The Power of Data

Predicting cycling performance with rider power output

Duncan Clark

March 20, 2018

Introduction

In the last 20 years cycling racing and training has evolved rapidly, with no small part due to the availability of on-board power meters, effectively giving cyclists a significantly more objective measure of their performance.

In the last 10 years power meters have become much more accessible to the cycling public, the the number of options ballooning and the prices tumbling, so much so that it is quite unusual for an amateur racer not to be measuring their power.

The book of reference for many cyclists starting their training with power has been the book of Allen and Coggan (2010). In this report, I will briefly review their proposed method of measuring form, highlighting subjective decisions on their part. I then look at my own training data for a period of 6 months and use multivariate statistical techniques to analyse the data, with the hope of finding a measure of form that more accurately predicts my performance. In my case the performance measure of choice is the power output over the first 5 minute interval of a 5 x5 minute interval session.

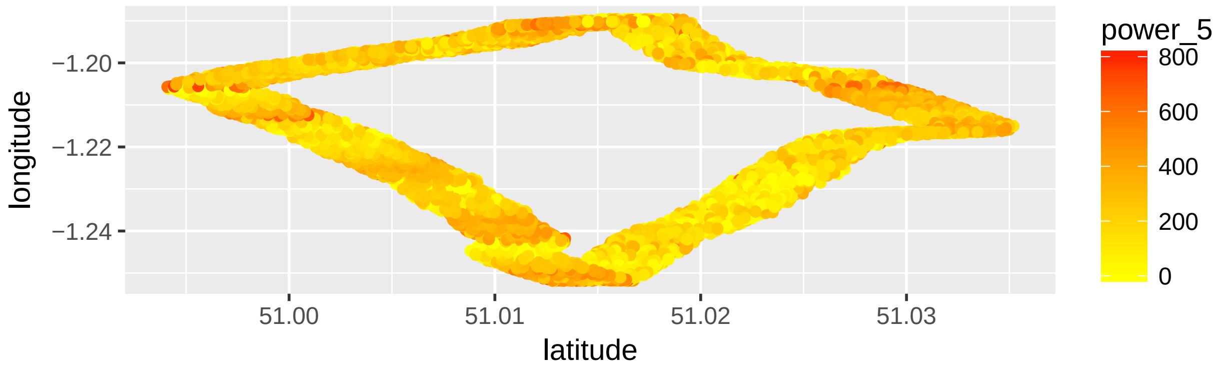
I also look at the dimension reduction of the information contained in a ride using PCA, and interpret the key components of a ride with this.

The Data

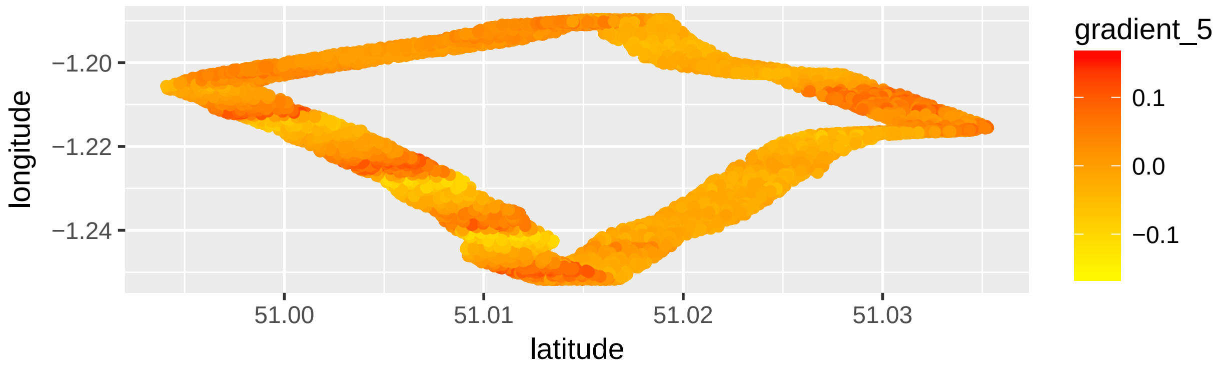
For this project I am focusing on data from a period of around 6 months from December 2016 to June 2017, through which I was training and racing. There are 183 activities spread out over 199 days, though some days have multiple activities. For each activity there is second by second data on following key variables: power, heart rate, speed, cadence, altitude, position. Data was collected on a Garmin Edge 520 bike computer, and parsed from .fit format using the R package found https://github.com/kuperov/fit. Missingness was low and usually only lasted a few seconds. Missing values were imputed to be equal to the value from the closest time that was not missing.

For illustration of what is possible with the data given the below figures are generated from the data file of a race which was laps of an around 7 mile circuit.

Plot of 5 second smoothed power and position in a circuit race

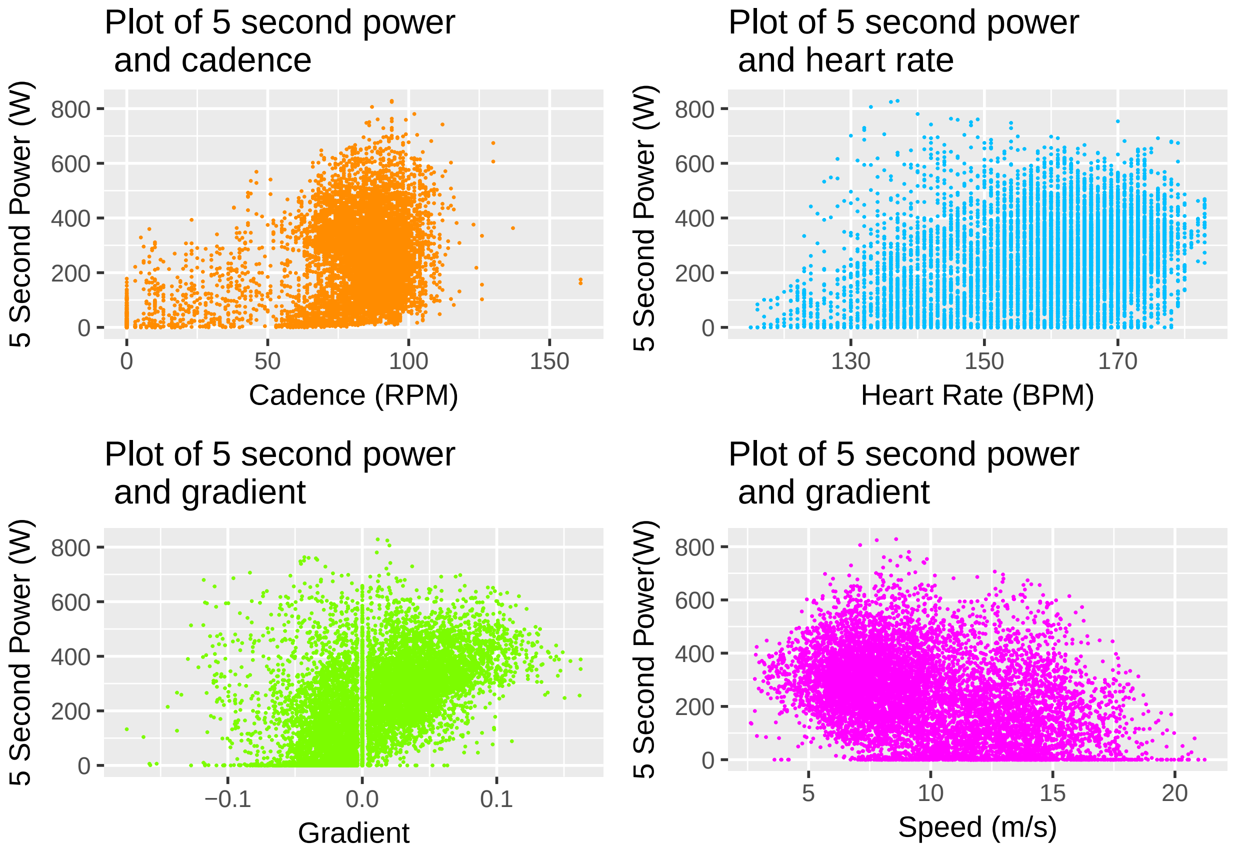


Plot of 5 second smoothed gradient and position in a circuit race



Although power data is often very noisy, it is clear that higher power output generally happened at different parts of the course. In this particular race, there was a flat section or descent in the lower right part of the plot resulting in lower power, and various rolling climbs which generally result in higher power outputs, accounting for the sections of orange. Also note that some red near corners in the course, this accounts for the concertina effect in bike racing, where the first few riders have a the optimum line through a corner, but the remaining peleton has to brake and then sprint out of the corner to stay in the group. This visualization is not that helpful in making quantitative interpretations of the ride, but it is promising that it matches up with our intuition of the how power output is spread through a race.

We can also briefly examine the relationships between a few key variables, e.g. gradient and power, cadence and power and heart rate and power. The hypothesis, being that higher power is produced up steeper hills, when spinning faster, and when the heart is beating faster. Though clearly it is likely not so simple as for example from experience whilst climbing, riders often maintain a lower cadence.


Plot of 5 second power and cadence

Plot of 5 second power and heart rate

Plot of 5 second power and gradient

Plot of 5 second power and gradient

These graphs are not particulary interesting, beyond showing possible slight positive relationships and underlining how noisy power data is. Perhaps the clearest is the gradient graph, since on gradients of above around 5%, there is minimum power output to keep from grinding to a standstill.

Clearly each observation of the data is not independent, in fact data for close times are likely to be highly correlated. As such this rules out applying many of the methods in this course to the data without some kind of transformation. During this period I often performed 5 minute interval training sessions with the first effort being an all out 5 minute effort. These sessions were fairly well spread, and are somewhat independent. However we expect previous training to effect these efforts, which may take time to wear off, thus efforts within a few weeks of each other are likely to be correlated, for this analysis, this was ignored for the purpose of exploring the methods presented in class.

Overview of FTP, TSS^{TM}

The human body uses two methods to covert stored fuel to energy, anaerobic and aerobic respiration. In short anaerobic does not require oxygen and is unsustainable efforts longer than around a minute, aerobic respiration is more sustainable. Presented in detail in Allen and Coggan (2010) is a cyclist's functional threshold power (FTP), which is defined as the power output that can be sustained for around an hour. It is thus named since it aims to represent the power output at which a rider may ride aerobically known as critical power. 1 hour seems a conveniently arbitrary choice of duration, but it seems anecdotally reasonable and has been shown to be good approximation to critical power in Oliveira et al. (2017).

In Allen and Coggan (2010), the training effect of any given ride is quantified through the ride's training stress score (TSS^{TM}), for which the concept of Normalized power (NP) is required. Since a stop start ride is harder than a steady ride of the same average power (AP), NP attempts to quantify this:

$$NP = \left(\sum_{t=1}^{n} p_t^4\right)^{1/4}$$

where p_t is the power at time t in ride

This is described as the power, which if riding steady at for the same duration of the ride, would produce the same training effect as the given ride e.g. a stop start ride at NP of $220W$ and AP $200W$ for an hour is supposed to give to the same effect as a completely steady ride with NP of $220W$ and AP of $220W$. 100 TSS is then defined, as the training effect of riding at FTP for one hour. Using the normalized power of a ride the TSS is defined as follows:

$$TSS = \frac{t \times \frac{NP}{FTP} \times NP}{FTP \times 3600} \times 100$$

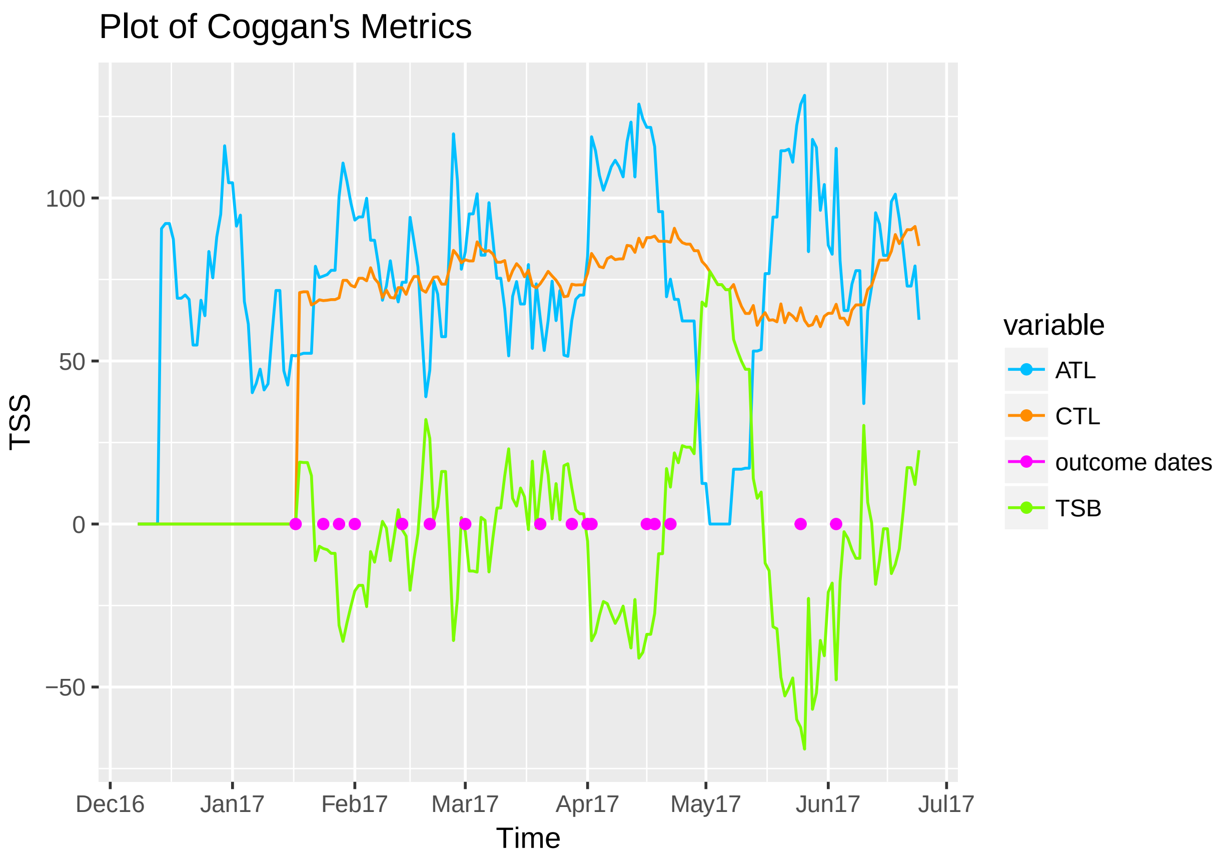
where t is time of ride in seconds

In order to quantify fitness changes over time, and to capture that elusive form, that cyclists aim for on race day. Acute training load (ATL) and chronic training load (CTL) are developed. ATL is the 7 day moving average of TSS, and CTL the 42 day average of TSS. We then define TSB as CTL - ATL. These are often referred to as CTL = fitness, ATL = fatigue, TSB = form. Thus predicting good performance days is done my having high fitness and low fatigue i.e. high form.

It seems much of the above was designed to match up with cycling coaches intuition, to get a number for TSS that feels about right to quantify how hard a ride was. Obviously there is nothing intrinsic about taking the 4th root of the sum of the 4th power. Picking rolling averages, and the 42 days and 7 days,also seems very reasonable; a workout more that 42 days ago seems unlikely to have much effect, but a workout less than 7 days ago is probably fatiguing your muscles somehow. Though these are very arbitrary numbers, TSB may perform well as a predictor on average for athletes, but on average is rarely useful. The question is can we come up with a method of prediction that relies less on physiological knowledge and more on exploiting the structure of training data.

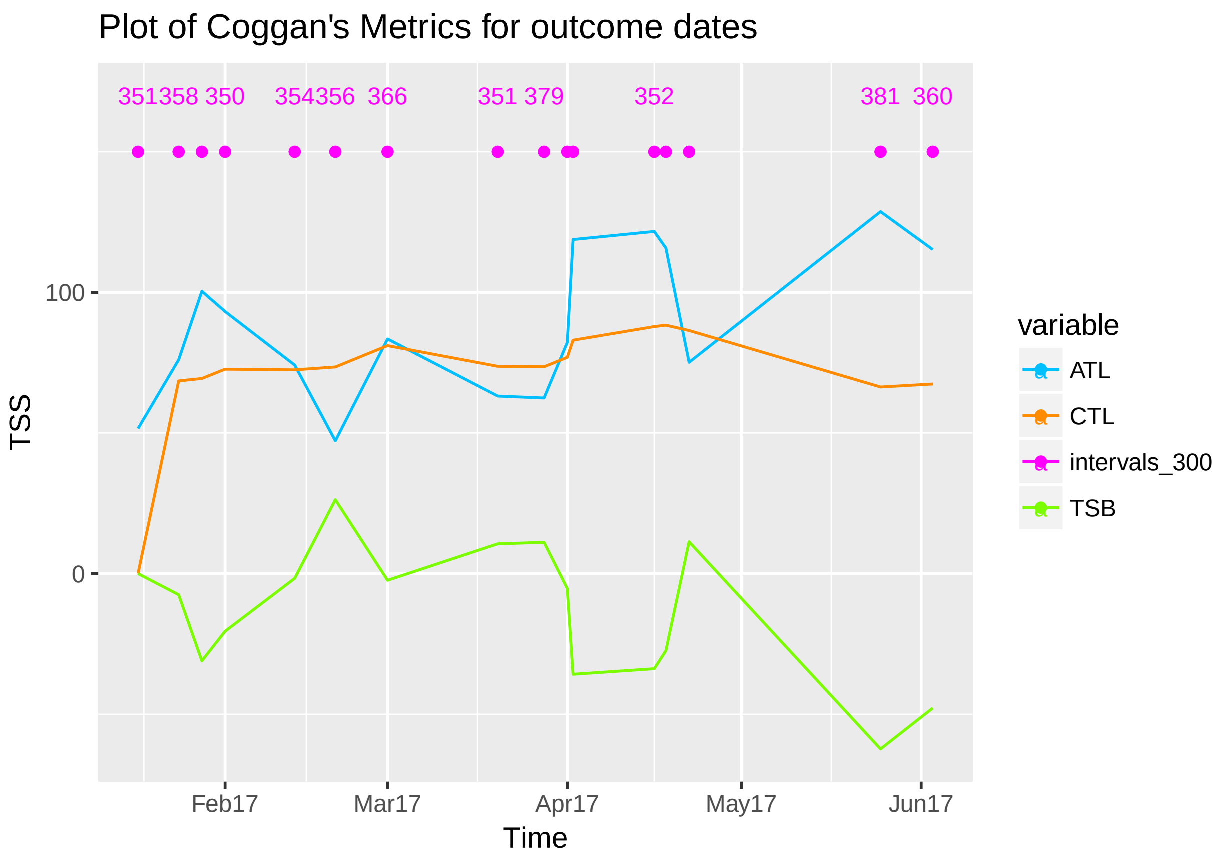
ATL and CTL as predictors

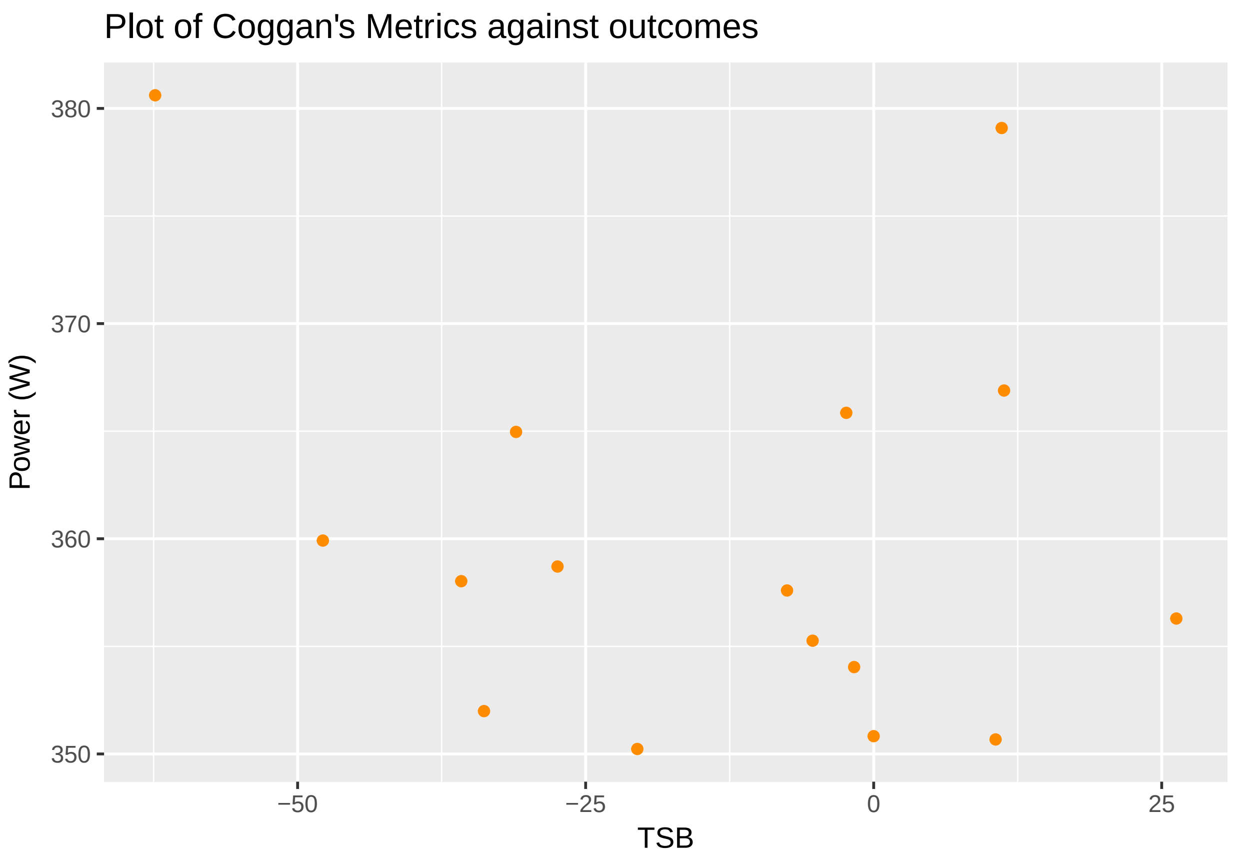
The natural thing to do, in light of the theory in Allen and Coggan (2010) is to use ATL and CTL to predict performance in the 5 minute intervals. The below graph shows the evolution of ATL,CTL and TSB over the period of interest, with the outcome dates represented by the pink dots.



This plot seems ties up with the intuition of fitness, form and fatigue. CTL changes slowly though increases after sustained periods of high ATL, and TSB dips in periods of high ATL. TSB is highest after a period of high ATL followed by a period of low ATL e.g. intense training followed by tapering for an event. In the context of my training and race program, the period from March to the middle of April was the start of racing season, with warmer weather and more intense training, leading to high sustained ATL and a corresponding increase in CTL. I then did not ride for 2 weeks at the end of April and beginning of May resulting in a dip in CTL, and the highest TSB values. I then continued intense training and racing, with June being a particularly heavy training month. Thus ATL and CTL metric seems to tie up with our basic training intuition. However, one should take this with a pinch of salt, since it seems unlikely that some underelying physiological fitness measure declined over 2 weeks of inactivity. However in the context of a trained athelete, where improvements might be small and hard to achieve, a CTL interpretation as “how much recent training has been done” is much more informative than simple fitness. I do not interpret such metrics for untrained athletes that are likely to see very rapid fitness gains as their body adapts to training.

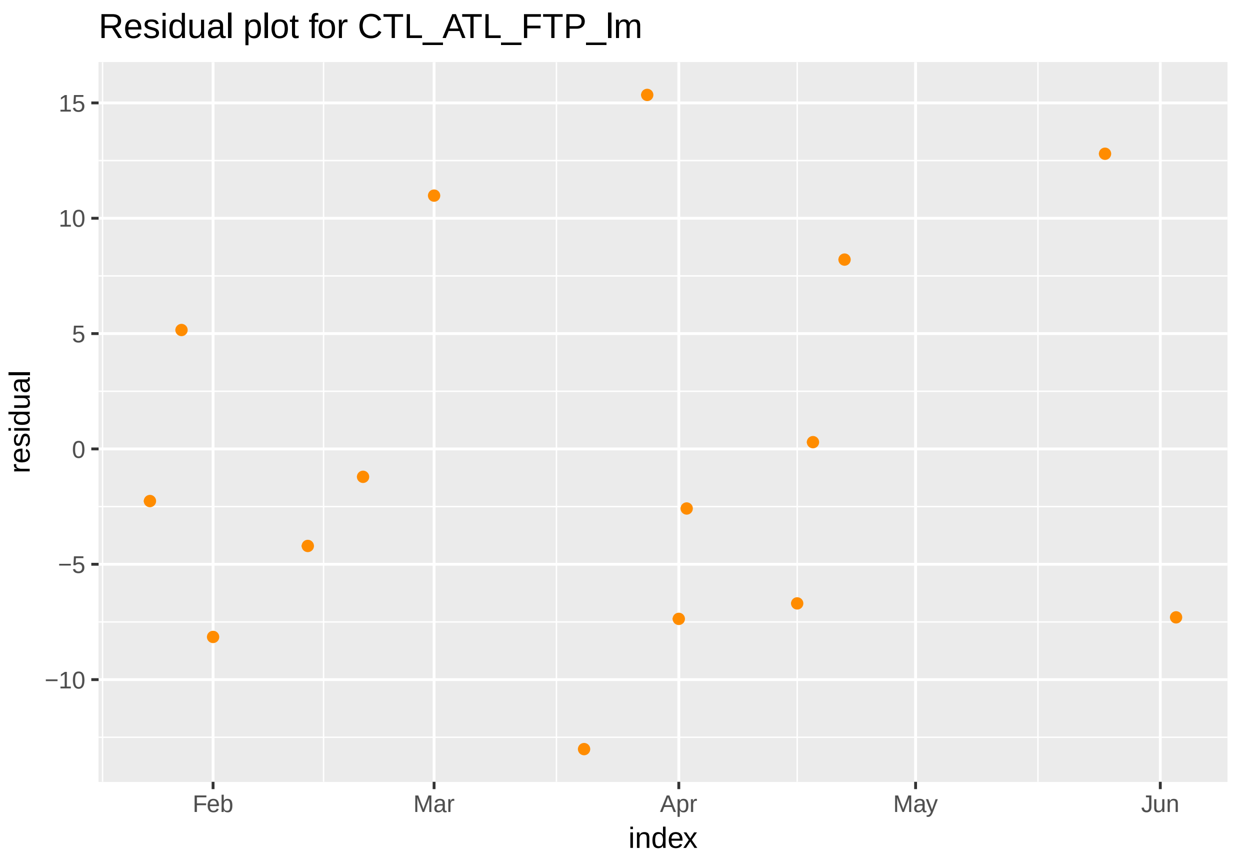
The below plot gives the ATL, CTL, for the dates of the outcome of interest with the wattages on the graph.


Plot of Coggan's Metrics for outcome dates
351358 350
354356 366
351 379
352
381 360
100
0
TSS
Feb17
Mar17
Apr17
May17
Jun17
Time
variable
ATL
CTL
intervals_300
TSB



By eye there seems to be little association between the metrics, so a linear regression should yield insignificant estimates. I fitted a simple linear regression to start with as follows:

```
##                 Estimate  Std. Error    t value   Pr(>|t|)
## (Intercept) 274.49864298 101.1298984  2.7143174 0.02013492
## CTL          -0.40451243   0.3668280 -1.1027305 0.29368545
## ATL           0.01226879   0.1125347  0.1090222 0.91514824
## FTP           0.38012040   0.3592223  1.0581761 0.31265952
```



The results of this model do not really help or tell us anything since all the estimates are non significant, and the estimates have very large standard errors.

I think it is unlikely that if a relationship did exist it would be linear, however if it did exist I would still expect it to show up in the linear model as significant. Although I do not expect to find significant results I also ran a kernelised regression model, in order to reduce any possible misspecification. The average partial derivative for this model were as follows, all were non significant.

```
##     Estimate Std_Error t.value P_value
## ATL   0.0021    0.0105  0.2029  0.8426
## CTL  -0.0083    0.0105 -0.7938  0.4428
## FTP   0.0237    0.0278  0.8510  0.4115
```

Again this does not really tell us anything since the significance levels are so low. Thus for prediting my 5 minute interval efforts the measures fo ATL and CTL are not useful for my small data of this period.

Past TSS and FTPs as predictors

CTL and ATL are rolling averages of TSS, my first approach was to attempt to predict the outcome using a regularised regression model, with the variables equal to lagged TSS, i.e. TSS 1 day ago, 2 days ago, 3 days ago and so on. I also included FTP but not lagged. I ran a ridge regression. This gives 199 variables to use for each observation, though many of these will be 0 i.e on day 63 there will only be 62 + 1 non zero variables to use. I used ridge as my regularisation of choice rather than Lasso, since I have the prior that all coefficients should be uniformly small, rather than the prior that many should be 0, which corresponds to Lasso since Lasso induces sparse solutions.

I chose the regularisation parameter using leave one out cross validation, since my data set was small this did not result in unreasonable run time.

Interpreting the results is hard, since I do not have a test set to test my solutions against, I report a summary of the number of days lagged of the negative and positive ridge coefficients, which may help to understand the effects of long term or short term TSS. Under the null that the training data is not associated with the outcomes, we would expect the distribution of the negative an positive lags to be similar.

For positive coefficients the lag summary is:

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##    1.00   47.00   73.00   77.26   97.00  166.00 
```

For negative coefficients the lag summary is:

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
##    3.00   34.75  111.50   95.33  133.00  177.00       1 
```

The summary measure might suggest that negative coefficients are distributed among days TSS that are further in the past. Though this seems very weak evidence and I do not make any inference from this.

My next idea was to run the ridge regression with the maximum lag set to 42 days since this is the maximium window of dat use in ATL and CTL. This also allows all the outcomes to have equal data, i.e. no observations have many of their variables set to 0.

For positive coefficients the lag summary for 42 days max lag is:

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
##    1.00   10.00   19.00   20.26   31.00   41.00       1 
```

For negative coefficients the lag summary for 42 days max lag is:

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##       2      16      24      23      32      42 
```

Once again perhaps negative coefficients are distributed in the higher lags, though again the distributions are so similar that I do not make any inference here, though I do note that this is the opposite to what might be expected based on the ATL and CTL theory presented by Allen and Coggan (2010).

ANT+ Session Info as predictors

ANT+ fit files, the type generated by the bike computer I used produce a session information summary for each ride. Some of the key variables are as follows:

- average and maximum cadence
- average and maximum power
- average and maximum heart rate
- normalized power
- duration
- calories
- total ascent/descent
- TSS
- time in 7 distinct heart rate zones
- time in 10 distinct power zones

It would seem reasonable that all these combinations of variables describe a ride well, and thus should the outcome i.e. the power output over a 5 minute interval be related to previous rides, we would hope that some combination of these variables would be able to explain this.

My initial idea was to include all variables with all possible lags and then regularise away the overfitting by using heavy ridge penalty term. I selected this penalty term again through leave one out cross validation.

For positive coefficients the lag summary is

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   52.00   90.00   89.08  128.00  176.00
```

For negative coefficients the lag summary is

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   30.00   73.00   79.85  125.00  176.00
```

Once again the lags seem similarly distributed between positive and negative coefficients, so no inference can be drawn.

I considered a summary of the position of each variable in the ranking of size of ridge coefficients, where position 1 is the highest ranked and positive, with the largest number rankings being negative coefficients. I show for the first 10 variables for brevity

```
## $avg_cadence
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     567    1502    2070    4572    9615   10120
##
## $avg_fractional_cadence
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       2      90     488    4807   10731   10891
##
## $avg_heart_rate
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1055    1846    2471    4980    9432    9967
##
## $avg_power
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1220    2061    2602    4896    9228    9941
##
## $avg_speed
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     339     752    1075    4150   10131   10507
##
## $event
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     498    5344    6135    6053    6926   10457
##
## $event_type
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    42.0   187.8   437.5  3834.8 10541.2 10865.0
##
## $first_lap_index
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##
##
## $intensity_factor
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      10     139     408    4244   10656   10885
## 
## $max_cadence
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##     934    1722    2287    4566    9503   10039
```

The coefficients are essentially similarly distributed accross ranks of ridge coefficients, so this model is uninformative. In particular there are a few variables that definitely should not have predicitive power that have skewed distributions e.g. average fractional cadence defined as the non whole number portion of the average cadence variable. The fact that this has a somewhat skewed distribution suggests this model has no meaning.

My next attempt to use this data was to use lagged power zone data in a regularised model to predict the outcome variable. The distribution of a rider's power output often broadly describe a ride, since power describes how hard the body is working. For example a long steady ride may have the same average power as a really hard sprinting session, but the distribution of the power data will be different. Thus this model attempts to capture something that normalised power and resulting TSS also tried to capture,without enforcing a formula not motivated by the data. I had high hopes for this model, since it seems to me if it is at all possible to predict performance, these variables should have sufficient granularity to characterise previous training well.

Once again I look at the distribution of the lags of the negative and positive coefficients.

For positive coefficients the lag summary is

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##    1.00   54.75   91.00   89.68  128.00  167.00
```

For negative coefficients the lag summary is

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##    1.00   29.00   67.00   77.03  122.50  176.00
```

It seems here the negative coefficients are distributed with lower lags - in line with the thinking of TSB.

I also ran this restricting to 42 days max lag which is perhaps a better model since we have this data for all outcomes.

For positive coefficients the lag summary is

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##    1.00   10.00   22.00   21.99   33.00   42.00
```

For negative coefficients the lag summary is

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##    1.00   11.00   20.00   21.13   30.00   41.00
```

Here the distributions look very similar, which makes it hard to evaluate how the timing of training might predict form.

However it is useful to look at the distribution of the coefficients split into different variable classes. Here we look at the position of the coefficients for each variable type, when the coefficients are ranked from most positive to most negative again.

```
## $time_in_power_zone_1
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##    2.00   14.75  388.50  244.90  405.75  422.00 
## 
## $time_in_power_zone_2
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##    77.0   122.5   215.5   194.5   263.5   312.0
```

```
##
## $time_in_power_zone_3
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##    74.0   122.5   188.0   194.5   268.0   329.0 
##
## $time_in_power_zone_4
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##    62.0   120.0   232.5   214.8   277.8   330.0 
##
## $time_in_power_zone_5
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##    57.0   134.2   244.5   219.8   310.0   348.0 
##
## $time_in_power_zone_6
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   43.00   99.25  258.50  221.83  331.75  371.00 
##
## $time_in_power_zone_7
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   29.00   58.75  290.00  227.86  361.00  386.00 
##
## $time_in_power_zone_8
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   41.00   83.75  196.00  206.43  337.00  376.00 
##
## $time_in_power_zone_9
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##     6.0    28.5   183.5   210.0   392.5   421.0
```

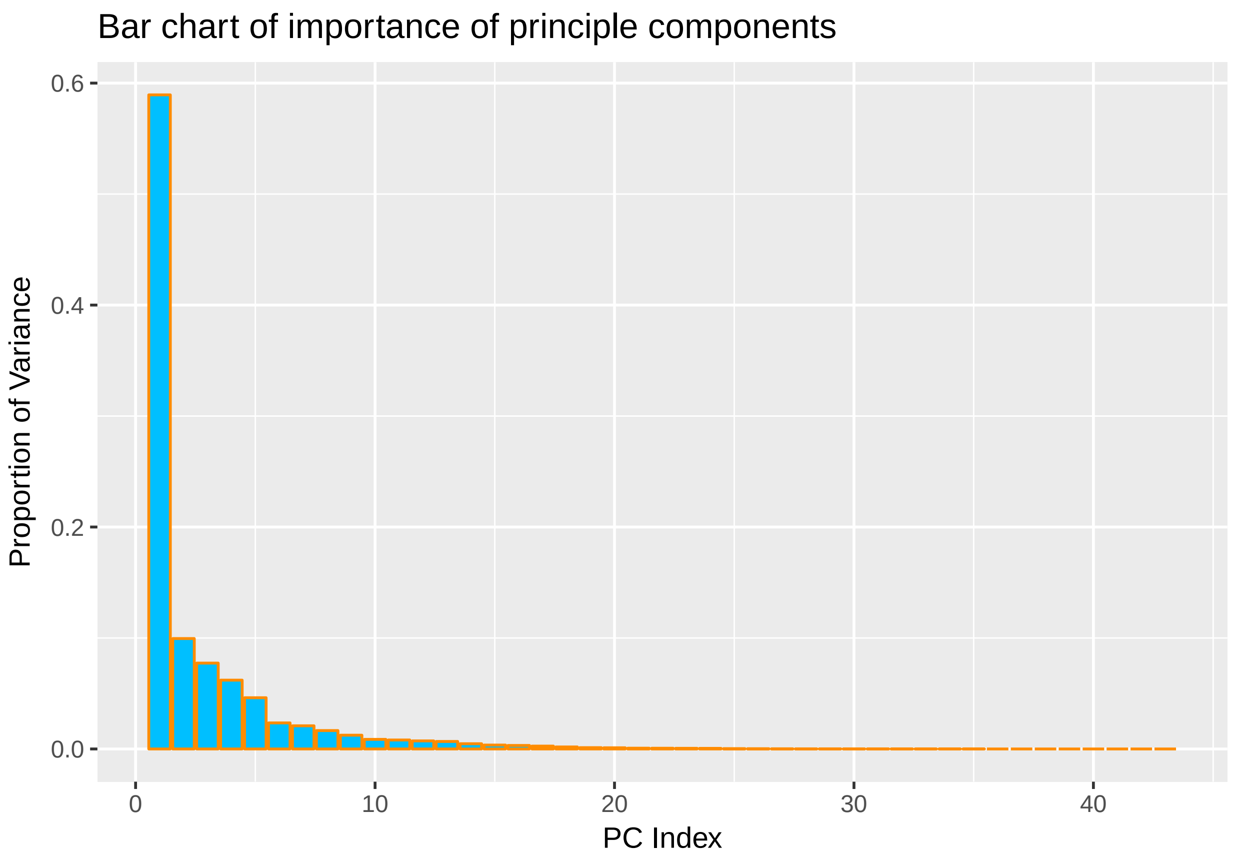
It seems that the higher zones have more of the larger coefficients, however once again this is certainly not enough to make any inferences. And it seems unlikely that such model will generalise well. I also note that due to the high penalty term included, the intercept is very close the outcome means, suggesting that this model would be very bad at predicting out of sample.

PCA dimension reduction for session data

Having concluded that with the small amount of testing data I have it is impossible to come to any conclusion about possible form and fitness effects on the testing intervals. I now turn briefly to try and understand the rides better through the session data. I believe the session data for each ride should explain a ride's characteristics quite well, though it is not obvious what the key components should be. Here I carry out PCA, and plot the results on a biplot to try and interpret the different categories of ride that I go on. I hope to find some sort of distinction between the 4 broad categories of ride that I go on :

- Steady
- Intervals
- Group Ride
- Races

So we run PCA and we get the following variances explained by the PCs explained in the below graph:



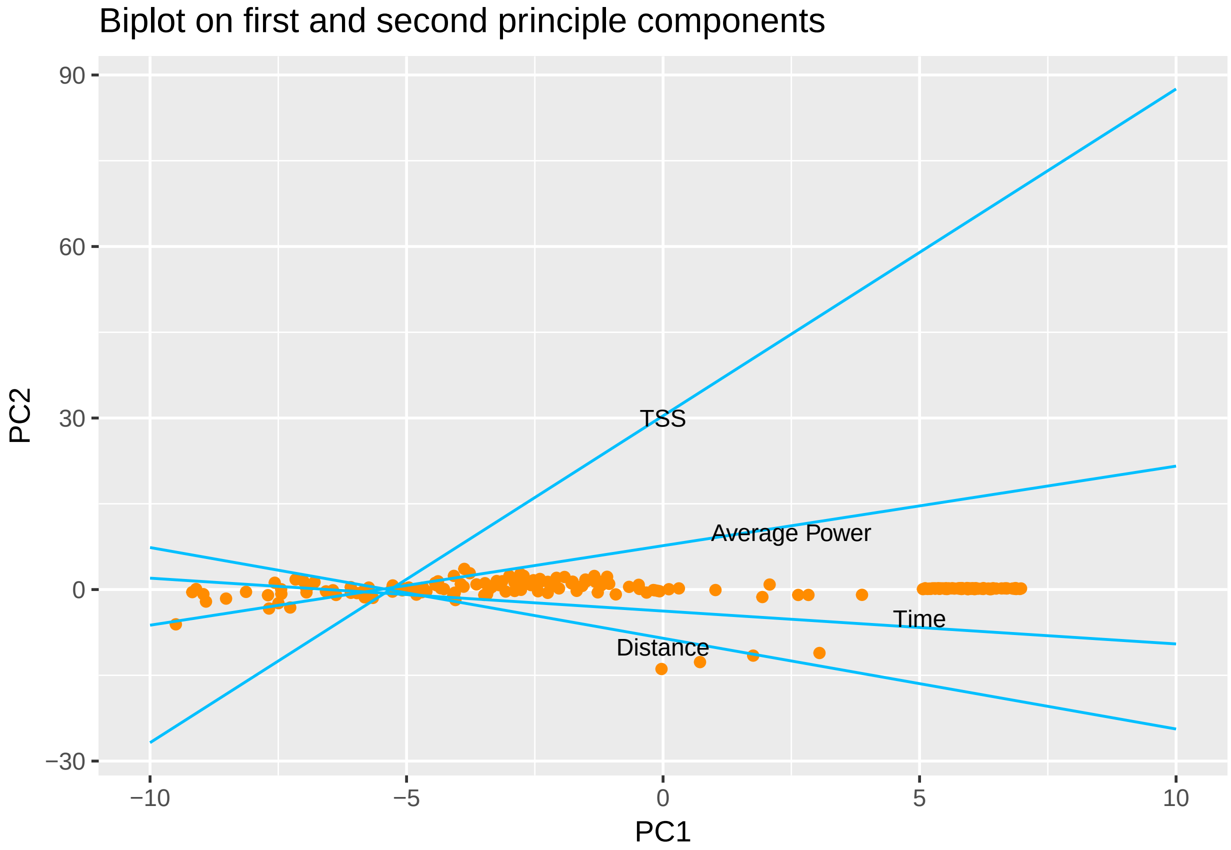
We note that the first principle component explains a large proportion of the variance between rides, with perhaps the next 4 also being useful before a drop off to PCs that do not explain very much variance.

Looking at the first 10 largest components of the PCs is informative, and gives us an idea of the type of variation each PC might be accouting for.

```
##                PC1_names  PC1_coefs           PC2_names   PC2_coefs
## 1             total_work -0.1888724   total_elapsed_time -0.26608813
## 2              max_speed -0.1888262        total_descent -0.25191765
## 3           total_cycles -0.1885463         total_ascent -0.24997975
## 4  training_stress_score -0.1881457     total_timer_time -0.20254650
## 5         total_distance -0.1862947       total_calories -0.14936302
## 6            max_cadence -0.1857137       total_distance -0.11737707
## 7       normalized_power -0.1831253    time_in_hr_zone_1 -0.11173476
## 8              max_power -0.1828096    time_in_hr_zone_2 -0.09464124
## 9         total_calories -0.1827540    time_in_hr_zone_3 -0.08930616
## 10             avg_speed -0.1824610 time_in_power_zone_2 -0.05704766
##                PC3_names   PC3_coefs
## 1   time_in_power_zone_9 -0.33569063
## 2  time_in_power_zone_10 -0.30651523
## 3   time_in_power_zone_8 -0.28477785
## 4      time_in_hr_zone_5 -0.26101653
## 5      time_in_hr_zone_6 -0.22232245
## 6   time_in_power_zone_7 -0.16660600
## 7   time_in_power_zone_6 -0.10125280
## 8      time_in_hr_zone_7 -0.09347499
## 9           total_ascent -0.07948186
## 10         total_descent -0.07433315
```

Intuitively PC1 seems to account for how easy a ride was in some sense, with large negative coeffiecients for coefficients denoting a hard,long or fast ride. PC2 accounts for hard rides, with large negative coefficients on time and low intensity power and heart rate zones. PC3 accounts for hard efforts within rides, with almost all large negative emphasis on high heart rate or power zones, this likley accounts for hard interval session or more likely races. Note that PC4 and beyond start to have much less clear interpretations so are not included here.

The following biplots help visualise the above principle component analysis. Note that directions are somewhat arbitrary in the below discussion, and the below biplots directions are not necessarily orrganised in the most intuitive manner.

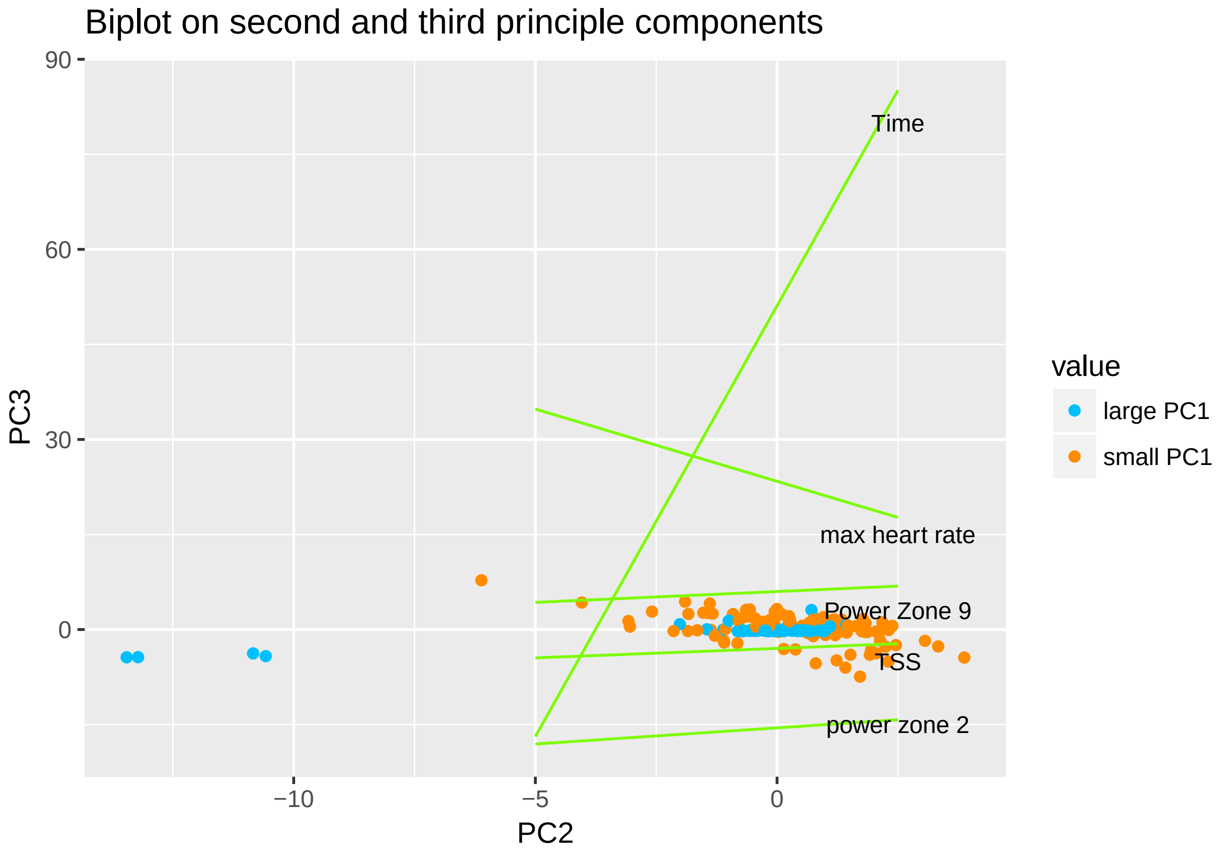


In the biplot PC1 accounts for almost all of the time elapsed and normalised power since they are almost parallel to the PC1 axis. Since time and average power could be considered the two key components of how hard a ride is this lends weight to the interpretation of PC1 accounting for some measure of how hard or easy and how long or short the ride was. Note that logn rides are more likely to have low normalised power and thus the noramlised power line and the time line in the above are in some sense in opposite directions.

Interestingly TSS seems to have a significant component in both PC1 and PC2, since TSS is the established metric for the training effect of a ride, perhaps PC1 cannot be interpeted as simply how hard a ride was but perhaps even how hard a ride was given that it was steady and low intensity, though this may be step too far. The fact that TSS has a significant component in both directions matches up with our intuition that training effect can be gotten through either long steady rides or shorter more intense rides.

We also note that there is a clear separation group on the right hand side, with high PC1 component, I believe these are likely to be group rides which tend to be low intensity and long and often involve a stop for coffee and cake or such like, which increases the time of the ride, without increasing the intensity at all.

I also plotted the second and third PCA biplot, with the colour of the points separating the points on whether they were above or below the mean value for components in the direction of PC1.



First we note the separation of the values on the left, this are likely the same 4 showing up on the first biplot, on the right on the distance line. I suspect these are long solo rides, since they do not have as large a time element like the group rides but do still have a large distance component, since they are often steady rides their TSS is low, and they have low PC3 components suggesting that PC3 may be a measure of how uneven the effort was on the ride.

Next we look at directions that have component in the PC3 direction, i.e. time and max heart rate. Since its hard to maintain a ride with high variance in power output for long, high time gives high PC3, and likewise a ride that has high variance in power output, is likely to be a race or interval session so is more likely to have a high maxmimum heartrate. So in this biplot we would interpret low value of PC3 being rides with high variance in power output, e.g. rides with many accerlartions of sprints.

Conclusions

Overall, generating a model that has coefficents that match up to any kind of prior intuition seems like a hard problem. I suspect that the outcome data in this case is just too noisy, with too few samples, to be able to predict performance with any degree of accuracy.

Performing dimension reduction on ride summaries however did produce interpretable results in terms of the key components characterising the variation between rides. This aided understanding and produced some nice plots, but does not aid understanding of the effect of training on the ability to produce high powered efforts.

Further Work

I believe there is a lot more scope for quantatative evaluation of cycling power data, however the key problem that I for see is lack of testing data. Here I attempted to use all out 5 minute efforts spaced through my cycling season, which did not give a useful model. The problem is, cyclists rarely perform the same test on a regular basis, so it is hard to track fitness and form and its effect on these efforts.

However with a large number of cyclists each performing all out efforts relatively frequently it might be possible to come up with a better model than TSB for predicting form. A good such source of data could be indoor training app data. Such apps prescribe structured workouts that users perform. With such data, as well as the users general riding data, variance of estimators would be greatly reduced making for better prediction.

On a different tack, as shown in chart in the data section, analysing this data spatially could also be interesting. Particularly when racing in flat stage, the last few minutes can be very hectic with teams positioning their sprinters in the best place to sprint for the win. Ultimately it would be exciting to see, if given the data from a World Tour sprint leadout train, we could optimise the timing and positioning of such a leadout to improve race results. Though I suspect this would get very complex very quickly as there is a network of such riders and teams all trying to do the same thing. This would be interesting to explore.

In general I would like to do a more thorough literature review of the area, since as the end of the quarter approached I was not able to read around as much as I would have liked.

There could be further work with analysing the ride summaries, since I only looked into using PCA to understand the components of different rides. With the possibility of classifying rides into certain types. In hindsight, this may have been a better avenue to explore with a view to using the methods described in the course, however, this was not my primary interest at the start of this project.

References

Allen, Hunter, and Andrew Coggan. 2010. *Training and Racing with a Power Meter.* VeloPress.

Oliveira, Luana F. de, Guilherme Yamaguchi1, Victor de S. Painelli1, Rafael P da Silva1, Lívia S. Gonçalves1, Bruno Gualano1, and Bryan Saunders1. 2017. "Comprehensive Reliability Analysis of a Workbased (420 kJ) Cycling Time-Trial in Recreationally-Trained Individuals." *Journal of Science and Cycling* 6 (1): 11–17.