B1701 - Assignment 2 - Melbourne Victory Women’s Team - Defender Recruitment Analysis

MC

2024-12-10

# Introduction

Melbourne Victory’s women’s football team was let down by their poor defence in the 23/24 season. Melbourne Victory scored the most goals in the league however lost or drew crucial games where their attack wasn’t on top form or was stopped by the other team’s defence. Addressing their weaknesses, especially in interceptions, progressively passing out from the back and being able to progressively carry the ball out from defence, are crucial to regaining dominance in the A-League.

The rise of data-driven recruitment [1] has transformed how teams address performance gaps. Data analysis allows clubs to make informed and objective decisions to identify players who align with their tactical and strategic needs. This is done using key performance indicators (KPIs) as it allows teams to compare athletes on metrics that they know are crucial for the athlete they are recruiting to have [2].

This study aims to:

1. Evaluate current squad deficiencies using KPIs to inform recruitment decisions
2. Identify five high-performing defenders using data from the Women’s World Cup 2023 who meet Melbourne Victory’s defensive requirements

These objectives support Melbourne Victory’s goals of strengthening their defensive unit and enhancing their overall league performance.

This report begins with the methodology (data sources, KPI development, and the analytical framework), presents the results with visual comparisons and discusses the findings in the context of Melbourne Victory’s strategic goals. The report concludes by summarising the aim of the objective and showing that the objectives have been met.

# Methodology

## Data Collection

This study’s primary data source is FBRef, which provides a comprehensive dataset of player performance metrics from domestic and international football. This dataset includes key variables such as passes completed, touches during a game, errors and may other performance indicators relevant to defender evaluation. These metrics were extracted from the Women’s World Cup 2023 data. This data was chosen as it provides a selection of players who have been selected to play for their national teams so have a strong level of ability [3].

## KPI Development

The key performance indicators (KPI) and weightings used for composite score used in this study were developed based on a combination of football analytics literature [4] and the answers from the questions asked to management. The eight key performance indicators selected to form a composite score are:

* Total interceptions per 90 (25%)
  + Interceptions are a crucial defensive action as it shows the defender’s ability to read the game and disrupt the opposition’s attacking play [5].
* Total progressive passes per 90 (20%)
  + Progressive passing is an essential metric in modern football [6] all positions of defender. It helps to show a defender’s contribution and ability to transition the ball from defence to attack.
* Total progressive carries per 90 (20%)
  + Similar to progressive passes, progressive carries assess a defender’s ability to drive the ball forward and initiate the attack. A defender who can carry the ball out effectively from defensive zones is important, especially in counter-attacking systems [7].
* Total passes completed per 90 [8] (15%)
  + Total passes per 90 was selected as a KPI because it helps to show how involved the defender is in build-up play. This metric is weighted lower due its supportive role in the overall assessment of the defender’s performance.
* Total touches per 90 (10%)
  + This is more of a general metric however has been chosen due to the nature of the defender wanted by the club. The club want a defender who is comfortable on the ball so having more touches of the ball per game suggests their comfort with the ball at their feet.
* Total clearances per 90 (5%)
  + Clearances are important for relieving pressure in high-stress situations [9]. While important, clearances are weighted lower due to their reactive nature rather than a productive one.
* Total blocks per 90 (3%)
  + Blocks represent a defender’s ability to prevent shots or passes. This KPI was assigned a lower weight as it is more situations and dependent on the type of attacking pressure faced. However, it remains valuable as it highlights a defender’s physical presence and anticipation.
* Average pass success percentage per 90 (2%)
  + Pass success percentage is weighted lower in this analysis due to the high weighting of total passes completed per 90, however, this metric provides additional insights into a defender’s decision-making and execution under pressure.

KPIs were normalised using min-max normalisation [10] between 0 and 1 to allow for fair analysis across all metrics and expressed per 90 minutes to ensure an equal comparison between players with varying amounts of playtime[11], especially as the World Cup is a knockout competition. Players with <91 minutes or older than 30 were excluded to ensure relevance to immediate recruitment needs.

# Analytical Approach

The analytical approach for identifying high-performing defenders includes data filtering, the use of composite scores, and the use of visualisations to display the results.

## Data Filtering

Women’s World Cup 2023 players were filtered to a minimum playtime threshold of 91 minutes. >=91 minutes was chosen as Melbourne Victory, needs a defender who can slot straight into the squad and not be a bench player. Additionally, this took out 1 quarter of the dataset of defenders with the 1st quartile equalling exactly 90 minutes. The age was also filtered to <=30 based on research on the average age of a woman footballer’s prime [12]. This approach excludes players with low playing time who may be either emerging talents or established players, which may impact the results.

## Composite Score Ranking

A composite score system was used to prioritise defenders based on their performance across multiple selected KPIs. Each KPI was weighted according to its relevance to Melbourne Victory’s defensive needs and scores were taken to form a single composite score for each player. This approach allows for easy comparisons between players and helps identify those who meet the team’s specific requirements.

## Visualisations

Visualisations, such as radar charts and scatter plots were used to communicate player performance. Visualisations are highly effective in presenting complex data in an accessible format.

# Data Pre-Processing and Cleaning

## Handling Missing Values

Missing data was addressed through imputation where possible. In the Women’s World Cup 2023 dataset, many players had missing values due to the player not creating an action to lead to an output in this cell. Any cell that included NA was filled with a 0.

## Data Integrity and Relevance and Outlier

Due to data being sourced from a reliable and well-maintained database of FBRef, the assumption is that the data is accurate and consistent therefore, as said before, any NA was put to 0 and outliers were treated as normal as these numbers could exist due to the ability of the athletes.

# Results

To evaluate the recruitment options for a new Melbourne Victory Women’s defender, the following data visualisation were created to analyse both team performance and potential recruits. These visualisation provide an approach to identify weaknesses, assess the new candidates, and ensure value for money in the decision-making process.

## Melbourne Victory Analysis

Firstly, a bar graph was created to rank the Melbourne Victory defenders based on composite score. The same KPIs and weights were used in both composite scores to ensure consistency. Using a bar graph is a simple but effective way to show how each player ranks and shows the overall suitability of the player based on the KPIs.

A graph with blue squares

AI-generated content may be incorrect.

The bar graph ranks Jamillia Rankin and Kayla Morrison as the top 2 players and Tori Hansen to be underperforming the most. Emma Checker is also underperforming compared to the top 3 with a gap of 0.249. This visual helps to simplify the comparisons and provides clarity on which defenders are performing best and contributing most to the teams strategic objectives and fitting the KPIs set by management.

Next, a radar chart of the Melbourne Victory defenders was created as this visualisation helps to show management each player’s strengths and weaknesses and where gaps need to be filled.

A diagram of a hexagon with colorful lines and dots

AI-generated content may be incorrect.

From the radar chart, we can see that Kayla Morrison who ranked second in composite score, is the most well-rounded player however lacks interceptions per 90 which is weighted highly in the composite score due to its importance to management. Tori Hansen, who was ranked lowest on composite score in the squad, has shown in this radar chart to have a strong pass percentage however not many of these are progressive passes which suggests she plays a simple pass to a defender. She also down not progressive the ball forward at all or intercepting. Emma Checker has the smallest area in this chart and ranks worst in terms of interceptions per 90.

## Potential New Defenders

Before moving onto the new potential defenders, I noticed that Melbourne Victory have a lower-than-average number of defenders. With only five defenders and playing a 4-3-3, there is little margin for a player to be out for injury or be rested to keep them fully fit. A suggestion would be to take two defenders from the list to add to the squad’s depth.

Now we have analysed the gaps in the Melbourne Victory squad, we move on to the new defenders. The management was unsure of which position to recruit. To work this out, I calculated the average composite score per position and visualised this on a jitter plot. With the composite being weighted off of what the team needed, this was an effective way to see which position performed best based on the metrics.

A graph with different colored dots

AI-generated content may be incorrect.

The jitter plot reveals that both left backs and right backs perform slightly better on average however the gaps are minimal. This allows the analysis to look at potentials spread across all positions of the back line and base the decision on the top composite score rather than having to filter out by specific position.

Similar to the Melbourne Victory defenders, a bar graph was created of the top 5 new potential defenders within budget to show who ranks best.

A graph of a number of players

AI-generated content may be incorrect.

Top of this ranking is Cecilia Salvai, the Italian centre-back who plays for Juventus. She played 180 minutes during the Women’s World Cup however played 1,387 minutes for her club in the Serie A. She picked up an injury in November however is almost back to full fitness and expected to return to the team. Katie Bowen ranks second on the list and is a New Zealand centre-back. She also plays in the Italian league, playing for Inter Milan and has played the majority of games this season so far. The link to New Zealand could be a benefit in this transfer.

A radar chart was also created to show the strengths and weaknesses of these players.

A diagram of a hexagon with colorful lines and numbers

AI-generated content may be incorrect.

Each player has a very similar radar plot however some have their added strengths. Salvai ranks highest in interceptions per 90 which is an important KPI for Melbourne Victory’s new defender and Noelle Maritz is a standout with the highest total progressive passes per 90 however does not have a very high pass percentage which would suggest that she attempts to play these passes a lot and they do work however the ball can also be given away a lot. CJ Bott leads on progressive carries per 90 which shows she is comfortable taking the ball forward and overall is well-rounded.

When buying a player, it is important to consider not only the quality of the player but also the cost of the player and whether they are worth the money. As noted by Andrews [13], “Because of the coronavirus pandemic and revenues down even at the richest clubs, finding the right player for the lowest price is fast becoming a necessity.” I calculated the value for money by dividing the market value by the composite score and the bar graph of the results is below.

A graph with different colored rectangular bars

AI-generated content may be incorrect.

This graph reveals that although the top 4 players rank the best, CJ Bott is an outlier. Coming in at 40,000 Euros (according to SoccerDonna [14]) compared to the top 4 who are 65,000 to 80,000, with a gap of only 0.06 in composite score, she provides the best value for money.

# Discussion

The findings from this analysis directly align with Melbourne Victory’s strategic goals of strengthening their defensive capabilities and improving their overall competitiveness. By identifying key performance gaps, particularly in interceptions and progressive ball play, the analysis highlights the defensive struggles that the team is facing. Players like Cecilia Salvai, Katie Bowen and CJ Bott would enhance the team’s ability to break down opposition plays and contribute to builder attacks, enabling a more balanced playing style with stability at the back.

Recruitment analytics in women’s football remains significantly behind men’s due to limited resources and incomplete data coverage. Websites like FBRef and Statsbomb and tools like WTDTool aim to bridge this gap by automating data collection across 30 leagues and approximately 12,000 players, generating alerts for players meeting predefined criteria. This helps address the inefficiencies in traditional scouting, where a lack of comprehensive metrics and league coverage has limited talent identification. Having these methods in place helps more teams and organisations harness the power of sports data analytics to help them make objective and data-driven decisions.

The methodology employed in this analysis has several strengths. The use of objective metrics provided a clear, data-driven approach which allows for reproducibility and removes subjective biases from the recruitment process. This approach also makes it possible to directly compare players across various KPIs. However, the main limitation of this analysis is the sole use of international data, particularly in terms of its direct applicability to the domestic league’s style and demands. Further research and additional data from player’s league stats could mitigate this limitation. Additionally, further research into how these players would interact with the current team would help with this analysis however is unavailable.

# Conclusion

In conclusion, the top 5 players found in this analysis are Cecilia Salvai (CB), Katie Bowen (CB), Noelle Maritz (LB), Jayde Riviere (RB) and CJ Bott (RB) based on their composite score and within the budget. 8 other players ranked above and between these players, however, were above budget. My top suggestion based on this analysis is Cecilia Salvai (CB) based on her defensive strengths, especially regarding interceptions which was the most important metric given by management. If the decision based off of my recommendation of increasing the squad’s depth was taken on board, my second recommendation would be CJ Bott (RB). For a low price tag, she offers a very high quality of defender and a different natural position as well.

Furthermore, this analysis highlights the value of data-driven recruitment and its growing role in supporting recruitment for teams like Melbourne Victory’s women’s team. It allows an objective way to identify key defensive weaknesses and recommend players to address these gaps to help them strengthen their squad and work towards regaining dominance.

# References

[1] R. De La Torre, L. O. Calvet, D. Lopez-Lopez, A. A. Juan, and S. Hatami, ‘Business Analytics in Sport Talent Acquisition: Methods, Experiences, and Open Research Opportunities’, Int. J. Bus. Anal., vol. 9, no. 1, pp. 1–20, Oct. 2021, doi: 10.4018/IJBAN.290406.

[2] M. D. Hughes and R. M. Bartlett, ‘The use of performance indicators in performance analysis’, J. Sports Sci., vol. 20, no. 10, pp. 739–754, Jan. 2002, doi: 10.1080/026404102320675602.

[3] J. J. Zhang and B. G. Pitts, Eds., The global football industry: marketing perspectives. in World association for sport management series. London ; New York, NY: Routledge, Taylor & Francis Group, 2018. [4] D. H. Kim et al., ‘Discovering primary indicators for evaluating defender’s technical performance using multivariate statistics in football games’, IJASSInternational J. Appl. Sports Sci., vol. 32, no. 1, pp. 27–36, Jun. 2020, doi: 10.24985/ijass.2020.32.1.27.

[5] D. Barreira, J. Garganta, J. C. Machado, and M. T. Anguera, ‘Repercussões da recuperação da posse de bola nos padrões de ataque de futebol de elite.’, Rev. Bras. Cineantropometria E Desempenho Hum., vol. 16, no. 1, pp. 36–46, Dec. 2013, doi: 10.5007/1980-0037.2014v16n1p36.

[6] B. Deb, J. Fernandez-Navarro, A. P. McRobert, and I. Jarman, ‘Finding repeatable progressive pass clusters and application in international football’, J. Sports Anal., vol. 9, no. 4, pp. 289–303, Feb. 2024, doi: 10.3233/JSA-220732.

[7] C. Pfaff et al., ‘Controlling ball progression in soccer’, Nov. 17, 2022, arXiv: arXiv:2210.16474. doi: 10.48550/arXiv.2210.16474.

[8] M. Hughes and I. Franks, ‘Analysis of passing sequences, shots and goals in soccer’, J. Sports Sci., vol. 23, no. 5, pp. 509–514, May 2005, doi: 10.1080/02640410410001716779.

[9] J. B. Taylor, S. D. Mellalieu, and N. James, ‘Behavioural comparisons of positional demands in professional soccer.’, Int. J. Perform. Anal. Sport, vol. 4, no. 1, pp. 81–97, Aug. 2004, doi: 10.1080/24748668.2004.11868294.

[10] S. G. K. Patro and K. K. Sahu, ‘Normalization: A Preprocessing Stage’, IARJSET, pp. 20–22, Mar. 2015, doi: 10.17148/IARJSET.2015.2305.

[11] C. Carling, J. Bloomfield, L. Nelsen, and T. Reilly, ‘The Role of Motion Analysis in Elite Soccer: Contemporary Performance Measurement Techniques and Work Rate Data’, Sports Med., vol. 38, no. 10, pp. 839–862, 2008, doi: 10.2165/00007256-200838100-00004.

[12] J. Barreira, ‘Age of Peak Performance of Elite Women’s Soccer Players’, 2016.

[13] C. Andrews, ‘Data-driven transfers are football’s new normal’, Eng. Technol., vol. 16, no. 8, pp. 104–110, Sep. 2021, doi: 10.1049/et.2021.0810.

[14] J. López-De-Armentia, ‘WTDTool: Women’s Talent Detection Tool’, in 2024 IEEE International Workshop on Sport, Technology and Research (STAR), Lecco, Italy: IEEE, Jul. 2024, pp. 144–149. doi: 10.1109/STAR62027.2024.10635969.

Appendixes

Questions to Management

1. What position and style of defender do you want?

Answer – We like to gain a progressive defender who is strong on the ball and able to make successful interceptions

1. What KPIs are most important for the defenders? Can you rank these KPIs as well for me please so I can understand what is most important to you. If wanting a centre-back and full-back, please can I know what is most important for each.

Answer – I have indicated above which factors I think are important in a player and I’m ot sure about weightings you can advice me on that. As for position, we would like you to assess which players are currently underperforming and our new recruit shold ideally be able to play in these positions.

1. What is the budget for these players (per player budget or total budget for 5 players?) and which leagues/nationality do you typically shop in?

Answer - £85k max for the defender, no geographical limit.