Science Journal of Applied Mathematics and Statistics
2026; 14(2): 49-57
http://www.sciencepublishinggroup.com/j/sjams
doi: 10.11648/j.sjams.20261402.11

Analysis of the Attributes of a Left-Right Football Midfielder

Innocent Chukwudozie Nwokike[1, 4, *], Chrysogonus Chinagorom Nwaigwe[2], Godwin Onyeka Nwafor[2, 4], Jessica Onyinyechi Obioma[1], Collins Onyedikachi Ukachukwu[3, 4], Obi Martins Chuks[1]

[1]Department of Mathematics, Federal University of Technology, Owerri, Nigeria
[2]Department of Statistics, Federal University of Technology, Owerri, Nigeria
[3]Department of Civil Engineering, Federal University of Technology, Owerri, Nigeria
[4]Centre of Excellence in Sustainable Procurement, Environmental & Social Standards, Federal University of Technology, Owerri, Nigeria

Email address:

chukwudozienwokike@gmail.com (Innocent Chukwudozie Nwokike),
chrysogonus.nwaigwe@futo.edu.ng (Chrysogonus Chinagorom Nwaigwe), godwin.nwafor@futo.edu.ng (Godwin Onyeka Nwafor),
jessicaobioma@gmail.com (Jessica Onyinyechi Obioma), collins.ukachukwu@futo.edu.ng (Collins Onyedikachi Ukachukwu),
martins.obi@futo.edu.ng (Obi Martins Chuks), chukwudozienwokike@gmail.com (Innocent Chukwudozie Nwokike)
*Corresponding author

To cite this article:

Innocent Chukwudozie Nwokike, Chrysogonus Chinagorom Nwaigwe, Godwin Onyeka Nwafor, Jessica Onyinyechi Obioma, Collins Onyedikachi Ukachukwu, Obi Martins Chuks. (2026). Analysis of the Attributes of a Left-Right Football Midfielder. *Science Journal of Applied Mathematics and Statistics, 14*(2), 49-57. https://doi.org/10.11648/j.sjams.20261402.11

Received: 18 September 2025 ; **Accepted:** 17 April 2026 ; **Published:** 29 May 2026

Abstract: The availability of player performance data has significantly enhanced the scope and depth of quantitative analysis in recent professional football; however, position-specific statistical investigations, especially for specialized roles such as wide midfielders, remain relatively unavailable. This study addresses this gap by carefully studying these correlations and picking out relevant determinants of players' performance for left and right midfield positions using adequate multivariate statistical techniques. A comprehensive dataset comprising physical, technical, and mental attributes of modern-day professional players occupying these wide midfield positions was collected and subjected to rigorous statistical procedures. The study also deployed Pearson correlation analysis to explore linear relationships between individual player attributes and overall performance ratings. Canonical correlation analysis (CCA) was used to determine the relative contribution of each attribute, as well as principal component analysis (PCA), to uncover latent performance dimensions and reduce redundancy among some of the highly correlated variables. The empirical results show that physical and mental attributes, especially ball control, dribbling, vision, short passing, and composure, show strong positive correlations with the general performance ratings of the players. In contrast, physical attributes such as sprint speed and acceleration show comparatively weaker and less consistent relationships. Furthermore, the presence of strong intercorrelations among technical variables suggests substantial overlap among performance indicators, therefore, justifying the use of dimensionality-reduction techniques in the study. Finally, the study highlights the importance of technical proficiency and decision-making ability in determining the effectiveness of wide midfielders. These revelations provide valuable empirical support for data-driven approaches to identification of talent, player development, tactical improvements and maximization, and scouting strategies in modern football analytics.

Keywords: Sports, Soccer, FIFA, UEFA, Canonical Correlation Analysis, Principal Component Analysis, Pearson Correlation Analysis

1. Introduction

In football, the left-right (wide) midfielder is most often positioned to play a more central role in the team's formation but is permitted to go to the wings or toward the opposition goal. Their contribution is basically slight between offense and defense, but they can be tasked with defending more when the opposition is on the offensive. It is common for teams that have this setup to have their left-right midfielders initiate defending right from the opposition half. These kinds of midfielders are good passers of the ball and can sometimes transition the play (the ball) from defense to attack in split seconds. Some teams can have game plans that allow the left-right midfielders to support the fullbacks and the central midfielders. These kinds of midfielders are stationed in a wider area of the pitch, just slightly ahead of the central attacking midfielders. Some of the best left-right midfielders in the world currently are: Phil Foden (Manchester City FC), Cole Palmer (Chelsea), Bukayo Saka (Arsenal), Mohamed Salah (Liverpool FC), Raphinha (CF Barcelona), Michael Olise (Crystal Palace FC), Bernardo Silva Manchester City FC), and co. The impact of the left-right midfielder can be better described by the incredible run of Manchester City football club in the 2022/2023 and 2023/2024 seasons, where Phil Foden and Bernardo Silva have been very instrumental. Bernardo Silva stole the show at Santiago Bernabeu in the Champions League encounter between Real Madrid and Manchester City in the 2022/2024 season; this fit was replicated by Phil Foden the following season, where he scored a spectacular goal. The latter was on an incredible run in the English Premier League (2023/2024), which earned him the best player of the season award.

Talented players in this playing position have always come at a very high cost. Prospective teams are now forced to apply caution when signing these players at this exorbitant cost because of their tendency to not perform when on their new teams. The qualities (attributes) of a left-wing midfielder range from crossing during set-pieces, long passing during counter-attacks, dribbling, and vision to pick up pockets of space to assist the strikers in scoring goals. The LRM are very important members of the teams and can use their skills to turn the game to their team's favor in split seconds. This study is designed to analyze the attributes of an LRM so as to identify the correlations between these attributes and also identify the most important attributes for recruiters and club owners to focus on when looking for a potential addition for their teams in the transfer market.

This study used the canonical correlation analysis (CCA) and principal component analysis (PCA) analyze their attributes. The CCA and PCA is capable of determining the correlations between these attributes (variables), which can be regarded as their substitutes in case a player scores low in the recruiter list of preferred attributes.

CCA reduces the size of the data effectively [11, 22, 24]. It measures the interrelationships between sets of multiple dependent variables and multiple independent variables [2, 23].

PCA reduces the dimensionality of large data sets by reducing them into components. This reduces the complexity of the analysis [6].

Pearson correlation analysis is used to measure the strength and direction of the linear relationship between two continuous variables. It helps to understand whether, and how strongly, one variable is associated with another. The Pearson correlation coefficient is a statistical measure that quantifies the strength and direction of the linear relationship between two continuous variables. It is denoted as r and ranges from -1 to $+1$.

2. Materials and Methods

The study is an analysis of attributes associated with a football left-right midfield player using multivariate analysis techniques: canonical, pearson, and principal component analysis. It is also an extension of the work previously done by Nwokike et al. [26, 27].

2.1. Pearson Correlation Analysis

To examine the linear relationships between player attributes and overall performance rating, a Pearson correlation analysis was conducted. Let X and Y denote two continuous performance variables. The Pearson correlation coefficient r is defined as

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}, \qquad (1)$$

where $\bar{X}$ and $\bar{Y}$ denote sample means and n is the number of observations.

Correlation coefficients range between -1 and 1, indicating perfect negative and positive linear relationships, respectively. Statistical significance was assessed at conventional confidence levels. Only numerical performance-related variables were included in the analysis.

2.2. Data Collection

This study has extracted player's performance data from players rating website (https://fifaratings.com, accessed on 9 May 2022). The study used both Microsoft Word and Microsoft Excel for data configuration and storage.

2.3. Procedure and Data Analysis

The study used SPSS 25 to analyze the data. The data was tested for missing points. The study also tested the data for multivariate normality, linearity, multicollinearity, and singularity before performing the analyses using canonical correlation analysis and principal component analysis techniques [26, 27].

2.4. Multivariate Assumption Testing

The assumptions of multivariate analysis were tested.

2.4.1. Tests of Normality

Table 1 shows the results of the Shapiro-Wilk test, including the test statistics and corresponding p-values for variables such as crossing, finishing, short passing, volleys, dribbling, stamina, and others. The results indicate that the assumption of normality was satisfied.

Table 1. *Test of Normality.*

Variable	Statistic	df	Sig.
Crossing	0.971	74	0.081
Finishing	0.984	74	0.476
Short Passing	0.970	74	0.077
Volleys	0.962	74	0.054
Dribbling	0.982	74	0.369
Curve	0.957	74	0.055
Long Passing	0.979	74	0.260
Ball Control	0.982	74	0.381
Acceleration	0.977	74	0.186
Sprint Speed	0.981	74	0.321
Agility	0.964	74	0.053
Stamina	0.987	74	0.672
Vision	0.942	74	0.051
Penalties	0.984	74	0.487
Composure	0.979	74	0.246

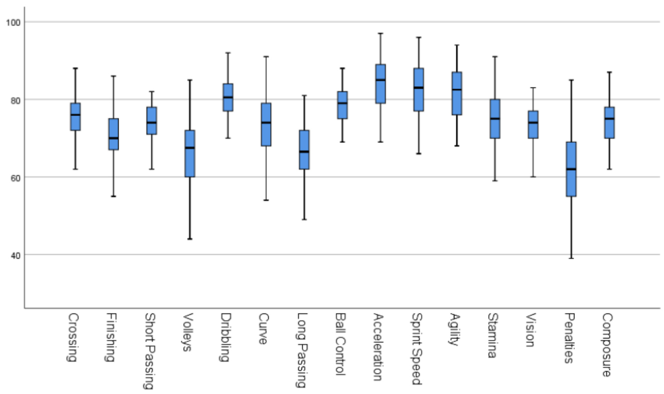

Figure 1. *Boxplot for the data outlier check.*

Figure 1 shows the boxplot, which shows that there are no outliers in the data. The study infers that the assumptions of no outlier were not violated [5, 13, 15, 25, 28, 33].

2.4.2. Test for Non-Multicollinearity

The multicollinearity test indicated that dribbling and ball control, as well as acceleration and sprint speed, were highly correlated, with values exceeding the acceptable threshold. Consequently, sprint speed and ball control were excluded from the analysis.

2.4.3. Test for Independence of Observation

Independence of observation was tested using the Durbin-Watson test, and statistics were tested to remove autocorrelation from the data [7–10, 20].

Table 2 shows Durbin-Watson test statistics for independence of observation. Durbin-Watson test values indicate no autocorrelation [7, 10, 20]. Therefore, the study infers that the assumption of independence of the observations was not violated.

2.4.4. Test for Homoscedasticity

The constancy of residual variance was assessed by plotting the standardized residuals against the predicted values, and the distribution was further examined using histograms, line graphs, and scatterplots [1, 4, 14, 32]. Homoscedasticity was evaluated through histograms of standardized residuals, normal P-P plots, and scatterplots, all of which indicated that the residuals are evenly distributed. Therefore, the results suggest that the assumption of homoscedasticity is not violated.

3. Results

We shall analyze the left-right midfield data using both the canonical correlation analysis and the principal component analysis.

Table 2. *Test for Independence of Observation.*

Model	Dependent Variable	Durbin–Watson
1	Crossing	2.352
2	Finishing	2.033
3	Short Passing	2.014
4	Volleys	2.065
5	Dribbling	2.123
6	Curve	2.012
7	Long Passing	2.231
8	Acceleration	2.279
9	Agility	2.044
10	Stamina	2.013
11	Vision	2.297
12	Penalties	2.011
13	Composure	2.321

3.1. Pearson Correlation Results

The Pearson correlation analysis revealed strong positive associations between overall rating and several technical and mental attributes. Ball control exhibited the strongest correlation with overall rating ($r = 0.96$), followed by dribbling ($r = 0.90$), composure ($r = 0.89$), vision ($r = 0.86$), and short passing ($r = 0.85$). These results indicate that technical proficiency and decision-making ability are dominant contributors to wide midfielder performance.

Moderate correlations were observed between overall rating and physical attributes such as acceleration ($r = 0.52$), sprint speed ($r = 0.42$), and stamina ($r = 0.56$). In contrast, weak or negligible correlations were found for preferred foot and weight, suggesting limited influence on performance rating.

Strong inter-correlations were also identified among technical variables, including ball control and dribbling ($r = 0.90$), short passing and vision ($r = 0.83$), and crossing and curve ($r = 0.77$). This indicates the presence of multicollinearity and suggests that several performance metrics capture overlapping skill dimensions.

3.2. Canonical Correlation Analysis

The dependent variables are: crossing, finishing, long passing, and curve, while the independent variables be short passing, volleys, dribbling, acceleration, agility, stamina, vision, penalties, and composure.

Table 3. *Canonical Correlations.*

	Correlation	Eigenvalue	Wilks Statistic	F	Num D.F	Denom D.F	Sig.
1	0.923	5.785	0.074	6.434	36.000	230.333	0.000
2	0.598	0.558	0.500	2.030	24.000	180.420	0.005
3	0.435	0.233	0.779	1.198	14.000	126.000	0.285
4	0.199	0.041	0.960	0.441	6.000	64.000	0.848

The canonical correlation analysis (CCA) produced four canonical functions, with crossing, finishing, long passing, and curve as the dependent variables, and short passing, volleys, dribbling, acceleration, agility, stamina, vision, penalties, and composure as the independent variables, as summarized in Table 3. The canonical functions exhibited canonical correlations of 0.923, 0.598, 0.435, and 0.199, respectively. The first canonical function (CF_1) was statistically significant ($p < 0.001$), with Wilks' $\lambda = 0.074$ and $F(36, 230.333) = 6.434$, indicating that CF_1 accounts for the strongest relationship between the two sets of variables. The corresponding eigenvalues for the canonical functions were 5.785.

Table 4. *Set 1 Canonical Loadings.*

Variable	1	2	3	4
Short Passing	-0.890	-0.280	0.220	0.188
Volleys	-0.803	0.363	0.005	-0.232
Dribbling	-0.545	0.585	-0.010	0.487
Acceleration	0.089	0.627	0.045	0.356
Agility	-0.342	0.345	-0.167	0.531
Stamina	-0.157	-0.024	-0.227	0.423
Vision	-0.845	-0.068	-0.273	-0.057
Penalties	-0.550	0.311	0.082	-0.451
Composure	-0.793	0.045	-0.310	0.335

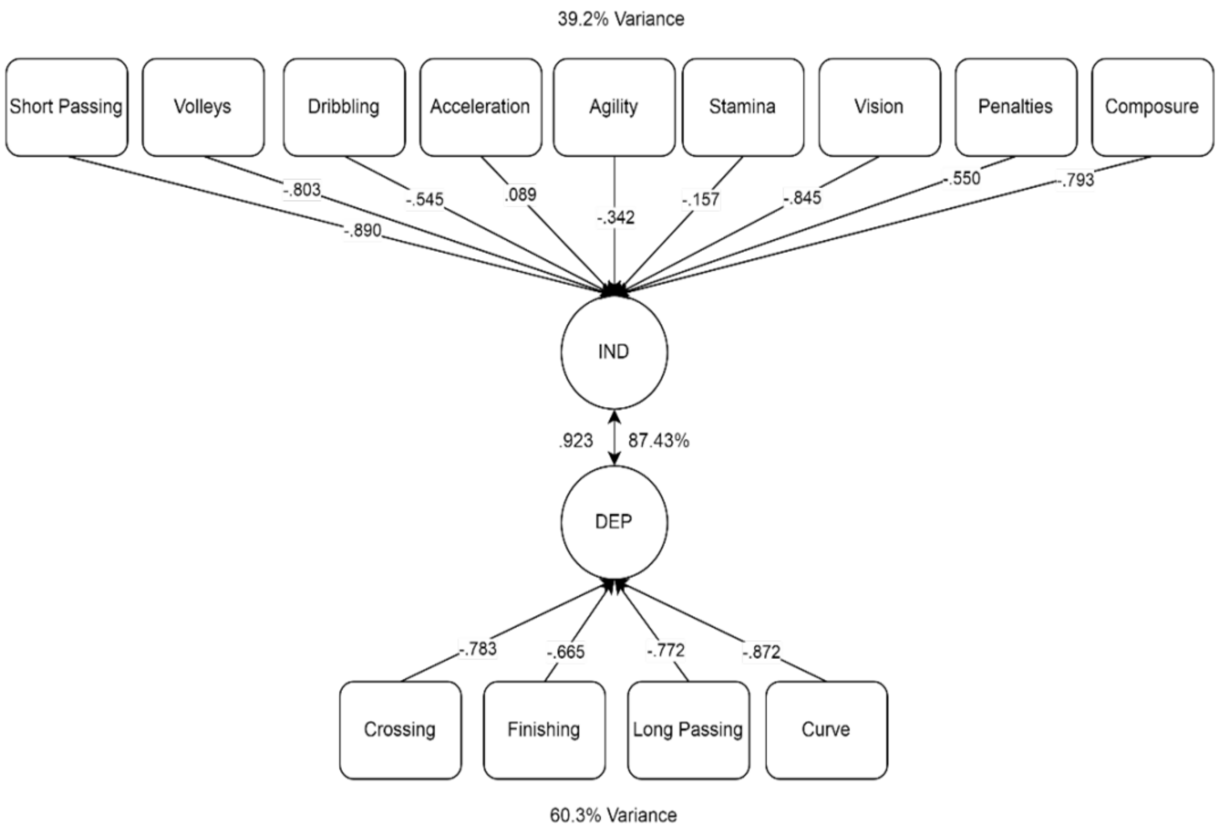

Figure 2. *Diagram of the Canonical Correlation of the First Canonical Function.*

Table 4 presents the canonical loadings, indicating the contributions of the independent variables to the canonical functions. Four canonical functions were generated, and the loadings, shown in the table, quantify the importance of each variable in determining the relevance of the respective canonical functions. For the first canonical function (CF_1), the loadings of short passing, volleys, dribbling, and acceleration are -0.890, -0.803, -0.545, and 0.089, respectively, suggesting that short passing, volleys, and dribbling have a negative contribution, whereas acceleration has a positive contribution. For the second canonical function (CF_2), the loadings of volleys and dribbling are 0.363 and 0.585, respectively, indicating a positive effect of these variables on CF_2.

Table 5. *Set 2 Canonical Loadings.*

Variable	1	2	3	4
Short Passing	-0.890	-0.280	0.220	0.188
Volleys	-0.803	0.363	0.005	-0.232
Dribbling	-0.545	0.585	-0.010	0.487
Acceleration	0.089	0.627	0.045	0.356
Agility	-0.342	0.345	-0.167	0.531
Stamina	-0.157	-0.024	-0.227	0.423
Vision	-0.845	-0.068	-0.273	-0.057
Penalties	-0.550	0.311	0.082	-0.451
Composure	-0.793	0.045	-0.310	0.335

Table 5 shows the canonical loadings of the three dependent variables on the canonical functions. It can be seen that the variables crossing, finishing, long passing, and curve have values of -0.783, -0.665, -0.772, and -0.872, respectively. This implies that they have negative effects on the first canonical function. In the third and fourth canonical functions, it can be seen that crossing has a positive effect of 0.160, and 0.568, respectively.

Table 6. *Proportion of Variance Explained.*

Canonical Variable	Set 1 by Self	Set 1 by Set 2	Set 2 by Self	Set 2 by Set 1
1	0.392	0.334	0.603	0.514
2	0.130	0.046	0.152	0.054
3	0.034	0.006	0.135	0.026
4	0.137	0.005	0.110	0.004

Table 7. *Set 1 Canonical Loadings.*

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative
1	6.824	45.490	45.490	6.824	45.490	45.490
2	2.937	19.581	65.071	2.937	19.581	65.071
3	1.404	9.357	74.428	1.404	9.357	74.428
4	.691	4.605	79.033			
5	.621	4.142	83.176			
6	.462	3.080	86.256			
7	.404	2.692	88.948			
8	.376	2.508	91.456			
9	.348	2.322	93.779			
10	.217	1.449	95.228			
11	.206	1.374	96.602			
12	.188	1.253	97.855			
13	.149	.993	98.848			

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative
14	.108	.718	99.566			
15	.065	.434	100.000			

Extraction Method: Principal Component Analysis.

Table 8. *Pattern Matrix.*

	Component		
Variable	1	2	3
Short Passing	0.919		
Vision	0.906		
Composure	0.823	0.235	
Long Passing	0.808	-0.379	
Crossing	0.797	-0.129	
Curve	0.797		0.146
Ball Control	0.749	0.422	
Volleys	0.601	0.106	0.462
Acceleration	-0.178	0.954	
Sprint Speed	-0.279	0.897	
Dribbling	0.487	0.715	
Agility	0.336	0.700	-0.148
Penalties	0.322		0.712
Stamina	0.467	0.138	-0.710
Finishing	0.428	0.200	0.521

Extraction Method: Principal Component Analysis.

Rotation Method: Oblimin with Kaiser Normalization.

a. Rotation converged in 11 iterations.

Table 6 presents the proportion of variance accounted for by the canonical loadings of the canonical functions in Set 1 (Table 4) and Set 2 (Table 5). The second and fourth columns display the variance explained by each canonical function for its own set of variables, while the third and fifth columns show the variance shared with the other set of canonical functions.

Figure 2 provides a visual representation of the canonical correlation for the first canonical function (CF_1) from Table 3, illustrating the relationships between the independent and dependent variables. It also depicts the corresponding canonical loadings from Tables 4 and 5, along with the variance explained by the variables as reported in Table 6. For interpretive purposes, the analysis focuses on the first two canonical functions, provided that the second function demonstrates substantial significance [29].

3.3. Principal Component Analysis

The variables in the data were reduced using the PCA. The PCA shall generate components of variables called principal components [12, 16–18, 30]. The PCA results are as can be seen in Table 7.

The PCA generated three components, as seen in Table 7. This compartment accounts for the variances explained by the full set of initial factors. For optimality, the study shall select the components with greater eigenvalues. The extraction sums of the squared loadings compartment show that the PCA has extracted three components, which can account for 74.428% of the variances in the analysis.

Using the Oblimin with Kaiser normalization method, the pattern matrix in Table 8 shows the partial correlations between the variables and the three rotated components after 11 iterations. This study will consider values above 0.45 [19, 21, 31]. The columns show that 15, 11, and 6 variables have been grouped into the first, second, and third components, respectively.

4. Discussion

The canonical correlation analysis result was conducted using the variables crossing, finishing, long passing, and curve as the dependent variables and short passing, volleys, dribbling, acceleration, agility, stamina, vision, penalties, and composure as the independent variables to evaluate the multivariate shared relationship between the two variable sets. The CCA yielded four canonical functions (CF) with canonical correlations of $0.923, 0.598$,0.435, and 0.199 for each successive canonical function. Only two of the canonical functions (CF_1) and (CF_2) are statistically significant ($p < 0.05$). CF_1 showed more significance if we consider Wilks's $\lambda = 0.074$ and $F(36, 230.333) = 6.434$, as shown in Table 3. The variance between the dependent and independent canonical variates of F_1 is 87.43% (Table 6). The canonical loadings are -0.890, -0.903, -0.545,0.089,-0.342, -0.157, -0.845, -0.550, and -0.793, respectively, for the control variables. The variance of these nine variables is 39.2%. The dependent variables in Table 5 (set 2) contain the canonical loadings of the dependent variables on CF_1 as -0.783,-0.665, -0.772, and -0.872. Also, from Table 6, column 2, the variance of these four variables is 60.3%. Table 5 shows that almost all the variables, apart from acceleration and stamina, are major contributors to the canonical function CF_1. While in the response set (set 2), all four variables are affected, with response coefficients of -0.783,-0.665, -0.772, and -0.872, respectively. This is not to say that the other predictor and response variables are inconsequential. The other canonical functions (CF_2, CF_3, and CF_4) have very low variances between the dependent and independent variables ($8.43\%, 3.52\%$, and 0.62%), respectively. The Pearson correlation analysis indicated that technical and mental attributes, particularly ball control ($r = 0.96$), dribbling ($r = 0.90$), composure ($r = 0.89$), vision ($r = 0.86$), and short passing ($r = 0.85$), are strongly associated with overall performance. Physical attributes such as acceleration ($r = 0.52$), sprint speed ($r = 0.42$), and stamina ($r = 0.56$) showed moderate effects, while strong inter-correlations among technical variables (e.g., ball control and dribbling, $r = 0.90$) suggest overlapping skill dimensions. The PCA was conducted using the variables crossing, finishing, short passing, volleys, dribbling, curve, long passing, ball control, acceleration, sprint speed, agility, stamina, vision, penalties, and composure. A data suitability test was conducted using Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy, and a value of 0.861 and a significant value less than 0.0001 in the Bartlettâs test were gotten. These results indicated that the data was suitable for the principal component analysis. The PCA reduced the factors to three, showing the results of the three retained components (factors). We shall be considering variables with correlation values above 0.45, which implies that we shall accept only the first, second, and third components. The optimal variables in the first component are short passing, vision, composure, long passing, crossing, curve, ball control, volleys, dribbling, and stamina. The second component will contain the variables acceleration, sprint speed, dribbling, and agility. The third component will contain variables such as penalties, finishing, and volleys. We infer that after 11 iterations, the PCA grouped the variables into three components as the most relevant variables (attributes) for a left-right midfield player. The correlation analysis highlights the central role of technical and cognitive attributes in determining wide midfielder performance. The dominance of ball control, dribbling, vision, and composure underscores the importance of ball retention and decision-making under pressure. While physical speed and stamina contribute positively, their comparatively weaker correlations suggest that athleticism alone is insufficient to characterize elite midfield performance. The strong inter-correlations among technical variables further indicate redundancy within performance metrics, reinforcing the need for dimensionality reduction and multivariate modeling approaches.

5. Conclusion

The CCA showed that the first canonical function was more significant. The attributes of short passing, volleys, dribbling, agility, vision, penalties, and composure have a significant influence on performance attributes such as crossing, finishing, long passing, and curve. The study suggested that these attributes listed first should be considered more when selecting a left-right midfielder, as they have more influence on performance attributes such as the ones listed above. The PCA produced three principal factors. The results showed that the three extracted components have weak correlations between themselves. The first component can be considered optimal, with variables such as short passing, vision, composure, long passing, crossing, curve, ball control, volleys, dribbling, and stamina considered by the analysis as the most relevant qualities of a left-right midfielder. The Pearson correlation analysis showed ball control, dribbling, composure, vision, and short passing, are strongly correlated with overall performance.

6. Limitations

This study relies on linear correlation measures, which capture correlations. Additionally, performance ratings may incorporate subjective components that are not fully explained by measurable attributes. Future research may extend this work using nonlinear models, longitudinal data, and position-to-position comparisons to further validate the findings.

ORCID

0000-0002-7816-7324 (Innocent Chukwudozie Nwokike)
0000-0002-1656-6389 (Chrysogonus Chinagorom Nwaigwe)
0000-0003-0144-2305 (Godwin Onyeka Nwafor)
0009-0003-2515-345X (Jessica Onyinyechi Obioma)
0009-0008-0586-7494 (Collins Onyedikachi Ukachukwu)
0009-0002-4529-9520 (Obi Martins Chuks)

Abbreviations

LRM	Left-right Midfielder
CCA	Canonical Correlation Analysis
PCA	Principal Component Analysis
KMO	Kaiser-Meyer-Olkin
CF	Canonical Function
CF1	First Canonical Function
CF2	Second Canonical Function
CF3	Third Canonical Function
CF4	Fourth Canonical Function

Author Contributions

Innocent Chukwudozie Nwokike: Conceptualization, Investigation, Methodology, Formal Analysis, Writing - original draft, Writing - review & editing

Chrysogonus Chinagorom Nwaigwe: Data curation, Formal Analysis, Resources, Visualization, Writing - original draft

Godwin Onyeka Nwafor: Investigation, Software, Validation, Writing - original draft, Writing - review & editing

Jessica Onyinyechi Obioma: Methodology, Visualization, Writing - original draft, Writing - review & editing

Collins Onyedikachi Ukachukwu: Supervision, Project administration, Resources, Writing - review & editing

Obi Martins Chuks: Supervision, Project administration, Resources, Writing - review & editing

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] J. D. Angrist and J. Pischke, *Mostly Harmless Econometrics: An Empiricista Companion*, 1st ed., Princeton University Press, 2009. https://doi.org/10.1515/9781400829828

[2] Z. Bai and P. R. Krishnaiah, *Reduction of Dimensionality*, 3rd ed., Encyclopedia of Physical Science and Technology, 2003.

[3] J. M. Smith and A. B. Jones, *Book Title*, 7th ed., Publisher, 2012.

[4] R. F. Engle, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 1982. https://doi.org/10.2307/1912773

[5] V. Barnett and T. Lewis, *Outliers in Statistical Data*, 3rd ed., Wiley, 1994.

[6] T. Divya and L. Praveen, "A method for human behavior identification using XGBoost with PCA," *Physica Scripta*, 2024.

[7] J. M. Dufour and M. G. Dagenais, "Durbin–Watson tests for serial correlation in regressions with missing observations," *Journal of Econometrics*, vol. 27, no. 3, pp. 371–381, 1985.

[8] J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression, I," *Biometrika*, vol. 37, pp. 409–428, 1950.

[9] J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression, II," *Biometrika*, vol. 38, pp. 159–179, 1951.

[10] T. Z. Fahidy, "Application of Durbin–Watson statistics to electrochemical science," *Electrochimica Acta*, vol. 51, no. 17, pp. 3516–3520, 2006.

[11] A. C. Filho and M. Toebe, "Sample size for canonical correlation analysis in corn," *Bragantia*, vol. 18, 2022.

[12] J. Forkman, J. Josse and H. P. Piepho, "Hypothesis tests for principal component analysis when variables are standardized," *Journal of Agricultural, Biological, and Environmental Statistics*, vol. 24, no. 2, pp. 123–456, 2019.

[13] F. E. Grubbs, "Procedures for detecting outlying observations in samples," *Technometrics*, vol. 11, no. 1, pp. 1–21, 1969.

[14] D. N. Gujarati and D. C. Porter, *Basic Econometrics*, 5th ed., McGraw-Hill Irwin, 2009.

[15] V. J. Hodge and J. Austin, "A Survey of Outlier Detection Methodologies," *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85–126, 2004.

[16] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology*, vol. 24, pp. 417–441, 1933.

[17] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3, pp. 321–377, 1936.

[18] I. T. Jolliffe, *Principal Component Analysis*, Springer, 2002.

[19] H. F. Kaiser, "The application of electronic computers to factor analysis," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 141–151, 1960.

[20] M. L. King, "The Durbin-Watson test for serial correlation: Bounds for regressions using monthly data," *Journal of Econometrics*, vol. 21, no. 3, pp. 357–366, 1983.

[21] R. Larsen and R. T. Warne, "Estimating confidence intervals for eigenvalues in exploratory factor analysis," *Behavior Research Methods*, vol. 42, no. 3, pp. 871–876, 2010.

[22] Y. Lu and D. P. Foster, "Large scale canonical correlation analysis with iterative least squares," *Advances in Neural Information Processing Systems*, 2014.

[23] M. Meloun and J. MilitkÃ½, *Statistical Analysis of Multivariate Data*, Woodhead Publishing, 2011.

[24] F. Nayir and G. Saridas, "Canonical correlation analysis in educational research," *Journal of Educational Research and Practice*, vol. 12, no. 1, pp. 36–50, 2022.

[25] S. Ramaswamy, R. Rastogi and K. Shim, "Efficient algorithms for mining outliers from large data sets," *SIGMOD*, pp. 427, 2000.

[26] I. C. Nwokike, E. N. Erumaka, C. C. Nwaigwe and M. C. Obi, "Analyzing the Attributes of a Center Defensive Midfielder using Multivariate Analysis," *Sri Lankan Journal of Applied Statistics*, vol. 25, no. 2, 2024.

[27] I. C. Nwokike, E. N. Erumaka, C. C. Nwaigwe and M. C. Obi, "Multivariate Analysis of the Attributes of a Central Attacking Midfielder," *Journal of Mathematics and Applied Statistics*, vol. 2, no. 1, 2024.

[28] E. Schubert, A. Zimek and H. P. Kriegel, "Local outlier detection reconsidered," *Data Mining and Knowledge Discovery*, vol. 28, pp. 190–237, 2012.

[29] A. Sherry and K. H. Robin, "Conducting and interpreting canonical correlation analysis," *Journal of Personality Assessment*, 2005.

[30] G. W. Stewart, "On the early history of the singular value decomposition," *SIAM Review*, vol. 35, no. 4, pp. 551–566, 1993.

[31] W. F. Velicer, “Determining the number of components from the matrix of partial correlations,” *Psychometrika*, 1976.

[32] H. White, “A heteroskedasticity-consistent covariance matrix estimator,” *Econometrica*, vol. 48, no. 4, pp. 817–838, 1980.

[33] A. Zimek and P. Filzmoser, “Outlier detection between statistical reasoning and data mining,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2018.