/2015/07/25/how-a-virginia-physicist-can-predict-the-tour-de-frances-outcome-from-4000-miles-away/.
[16] http://edition.cnn.com/videos/world/2015/07/26/exp-using-physics-to-predict-tour-de-france-stages.cnn.