Project #1 Report

Andrew Muth, Jerry Zikun Chen

25/11/2019

Summary

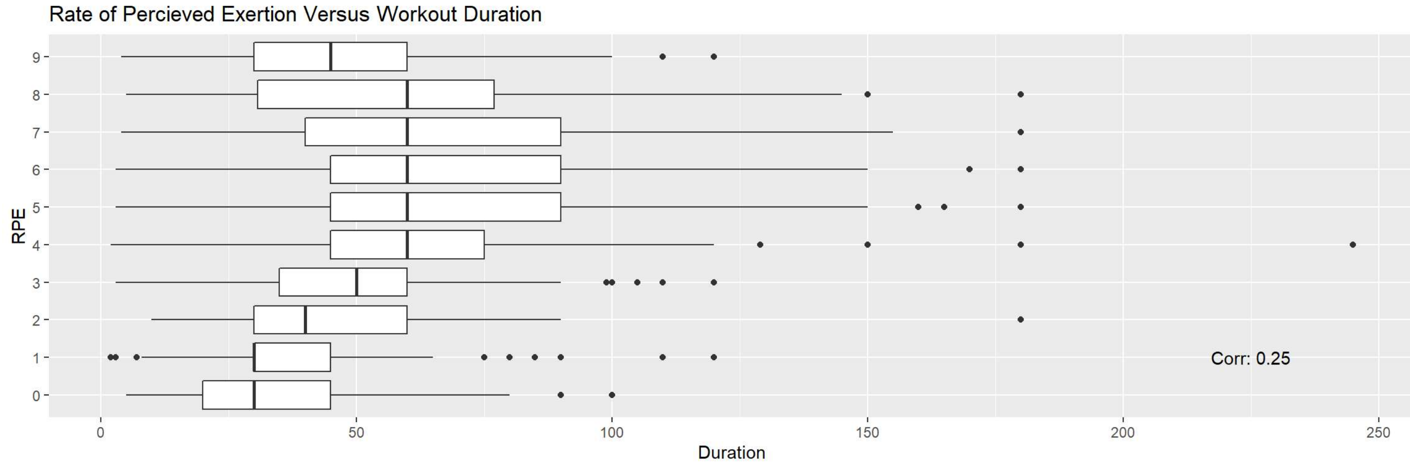
As part of their daily routine the Canadian Women's Rugby team answers a daily survey concerning their overall wellness. This survey consists primarily of ordinal and binary response data concerning their mental and physical well being as well as their sleeping and nutritional habits. The women also complete a survey after each training session. As part of a broader change in high performance sport, Rugby Canada is interested in whether this data can be used to improve player performance.

We consider the two datasets separately as there was inadequate data regarding the training load that athletes carry into competition to conduct a meaningful analysis. Instead we offer what we believe is an improvement over current practices in measuring session load. Our measure of session load is a pseudo heart rate-based measure which can be implemented immediately using only the data currently being collected. It is also immediately extendable if we were to gain access to the athletes' heart rate data.

The wellness survey is conducted both in and out of competition and thus there was enough data to study the effects of athlete wellness on in game performance. We the extend this analysis to offer interpretable measures of athlete well-being which can be used on a daily business.

Measuring Training Load

After each training session the duration of the session is recorded, and the athlete is asked to give a rate of perceived exertion of the training session on an ordinal scale from 0-9. The session load is then calculated as the RPE multiplied by the duration of the session. We consider this to be a crude measure of session load for two reasons. The first is that RPE is both a subjective and ordinal score and as such may differ across athletes and the calculation of session load implicitly assumes that RPE increases linearly with the underlying exertion of the athlete. Our second object is that we are unconvinced that athletes either do or are able to accurately account for the length of the training session when reporting RPE. Consider the following hypothetical: training session 1 is a five-minute run and training session 2 is a 90 min run. If the athlete is told to run both workouts as hard, they can they are likely to rate both sessions as a 9 RPE. However, if we were able to measure the athlete's heart rate for both sessions, we would see that the athlete's average heart rate for the five minute run was higher than for the 90 minute run and as such a more objective measure of exertion (here heart rate) is able to distinguish between the two sessions whereas RPE is not. We believe the graph below shows that our concerns are justified.



If athletes were able to accurately account for the duration of the training session in reporting RPE we would see the mean duration of the workout decrease as RPE increased. Additionally, while RPE is ordinal we would expect to see a negative correlation.

There already exists a broad literature on different methods for measuring workout intensity. A good overview of some of the more popular methods can be found in "Rationale and resources for teaching mathematical modeling of athletic training and performance" by David C. Clarke and Philip F. Skiba. We would also like to extend a thank you to Ming for sharing this resource with us.

While we did not have access to it, we know that Rugby Canada collects heart rate data for the athletes under consideration. We therefore, sought to create a pseudo heart rate-based measure of workout intensity. To do this we looked to combine summated heart rate zone scores with Eric Banister's training impulse (TRIMP). That is, based on the duration of the training session and the RPE we looked to use summated heart rate zone scores to estimate the athlete's mean heart rate for the session which we then use along with the session duration to get an estimate of the TRIMP of the session.

For those unfamiliar Eric Banister's training impulse is given by,

$$TRIMP = t \times \kappa \times FHHR$$

$$FHHR = \frac{HR_{average} - HR_{rest}}{HR_{max} - HR_{rest}}$$

where t is the session duration in minutes, $FHHR$ is the fraction of heart rate reserve and $\kappa = 0.64e^{1.92 \times FHHR}$ for men and $\kappa = 0.86e^{1.67 \times FHHR}$ for women.

There are many variants of the summated heart rate zone scores. However, we chose to adopt the one used by Rowing Canada Aviron. Both because of our familiarity with it and because in addition to heart rates Rowing Canada Aviron's system also offers training session durations. It should also be noted that the physiology between rugby players and rowers is similar.

THE SYSTEM OF TRAINING INTENSITY

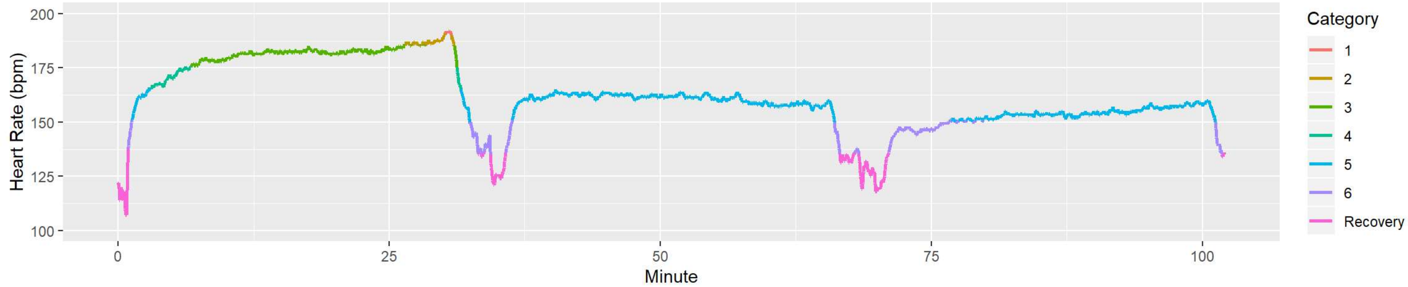
INTENSITY CATEGORY	APPROX HEART RATE RANGE	DURATION ONE PEICE (MIN)	RATIO WORK:REST	GOALS OF THE TRAINING INTENSITY	PRACTICAL EXAMPLES (SR = STROKE RATE)	LACTATE LEVEL (mmol/l)
I	Max H.R	0.5 - 1.5	1:4 - 1:5	-ANAEROBIC CAPACITY -TRANSPORTATION=DEVELOPMENT OF CARDIOPULMINARY SYSTEM -ABILITY+FEELING AF START	-1 - 6 X 500M (WITH START) - INTERVAL TRAINING (SHORT PEICES) SERIES OF 30 - 60 STROKES OR SERIES OF 1 - 2 MIN SR: > RACE SR	>10
II	Max H.R	2 - 7	1:2 - 1:3	-RACE ENDURANCE -TRANSPORTATION=DEVELOPMENT OF CARDIOPULMINARY SYSTEM -RACE SPEED FEELING	-RACE OVER 1500 - 2000 M - 6 X 2 MIN - 3 X 1000 M - 5 X 750 M SR: RACE- SR	8-14
III	Max H.R	6 - 10	2:1 - 1:2	-DEVELOPMENT OF AEROBIC CAPACITY -STRENGHT ENDURANCE -TACTICS -TECHNIQUE	-4 X 7 MIN -3 X 2000M CONSTANT SPEED -5 X 5 MIN STRENGTH - ENDURANCE WATER	5-8
IV	165-175	10 - 45	4:1	ANAEROBIC THRESHOLD -DEVELOPMENT OF AEROBIC CAPACITY -EFFICIENCY -STRENGTH ENDURANCE	-2 X 20 MIN WITH SR- CHANGE -3 X 5 KM TIME - CONTROL -10 KM HEAD - RACE - 3 X 12 MIN STRENGTH ENDURANCE WATER	~4
V	150-165	30-90	-	-BASIC ENDURANCE -UTILIZATION OF AEROBIC -CAPACITY -MAINTENANCE -TECHNIQUE	30-90 MIN STEADY STATE SR: 10 - 12 LESS THAN RACE - SR	~3
VI	135-150	>45	-	-UTILIZATION OF AEROBIC -CAPACITY -REGENERATION -MAINTENANCE TECHNIQUE	-45 - 120 MIN STEADY STATE SR: 18-24 / MIN	<2

(FROM: V. NOLTE / A.MORROW / B.RICHARDSON / A.ROAF)

Later we will need heart rate intervals for categories 3, 2, and 1. We chose to partition the interval (175, 200] as (175, 185], (185, 190], and (190, 200]. This choice was subjective and based on past experience working with this system. For those unfamiliar with heart rate data or heart rate training zones we offer the following examples of it.

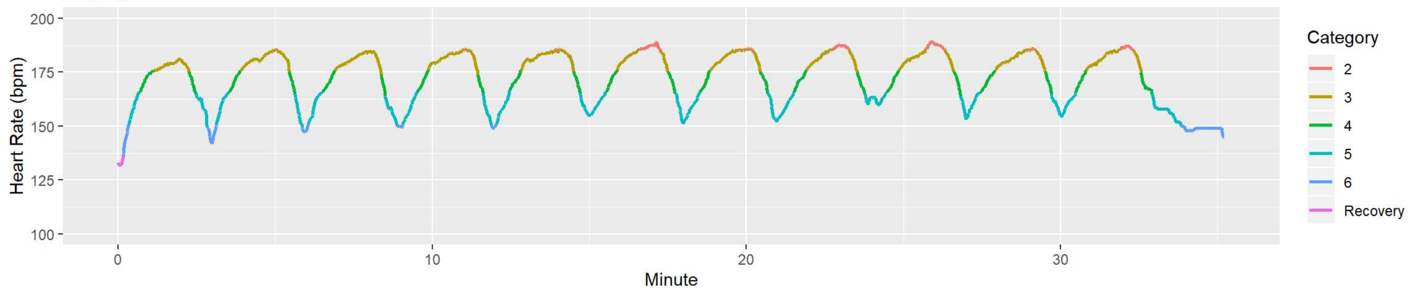
3x30min Ergometer

Training Load: 198.53



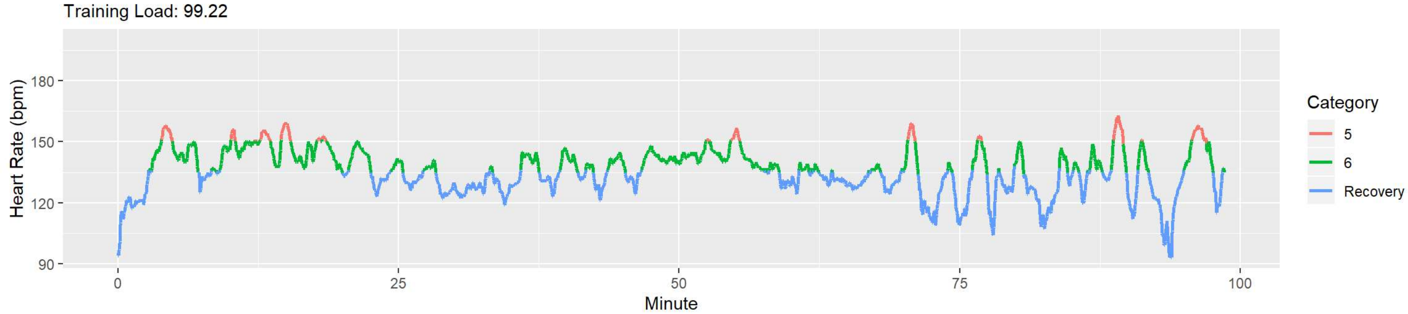
Intervals Ergometer

Training Load: 85.94

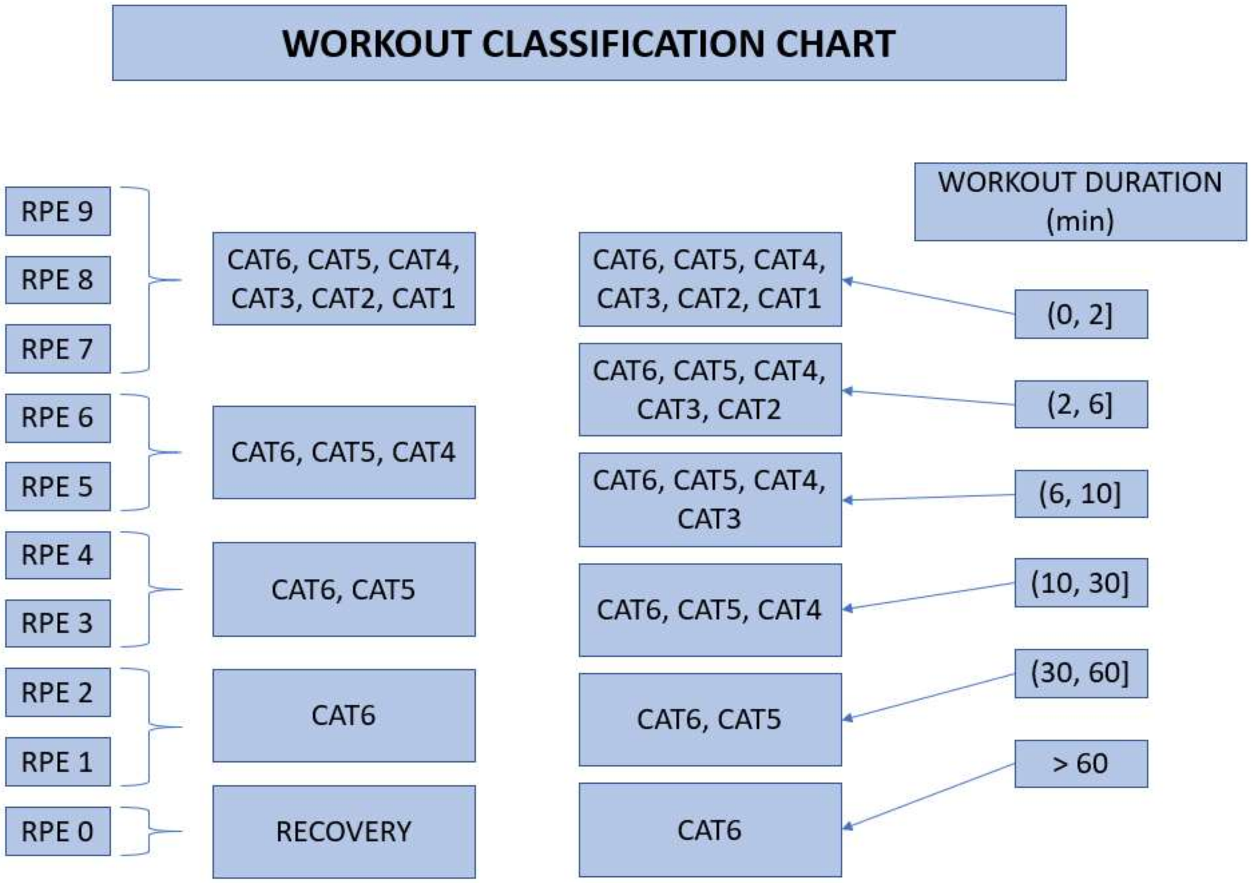


50km Bike

Training Load: 99.22



Using the heart rate training zones and session durations used by Rowing Canada Aviron we came up with the following classification chart for the possible training categories of the each recorded workout based on RPE and duration.



We then classified the session as the intersection of the possible training categories for both the RPE and duration of the sessions. The average heart rate of the session was then estimated as the average of the median heart rate for each training category, and the load of the session was calculated according to TRIMP.

Training load has a cumulative effect on the body and hence physical performance as well. Thus to meaningfully study the effects of training load on athletic performance it is necessary to have several days (or even a week) of reported session load leading up to a game or physical test Unfortunately, the RPE data leading up to the tournaments was not complete enough to conduct any meaningful analysis as to how our measure of session load compares to the measure currently used by Rugby Canada. However, as extensive work has been done on constructing heart rate-based measures of session load we believe that our measure of session load is likely an improvement on the measure currently used. We suggest that our method be adopted alongside the current measure of session load so that a meaningful analysis of the two measures may conducted once there is sufficient data.

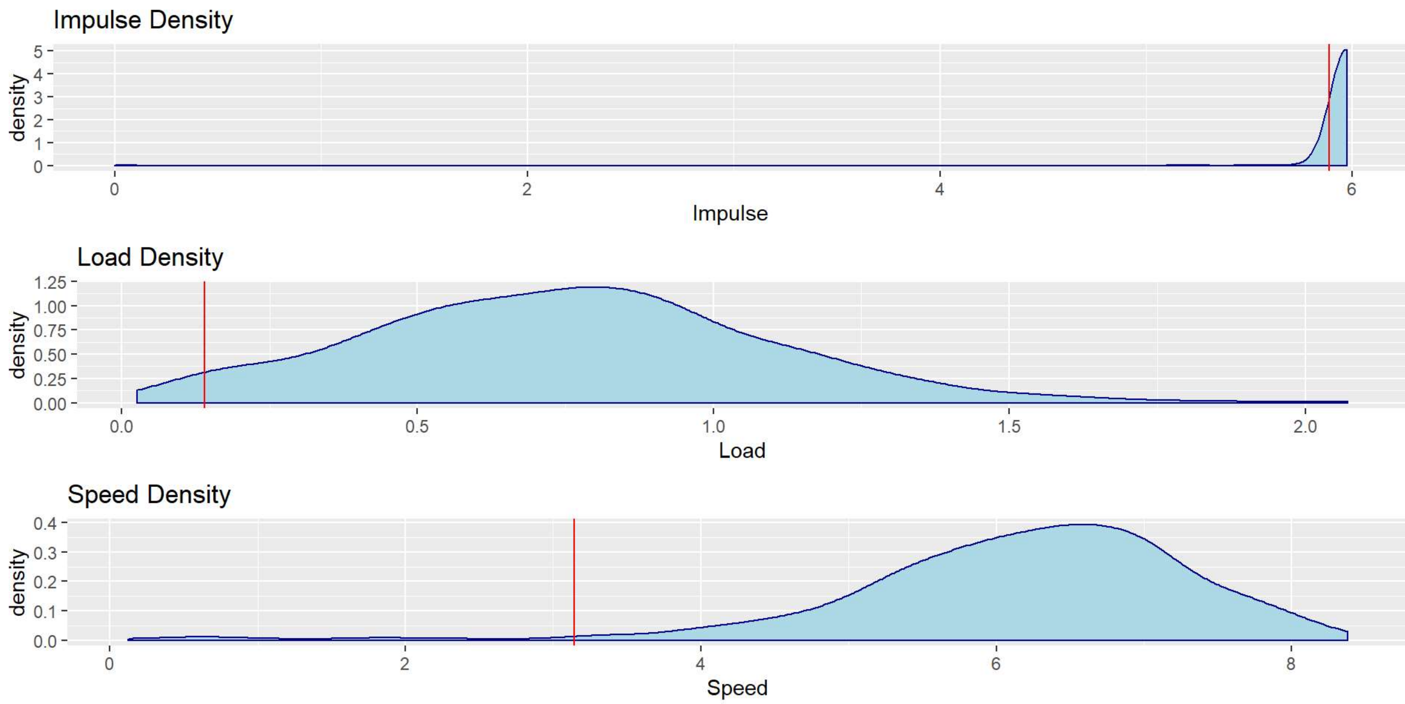
Measuring Wellness

The daily wellness survey that athletes consists of ordinal and binary data covering mental and physical health, nutrition, and sleep habits. Prior to any analysis it was decided to “normalize” ordinal responses for each athlete by treating the data as if it were interval data and then by dividing the response for each athlete by that athlete’s mean response for that question. This allows for more general comparisons of response data across athletes.

In constructing our measures of wellness, we wanted the resulting measure to be a function solely of the daily wellness survey so that it could be incorporated into daily training decisions. We also wanted the resulting measure to be easily interpretable by the training staff so that it may realistically be implemented.

Women's Rugby Canada has their athletes complete a physical battery of tests on a regular basis. This battery consists of a 40-yard dash as well as several other tests to measure the athlete's explosiveness. In investigating the impact of an athlete's wellness on their ability to perform we aimed to create metrics from the GPS data which could serve as proxies for this regular testing. To this end we chose to look at the impact of an athlete's wellness on their maximum observed acceleration load, acceleration impulse, and average speed over five seconds in each game (as a stand in for a 40-yard dash).

To avoid the case where an athlete only briefly entered the game and was not very active, we only consider cases which exceed the second quantile for each metric under consideration. The need to omit some of the game data is made apparent by the graph below.



To estimate the effect of the various wellness variables on the three-performance metrics we fit a random intercepts model to account for the variation among players. It should be noted however, that the data is of a more complex hierarchical structure in that observations at the individual level occur within games which in turn occur within a tournament. Unfortunately, fitting such a model was not possible due to limitations with the data. One potential workaround to this problem would be to include variables for the number of games played on a given day or during a tournament. We elected not to include such variables as the resulting model would not be able to be used to offer measures of wellness outside of competition. The resulting wellness scores are then just the athlete's predicted performance on these tests based on their responses to the morning wellness survey.

We also chose not to fit the model with an intercept. In omitting an intercept from the model, the random effects for each player may be thought of that player's "baseline" score, rather than how each player varies around a constant mean (it also increases the size of the random effects). This increases model interpretability among the training staff.

Before going over the resulting measures of wellness we look at the random and fixed effects of the 40-yard dash model. Note that p-values for mixed effects models do not make sense in the classical sense. The p-values presented below are obtained via Satterthwaite's method and implementation is done through the "lmerTest" package.

Fixed Effects for Speed Test

	Estimate	Std. Error	df	t value	Pr(>|t|)
Fatigue	0.16	0.36	188.46	0.46	0.65
Soreness	-0.37	0.28	192.33	-1.30	0.20
Desire	0.38	0.24	193.48	1.59	0.11
Irritability	0.05	0.48	192.97	0.10	0.92
BedTime	-0.01	0.02	188.39	-0.62	0.54
SleepHours	0.49	0.07	136.01	7.33	0.00
SleepQuality	-0.54	0.35	183.47	-1.56	0.12
Pain	-0.16	0.21	189.31	-0.77	0.44
Illness	-0.27	0.26	193.17	-1.03	0.30
Menstruation	0.04	0.23	182.40	0.18	0.86
Nutrition	1.76	0.48	143.28	3.63	0.00
NutritionAdjustment	-0.39	0.37	123.91	-1.05	0.29
USG	0.14	0.36	194.48	0.38	0.71
TrainingReadiness	0.69	0.57	22.94	1.21	0.24

As we can see most of the variables affect speed in the way that one would expect. For example, nutrition and soreness work to increase and decreases speed respectively. Others such as fatigue seem to effect performance opposite to what one would expect, however fatigue does have a large standard error relative to its point estimate. Indeed, it seems as though the model can likely be further reduced. We opted not to reduce any of the mixed effect models on grounds that that should be done with someone with domain expertise to avoid removing a variable which appears insignificant, but for which there is a plausible causal relationship.

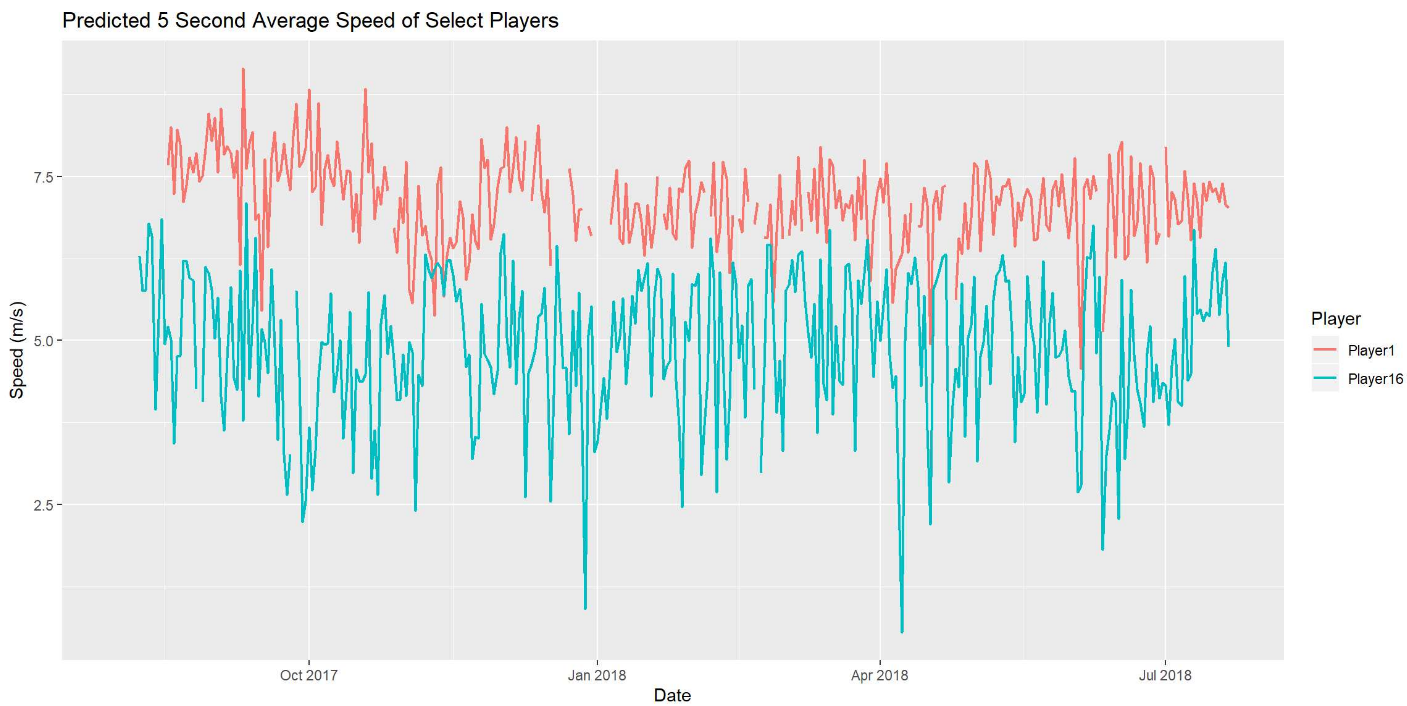
Random Player Effects for Speed Test

	Player Effect	Std. Error
Player1	1.67	0.11
Player2	1.06	0.11
Player3	0.97	0.11
Player4	0.80	0.11
Player5	1.78	0.11
Player6	0.21	0.11
Player7	0.61	0.11
Player8	0.58	0.11
Player9	1.00	0.11

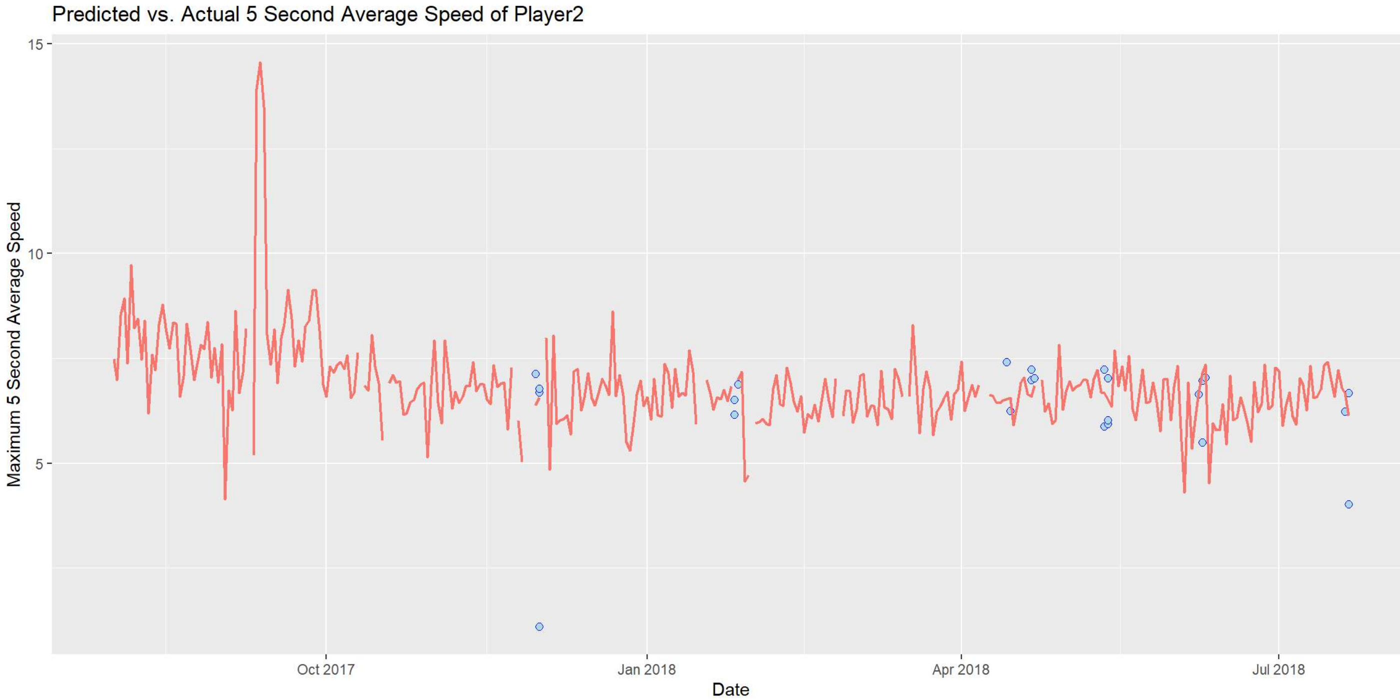
	Player Effect	Std. Error
Player10	1.07	0.11
Player11	0.73	0.11
Player12	-0.02	0.11
Player13	0.18	0.11
Player14	-0.73	0.11
Player15	0.91	0.11
Player16	-0.92	0.11
Player17	-1.03	0.11

The random effects for the players may be thought of as each player's "baseline" although it should be noted that because the ordinal data was standardized to have mean one for each player instead of mean zero these effects are not how fast each player would run 40 yards absent any fixed effects. To check the validity of this approach the athlete's coach should be consulted to see if faster (slower) players do indeed have higher (lower) random effects.

We now present the predicted wellness scores as well as a comparison between predicted and observed wellness for select athletes.

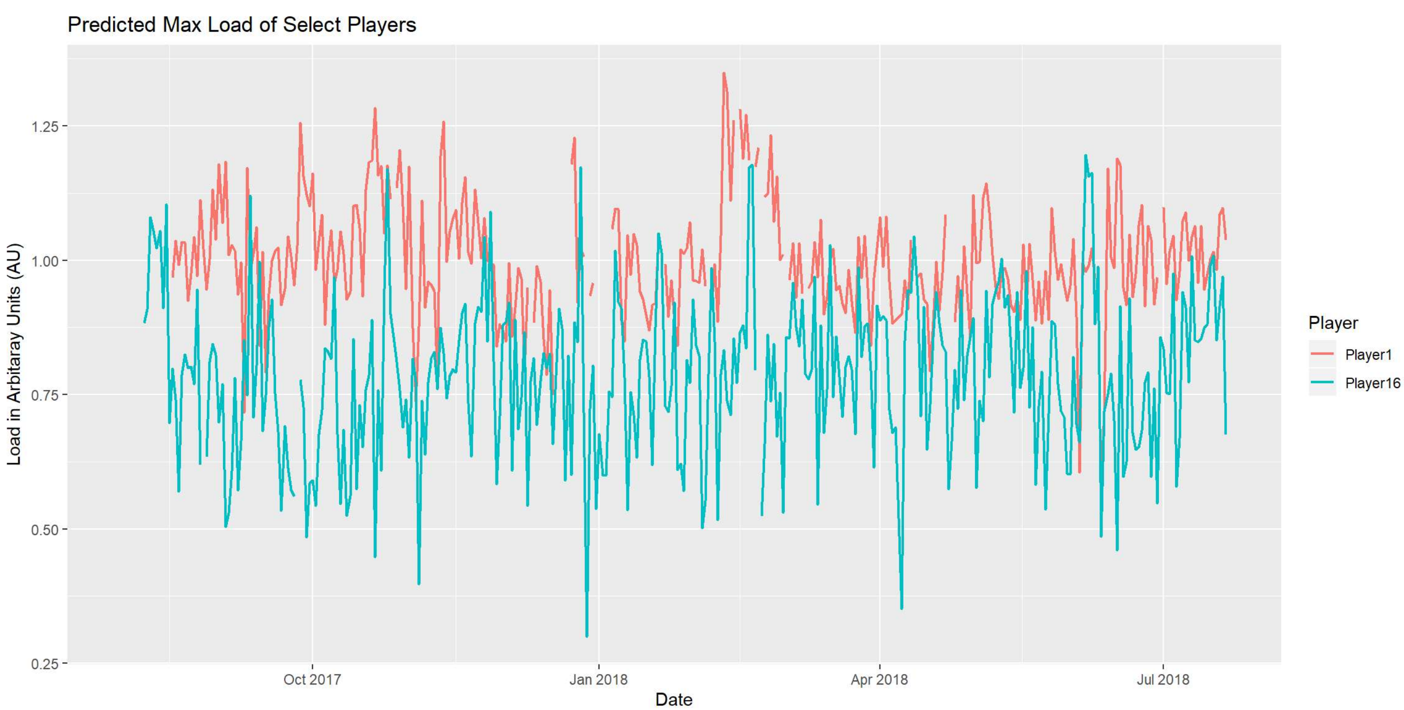


We see that player 1 is predicted to be faster than player 16. The predicted speed of both players looks like a white noise process with a constant mean which would be consistent with a constant daily training load. The model does however occasionally predict very slow speeds for player 16 indicating that it may be beneficial to adopt a higher percentile as a cut-off. The athletes' coach should be asked to comment on these findings as well as on the interpretability of the model's output. For example would it be more useful to have a rolling mean of the estimates rather than the estimates for each day?

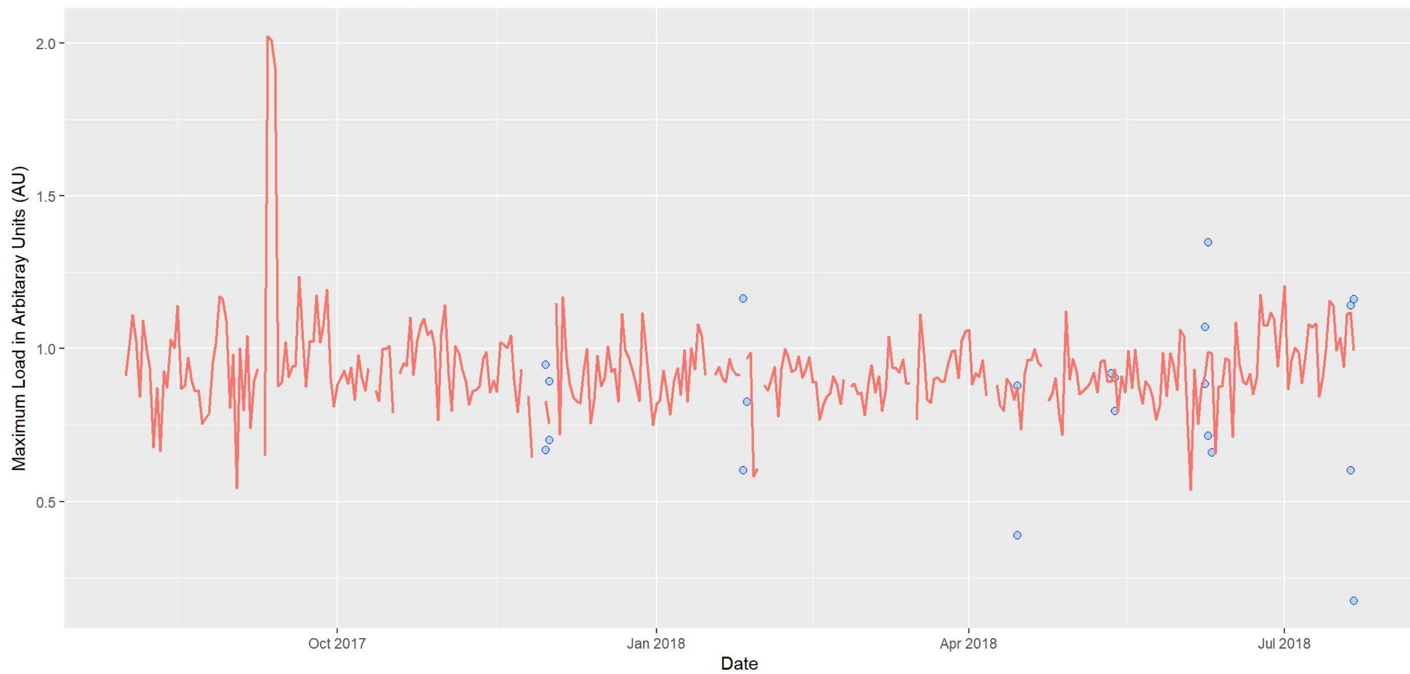


We can see that the predicted estimates for player 2 closely align with the fitted values. Again, there are cases where the model significantly overestimates the athlete's speed suggesting that a higher cut-off value may offer an improvement. Additionally, the model predicts a speed of $15m/s$ for this athlete sometime in September 2017. This is obviously not possible and thus care should be taken to ensure that if this model is put into use such unrealistic predictions are not commonplace.

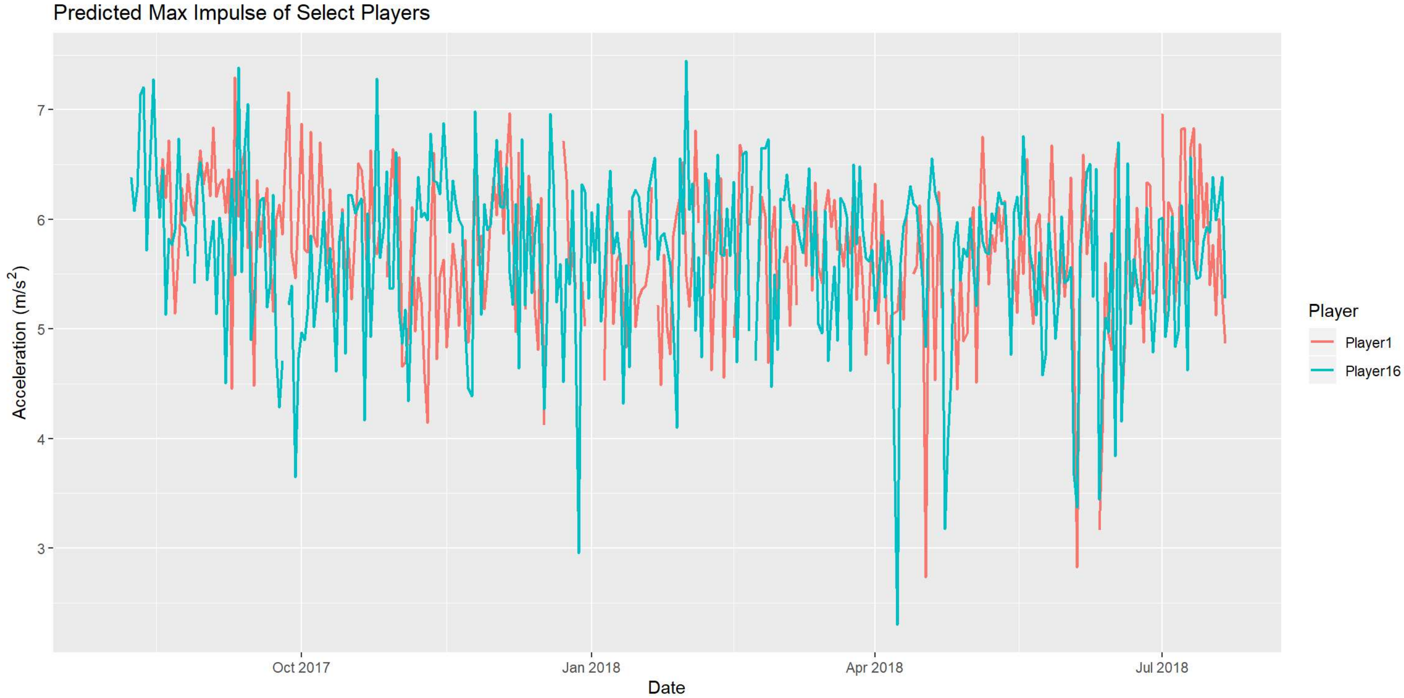
Our findings and thoughts on the use impulse and load models are more or less identical to that for the speed model. As such we offer the same two graphs as above for each of the two additional models primarily for the reader and refrain from commenting to avoid repeating our previous analysis.

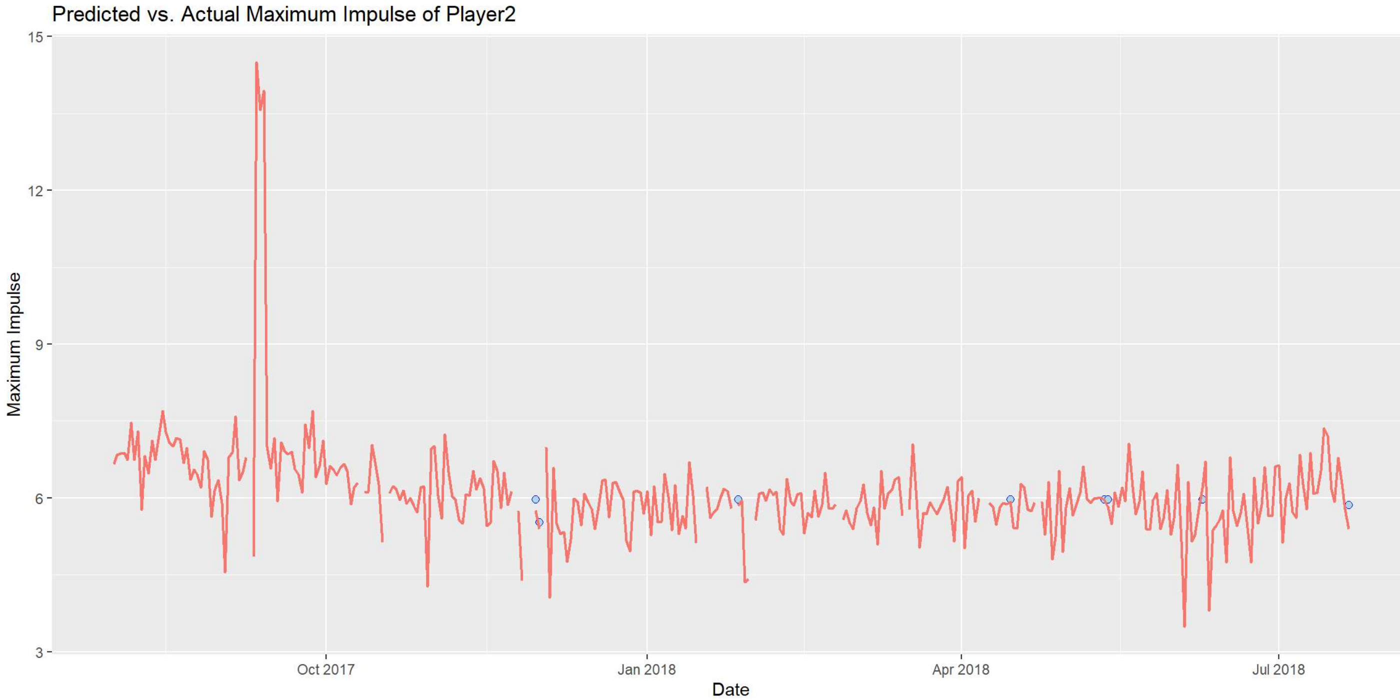


Predicted vs. Actual Load in Arbitaray Units for Player2



Predicted Max Impulse of Select Players





Conclusion

We applaud Rugby Canada both for being open to how statistics and data science can help them improve their training and game performance, and for monitoring their athletes' welling being. There are still many areas of high-performance sport that are not concerned with one or both of these aspects of modern sport. When coming up with our measures of training load and wellness we kept these concerns in the front of mind, aiming to create easily interpretable, relevant, and accurate metrics which can be used both in and out of competition.

Our proposed measure of session load is loosely based on well established models which quantify session load through heart rate. As such, we are confident that it will offer a more accurate measure of training load for Rugby Canada's athletes. Additionally, we believe that our measure somewhat corrects for the issues raised in regarding athletes' ability to account for session duration when reporting rate of perceived exertion. Finally, in the event that heart rate data becomes available, or that individual athletes have access to their own heart rate data, the code used in our measure of session load is immediately reproduceable and can be used to calculate Eric Banister's training impulse score.

To quantify wellness, we have taken ordinal and binary survey produced easily interpretable numerical scores relating to different physiological aspects of rugby. Additionally, these scores are available daily both in and out of competition and can thus be used to help monitor the health of the athletes throughout the training cycle. Admittedly, our scores are somewhat unpolished at this time, but can be further refined with the help of the training staff.

Unfortunately, due to the lack of session load data leading up to competition we were unable to study the effects of our session load scores and wellness scores on athletic performance. Future work should include such an analysis.