**Ultrax Team 2 Executive Summary**

**Patrick Bernius, Liam Egan, Isabelle Esler, Connor Toomey**

App Link: https://mdpproj2-seedwsstjr4cnsjvdgg4fy.streamlit.app/Home

## Problem Framing

The challenge addressed in this week was how to reliably identify and quantify the Most Demanding Periods (MDPs) of football/soccer training sessions using high-frequency athlete tracking data. Traditional workload metrics based on single variables such as speed or heart rate often fail to capture the true physiological cost of training, particularly during short, acceleration-driven efforts. This study frames metabolic power as a unified workload signal and develops a scalable method to detect MDPs and summarize session intensity across multiple time scales.

## Methodology

MDPs were identified by evaluating key exertion-related features across multiple training sessions for football/soccer athletes. Initial exploration of the data set revealed several relevant variables, including Timestamps (ms), Speed (m/s), Odometer (m), Acceleration (m/s²), Heart Rate (bpm), and Metabolic Power (MP). Due to its composite kinematic nature (Hoppe, Baumgart, Slomka, Polglaze, & Freiwald, 2017) and its established alignment with athlete-perceived exertion and fatigue (Alghannam, 2012), MP offers a more comprehensive foundation for modeling workload and detecting MDPs than any single variable (e.g. speed, acceleration, or heart rate) alone. Although MP was provided as a unitless value in the dataset, prior literature indicates that MP can be derived from a combination of speed, positive and negative acceleration, and cadence (Moore, Aguiar, Ducharme, & Tudor-Locke, 2021).

Data cleaning was conducted to isolate relevant time windows and ensure analytical quality. Four training dates from late October to early November 2020 were selected for initial exploration. In total, forty-eight raw data files were merged and included in the full analysis pipeline. As such, during the cleaning process, the complete dataset was reduced from approximately 2.5 million to 250,000 rows by removing entirely null columns (e.g., Altitude) and columns with systematic missingness (e.g., millisecond rows with absent heart rate values). Millisecond timestamps were converted into standardized date–time formats to enhance within-session visualization and enable robust cross-session alignment.

After preliminary cleaning and exploration, the forty-eight files were partitioned into training and testing sets in order to calibrate the derived MP and intensity equations. Thirty files were used to develop and validate the MP formulation, and the remaining eighteen files were reserved for testing to evaluate generalizability across different athletes and sessions. Equations for MP and session intensity were derived to construct an Intensity Index capable of identifying MDPs (see Appendix B). The MP equation, formulated from speed, acceleration, and cadence terms, was defined as:

$$MP(t) = 2 + 27v(t) - 0.2v(t)^2 - 15a^{+}(t) - 22a^{-}(t) + 23a^{+}(t)v(t) - 2.6a^{-}(t)v(t) - 0.02cadence(t)$$

To evaluate the temporal variability of MP, sliding windows of 10, 20, and 30 seconds were applied across each session. MP values within each window were used to determine window-specific MDPs. These windowed maxima were then aggregated to compute three session-level attributes: explosiveness (E), repeatability (R), and volume (V), which serve as components of the Session Intensity Index. The resulting intensity equation is expressed as:

$$SI = \omega_E E + \omega_R R + \omega_V V, \quad \omega_E + \omega_R + \omega_V = 1$$

Both equations were integrated into a deployed Streamlit application designed to visualize MP-derived metrics across players and sessions and to facilitate interactive identification of MDPs.

## Insights

Analysis of metabolic-power–derived features across sessions revealed several notable patterns regarding athlete workload and the characteristics of Most Demanding Periods (MDPs). First, MP proved to be a more sensitive indicator of high-intensity activity than traditional metrics such as speed, acceleration, or heart rate. In particular, accelerative and decelerative actions contributed disproportionately to MP spikes, confirming that the energetic cost of rapid directional or velocity changes is substantially higher than that of steady running. As a result, many MDPs occurred during short, dense bursts of accelerations rather than during periods of peak speed, underscoring the value of MP in capturing the true physiological stress of training.

Windowed MP analysis further showed that MDPs were highly dependent on the chosen temporal scale. Ten-second windows highlighted explosive, short-duration loads, whereas 20- and 30-second windows revealed sustained periods of elevated energetic demand. These distinctions aligned naturally with the derived explosiveness (E), repeatability (R), and volume (V) dimensions of the Intensity Index. Sessions characterized by high E but low R indicated sharp, isolated bursts of intensity. In contrast, sessions with elevated R values demonstrated repeated exposure to near-maximal metabolic load, indicating higher cumulative stress on the athlete. Volume (V) provided a broader perspective on overall session demand, capturing the extent to which energetic cost was sustained across longer windows.

Across the training files analyzed, athletes displayed distinct MDP signatures. Some athletes exhibited pronounced peaks concentrated in short intervals, whereas others showed moderate but consistent metabolic demand over extended windows. For example, Player 25 was consistently in the top 25 for Most Explosive Efforts, Best Sustained Effort, Biggest Workloads, and Hardest Sessions Overall across these training sets. These individualized patterns highlight the potential for MP-based modeling to capture player-specific exertion profiles, which can inform tailored training prescriptions and load-management strategies.

Finally, cross-session comparison suggested that MDP patterns were not random but exhibited a degree of stability within athletes across similar session types. This consistency indicates that the MP-based Intensity Index can serve as a reliable indicator of an athlete's typical exertion profile and may support early identification of deviations that signal fatigue accumulation or increased workload risk.

The integration of the derived equations into an interactive Streamlit application allowed these insights to be visualized clearly, enabling practitioners to inspect MP trajectories, identify MDPs, and compare session intensity across players. This tool demonstrates the practical utility of MP-driven analytics in real-time or retrospective performance monitoring and provides a foundation for ongoing development of fatigue and workload-modeling systems.

## References:

Alghannam, A. F. (2012, June). *Metabolic limitations of performance and fatigue in football*. Asian journal of sports medicine. https://pmc.ncbi.nlm.nih.gov/articles/PMC3426724/

Hoppe, M. W., Baumgart, C., Slomka, M., Polglaze, T., & Freiwald, J. (2017, August 1). *Variability of metabolic power data in elite soccer players during pre-season matches*. Journal of human kinetics. https://pmc.ncbi.nlm.nih.gov/articles/PMC5548171/

Moore, C. C., Aguiar, E. J., Ducharme, S. W., & Tudor-Locke, C. (2021, January). *Development of a cadence-based metabolic equation for walking*. Medicine and science in sports and exercise. https://pmc.ncbi.nlm.nih.gov/articles/PMC7896743/

# Appendices:

# Appendix A (Figures):

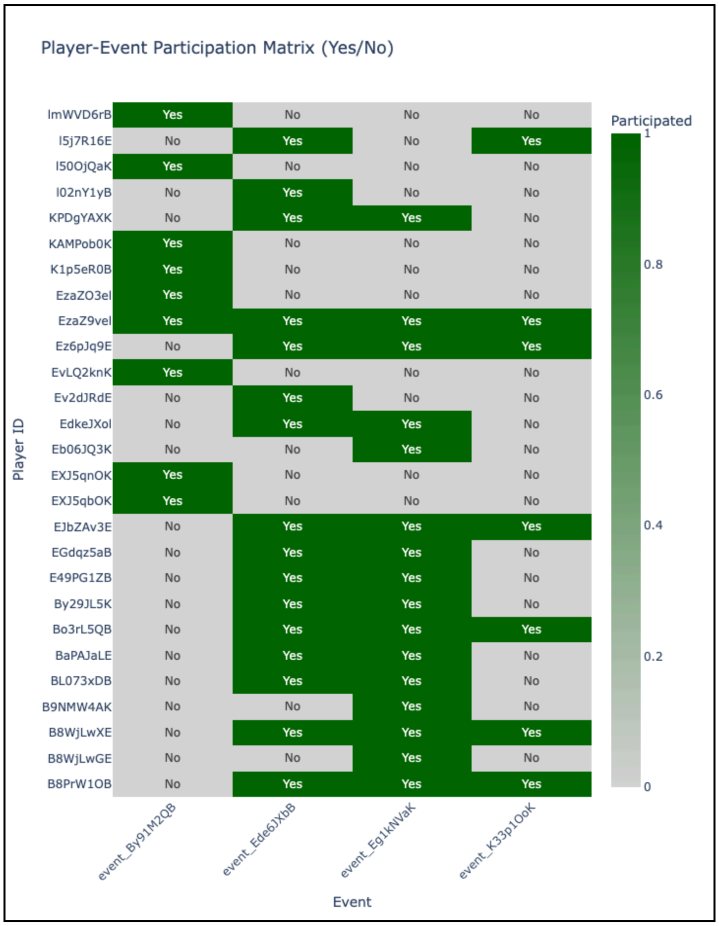


Figure 1. Player Participation Matrix

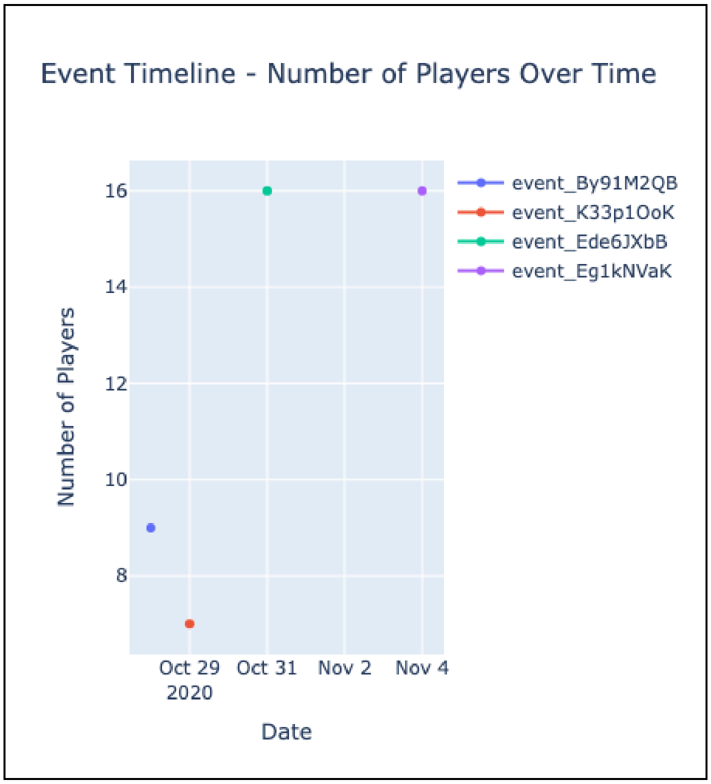


Figure 2. Number of Participants over time

# Appendix B Methodology

## B.1 Data Inputs

Player tracking data consisted of time-stamped measures of speed, acceleration, and cadence sampled at a fixed frequency. Acceleration was decomposed into positive ($a^{+}$) and negative ($a^{-}$) components prior to analysis.

## B.2 Metabolic Power Estimation

Instantaneous metabolic power (MP) was computed at each time step using a regression-based formulation incorporating speed, acceleration, acceleration–speed interaction terms, and cadence:

$$MP(t) = 2 + 27v(t) - 0.2v(t)^{2} - 15a^{+}(t) - 22a^{-}(t) + 23a^{+}(t)v(t) - 2.6a^{-}(t)v(t) - 0.02cadence(t)$$

where $v(t)$ denotes speed, $a^{+}(t)$ positive acceleration, $a^{-}(t)$ negative acceleration, and cadence step frequency. Metabolic power was expressed in watts per kilogram ($W \cdot kg^{-1}$).

## B.3 Sliding Window Processing

To characterize sustained physical demand, rolling mean metabolic power was computed using a sliding window of fixed duration:

$$\overline{MP}^{(w)}(t) = \frac{1}{N_{w}} \sum_{i=t}^{t+N_{w-1}} MP(i)$$

where $N_w$ is the number of samples within the window. Window durations of 10, 20, and 30 seconds were evaluated.

## B.4 Most Demanding Period (MDP)

For each window duration, the Most Demanding Period was defined as the maximum rolling mean metabolic power observed during the session:

$$MDP^{(w)} = max(\overline{MP}^{(w)}(t))$$

This yielded session-level values $MDP^{(10)}$, $MDP^{(20)}$, and $MDP^{(30)}$.

## B.5 Raw Load Metrics

### B. 5.1 Volume

Session volume was defined as mean metabolic power multiplied by session duration:

$$V^{raw} = \overline{MP} * T$$

where T denotes session duration.

### B.5.2 Explosiveness

Explosiveness was defined as the peak short-duration metabolic demand:

$$E^{raw} = MDP^{(10)}$$

### B.5.3 Repeatability

Repeatability was defined as the average of medium- and longer-duration demanding periods:

$$R^{raw} = \frac{MDP^{(20)}+MDP^{(30)}}{2}$$

## B.6 Normalization

Raw load metrics were normalized using z-score normalization:

$$X = \frac{X^{raw} - \mu_X}{\sigma_X}, \quad X \in \{E, R, V\}$$

where $\mu_X$ and $\sigma_X$ were computed over the selected reference population.

## B.7 Composite Session Intensity

A composite session intensity index was computed as a weighted linear combination of normalized components:

$$SI = \omega_E E + \omega_R R + \omega_V V, \quad \omega_E + \omega_R + \omega_V = 1$$

Weights were selected empirically and held constant across analyses.

## B.8 Implementation Summary

The analysis pipeline consisted of:

1. Computing instantaneous metabolic power from tracking data
2. Applying sliding windows to derive rolling mean MP
3. Extracting MDP values for multiple window durations
4. Computing raw volume, explosiveness, and repeatability
5. Normalizing metrics via z-scores
6. Aggregating components into a composite intensity index