A number of different ways to break this problem down, but basically we want to know what the chances of sprinters vs climbers are given the difficulty of the climbing on a stage.

What I have found is that looking at the toughest climb on the stage (by model category) is a very strong way of explaining which types of riders go to the end (R^2 ~0.4 of toughest climb on the average sprinter vs climber PC).

Looking at odds for sprinters and climbers to finish top 10 in a stage based on the most difficult climb or most difficult climb in the last 50 km also is very predictive.

Basically the 50/50 point in sprinters vs climbers odds comes around 4.0 difficulty. Sprinters are 4x climbers at 2.0, 2x at 3.0, even at 4.0, 0.5x at 5.0, and 0.25x at 6.0.

If the last 50 km has a climb of 6.0 or tougher it’s very unlikely the sprinters survive.

Looking at classics guys (PC3 or PC4 depending on iteration), they have equal t10 odds as climbers at 5.0 and equal odds at sprinters at around 3.5.

Some examples, Wilunga Hill is about a 4.0. Cote de Mutigny in St3 2019 TDF was about 3.0, St5 2019 TDF had a 4.5 climb, St8 2019 TDF had a 4.0, St 12 2019 TDF (Ancizan) was 12.0, Col de la Sentinelle (St17 2019 TDF in Gap) was a 3.0. The 2019 TDF had St1, 4, 7, and 11 as pure sprinter stages. None had a climb over 2.0.

2019 Vuelta

St2 which shredded everyone had a 4.6

St3 which was sprinters maxed out at 3.3

St4 which is projected sprinters again is 2.6

St8 which is a flat finish has a 6.4 climb in last 50km – very unlikely sprinters survive that one

**Modelling individual riders**

We construct a GAM for each rider with sufficient information which tests the impact of max\_effort (summit\_finish or flat finish) and limit\_effort (from above) on odds of finishing position (binomial model).

Uses five values for K to find best fit without over-fitting

Includes adjustment for strength of field inside the finish position eg, TDF top 12, 2.2 event is top 3

Weights included which decay performance slowly