Evaluation of bike ride metrics using linear regression and mle analysis

Malcolm White
3/27/2014

Overview

Goal:

Measure fitness / evaluate progress

Measure ride intensity

Challenges:

Existing models assume steady state, but data is transient

Data is noisy / messy (gaps in hr data, e.g.)

Tools:

statsmodels, pandas, scipy, numpy, matplotlib

Cleaning data (pandas and numpy)

Outliers:

Applied winsorization at 98th power percentile

Missing data:

Filled missing HR values with valid neighbor

Misleading data:

Clipped empty values at beginning / end of ride

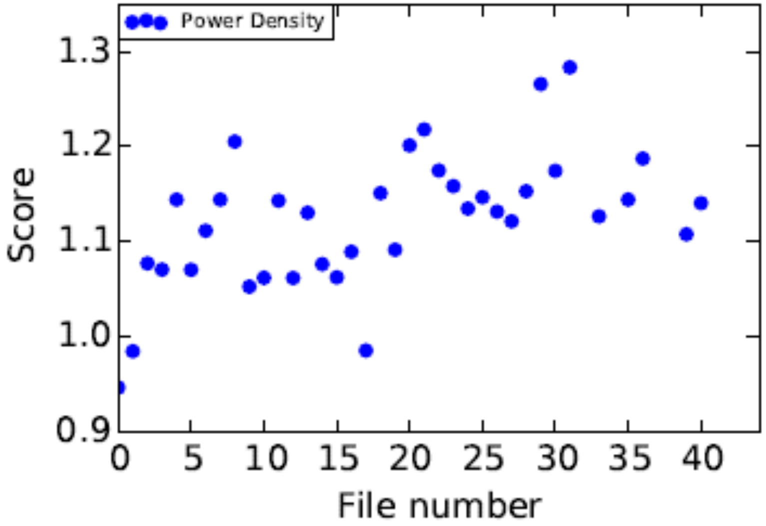
Noise:

Compressed data into series of local means

Fitness Parameter 1 (Power density)

Amount of power generated by number of heartbeats during “active state,” where “active state” is any time-shot with power > 0

$$\text{Density Parameter} = \frac{\text{Mean power}}{\text{Mean HR}}$$



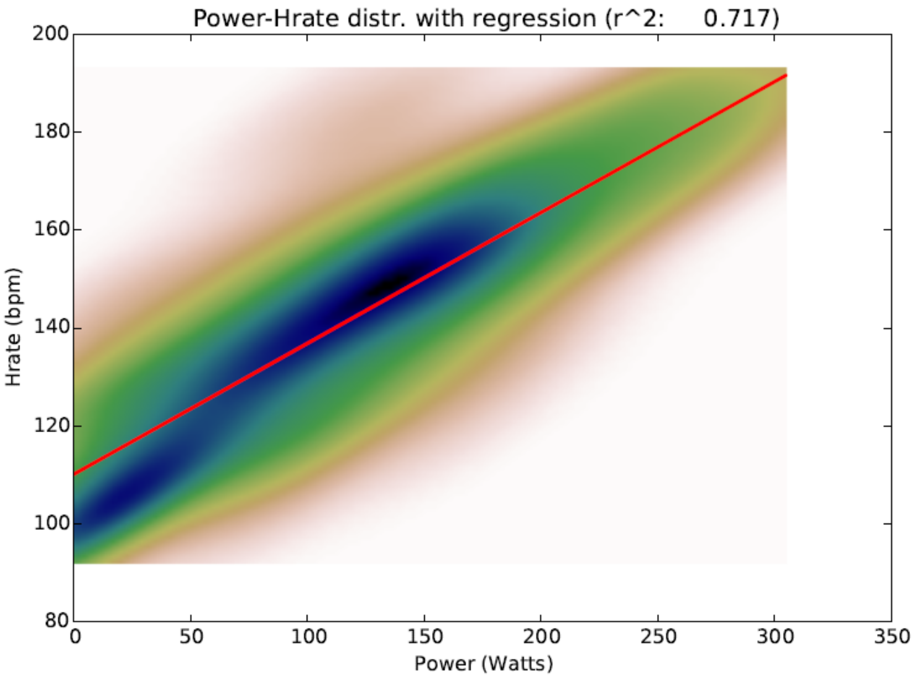
Regression for Fitness Parameter 2

Shifted HR forward in time

Put HR, power in 75 sec. time-boxes

Used mean of boxes for ols regression

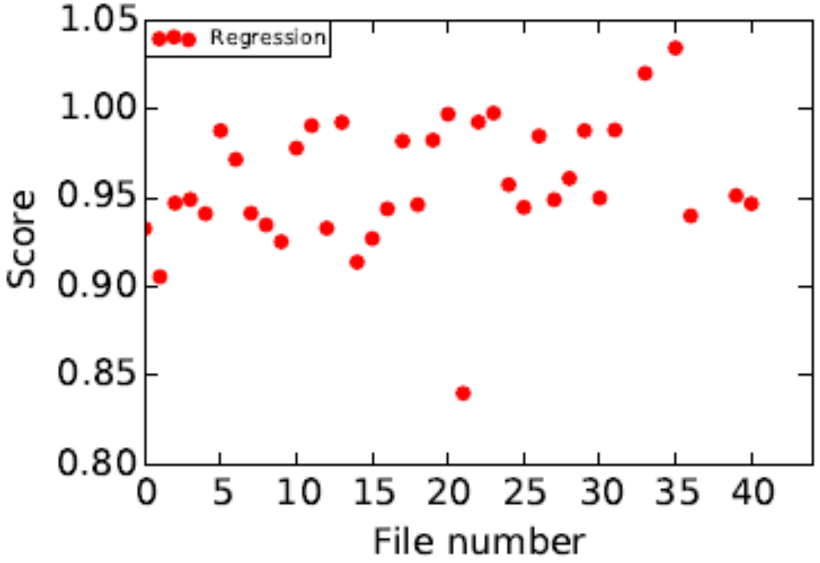
Mapped density with gaussian kde



Fitness Parameter 2 (Predicted HR)

Parameter defined by a scaling term divided by the heart-rate at a given power (pwr) as predicted by the HR-power regression

$$\text{Regr. Parameter} = \frac{150}{m*pwr + b}$$



MLE for Fitness Parameter 3

Originally considered ARX model

- eg: *HR[t] = (a1*HR[t-1] + a2*HR[t-2]) + b*Power[t] + error*

Provided good fit, but results not immediately useful

- For a given t, most of HR[t] was decided by HR[t-1]
  - Low b parameter, a1 parameter close to 1

If possible, best case only considers Power-HR relationship

- 0th-order ARX (essentially a mle fit)
- *HR[t] = b*Power[t] + error*

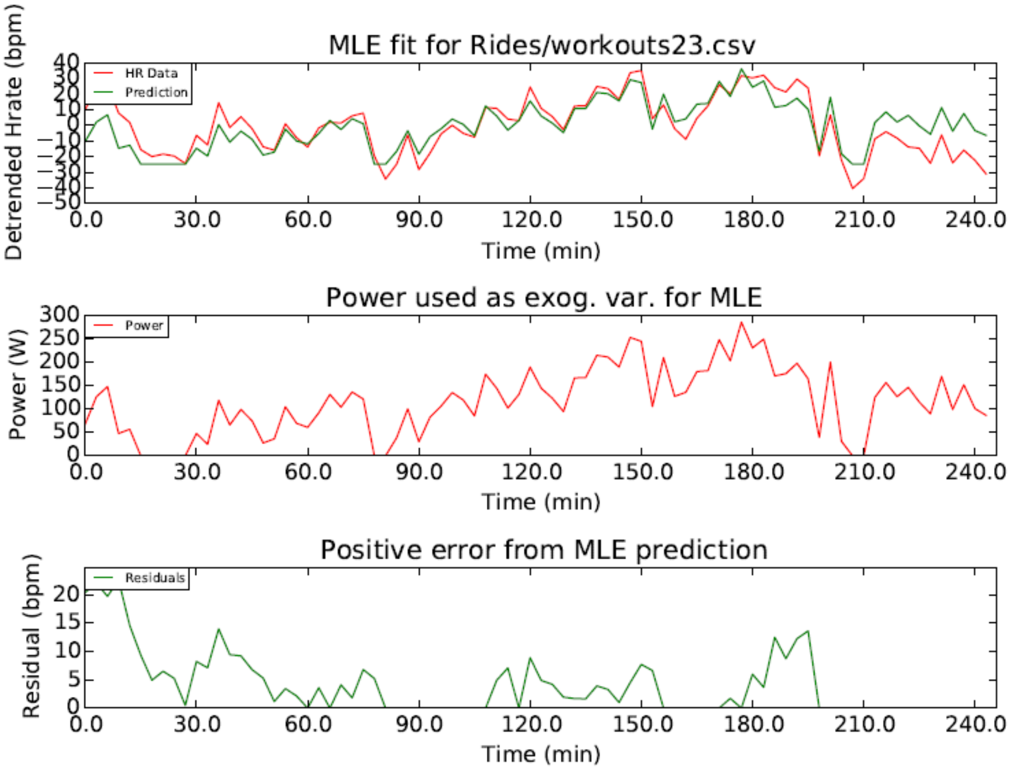
MLE for Fitness Parameter 3

Trying to simulate steady state at each index

Used 3 min time-boxes to erase “HR memory”

Use mle to fit:

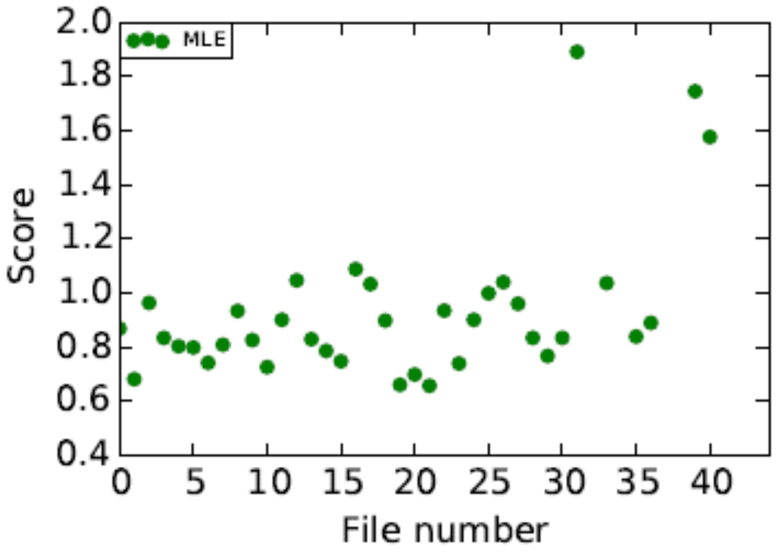
[HR] = [Power] * coeff



Fitness Parameter 3 (MLE fit)

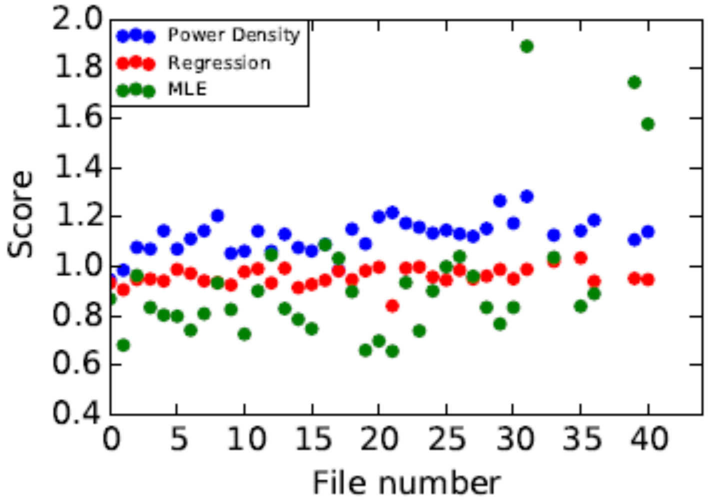
Scaling term divided by the fitted coefficient for mle fit of HR-power in time

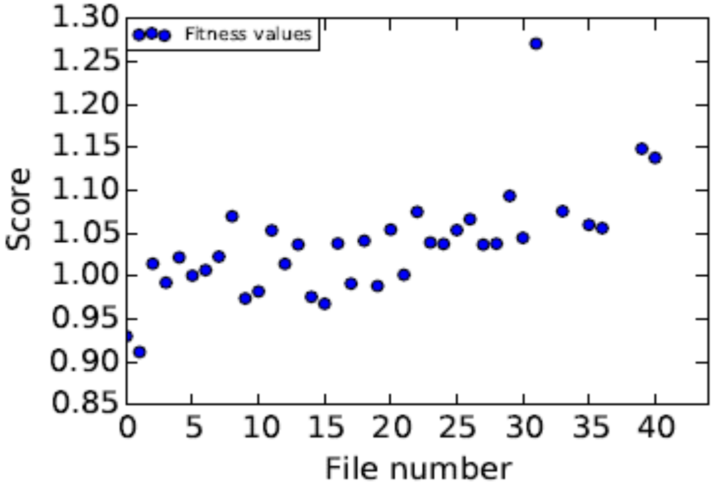
$$\text{MLE Parameter} = \frac{0.2}{\text{MLE fit coefficient}}$$



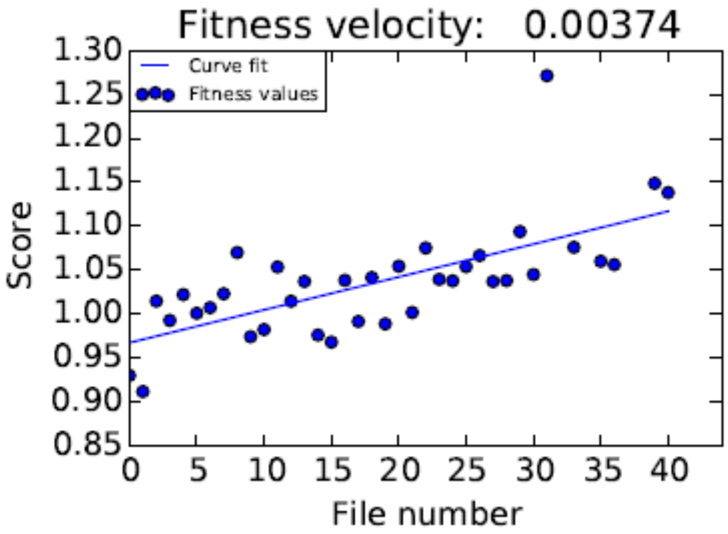
Fitness Parameter Summary

Fitness Score = 0.5*Density + 0.35*Regression + 0.15*MLE

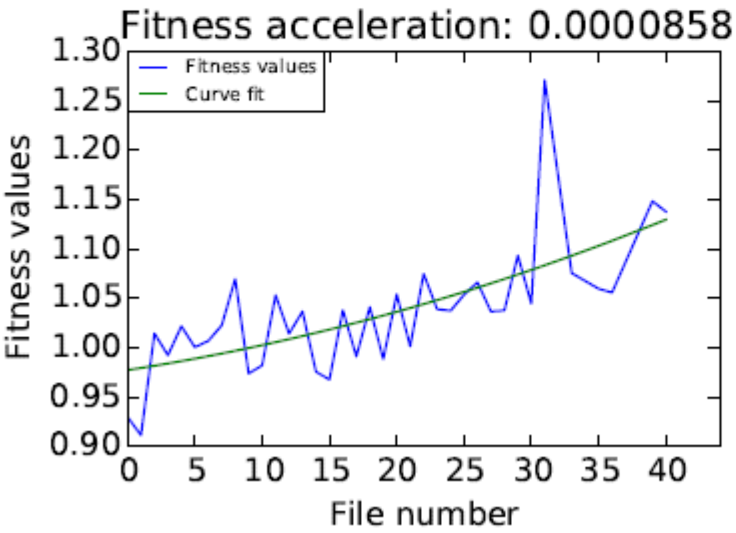




Fitness Parameter Analysis



Fitness velocity measures progress rate



Acceleration suggests diminishing / accelerating returns

Ride Intensity Analysis (preliminary)

Balancing workout intensity is important to prevent under / over training

Intensity may be correlated to super-linear HR responses to power impulse

Error from fit of linear model to hr-power time series may be informative

MLE used for current model, considering switch to ols