Volume 07 Issue 01 (2026) Pages 09-14

INSPIREE: INDONESIAN SPORT INNOVATION REVIEW

ISSN 2746-6965 (Online), 2774-2520 (Print)
Journal Homepage: https://inspiree.review/index.php/inspiree

ORIGINAL RESEARCH ARTICLES

OPEN ACCESS

Association Between Expected Goals (xG) and Goal Scoring Across Club and National Team Competitions in Elite Football

https://doi.org/10.53905/inspiree.v7i01.161

Cristian Murillo García[*1abcde]

[1]Corporación Universitaria del Caribe - CECAR, Colombia.

A B S T R A C T

The purpose of the study. The purpose of this study was to analyze the association between xG and goals scored in soccer matches.
Materials and methods. This study was quantitative, correlational, and longitudinal. 333 league matches were analyzed. The dataset was obtained from Sofascore, a website that offers statistics and results from various sporting events. The hypothesis was put forward that there is a positive correlation between xG and goals scored. The sample consisted of 333 official matches from international competitions, including the 2013/24 Champions League, the 2024 Copa Libertadores, Euro 2024, and the 2024 Copa América. These tournaments were selected due to their significant international relevance.
Results. The results of this research confirm the validity of using xG as an indicator for analyzing a team's offensive performance, although its predictive capacity tends to vary depending on the context or competition. In clubs, players have more continuity because they train together, which makes it easier to establish a playing pattern, so this may be a reason for higher xG compared to national team tournaments.
Conclusions. The results obtained in this study confirm the relationship between xG and goals scored in football matches, thus establishing it as a valid indicator for measuring offensive performance in football. It was found that in club tournaments, the amount of association, in the Copa Libertadores (r = 0.537) having the best association between the variables, in the Champions League the association was lower (r = 0.403). Meanwhile, in national team tournaments, the values for the Copa América and Euro were (r = 0.475) and (r = 0.479), respectively, where some similarity can be observed in the association of variables, which is attributed to the poor group cohesion and poor tactical fit of the national teams compared to the clubs.

Keywords: expected goals (xg); football performance analysis; offensive performance indicators; goal scoring efficiency; match analytics; elite football competitions.

A R T I C L E I N F O

EDITED BY
Dr. Samsurijal Hasan, MM
Universitas Pahlawan, Indonesia..

Assoc. Prof. Herli Pardilla, Ed.D.
STOK Bina Guna, Indonesia.

ARTICLE HISTORY
Received : September 27, 2025
Accepted : January 20, 2026.
Published: January 27, 2026.

CITATION
Murillo García, C. (2026). Association Between Expected Goals (xG) and Goal Scoring Across Club and National Team Competitions in Elite Football. *INSPIREE: Indonesian Sport Innovation Review*, 7(01), 09-14.
https://doi.org/10.53905/inspiree.v7i01.161

INTRODUCTION

Football is the most recognized sport in the world and one of the highest-grossing sports for FIFA. In the 2022/23 season, the English Premier League reached approximately three billion pounds in broadcast and commercial rights (Bradlow et al., 2017). This significant financial investment requires an approach to performance optimization using innovative methods (Davis & Robberechts, 2020). As the sport's global growth increases, so does the need for strategies and technological tools to support team and player performance.

As football's global growth increases, so does the need to implement technological tools for group and collective performance. Thus, offensive performance analysis has taken on significant importance and evolution through the use of metrics, including expected goal difference (xG), which has emerged as a tool to evaluate the quality of chances generated by a player or team (Fernández-Navarro et al., 2019).

xG measures the probability of a shot resulting in a goal. It is based on variables such as shot position, distance from the goal, and opposing defensive pressure (Hoernig & Link, 2017). Furthermore, this tool allows for the analysis of a team's offensive efficiency and comparison of expected and actual performance, thus gaining strength as a valid tool for coaches, analysts, and sports scientists (Hoernig & Link, 2017).

Despite the rise of xG as an analytical tool, many questions remain about its predictive capacity in different competitive

[abcde]Authors'Contribution: a-Study design; b-Data collection; c-Statistical analysis; d-Manuscript preparation; e-Funds collection.
Corresponding Author: Cristian Murillo García, e-mail: cristian.murillog@cecar.edu.co

© 2026 The Author. This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

contexts. One study found that the association between xG and goals scored depends on the tournament, level, and dynamics of the game (Gómez et al., 2018).

On the other hand, advances in sports analysis have led to the development of metrics that help teams manage player performance. A metric such as expected goals (xG) has become an integral part of clubs' decision-making processes (Zeng & Pan, 2021). However, a comprehensive approach to evaluating players and teams is lacking, as well-known metrics do not provide a complete view of player contributions or team effectiveness (Mead et al., 2023).

Consequently, this gap stems from the absence of crucial characteristics and the inherent limitations of xG, performance analysts need to recognize the importance of context; relying solely on one metric for statistical analysis can be detrimental (Brechot & Flepp, 2020). While those using xG recognize its advantages, it is also necessary to acknowledge its limitations. While the data used in any model is important, how that data is interpreted is vital. A comprehensive understanding that encompasses both the advantages and disadvantages of xG is necessary for maximum development (Anzer & Bauer, 2021).

Although football is a very popular sport, historically there hasn't been much of a connection between sport and statistical data analysis. However, the xG metric mitigates this by using statistical methods to estimate the probability of a shot resulting in a goal (Brechot & Flepp, 2020). Furthermore, studies have explored the application of xG models to incorporate different characteristics, based on two key variables: distance from goal and shot angle (Antipov & Pokryshevskaya, 2020). One study found that using statistics from the 2012/13 English and German league seasons, the study involved dividing the pitch into eight zones and analyzing the probability of scoring based on shots taken from each zone. The results showed that distance and angle significantly influence the probability of scoring a goal (Bradlow et al., 2017).

Similarly, another study aimed to investigate the impact of shooting distance and angle. This study was based on events presented in a football league during the 2017/18 season. This study used a probabilistic approach to analyze goal-scoring opportunities. This study highlighted the complexity and potential of xG modeling in the analysis (Anzer & Bauer, 2021). Shot type must be taken into account, which provides contextual information about the shots. Shots can be divided into: body segment used (right foot/left foot or head) and the game situation in which the shot occurred. The authors incorporated these characteristics into their model and found that both aspects influence shot effectiveness (Fairchild et al., 2018). Furthermore, they found that free kicks and penalties are more likely to result in goals compared to shots from open play and kicks from feet; conversely, headers are less likely to result in goals (Schulze et al., 2017). Based on the studies reviewed above, the hypothesis was proposed: there is a significant relationship between xG and goals scored in matches. The purpose of this study was to analyze the association between xG and goals scored in soccer matches.

MATERIALS AND METHODS

This study was quantitative, correlational, and longitudinal. 333 league matches were analyzed. The dataset was obtained from Sofascore, a website that offers statistics and results from various sporting events. The hypothesis was put forward that there is a positive correlation between xG and goals scored.

The sample consisted of 333 official matches from international competitions, including the 2013/24 Champions League, the 2024 Copa Libertadores, Euro 2024, and the 2024 Copa América. These tournaments were selected due to their significant international relevance.

Study Design

Two main variables were analyzed:

Independent variable:

xG, which corresponds to a metric used to estimate the probability of a shot ending in a goal, which is based on factors such as: distance, shot angle, defensive pressure, assist, among others.

Dependent variable:

Goals scored, the goals achieved in each match. The data used in this research was taken from the SofaScore platform, which provides data for analyzing the performance of certain sports and also real-time statistics.

Statistical Analysis

For data analysis, SPSS statistical software was used, the matches, goals scored and xG of each match were processed, followed by the creation of a table of paired samples to determine if there were significant differences between the means and a correlation to indicate the degree of relationship of the variables used. The data was obtained from match logs available on the SofaScore website. It was then exported to a database for cleaning.

Ethical Considerations

This study utilized secondary data derived exclusively from publicly accessible match statistics available on the SofaScore platform and did not involve direct interaction with human participants, biological samples, or the collection of identifiable personal information. Therefore, the study was classified as minimal-risk research. Ethical clearance was obtained from the Research Ethics Committee (Institutional Review Board) of Corporación Universitaria del Caribe (CECAR), Colombia, under approval number: CEC-IRB/SPORTSCI/2025/091. All research procedures were conducted in accordance with institutional guidelines and internationally accepted ethical standards for research, ensuring transparency, data integrity, and responsible use of open-access data.

RESULTS

The table shows the results of the average xG and the number of goals scored in each competition, which shows that there

© 2026 The Author. This article is licensed CC BY SA 4.0.
Visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Cristian Murillo García.

are no significant differences in the means (Table 1). On the one hand, the xG in the Copa Libertadores averaged 2.30 ± 0.95, on the goals side the average was 2.40 ± 1.43, in the Champions League the average xG was 3.00 ± 1.04 and the average number of goals was 2.99 ± 1.70. Turning to the national team tournaments, the Copa América had an average of 2.41 ± 0.84, but with a lower scoring average (2.18 ± 1.55), in the Euro an average of 2.61 ± 0.91 was obtained in terms of xG and the goal average was 2.29 ± 1.36.

Table 1. Paired Samples Statistics

Pair	Variable	Mean	N	Std. Deviation	Std. Error Mean
1	xG Copa Libertadores	2.30	125	0.95	0.08
	Libertadores Goals	2.40	125	1.43	0.12
2	xG Champions League	3.00	125	1.04	0.09
	Champions Goals	2.99	125	1.70	0.15
3	xG Copa América	2.41	32	0.84	0.14
	Copa América Goals	2.18	32	1.55	0.27
4	xG Euro	2.61	51	0.91	0.12
	Euro Goals	2.29	51	1.36	0.19

In the table you can see the results of the average xG and the number of goals scored.

Table 2. Paired sample correlations

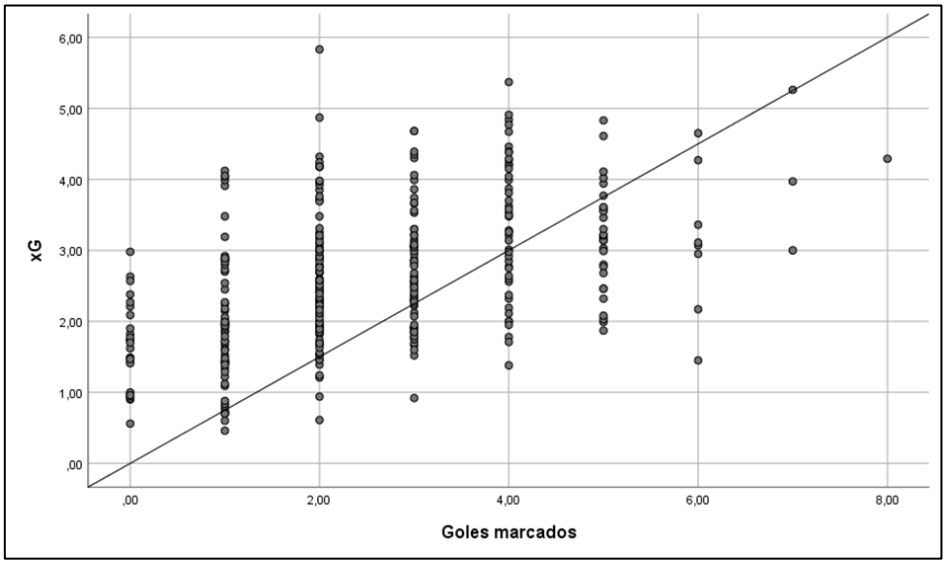
Pair	Variables	N	Correlation	Sig. (2-tailed)
1	xG Copa Libertadores & Libertadores Goals	125	0.537	< .001
2	xG Champions League & Champions League Goals	125	0.403	< .001
3	xG Copa América & Copa América Goals	32	0.475	.006
4	xG Euro & Euro Goals	51	0.479	< .001

The table indicates a positive correlation. The asymptotic significance had a value <0.00, so the researcher's hypothesis is accepted, in other words, it is accepted that the higher the xG, the greater the number of goals scored. In addition, a correlation coefficient of r = 0.488 was obtained, which indicates that there is a moderate and positive relationship between the two variables analyzed. The significance (p = 0.000) confirms that it is significant, indicating that xG is a valid predictor to predict the number of goals in the sample analyzed. Despite being a moderate correlation, since it is not high, it is stated that there are other factors that influence the scoring of goals, such as the quality of the shot, the quality of the goalkeeper, the performance of the defense, among others.

Table 3. Correlation Between Expected Goals (xG) and Goals Scored

Variable		xG	Goals Scored
xG	Pearson Correlation	1.000	0.488**
	Sig. (2-tailed)	—	< .001
	N	333	333
Goals Scored	Pearson Correlation	0.488**	1.000
	Sig. (2-tailed)	< .001	—
	N	333	333

In the graph, you can see the clear positive linear relationship in the data, as the xG increased, so did the number of goals scored. Therefore, it can be established that the number of goals scored was directly related to the xG or that they were directly proportional variables (Figure 1). The dispersion indicates that, although there is a relationship between the two variables analyzed, it is not entirely perfect. There are points above and below the regression line, indicating that some clubs or national teams exceeded or failed to achieve the xG in some matches. This can be attributed to factors such as a lack of finishing, the performance of the opposing goalkeeper, or poor or excessive effectiveness in creating scoring opportunities.



© 2026 The Author. This article is licensed CC BY SA 4.0.
Visit Creative Commons Attribution-ShareAlike 4.0 International License.

Figure 1. Simple dispersion

The results of this research confirm the validity of using xG as an indicator for analyzing a team's offensive performance, although its predictive capacity tends to vary depending on the context or competition. In clubs, players have more continuity because they train together, which makes it easier to establish a playing pattern, so this may be a reason for higher xG compared to national team tournaments.

DISCUSSION

The study on the modern statistical metric of expected goals (xG) and its impact on goals scored yielded several significant findings (Khrapach & Siryi, 2024). One of the main results of this research was the strong association between goals scored and xG (Khrapach & Siryi, 2024). One study agreed with this one's findings, as teams with higher xG averages had higher goal averages over a season (Sabin, 2021).

On the other hand, a study determined that xG predicts a team's performance with a high level of reliability, where teams that had high xG in previous matches tended to score a large number of goals in subsequent matches, so this tool helps coaches and analysts to carry out strategies for decision-making at a tactical level (Mead et al., 2023). Another study indicated that due to the association of xG with goals scored, coaches could use the information provided by xG to adjust their offensive tactical work, planning sessions where players are helped to find more advantageous positions to score (Zi & Gao, 2023).

The xG is a useful indicator to indicate the number of goals scored by teams, this helps to improve individual and collective performance objectively (Anzer & Bauer, 2021). In training sessions, coaches can use xG data to design drills that simulate high xG scenarios (Cavus & Biecek, 2022). For example, creating drills that focus on finishing from central positions inside the penalty area can help players develop their ability to convert high-quality chances. Additionally, xG data can be used to assess performance, providing clear, data-driven goals to work toward in training.

Although xG is a useful predictor of offensive performance, there are other factors that can influence scoring in a match. Some factors that can influence goals are:

Individual quality:

Teams with outstanding forwards have a greater chance of exceeding their xG, on the other hand, the lower the quality of the players, the greater the difficulty in scoring.

Opposing goalkeeper performance:

When a goalkeeper performs well, he or she is characterized by his or her ability to prevent goals against him or her.

Defensive tactics:

Teams that place a high priority on defensive solidity are able to reduce their opponent's options and therefore their xG.

The findings of this research have practical applications for analyzing performance and making tactical decisions. Coaches and analysts can use the xG metric not only to evaluate the offensive performance of their own or opposing teams, but also to identify areas for improvement in finishing plays and creating scoring opportunities.

CONCLUSION

The results obtained in this study confirm the relationship between xG and goals scored in football matches, thus establishing it as a valid indicator for measuring offensive performance in football. It was found that in club tournaments, the amount of association, in the Copa Libertadores ($r = 0.537$) having the best association between the variables, in the Champions League the association was lower ($r = 0.403$). Meanwhile, in national team tournaments, the values for the Copa América and Euro were ($r = 0.475$) and ($r = 0.479$), respectively, where some similarity can be observed in the association of variables, which is attributed to the poor group cohesion and poor tactical fit of the national teams compared to the clubs. The use of xG is recommended for analyzing teams' offensive performance. Another point worth highlighting is the variation in xG depending on the tournament. Therefore, it is recommended that future studies incorporate more variables to better understand goal scoring.

CONFLICT OF INTEREST

The author declares no conflicts of interest.

REFERENCES

Antipov, E., & Pokryshevskaya, E. (2020). Interpretable machine learning for demand modeling with high-dimensional data using gradient boosting machines and shapley values. J. Rev. Pricing Manage, 19(7), 355-364. https://doi.org/10.1057/s41272-020-00236-4

Anzer, G., & Bauer, P. (2021a). A Goal Scoring Probability Model for Shots Based on Synchronized Positional and Event Data in Football (Soccer). Front. Sports Act. Living, 3(62), 75-82. https://doi.org/10.3389/fspor.2021.624475

Anzer, G., & Bauer, P. (2021b). A goal scoring probability model for shots based on synchronized positional and event data in football (soccer). Frontiers in sports and active living, 3(21), 624-475. https://doi.org/10.3389/fspor.2021.624475

Bradlow, E., Gangwar, M., Kopalle, P., & Voleti, S. (2017). The role of big data and predictive analytics in retailing. Journal of Retailing, 5(93), 79-95. https://doi.org/10.1016/j.jretai.2016.12.004

Brechot, M., & Flepp, L. (2020a). Dealing with randomness in match outcomes: How to rethink performance evaluation in European club football using expected goals. J. Sport. Econ, 4(21), 335-362. https://doi.org/10.1177/15270025198979

© 2026 The Author. This article is licensed CC BY SA 4.0.
Visit Creative Commons Attribution-ShareAlike 4.0 International License.

Brechot, M., & Flepp, R. (2020b). Dealing With Randomness in Match Outcomes: How to Rethink Performance Evaluation in European Club Football Using Expected Goals. Journal of Sports Economics, 21(4), 335-362. https://doi.org/10.1177/1527002519897962

Cavus, M., & Biecek, P. (2022). Explainable expected goal models for performance analysis in football analytics. Data Science and Advanced Analytics, 4(9), 1-9. https://doi.org/10.1109/DSAA54385.2022.10032440

Davis, J., & Robberechts, P. (2020). How data availability affects the ability to learn good xG models. International Workshop of Machine Learning and Data Mining for Sports Analytics (Ghent), 13(24), 17-27. https://doi.org/10.1007/978-3-030-64912-8_2

Fairchild, A., Pelechrinis, K., & Kokkodis, M. (2018). Spatial analysis of shots in MLS: A model for expected goals and fractal dimensionality. Journal of Sports Analytics, 4(3), 165-174. https://doi.org/10.3233/JSA-170207

Fernández-Navarro, J., Fradua, L., Zubillaga, A., Ford, P., & McRobert, A. (2019). Influence of contextual variables on styles of play in soccer. International Journal of Performance Analysis in Sport, 19(6), 747-759. https://doi.org/10.1080/24748668.2019.1684773

Gómez, A., Mitrotasios, M., Armatas, V., & Lago-Peñas, C. (2018). Analysis of playing styles according to team quality and match location in Greek professional soccer. International Journal of Performance Analysis in Sport, 18(3), 986-997. https://doi.org/10.1080/24748668.2018.1539382

Hoernig, M., & Link, D. (2017). Individual ball possession in socce. PLoS ONE, 12(7), 1-15. https://doi.org/10.1371/journal.pone.0179953

Khrapach, V., & Siryi, O. (2024). Statistical Metric Expected Goals in Football and Its Impact on Scoring Performance: A review article. Health Technologies, 2(3), 47-54. https://doi.org/10.58962/HT.2024.2.3.47-54

Mead, J., O'Hare, A., & McMenemy, P. (2023a). Expected goals in football: Improving model performance and demonstrating value. PLoS ONE, 18(5), 20-28. https://doi.org/10.1371/journal.pone.0282295

Mead, J., O'Hare, A., & McMenemy, P. (2023b). Expected goals in football: Improving model performance and demonstrating value. PLoS ONE, 18(4), 45-55. https://doi.org/10.1371/journal.pone.0282295

Sabin, R. (2021). Estimating player value in American football using plus-minus models. Journal of quantitative analysis in sports, 4(17), 313-364. https://doi.org/doi.org/10.1515/jqas-2020-0033

Schulze, E., Mendes, B., Maurício, N., Furtado, B., Cesário, N., Carriço, S., & Meyer, T. (2017). Effects of positional variables on shooting outcome in elite football. Science and Medicine in Football, 2(2), 93-100. https://doi.org/10.1080/24733938.2017.1383628

Zeng, Z., & Pan, B. (2021). A Machine Learning Model to Predict Player's Positions based on Performance. icSPORTS, 1(17), 36-42. https://www.scitepress.org/PublishedPapers/2021/106533/106533.pdf

Zi, M., & Gao, D. (2023). A study of football goal trajectories of trained players through a kinetic model. Journal of engineering-joe, 12(2), 4-6.

© 2026 The Author. This article is licensed CC BY SA 4.0.
Visit Creative Commons Attribution-ShareAlike 4.0 International License.

Author information

Information about the authors/Author Biographies:

Author Information	
Cristian Murillo García (Author 1)	https://orcid.org/0009-0009-9096-421X
	Affiliation: Corporación Universitaria del Caribe - CECAR.
	Address: Sincelejo, Colombia.
	Disciplines: Sport Science
	Skills And Expertise: Sport Science
	Authors' Contribution: abcde
	Contact e-Mail: cristian.murillog@cecar.edu.co

© 2026 The Author. This article is licensed CC BY SA 4.0.
Visit Creative Commons Attribution-ShareAlike 4.0 International License.