Age of Peak Performance among Soccer Players in Sweden

Rasmus Säfvenberg[1*], Anders Nordgaard[1 [0000-0001-9385-5443]], Ola Lidmark Eriksson[2], Niklas Carlsson[1 [0000-0003-1367-1594]], and Patrick Lambrix[1 [0000-0002-9084-0470]]

[1] Linköping University, Linköping, Sweden
[2] Football Analytics Sweden AB, Stockholm, Sweden

Abstract. Performance and age have an intertwined relationship in soccer. However, despite the sport's popularity and the great value of insight into this topic to decision-making personnel, this is a fairly unexplored topic, particularly for lower-ranked leagues. In this paper, we use a novel performance metric to investigate the peak age among soccer players in Sweden for goalkeepers, defenders, midfielders, and forwards. Each playing position is analyzed separately using univariate and bivariate methods and is also modeled by a hierarchical Bayesian model with player-specific age trajectories.
The results indicate an average overall peak age between 25 and 27. Forwards typically peak at 25, while defenders and midfielders reach their peak performance between the ages of 25 and 27. For goalkeepers, the peak generally occurs by age 27. The performance decline post-peak is the steepest among forwards and midfielders. Defenders and goalkeepers see a long-lasting and slow decline.

1 Introduction

In the global industry that soccer has become, it is important for teams to acquire the right players for the right position and for the right price. Much research in soccer analytics has recently been performed to define performance metrics to value a player's performance. Some examples are VAEP (Valuing Actions by Estimating Probabilities) [7], expected threat [22], action-value functions [18], plus-minus [15], and expected passes [2].

Studies have shown, however, that performance varies as a function of age and this relationship can be captured by an aging curve which tends to be inverted-U shaped with a peak, as well as possibly varying levels of slopes before and after the peak [3]. Insight into the aging curves and expected performance trajectories is, therefore, of interest for player acquisition [10], the development and improvement of training regimes and management of the squad [14].

* This is the author's version. It is posted here for your personal use. Not for redistribution. The definitive version was published in Proceedings of the First International Sports Analytics Conference and Exhibition (ISACE 2024).

Investigating the relationship between performance metrics and age has been performed in different sports such as tennis [16], baseball [4], basketball [26], and ice hockey [5]. Sports with a high emphasis on agility, power, and strength, tend to have athletes that peak in their early or mid-20s while finesse and endurance-based sports generally perform optimally in their late-20s.

Studies analyzing the age-performance relationship in *s*occer have utilized different metrics, methods and data for their analyses.

In [8], the relationship between age and performance for field players from the highest tiers in England, Germany, Italy, and Spain during five consecutive seasons, starting from 2010-2011, was investigated. As performance metric, the performance ratings from WhoScored.com were used. Two polynomial regression models with either population-level ("fixed") or group-level ("random") effects were developed. The peak age occurred between ages 25 and 27. Forwards peaked the earliest, followed by midfielders and defenders.

In [14], players in the UEFA Champions League during 1992–1993 to 2017–2018 and how their market valuation related to age, were analyzed. The variables used in the study were number of seasons in the club, number of Champions Leagues won, and team performance. For data analysis, one-way ANOVA and a linear regression were used. Their findings revealed peak market value in the 26-30 age group, with center backs and goalkeepers peaking later than attacking players.

The same authors investigated physical and technical performance among Bundesliga players during the 2012-2013 to 2014-2015 seasons [21]. The variables used in the study were total distance covered, number of fast runs, number of sprints, and percentage of successful passes. They observed that players aged 30 or older exhibited significantly reduced physical performance across all positions except wide midfielders, yet ability to make successful passes increased with age.

In another study, [19] used the nomination for individual awards such as the Ballon d'Or, as a proxy of performance. Data regarding 685 retired players from 1956 to 2019 was used. Using this proxy, the overall peak age ranges from 27 to 28, showing variations across positions, with forwards peaking earlier (26) and goalkeepers reaching their peak later (31).

As existing work in soccer on the age of peak performance of players has focused on top leagues in Europe, it is not clear whether these results can be used for lower tier leagues. Lower tier leagues may not be seen as the final destination for many players and thus the makeup of the teams in the lower leagues tends to consist of younger players aiming for top leagues, players for which the lower tier league is their highest achievable level, and end-of-career players that have played in top leagues earlier. As an example, we use the top Swedish league, Allsvenskan, which in November 2023 ranked 23$^{\text{rd}}$ among European leagues according to [25]. In Swedish soccer and similar nations, a notable trend involves players migrating to higher-ranked foreign leagues. For example, reports indicate that many Swedish players relocate abroad before turning 23 [9].

This paper estimates at which age male soccer players in Sweden perform optimally, and explores if findings from top-tier leagues align with a lower-tier league. We note, however, that previous approaches have used different metrics of performance, which do not fit a lower-tier league or for which data is not available. For instance, the features used in [21] are nowadays part of modern metrics. Further, there are no player ratings from WhoScored.com (as used by [8]) for Swedish leagues, few teams from the Swedish leagues play in the UEFA Champions League (as used in [14]), and players in the Swedish leagues would not receive individual awards such as the Ballon d'Or (as used in [19]). Therefore, we needed to use a metric that can be measured using the available data. As more and more game data is available nowadays, we decided to use event data and use modern performance metrics. The data we use for the analysis covers event data from the three highest Swedish leagues and is described in Sect. 2. We quantify player performance by constructing a novel metric that combines modern metrics and employing Bayesian hierarchical modeling for estimation (Sect. 3). For each position (goalkeepers, defenders, midfielders, and forwards) a separate estimation is performed. We present results in Sect. 4, a discussion in Sect. 5, and our conclusions in Sect. 6.

2 Data

In this paper, two datasets were used. The first and main dataset contains event data for the three highest leagues in the Swedish male soccer hierarchy between seasons 2015 and 2021, 2016 and 2021, as well as 2020 and 2021 for the first (Allsvenskan), second (Superettan), and third (Ettan Södra and Ettan Norra) highest levels, respectively. This data was provided by Football Analytics Sweden AB. The other dataset contains information about age and playing position and was gathered from Transfermarkt.com. In the data, there were unique identifiers for each game and event. The game-level data contained information regarding home and away teams, final score, and date, while the event-level data contained information about the event type, the player(s) and their teams involved in the event, as well as the result, body part, time, start- and end-coordinates of each event (Tables 1 and 2).

3 Method

The method consisted of three major steps: (1) data preprocessing, (2) computing a performance metric, and (3) estimating the peak age of performance.

3.1 Data preprocessing

As a first step, we transformed the data from its original long format representation (where information about an event could be distributed in different rows) to its representation in the Soccer Player Action Description Language (SPADL, [7]). Events not present in the SPADL representation were excluded. Based on the coordinates we added information about throw-ins and short corners. In addition, for the earliest seasons, the data did not record explicit events for corners

Table 1: Attributes in the SPADL data representation.

Attribute	Description
Match id	Unique id of the match.
Action id	Unique id of the action within the match.
Period id	The id of the period/half.
Seconds	Time of the action.
Player id	The id of the player performing the action.
Team id	The team id of the player.
X start	X coordinate at the start of the action.
Y start	Y coordinate at the start of the action.
X end	X coordinate at the end of the action.
Y end	Y coordinate at the end of the action.
Action	The action being performed by the player.
Result	The result of the action.
Body part	The part of the body used to perform the action.

and free kicks. Corner kicks were added based on the coordinates of the shot and free kicks based on the fact that they are preceded by a foul. Further, we needed to deal with some inconsistencies regarding minutes played as well as matching player names from the different data sources.

The players were divided into 4 distinct groups according to position: forwards (strikers, wingers), midfielders (wide, defensive, central, attacking midfielders), defenders (left, right, center backs), and goalkeepers. We removed players with less than 450 minutes played in a season (equivalent to 5 full games) to filter out players having extreme performances per minute while rarely playing. The filtering led to 308 observations (out of originally 535) for 101 goalkeepers, 1422 (originally 2089) for 422 defenders, 1328 (originally 2170) for 419 midfielders, and 974 (originally 1596) for 317 forwards.

The age of the players was recorded as a continuous variable and computed once per season with the date of measurement on June 1^{th} each calendar year, as this date approximately aligns with the beginning of the summer intermission in Swedish soccer. Fig. 1 shows the distribution of age per group.

3.2 Performance metric

In contrast to previous methods where a proxy or a few specific variables were used, we decided to define a performance metric based on several state-of-the-art and currently popular metrics.

The first component in our metric is based on the VAEP (Valuing Actions by Estimating Probabilities) framework [7] which aims to estimate offensive and defensive values of events to determine the impact on scoring or conceding a goal within the next 10 events. VAEP values are computed in two stages: one for the offensive aspect where the outcome is a Bernoulli random variable with a value of 1 if the team scores within the next 10 events and 0 otherwise, and another for the defensive aspect where another Bernoulli distributed random variable is the

Table 2: Action types in the SPADL representation. Adapted from [7].

Action type	**Description**	**Result**
Bad touch	Player makes a bad touch and loses the ball.	Fail, Own goal[a]
Clearance	Player clearance.	Success
Cross	Cross into the box.	Success[b], Fail, Offside
Crossed corner	Corner crossed into the box.	Success[b], Fail
Crossed free kick	Free kick crossed into the penalty box.	Success[b], Fail, Offside
Dribble	Player dribbles at least 3 meters with the ball.	Success
Foul	Foul.	Fail, Red card, Yellow card
Free kick shot	Direct free kick on goal.	Success[c], Fail
Goal kick	Goal kick.	Success
Goalkeeper claim	Keeper catches a cross.	Success[d], Fail
Goalkeeper pick-up	Keeper picks up the ball.	Success
Goalkeeper punch	Keeper punches the ball clear.	Success
Goalkeeper save	Keeper saves a shot on goal.	Success
Interception	Interception of the ball.	Success[e], Fail
Pass	Normal pass in open play.	Success[b], Fail, Offside
Penalty shot	Penalty shot.	Success[c], Fail
Short corner	Short corner.	Success[b], Fail
Short free kick	Short free kick.	Success[b], Fail, Offside
Shot	Shot attempt not from penalty or free kick.	Success[c], Fail
Tackle	Tackle on the ball.	Success[e], Fail
Take-on	Attempt to dribble past opponent.	Success[f], Fail
Throw-in	Throw-in.	Success[b], Fail

[a] Was originally the result of a shot but now is the result of a bad touch.
[b] Reaches teammate. [c] Goal. [d] Does not drop the ball. [e] Regains possession.
[f] Keeps possession.

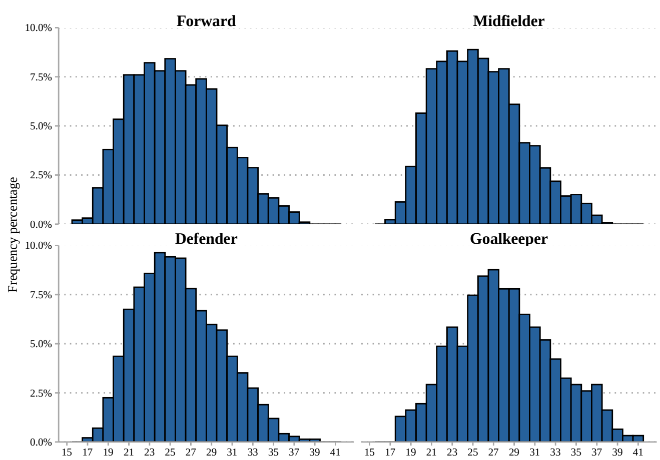

Fig. 1: Distribution of age by position.

outcome if a team concedes within the next 10 events (1) or not (0). The VAEP value is then the difference between offensive and negative defensive value. Since it is framed as a binary classification task, the choice of classification method is up to the user, and in this case, the choice is an eXtreme Gradient Boosting (XGBoost) model. As the VAEP framework estimates two separate probabilities, each aspect can be used as a metric. Furthermore, the offensive value in VAEP tends to be dominated by goal-scoring as they often receive probabilities close to 1 while the other probabilities lie closer to 0. Thus, it may be desirable to split the offensive value into two sub-components: shot offensive value and non-shot offensive value, since the components have different importance for different positions.

Another metric that has attained high popularity is xT (Expected Threat, method originally proposed by S Rudd, reproposed and coined by K Singh) [22], which is used to evaluate how a choice of pass impacts the team's probability of scoring. This is computed using Markov chains where the pitch is represented by a grid of 12×8 zones, where each zone is assigned a value based on how probable a goal is from the zone. The xT in a zone depends on the probability of scoring from that zone by a shot as well as on the xT of other zones to which the ball can be moved. In general, the most threatening areas on the pitch are within the penalty box of the opposing team. The threat of an action is defined as the difference in xT between the start zone and the end zone of the action.

As goalkeepers have a unique role in the game, we add a unique component to the performance evaluation. One suggested metric for this is that of post-shot expected goals (PSxG) [13], which evaluates how probable a goal is from a shot on target. The estimation of this value is done by utilizing information regarding, e.g., shot location, angle, direction, and goalkeeper position. Further, we also consider the difference between post-shot expected goals (PSxG) and actual goals allowed (GA), i.e., PSxG - GA. A positive value indicates that a goalkeeper saved more shots than expected.

After computing these metrics for each event in all games the values were then summed per player and season. The season-level values for each metric were then divided by the total minutes played during the season to adjust for players playing a different number of minutes during a season and then multiplied by 90 (length of a soccer game). We chose season-summarized statistics as individual games have larger variability than a full season. After computing the minute-adjusted metrics each respective metric was standardized to have mean zero and unit variance. Furthermore, players who had participated in only one season were removed as their peak age would be the age at that particular season.

Using the above metrics, we define a performance metric for goalkeepers y_{GK} and one for field players y_F as follows:

$$y_{GK} = \text{PSxG-GA} + \text{Non-shot offensive value} + \text{Defensive value} + \text{xT}, \quad (1)$$

$$y_F = \text{Shot offensive value} + \text{Non-shot offensive value} + \text{Defensive value} + \text{xT}. \quad (2)$$

For goalkeepers we do not use the shot offensive value as they rarely shoot the ball toward the other team's goal. For field players we do not use PSxG-GA.

3.3 Estimation

In this step we investigate the relationship between performance and age. As the ages of players were unevenly distributed, with fewer young and old players, we utilized a model with partial pooling, which uses shrinkage to combine information from both rare and common ages for a more informative estimation. We developed a mixed model consisting of performance as the outcome with a thin-plate smoothing spline as well as player-level intercepts and slopes. This choice of spline was based on its computational efficiency, the circumvention of the knot placement issues that tend to arise when fitting splines, and does not impact the inference and diagnostics of models [27]. The impetus behind the inclusion of player-level intercepts and slopes was that players are not expected to share the same peak or aging trajectory. Therefore, we allow for each player an intercept that adjusts the aging curve vertically and a slope that rotates and stretches the curve to match the trajectory [23]. More specifically, a hierarchical Bayesian model was implemented and is fit for each position p with sample size n_p as in Fig. 2, where s is a smoothing spline and (τ_μ, τ_σ) are the standard deviations of μ and σ, respectively, that model the variability and smoothness of the respective term. The half-Cauchy distribution corresponds to the right half of a symmetrical Cauchy distribution, while LKJcorr is the log-likelihood that, in this case, serves as a uniformly distributed prior (with default parameter value $\eta = 1$) for the correlation matrix $\mathbf{R}$ [17]. The priors for the intercepts α, both population and player-level, as well as the player-level slope β were set as generic weakly informative prior as described in [12]. Similarly, the priors for (τ_μ, τ_σ) were also set as weakly informative based on [11]. Finally, the priors for the player-level standard deviations $(\nu_{\alpha_\mu}, \nu_{\beta_\mu})$ were set to reflect the prior belief that, due to the data structure, with few observations per player and a larger set of players, a high standard deviation is less likely, and thus more centered near the lower limit of zero. The models were estimated using four Markov chains, each with 20 000 Hamiltonian Monte Carlo iterations and adapt delta of 0.99 from the brms package [6].

4 Results

With the model structure previously defined, we obtain both a population-level smooth spline for age as well as player-level intercept and slope that can then be used to estimate the peak age. Specifically, each player has an age where their performance is at its maximum given by the season in which they participated. Before showcasing the results of the model, we should also shed some light on the quality of the model. In particular, we can investigate the posterior predictive distribution of each model and verify if draws from the distribution align with the original data. This is visualized in Fig. 3 where 100 such draws from each model are compared to the data. From the respective distributions, we can see that each model captures the general structure of the original data, although some deviations can be observed. The draws for midfielders and defenders are most similar to the original data, while the data for forward and goalkeepers

$$
\begin{aligned}
y_i &\sim \text{Normal}(\mu_i, \sigma_i), \quad i = 1, \ldots, n_p \\
\mu_i &= \alpha_{\mu,\text{player}[i]} + \beta_{\mu,\text{player}[i]} + s(\text{Age}_i, \tau_\mu) \\
\sigma_i &= \alpha_\sigma + s(\text{Age}_i, \tau_\sigma) \\
\begin{bmatrix} \alpha_{\mu,\text{player}} \\ \beta_{\mu,\text{player}} \end{bmatrix} &\sim \text{MVNormal}\left(\begin{bmatrix} \alpha_\mu \\ \beta_\mu \end{bmatrix}, \mathbf{\Sigma}_\mu \right) \\
\mathbf{\Sigma}_\mu &= \begin{bmatrix} \nu_{\alpha_\mu} & 0 \\ 0 & \nu_{\beta_\mu} \end{bmatrix} \mathbf{R}_\mu \begin{bmatrix} \nu_{\alpha_\mu} & 0 \\ 0 & \nu_{\beta_\mu} \end{bmatrix} \\
(\alpha_\mu, \alpha_\sigma, \beta_\mu) &\sim \text{Normal}(0, 1) \\
(\tau_\mu, \tau_\sigma) &\sim \text{HalfCauchy}(0, 1) \\
\nu_{\alpha_\mu} &\sim \text{HalfCauchy}(0, 0.5) \\
\nu_{\beta_\mu} &\sim \text{HalfCauchy}(0, 0.1) \\
\mathbf{R}_\mu &\sim \text{LKJcorr}(1)
\end{aligned}
$$

Fig. 2: Model.

is somewhat lower than the simulations suggest. Some bends in the curvature can also be noted for goalkeepers for the higher positive values. In general, these findings indicate that the models capture the overall performance pattern present in the data for each position group.

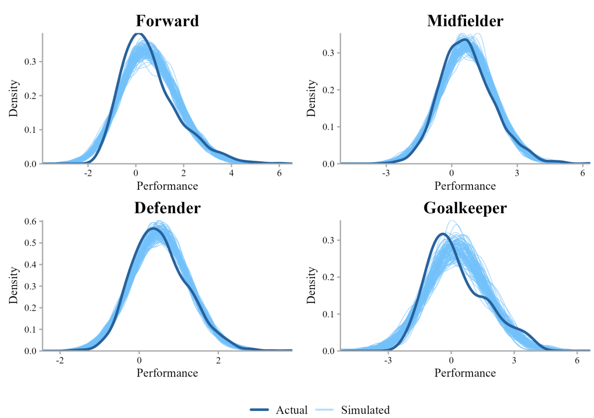

Fig. 3: Comparison of 100 simulated draws from the posterior predictive distribution and the performance for each position.

The conditional effects for each position can be seen in Fig. 4 while the summary statistics of the player-level fits are described in Table 3. From the figure, it can be discerned that the aging curve varies by position, with forwards having the sharpest decline post-peak while defenders have the slowest decline. Among forwards, the peak was found around age 25 which is preceded by a steeper in-

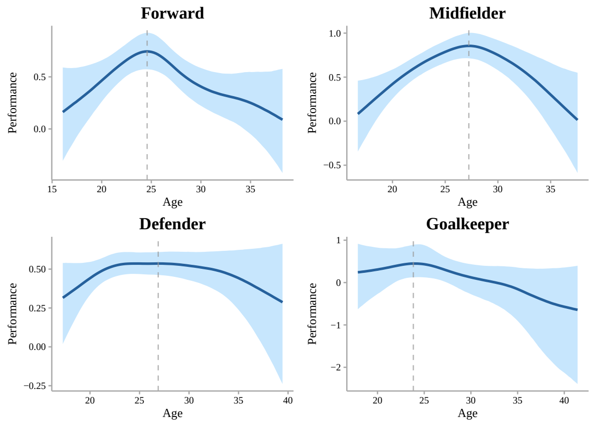

Fig. 4: Conditional effects of age by position. The vertical lines indicate the age of peak performance.

crease toward the peak and then followed by a sharp decrease in performance after the apex. In addition, the post-peak slope decreases after age 30, although this is coupled with an increasing standard error. Midfielders and defenders attain a later peak than forwards, between 26 and 28 years of age. The aging curve of midfielders exhibits a shape akin to a second-degree polynomial, where the slope before and after the maximum is similar. For defenders, the aging curve remains largely constant for many years, with a near-horizontal slope starting at age 23 and ending at age 30. Moreover, the performance curve for defenders has the lowest range, although the standard error is higher. Goalkeepers achieved the earliest peak, by age 24 according to the smoothing splines, which is then followed by a fairly constant decrease.

Table 3: Summary statistics of peak age from the hierarchical model.

	Percentile			**Mean**	**Std. Dev.**
	25^{th}	50^{th}	75^{th}		
Forwards	23.0	24.6	27.2	25.1	3.25
Midfielders	22.9	25.8	27.4	25.4	3.15
Defenders	23.1	24.7	27.8	25.6	3.43
Goalkeepers	23.9	25.6	29.1	26.6	3.95

When considering the summary statistics, it can be noted that the mean age of peak performance was between 25.1 and 25.6 for field players and 26.6 for goalkeepers. The quantiles also range from approximately 23 to 28 for field players while goalkeepers are approximately between 24 and 29. Moreover, the standard deviation obtained is somewhat high relative to the mean, mainly due to the players available in the data having a large dispersion between the youngest and oldest players for each position as well as few observations per player. Moreover, since each player is given an individual maximum point and there are players in the data with only a few seasons, the age at which they peak may not be their actual peak as it may lay outside the range of the recorded data. For instance, the data contains players for which the observations cover ages 16-18 or 35-40. Consequently, the peak age for these players may be far after or before their actual peak, which in turn increases the standard deviation seen in Table 3.

5 Discussion

The results of this paper indicate that professional soccer players in Sweden peak between the ages of 24 and 27, which also varies by playing position. Interestingly, although different performance metrics were used and, as explained in the introduction, the situation for top leagues and lower-tier leagues differs, the peak ages seem to be similar for the Swedish league and the results reported for top leagues in previous work.

We note that the results stand in agreement with the general physical and skill-based profile of athletes. Soccer is a mixed-skill sport with high requirements on both physical and physiological abilities. With the high endurance expected of soccer players, where field players on average run 10-12 kilometers and goalkeepers 4 kilometers, at an intensity near the anaerobic threshold of 80-90% of maximal heart rate [24], a tendency towards higher peak age may be expected. However, as performance in soccer also relies on aspects such as high-intensity running, jumping, and tackling, this should in turn lead to a lower peak age, as reported in sports such as track and field as well as swimming [8]. Moreover, the apparent differences between positions also highlight different expectations and emphases. In particular, forwards had a rapid decline post-peak while defenders and goalkeepers had a lesser decline. This can be explained by the higher number of high-intensity sprints and duels forwards engage in compared to, e.g., defenders [21]. Similarly, the length of near-peak performance of defenders also highlights a possible shift in playing style where more emphasis is placed on anticipation, positioning, and decision-making, which are all gained through experience rather than physical ability. The leverage of experience is what can enable a player to maintain high performance after their peak. However, there is a trade-off between gained experience and a decline in physiological capacity that typically results in a dominance of physiology rather than experience. This is particularly evident among forwards. Perhaps somewhat peculiar was that the peak age among goalkeepers occurred by age 25 according to the conditional age effect, though the mean peak age was also found at 26.6. As [20] notes, goalkeepers are the position with the lowest physiological expectations and risk of injury, which tends to lead to longer playing careers and can thus lead to longer periods of near-peak performance. However, this was not the case for our performance metric.

There are some limitations for this study. First, as a team sport, it is difficult to fully isolate and evaluate players without team and teammate quality impacting the results. Next, the data does not consist of information about full careers for the players. Instead, some players only have a few, e.g., one or two seasons where data is recorded, potentially giving a false representation of their actual peak. Specifically, as some players only have data for low (below 20) or high ages (above 35), the variance of the results will subsequently be inflated with each player being given a peak age. This also ties in with the stature of the leagues investigated, since the Swedish soccer leagues belong to the lower tier, the best players tend to be sold to clubs in higher-ranked leagues, leading to potential bias. Similarly, older players playing in higher-ranked leagues may choose to move to a lower-ranked league in the final year(s) of their careers, which also may skew the results. This may in turn also impact the modeling capability, as it will not be able to fully recover each player's aging trajectory due to a lack of data. Lastly, the positional grouping used here may not fully represent the intricacies within a team as the same position in one team can have vastly different roles and expectations in another team. For instance, a wing-back for a worse team may be expected to mainly contribute defensively while for a top team, the expectations may be more shifted toward offensive contribution.

6 Conclusion

This paper presents a study on age of peak performance for players in the top three Swedish leagues where performance is measured based on state-of-the-art metrics. As previous work dealt with top leagues, the Swedish leagues were chosen as representatives of a lower-tier league that has different prerequisites, in terms of large player turnover and different transfer patterns, than the top leagues. The results suggest that soccer players in these leagues attain peak performance between the ages of 25 and 27. For the given playing positions, forwards had the earliest peak at age 25, followed by midfielders and defenders. Goalkeepers had the highest average peak age.

Future work may take into account a more fine-grained classification of the players based on player roles. However, finding the right number of roles (e.g., [1] proposes 21 roles) and still retaining enough data per role may be a challenge. Another interesting problem is to try to relate performances in different leagues which would allow us to use data from different leagues in the computation of peak performances. Also, the use of other performance metrics for lower-tier leagues is an avenue for further work, as well as applying our metric to top leagues.

References

1. Aalbers, B., Haaren, J.V.: Distinguishing between roles of football players in play-by-play match event data. In: 5th Workshop on Machine Learning and Data Mining for Sports Analytics. pp. 31–41 (2018). https://doi.org/10.1007/978-3-030-17274-9_3

2. Anzer, G., Bauer, P.: Expected passes. Data Mining and Knowledge Discovery **36**(1), 295–317 (2022). https://doi.org/10.1007/S10618-021-00810-3
3. Berthelot, G., Johnson, S., Noirez, P., Antero, J., Marck, A., Desgorces, F.D., Pifferi, F., Carter, P.A., Spedding, M., Manoux, A.S., Toussaint, J.F.: The age-performance relationship in the general population and strategies to delay age related decline in performance. Archives of Public Health **77**, 51 (2019). https://doi.org/10.1186/s13690-019-0375-8
4. Bradbury, J.C.: Peak athletic performance and ageing: Evidence from baseball. Journal of Sports Sciences **27**(6), 599–610 (2009). https://doi.org/10.1080/02640410802691348
5. Brander, J.A., Egan, E.J., Yeung, L.: Estimating the effects of age on NHL player performance. Journal of Quantitative Analysis in Sports **10**(2), 241–259 (2014). https://doi.org/10.1515/jqas-2013-0085
6. Bürkner, P.C.: brms: An R package for Bayesian multilevel models using Stan. Journal of statistical software **80**, 1–28 (2017). https://doi.org/10.18637/jss.v080.i01
7. Decroos, T., Bransen, L., Van Haaren, J., Davis, J.: Actions speak louder than goals: Valuing player actions in soccer. In: 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 1851–1861 (2019). https://doi.org/10.1145/3292500.3330758
8. Dendir, S.: When do soccer players peak? a note. Journal of Sports Analytics **2**(2), 89–105 (2016). https://doi.org/10.3233/JSA-160021
9. Eliasson, A.: The European football market, globalization and mobility among players. Soccer & Society **10**(3-4), 386–397 (2009). https://doi.org/10.1080/14660970902771449
10. Frick, B.: Performance, salaries, and contract length: Empirical evidence from german soccer. International Journal of Sport Finance **6**(2), 87–118 (2011)
11. Gelman, A.: Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Analysis **1**(3), 515–534 (2006). https://doi.org/10.1214/06-BA117A
12. Gelman, A., Vehtari, A., Simpson, D., Margossian, C.C., Carpenter, B., Yao, Y., Kennedy, L., Gabry, J., Bürkner, P.C., Modrák, M.: Bayesian workflow. arXiv:2011.01808 (2020), https://doi.org/10.48550/arXiv.2011.01808
13. Goodman, M.: A new way to measure keepers' shot stopping: Post-shot expected goals (2018), https://statsbomb.com/articles/soccer/a-new-way-to-measure-keepers-shot-stopping-post-shot-expected-goals/
14. Kalén, A., Rey, E., Sal de Rellán-Guerra, A., Lago-Peñas, C.: Are soccer players older now than before? aging trends and market value in the last three decades of the uefa champions league. Frontiers in Psychology **10**, 76 (2019). https://doi.org/10.3389/fpsyg.2019.00076
15. Kharrat, T., McHale, I.G., Peña, J.L.: Plus–minus player ratings for soccer. European Journal of Operational Research **283**(2), 726–736 (2020). https://doi.org/10.1016/j.ejor.2019.11.026
16. Kovalchik, S.A.: The older they rise the younger they fall: Age and performance trends in men's professional tennis from 1991 to 2012. Journal of Quantitative Analysis in Sports **10**(2), 99–107 (2014). https://doi.org/10.1515/jqas-2013-0091
17. Lewandowski, D., Kurowicka, D., Joe, H.: Generating random correlation matrices based on vines and extended onion method. Journal of Multivariate Analysis **100**(9), 1989–2001 (2009). https://doi.org/10.1016/j.jmva.2009.04.008

18. Liu, G., Luo, Y., Schulte, O., Kharrat, T.: Deep soccer analytics: learning an action-value function for evaluating soccer players. Data Mining and Knowledge Discovery **34**(5), 1531–1559 (2020). https://doi.org/10.1007/S10618-020-00705-9
19. Oterhals, G., Lorås, H., Pedersen, A.V.: Age at nomination among soccer players nominated for major international individual awards: A better proxy for the age of peak individual soccer performance? Frontiers in Psychology **12**(661523) (2021). https://doi.org/10.3389/fpsyg.2021.661523
20. Perez-Arroniz, M., Calleja-González, J., Zabala-Lili, J., Zubillaga, A.: The soccer goalkeeper profile: bibliographic review. The Physician and Sportsmedicine pp. 1–10 (2022). https://doi.org/10.1080/00913847.2022.2040889
21. Sal de Rellán-Guerra, A., Rey, E., Kalén, A., Lago-Peñas, C.: Age-related physical and technical match performance changes in elite soccer players. Scandinavian Journal of Medicine & Science in Sports **29**(9), 1421–1427 (2019). https://doi.org/10.1111/sms.13463
22. Singh, K.: Introducing expected threat (xt) (2019), https://karun.in/blog/expected-threat.html, (Accessed: 3 February 2022).
23. Sóskuthy, M.: Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction. arXiv:1703.05339 (2017), https://doi.org/10.48550/arXiv.1703.05339
24. Stølen, T., Chamari, K., Castagna, C., Wisløff, U.: Physiology of soccer. Sports Medicine **35**(6), 501–536 (2005). https://doi.org/10.2165/00007256-200535060-00004
25. UEFA.com: Country coefficients (2023), https://www.uefa.com/nationalassociations/uefarankings/country/#/yr/2024
26. Vaci, N., Cocić, D., Gula, B., Bilalić, M.: Large data and bayesian modeling - aging curves of nba players. Behavior Research Methods **51**(4), 1544–1564 (2019). https://doi.org/10.3758/s13428-018-1183-8
27. Wood, S.N.: Thin plate regression splines. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **65**(1), 95–114 (2003). https://doi.org/10.1111/1467-9868.00374