# USOPC Athlete Wellness and Load Interview Report

Perry Battles

May 2025

## 1 Key Takeaways

- EDA demonstrated that average session heart rate and practice load likely express much of the same information. If collecting both is logistically intractable or difficult, not much is lost by sacrificing one. Because of the familiarity of heart rate data and zones, it may be more useful to present these values to coaches and athletes.
- Sleep has by far the most outsized effect in promoting athlete wellness by reducing fatigue and stress and increasing mood and motivation.
- Workload does not appear to affect sleep quality within the ranges in which it has been administered, with the exception of a potential slight effect for practice load.
- Soreness was negatively associated with sleep to a much greater degree than with any other variable. Competition load may positively influence soreness.
- There was an association between greater strength and conditioning load and lower levels of fatigue and stress. However, this effect is likely very slight and possibly confounded by position-based factors; for example, it may be the case that positions that engage in more strength and conditioning perform less other training.
- To promote athlete well-being, facilitating proper restful sleep should be considered one of the highest priorities, especially during circumstances of travel and international competition. Very secondarily, the inclusion of resistance training may help to reduce levels of fatigue and stress. Finally, since soreness is associated with high competition load, techniques to mitigate this trend may be beneficial for performance. For example, during periods of dense competition frequency when acute performance is prioritized over chronic adaptation, utilizing recovery modalities that reduce soreness may enhance performance.

Each of these points is treated much more extensively in the extensive report, submitted alongside this document. For a more complete treatment of the analytical methods and results, please refer to it.

## 2 Introduction and Overview

This report is intended to be read by a sports science team to provide an overview of key patterns in the data and directions for action to facilitate athlete health and wellness and, by extension, performance.

The first question provoked by the data concerned the position groups; since the athletes were aggregated already into different roles, were any differences in workload or wellness data simply a function of group? This question hasn't been answered extensively enough by the present analysis, but my tentative conclusion is that, while there are likely strong position-based trends as demonstrated by the exploratory data analysis, there are probably also some outlier cases in different positions in which a more nuanced model of workload and wellness is useful, and also some conclusions that apply across groups.

The approach I started with was to perform a grid search over several different models and hyperparameters in the data. However, once I had more carefully inspected the distributions of the wellness outcome variables and performed the Shapiro-Wilk test, I concluded that they were roughly gamma distributed and warranted analysis using a model whose assumptions were reflected by the data. Ultimately, I settled on using the gamma generalized linear model, which I implemented in R. This led to the insights enumerated above.

# 3 Methods

## 3.1 Exploratory Data Analysis and Data Preprocessing

EDA began by simply visualizing the distributions of the data to determine what models and transformations might be needed to make inferences with some conviction that the assumptions of the analysis are well-met. After determining that the data were organized essentially as a time series for each athlete/variable combo and were also very sparse, I concluded that they should be aggregated on a week-by-week basis to facilitate analysis, as trends that are present after grouping are often absent when the data are considered only by individual calendar day. I also looked at distributions of the variables by position group and concluded that there were some marked differences between them. However, I chose not to include the athlete's position group as a predictor in subsequent modeling approaches; in hindsight, I believe that this was correct, since the differences between the groups should be reflected in the metrics provided, assuming that they contained information that is relevant to the outcomes of interest.

## 3.2 Gamma Generalized Linear Modeling

Because none of the wellness variables were normally distributed according to the Shapiro-Wilk test and appeared to be roughly gamma distributed, I moved forward with a gamma generalized linear model, which loosens some of the statistical assumptions of standard linear regression. I selected some feature subsets that I thought would be informative in modeling each outcome based on known physiological principles and relationships, and moved forward.

# 4 Challenges

The primary challenge I ran into in analyzing these data was the fact that there were matrix multiplication errors in the underlying `NumPy` and R `glm` code that caused fits of the gamma generalized linear model to fail. Getting around this took some doing; after reading a bit on StackOverflow, it seemed to me that the primary issue was some large numbers being produced in one or more steps of the math underlying the model fit. After being uncertain how to address this in Python, I pivoted over to using R's `glm` package, initially did not have much luck, but ultimately stumbled upon the log linking function as the appropriate choice. This allowed the models to converge and get a proper fit.

Another challenge was that the data do not precisely fit the assumptions of many standard models; values that are arbitarily capped at 0 and 100 are not described very well by many existing distributions. This meant that, although the gamma seemed to be a good approximation of the underlying distribution, it couldn't be technically correct since a gamma-distributed random variable can't actually assume values equal to 0, and won't be capped at some arbitrary value. I don't have a ready-made solution to deal with that problem, but it is something that could probably be addressed by utilizing another form of model, such as a multilayer perceptron or similar architecture.

# 5 Conclusions

The primary conclusion I've drawn from this analysis is that managing workload is tremendously secondary to sleep in facilitating athlete performance for most intents and purposes, providing we are operating within the confines of reasonable training. If we had to choose between carefully trying to monitor athlete loading and attending to sleep, we should probably choose to attend to sleep I could see this becoming particularly relevant during travel, particularly where international competition is concerned.

I also thought it was very interesting that strength and conditioning load was associated with lower levels of stress and fatigue. I'm not convinced that this isn't primarily a function of athlete position group, since it seems that there is some kind of "keeper" position group that doesn't cover nearly as much distance and also probably performs more resistance training. In that case, it may be that higher volumes of lifting co-occur with lower volumes of other training. Very speculatively, it could be that there are even temperamental factors that predispose an athlete in this position group to perceive less stress and fatigue. The main point here is that this mild trend is present, but it isn't clear exactly what action should be taken on the basis of the trend. One avenue forward would be to increase the level of strength and conditioning modestly across the entire athlete population and see if some beneficial change results – that could move in the direction of answering the question of a potential causal relationship there.

# 6 Directions for Future Work

Originally, I wanted to have an analytical pipeline written in pure Python that performed a grid search across several different models and hyperparameter combinations. Some of these models (e.g., random forest regressors) don't produce coefficient estimates in the traditional sense or may not necessarily make assumptions that fit the data very well, but can still be very predictive, possibly even more so than the generalized linear model chosen here. Others, like the multilayer perceptron, could do a great job predicting values of the outcome variable but might not necessarily be very readily interpretable. However, there are ways around this: a few solutions that come to mind are Shapley values, as well as visualizing outcome variable surfaces when varying two inputs at a time while holding the other predictors constant. One can also use principal component analysis to find eigenvectors in the inputs, reduce the predictor space to be two-dimensional, and then visualize changes in the outcome variable that way.

Another question I would like to more extensively answer is how much of the variation in the wellness and workload variables simply depends upon player position group. Although this may not change much from a modeling standpoint, it could have some important practical applications, since a player's position could be a good heuristic for how we need to approach managing their workload. That being said, it could be that there are underlying player positions that do not correspond to the given groups, and discovering those could give us more insight into how to coach these players. Manifold learning could be a solution to this problem; it would be an easy task to implement this with the present dataset, and it could reveal interesting patterns that could improve interventions with athletes.