Chelsea

Ashwin Ramaseshan

March 2025

1 Introduction

Modern football demands a precise balance between performance, recovery, and injury prevention. With player availability and physical output closely linked to match outcomes, clubs are increasingly leveraging data science to make informed decisions about load management. The 2025 Chelsea FC Performance Insights Vizathon challenged participants to develop a metric that can unify disparate performance indicators into a meaningful, interpretable score that guides these high-stakes decisions.

This report presents the development of the **Load Balance Efficiency Score (LBES)**—a composite metric that quantifies how well a player's match demands, training intensity, and recovery quality are aligned. The LBES framework empowers performance analysts to flag imbalance, forecast injury risk, and optimize training strategies. The end-to-end solution includes data processing, model development, and a Power BI dashboard for interactive visualization.

2 Project Background

The LBES model was developed as part of the Chelsea FC Vizathon running from March 17 to April 7, 2025. Participants were provided access to anonymized player datasets representing match load (via GPS), training capability (via movement assessments), and recovery quality (via wellness indicators). The project aimed to achieve the following:

- **Quantify player load balance** using a unified metric scaled 0–100.
- **Normalize inputs** across multiple data sources for fair comparison.
- **Detect high-risk patterns** using trend analysis and predictive modeling.
- **Visualize insights** using Power BI for real-world applicability.

The LBES was designed not only as an academic solution but as a practical tool to be used by football clubs for day-to-day load management, early intervention, and performance planning.

3 Dataset Overview

Three core datasets were used in this project:

- **GPS Data:** Captures match-related output such as distance, sprinting, and accelerations.
- **Recovery Status:** Includes subjective and objective measures of fatigue, sleep, and wellness.
- **Physical Capability:** Compares player performance in physical tests (jump, agility, strength) to benchmarks.

All features were normalized to a common 0–100 scale to ensure interpretability and eliminate bias during composite scoring.

4 Methodology

This project followed a structured methodology to transform raw performance data into meaningful insights through normalization, scoring models, predictive analysis, and visualization. The pipeline includes:

1. **Normalization:** Bringing all metrics onto a 0–100 scale using Min-Max scaling for interpretability and comparability.
2. **Component Score Calculation:** Generating three key scores — Match Load Score (MLS), Recovery Score (RS), and Training Load Score (TLS).
3. **LBES Construction:** Combining all three scores using weighted averages to create the Load Balance Efficiency Score.
4. **Advanced Analytics:** Applying trend analysis, clustering, risk prediction, and recovery estimation models.
5. **Power BI Visualization:** Displaying results interactively to support coaching and medical decisions.

5 Data Normalization

Each dataset was standardized using min-max normalization:

$$\text{Normalized Value} = \frac{X - \min(X)}{\max(X) - \min(X)} \times 100 \tag{1}$$

This ensured that all input features, regardless of units, were transformed to a uniform scale (0–100). This enabled fair comparison across match, training, and recovery metrics.

5.1 GPS Data

The following GPS-derived metrics were normalized:

- **Distance Covered**
- **High-Speed Running** (distance over 21, 24, and 27 km/h)
- **Accelerations** (combined thresholds over 2.5, 3.5, and 4.5 m/s^2)
- **Peak Speed**

5.2 Recovery Data

Wellness metrics were pivoted and normalized:

- **Sleep Composite**
- **Subjective Wellness**
- **Soreness**
- **EMBOSS Score (Exercise-induced Muscle Balance, Oxygenation, and Subjective Symptoms)**

These were used to compute a unified Recovery Score (RS) for each session.

5.3 Physical Capability

Physical capability benchmarks were normalized for selected movements:

- Jump Test
- Agility Test
- Upper Body Test

Only records with valid benchmark percentages were included. The average benchmark score represented training capability.

5.4 Score Calculation

Each normalized metric was then used to compute component scores:

- **Match Load Score (MLS)**:

  $$0.30 \times \text{Distance} + 0.25 \times \text{High Speed} + 0.25 \times \text{Accelerations} + 0.20 \times \text{Peak Speed}$$

- **Recovery Score (RS)**:

  $$0.25 \times \text{Sleep} + 0.25 \times \text{Subjective} + 0.25 \times \text{Soreness} + 0.25 \times \text{EMBOSS}$$

- **Training Load Score (TLS)**:

  $$\text{Average Benchmark Score across jump/agility/strength}$$

5.5 LBES Score

The Load Balance Efficiency Score (LBES) was defined as:

$$\text{LBES} = 0.35\times\text{Recovery Score}+0.35\times\text{Training Score}+0.30\times\text{Match Load Score}$$

This score quantifies how well an athlete's physical load and recovery are balanced. Higher scores (closer to 100) indicate optimal balance, while lower scores (¡50) indicate high risk.

6 Exploratory Insights & Trend Analysis

To establish a foundational understanding of Chelsea FC's player load balance over time, we performed exploratory data analysis on three core components: match load, training intensity, and recovery status. These insights were crucial for understanding patterns in player stress and performance across seasons.

6.1 Training vs. Recovery vs. Match Load (Yearly Comparison)

We examined the average match score, training score, and recovery score over each year to understand how player workloads and recovery strategies evolved across time. This allowed us to spot macro-level imbalances or improvements in load management practices.

Key Insight: While training scores remained relatively stable, recovery scores dropped slightly in 2024 compared to 2023. Match loads fluctuated more dramatically, indicating shifts in competitive intensity or player deployment.

[Insert Power BI Chart: "Training vs Recovery vs Match Load (year)"]

6.2 LBES Line Chart

We plotted the Load Balance Efficiency Score (LBES) over time to capture fluctuations in overall player health and balance. LBES is a weighted combination of match, training, and recovery metrics that reflects how well an athlete is managing their workload.

Key Insight: Several significant drops were observed, potentially indicating high fatigue, poor recovery, or schedule congestion.

[Insert Power BI Chart: "LBES Line Chart"]

6.3 LBES vs. 3-Day and 5-Day Moving Averages

To smooth short-term fluctuations and observe trend momentum, we plotted 3-day and 5-day moving averages of LBES alongside the original signal.

Key Insight: While daily LBES can be volatile, the moving averages revealed underlying stability trends and helped flag periods of sustained risk.

[Insert Power BI Chart: "line chart lbes.pbix"]

7 Spike and Drop Detection in Load Balance Efficiency

To proactively monitor athlete health, we implemented a spike and drop detection system based on the Load Balance Efficiency Score (LBES). The goal was to flag sudden and significant changes in player load, which often precede injuries or performance dips.

7.1 Methodology

We calculated the percentage change in LBES from one session to the next:

$$\text{LBES\_Pct\_Change} = \left(\frac{LBES_t - LBES_{t-1}}{LBES_{t-1}} \right) \times 100$$

A threshold of $\pm 20\%$ was used to classify changes:

- **Spike:** A sudden increase in LBES exceeding 20%
- **Drop:** A sudden decrease in LBES exceeding 20%
- **Normal:** All other variations

Each data point was labeled accordingly as `Spike`, `Drop`, or `Normal`, and these were used to group and visualize periods of risk.

7.2 Visualizing Fluctuations in LBES

We created a scatter plot showing the distribution of LBES values by year and change type. Each point was colored based on its classification (Spike, Drop, or Normal), helping identify trends in instability.

Key Insight: Most LBES changes were classified as normal, but a few distinct spikes and drops stood out and warrant closer player-specific analysis.

[Insert Power BI Chart: "spikes.pbix"]

8 Clustering Load Profiles

To uncover underlying patterns in player workload and recovery behavior, we used unsupervised learning techniques—specifically **K-Means Clustering**—to segment players based on their load profiles. This allowed us to group similar sessions and identify load management trends.

8.1 Features Used for Clustering

We selected three normalized features to cluster each session:

- **Match Score (MLS)** – Represents match-related load.
- **Training Score (TLS)** – Quantifies physical capability metrics from gym sessions.
- **Recovery Score (RS)** – Measures player readiness and physical recovery.

These features were standardized using `StandardScaler` and fed into a KMeans model with $K = 3$, which was empirically chosen to balance granularity and interpretability.

8.2 Cluster Interpretation

The resulting clusters provided meaningful insights into load patterns:

- **Cluster 0: Balanced Load**
  Sessions with moderate match, training, and recovery scores.
- **Cluster 1: Low Match/Recovery Load**
  Sessions with high match score but low recovery—highlighting possible overtraining risk.
- **Cluster 2: High Match Load, Low Recovery**
  A critical cluster where players experienced elevated match intensity with insufficient recovery—requiring intervention.

8.3 Visualization of Clusters

The clusters were visualized using a 2D scatter plot with `Match Score` on the X-axis and `Recovery Score` on the Y-axis. Each point was color-coded based on its cluster label and annotated in the legend.

[Insert Python-generated Plot: "lbes_clusters_plot.png"]

9 Injury Risk Prediction

One of the most critical goals in sports performance analytics is to proactively identify athletes at high risk of injury. To achieve this, we designed a binary classification model that predicts whether a given session places a player at **high injury risk** based on their recent load and wellness trends.

9.1 Problem Framing

The model predicts a binary outcome:

- **Safe (0)**: LBES (Load Balance Efficiency Score) ≥ 50
- **High Risk (1)**: LBES < 50

Target Variable: A risk label generated using the LBES thresholds.
Features Used: We engineered a variety of time-series features including:

- Raw scores: match_score, training_score, recovery_score, LBES
- Historical aggregates: 3-day and 5-day moving averages
- Temporal changes: daily change, previous day's score
- Trends: slope over past 3 days

9.2 Model Selection and Training

We used a **Random Forest Classifier** for its ability to model nonlinear interactions, provide feature importance, and handle small-to-medium datasets effectively.

The dataset was split into an 80/20 training-test set. The model was trained using the engineered features and evaluated using a confusion matrix and classification report.

9.3 Performance and Evaluation

[Insert Confusion Matrix Image Here]
e.g., "confusion_matrix_injury_risk.png"

Key Metrics:

- **Accuracy:** 87%
- **Recall (High Risk):** 50% – Correctly identified 50% of at-risk sessions
- **Precision (High Risk):** 100% – No false positives

9.4 Feature Importance

To understand what drives the model's predictions, we extracted feature importances from the trained Random Forest model.

[Insert Power BI Bar Chart: Feature Importance]
e.g., "Feature Importance – Injury Risk Model.pbix"

Top Contributors:

- **LBES**: Strongest predictor of injury risk.
- **Match and Recovery Scores**: Key contextual indicators of stress and readiness.
- **Short-term trends**: 3-day averages and recent changes showed high predictive value.

9.5 Practical Application

This model can serve as a monitoring system for coaches and performance staff to flag potential red zones before a player experiences physical breakdown. By combining historical load, recovery, and wellness metrics, it empowers data-driven decision-making to optimize player availability and safety.

10 Time-to-Recovery Estimator

Understanding how long it takes a player to return to optimal readiness after a period of underperformance or excessive load is essential in managing player health and workload. In this section, we develop a time-based analysis to measure individual recovery windows based on the LBES score trajectory.

10.1 Methodology

We define two key thresholds:

- **Dip Threshold (LBES $<$ 50)**: Indicates a significant drop in player readiness or performance.
- **Recovery Threshold (LBES $\geq$ 60)**: Marks the point at which the player is considered to have recovered.

Our algorithm scans the time series of LBES scores to identify:

- The date of each LBES dip below 50.
- The first subsequent date when the LBES rebounds to 60 or higher.

The number of days between these two events is computed as the **recovery duration**.

10.2 Results

We identified multiple dips in LBES values across the season and tracked the time required for recovery. The table below summarizes key recovery windows:

Dip Date	Recovery Date	Recovery Days
08/11/2023	09/11/2023	31
10/09/2023	01/06/2024	89
01/07/2024	08/03/2024	209
08/04/2024	08/06/2024	2
09/04/2024	10/10/2024	36

10.3 Insights and Practical Use

- **Variation in Recovery:** The number of recovery days ranged from 2 to over 200, highlighting variability in player responses.
- **Use in Scheduling:** This information can guide return-to-play decisions and training load adjustments.
- **Injury Prevention:** Extended recovery times may signal underlying fatigue or recovery inefficiencies.

This analysis adds temporal context to the LBES framework, helping staff not just assess risk, but also understand how long players typically take to bounce back — allowing for more precise and individualized recovery planning.

11 Discussion and Insights

The development of the Load Balance Efficiency Score (LBES) and its supporting models offers a practical, interpretable, and actionable solution for sports performance teams managing complex data from multiple sources. Through a unified framework, we demonstrated how physical output, recovery, and training capacity can be synthesized to support real-time decision-making.

11.1 Key Insights Across the Models

- **LBES as a holistic readiness score:** LBES successfully combines match, training, and recovery signals into a single metric that is both interpretable and scalable.
- **Trend monitoring exposed volatility:** Our line and moving average plots showed significant fluctuations in LBES, including sudden drops that often persisted for long durations—some players took over 200 days to recover.

- **Spike/drop detection added early warnings:** The ability to flag ¿20% changes in LBES allows for early identification of overtraining or inadequate recovery.
- **Clustering revealed distinct load profiles:** KMeans helped segment sessions into interpretable groups like balanced, high match-low recovery, and under-recovered load patterns.
- **Injury risk model emphasized LBES importance:** The Random Forest model reinforced LBES as the most important predictive feature, validating its role as a leading indicator.
- **Personalized load recommendations created tangible targets:** By reverse-engineering periods of peak performance, we generated individualized benchmarks for optimal match, training, and recovery scores.

11.2 Implications for Performance and Medical Staff

This system has several real-world applications:

- **Daily Monitoring:** The LBES score can be used to monitor athletes daily, triggering alerts when scores fall below safe thresholds.
- **Load Planning:** Personalized recommendations support smarter workload periodization and more individualized interventions.
- **Injury Prevention:** By tracking dips in LBES and lagging recovery durations, teams can proactively manage player availability.

11.3 Limitations

Despite the utility of the framework, some limitations remain:

- **No player-level granularity:** The anonymized dataset lacked player IDs, preventing individualized timelines and model evaluation.
- **No injury labels:** True injury data was not included, so the "injury risk" label is based on LBES thresholds rather than medical diagnoses.
- **Single-season snapshot:** The models were trained on a single season, so longitudinal validation remains a future step.

11.4 Future Work

- Integrate positional and player identity data to build personalized risk profiles.
- Add live injury tracking data to validate and refine risk prediction models.

- Incorporate external variables such as travel, match schedule congestion, and sleep quality from wearables.
- Embed the LBES framework into a real-time monitoring tool for staff use.

Overall, the LBES framework provides a scalable and explainable solution that empowers Chelsea FC and other elite clubs to monitor, intervene, and optimize player performance more effectively.

12 Conclusion

The Chelsea FC Vizathon provided a unique opportunity to address one of the most pressing challenges in modern football — balancing player load, training demands, and recovery to optimize performance and reduce injury risk.

In response, we developed the **Load Balance Efficiency Score (LBES)**, a unified, interpretable metric that captures the dynamic relationship between match intensity, recovery readiness, and training capacity. LBES served as the backbone of multiple advanced analytics models, including trend detection, risk prediction, recovery forecasting, and personalized workload recommendations.

Our work demonstrates the power of combining statistical modeling, time-series analysis, and domain-specific expertise to create a practical tool for real-world decision-making. Through detailed analysis and Power BI visualizations, we uncovered actionable insights that support player management across a competitive season.

By integrating LBES into daily workflows, clubs like Chelsea FC can shift from reactive to proactive performance planning — enabling more informed decisions, earlier interventions, and ultimately healthier, more available players.

The framework presented here is modular, scalable, and adaptable — laying the groundwork for future development of intelligent monitoring systems in elite football.