

# Technical profiling of football players

Drs Raffaele Poli, Loic Ravenel and Roger Besson

## 1. Introduction

This Monthly Report presents the method developed by the CIES Football Observatory to determine the technical profile of footballers, ranking them in different groups and establishing hierarchies within the latter. To do this, we have based ourselves on the data referring to the technical gestures undertaken by players (shots, crosses, interceptions, etc.) produced by our partners from InStat.

The analysis was carried out on a sample of 7,215 footballers having played at least 750 domestic league minutes during the 2021 or 2021/22 seasons (up until the end of January 2022) within 36 top or second division leagues from UEFA member associations. Due to the specificity of their position, goalkeepers have not been included in the study.

Figure 1: leagues taken into account and number of players

	Bundesliga (AUT)	138 players	<div style="width: 100%;"></div>
	Pro League (BEL)	236 players	<div style="width: 100%;"></div>
	Premier League (BLR)	235 players	<div style="width: 100%;"></div>
	First League (BUL)	166 players	<div style="width: 100%;"></div>
	1. HNL (CRO)	122 players	<div style="width: 100%;"></div>
	1. Division (CYP)	127 players	<div style="width: 100%;"></div>
	First League (CZE)	191 players	<div style="width: 100%;"></div>
	Superliga (DEN)	129 players	<div style="width: 100%;"></div>
	Premier League (ENG)	257 players	<div style="width: 100%;"></div>
	Championship (ENG/2)	348 players	<div style="width: 100%;"></div>
	Liga (ESP)	257 players	<div style="width: 100%;"></div>
	Segunda División (ESP/2)	309 players	<div style="width: 100%;"></div>
	Veikkausliiga (FIN)	181 players	<div style="width: 100%;"></div>
	Ligue 1 (FRA)	260 players	<div style="width: 100%;"></div>
	Ligue 2 (FRA/2)	253 players	<div style="width: 100%;"></div>
	Bundesliga (GER)	217 players	<div style="width: 100%;"></div>
	2. Bundesliga (GER/2)	211 players	<div style="width: 100%;"></div>
	Super League (GRE)	163 players	<div style="width: 100%;"></div>
	NB I (HUN)	123 players	<div style="width: 100%;"></div>
	Ligat ha'Al (ISR)	165 players	<div style="width: 100%;"></div>
	Serie A (ITA)	269 players	<div style="width: 100%;"></div>
	Serie B (ITA/2)	245 players	<div style="width: 100%;"></div>
	Eredivisie (NED)	211 players	<div style="width: 100%;"></div>
	Eliteserien (NOR)	230 players	<div style="width: 100%;"></div>
	Ekstraklasa (POL)	202 players	<div style="width: 100%;"></div>
	Primeira Liga (POR)	213 players	<div style="width: 100%;"></div>
	Liga I (ROM)	199 players	<div style="width: 100%;"></div>
	Premier League (RUS)	187 players	<div style="width: 100%;"></div>
	Premiership (SCO)	145 players	<div style="width: 100%;"></div>
	Super Liga (SRB)	202 players	<div style="width: 100%;"></div>
	Super League (SUI)	113 players	<div style="width: 100%;"></div>
	Super Liga (SVK)	126 players	<div style="width: 100%;"></div>
	1. SNL (SVN)	107 players	<div style="width: 100%;"></div>
	Allsvenskan (SWE)	237 players	<div style="width: 100%;"></div>
	Süper Lig (TUR)	268 players	<div style="width: 100%;"></div>
	Premier League (UKR)	184 players	<div style="width: 100%;"></div>

## 2. Variables and classification axes

Among the very numerous indicators collected by InStat, eleven were selected to determine players' technical profile. They refer to both defensive and offensive actions. The selection was carried out to limit redundancies and eliminating variables that focus too closely on a few individuals due to the low number of actions. For example, shots were preferred to goals, the first being closely correlated to the latter, while more spread out among different players.

So as to determine technical profiles independently of the level and style of play of employer clubs, the values attributed to players for each of the eleven variables selected were defined by referring to the average value of the other team members, i.e. as a ratio between the player's value and that of teammates (excluding goalkeepers).

For example, a value equal to two on the levels of shots indicates that the player has shot twice as much as his partners. In this manner, a footballer playing in a team at a low level does not see its values structurally diminished in comparison to players of more competitive teams. Thus, we can actually analyse a player's game profile rather than a performance that is strongly linked to the team's overall strength.

From these relative values, we have performed a Principal Component Analysis (PCA). The results are expressed visually in Figure 3 in the form of a factorial display with the eleven variables selected represented by arrows. The longer the arrow and the closer it is to an axis, the more the variable in question is important in the definition of the latter.

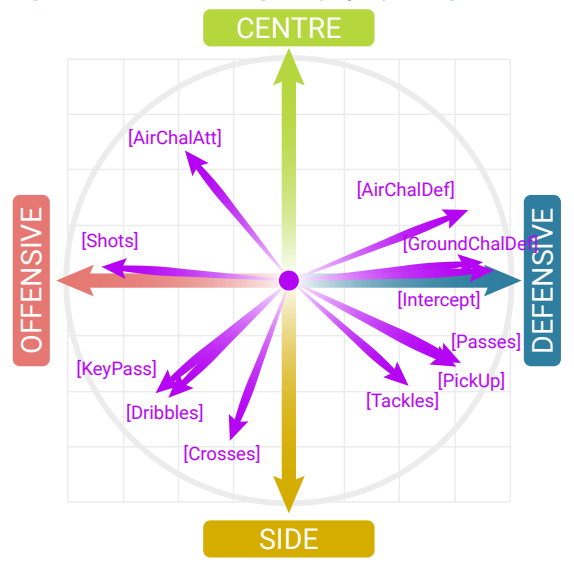
The variable "interception" is strongly involved in the formation of the horizontal axis, as are the shots towards his opponent. This axis therefore defines the defensive or offensive tendencies of players. The two variables that are the most telling from the point of view of the vertical axis are crosses, that are especially the province of wingers, and aerial duels, that are principally the work of centre forwards and centre backs. This axis tends thus to refer to the different positioning of players in the same area of play (defence, midfield, or attack).

The two principal axes explain 70% of the total variance, the defensive-offensive axis by itself explaining almost half of the latter. This signifies that the eleven values selected, as well as their relativisation with regard to teammates, allows us to account to a large extent for the differences in the technical profiles of footballers.

Figure 2: game indicator selected for profiling

Vocation	Variable	Id
Defensive	Air challenge won - defense	[AirChalDef]
	Interceptions	[Intercept]
	Picking-ups	[PickUp]
	Ground Challenge won - defense	[GroundChalDef]
	Successful tackles	[Tackles]
Neutral	Passes	[Passes]
Offensive	Air Challenge won - attack	[AirChalAtt]
	Crosses	[Crosses]
	Successful dribbles	[Dribbles]
	Key passes	[KeyPass]
	Shots	[Shots]

Figure 3: the factorial design for player profiling



### 3. Player proximity

This profiling method permits the calculation of distances between players. Using a reference footballer, it is possible to determine the players who are statistically the closest to him. For example, among big-5 league footballers, the closest footballer to Kylian Mbappé from the perspective of the technical actions performed is James Maddison (Leicester City). If we restrict the analysis to French Ligue 1 players, Stephy Mavididi (Montpellier HSC) is closest to the world champion.

This exercise can be carried out by using any player as a reference. For example, always at big-5 league level, Dušan Vlahović is the player who is closest to Erling Haaland, Paulo Dybala to Lionel Messi, Romain Faivre to Neymar Júnior, Dominik Szoboszlai to Kevin de Bruyne, Sadio Mané to Raheem Sterling, Remo Freuler to Jorginho Frello or Jonathan Tah to Virgil van Dijk.

### 4. Player classification

Aside from the calculation of statistical distances between players, the k-medoids algorithm allows us to rank them in groups. This method, derived from k-means, is based on the choice of reference players that serve as archetypes for the elaboration of groups to which all individuals will be attributed through successive statistical proximity. Six reference players with different positions and profiles were selected for this report.

Figure 4: five big-5 and Ligue 1 players with a technical profile most similar to Kylian Mbappé

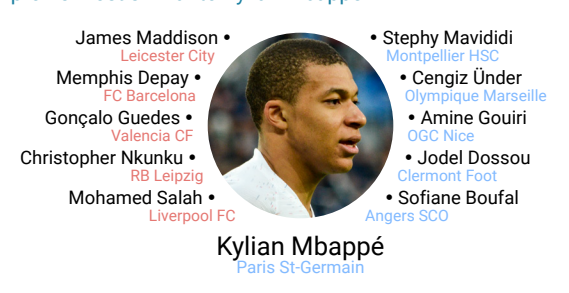


Figure 5: player archetypes used in the elaboration of groups

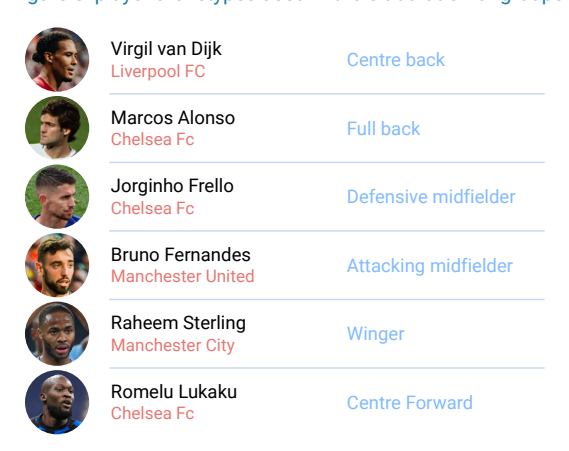


Figure 6 illustrates the dominant technical gestures for each of the classes constructed from the six reference players selected. For example, players from the class based on Virgil van Dijk, win, on average, 2.62 defensive aerial duels more than their teammates, those of the class based on Marcos Alonso cross 2.62 more times, and so on.

### Virgil van Dijk Class

The Van Dijk class regroups players whose values at the level of all defensive variables are above those of their teammates. They are principally centre backs who set themselves apart by their strong presence in duels, both on the ground and in the air. The class accounts for 21.7% of the players of our sample.

### Marcos Alonso Class

The Marcos Alonso class also identifies players with a defensive-oriented vocation. However, they are also very active offensively with regard to crosses. They are mainly full backs or wing backs within a "3-5-2" tactical formation. This class regroups 15.7% of the footballers taken into account.

### Jorginho Frello Class

The Jorginho Frello class also brings together players with a defensive vocation, but who are also relatively active offensively. They are generally central midfielders, both defending and box-to-box. This class is not only the one which regroups the most players, 26.5% of the total sample analysed, but also the most heterogeneous.

### Bruno Fernandes Class

The Bruno Fernandes class regroups players who are more active in attack than in defence. Their speciality resides in the ability to create opportunities for teammates, as well as in their importance in animating attack generally (dribbles, shooting, crosses, passes, etc.). This class is the least numerous: it regroups only 8.6% of players.

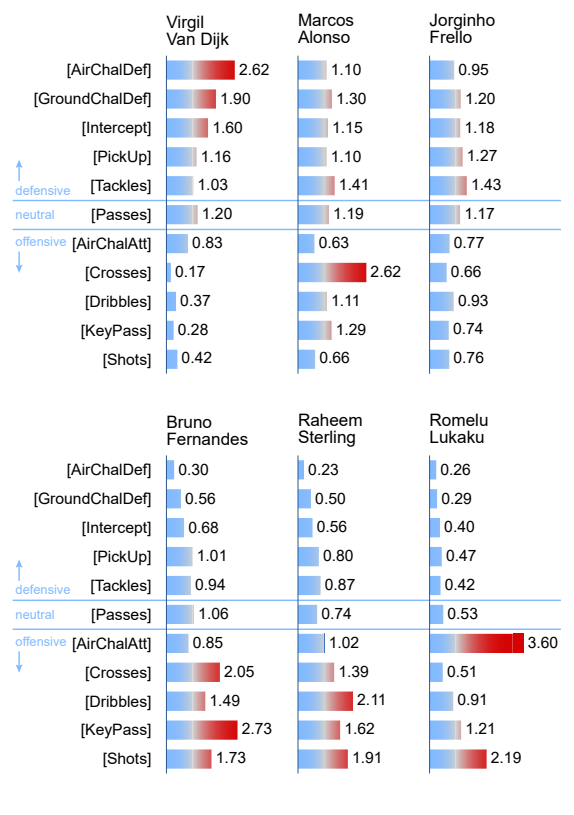
### Raheem Sterling Class

The Raheem Sterling class identifies players with a similar profile to that of Bruno Fernandes, but differentiating themselves by a greater propensity for shots and dribbles, and a lower propensity for key passes and crosses. This class is quite numerous as it accounts for 18.0% of the players of our sample.

### Romelu Lukaku Class

Finally, the Romelu Lukaku class picks out, above all, footballers playing as centre forwards. Players in this category are rarely in the thick of the action and concentrate their efforts on two speciality areas: finishing and offensive aerial duels. This class makes up 9.8% of footballers included in the analysis.

Figure 6: relative average frequency of technical gestures by class (with respect to teammates)



## 5. A tiering of payers according to the profiles defined

Any tiering of players comes up not only against the very often underestimated problem of the impact on individual performance of the collective strength's differential between opponents, but also against the difficulty of establishing groups of footballers with a style of play sufficiently close for a comparison to make sense.

The relativisation of performance indicators in comparison to teammates and the creation of player archetypes based on reference footballers, are both a means of limiting these problems without, however, resolving them completely. Indeed, any class, as homogenous as it may be, always contains margins where players with an atypical or inter-class profile are situated.

One solution resides in increasing the number of classes in comparison to the initial scheme. This can be done by determining, through the statistical tool of the silhouette value, the degree by which the class they are part of actually represents them. In doing this, players fitting well with their class remain in them, while new classes can be created from players with an inter-class profile.

For example, while a player such as Thomas Partey from Arsenal is well defined as a member of the Jorginho class, Nemanja Matić is situated in an intermediary position between this class and Virgil van Dijk's one. We can thus establish a new class bringing together all the individuals statistically closest to Matić than to Jorginho or van Dijk. In the end, two intermediary classes (Matić and Trippier) were added to the six initial ones.

Figure 7: the eight classes used for ranking

Classes	Most representative variable
1 Virgil Van Dijk	1 Air challenges won - defensive 2 Ground challenges won - defensive 3 Interceptions
2 Nemanja Matić	1 Air challenges won - defensive 2 Air challenges won - attacking 3 Interceptions
3 Marcos Alonso	1 Crosses 2 Successful tackles 3 Key passes accurate
4 Kieran Trippier	1 Crosses 2 Successful tackles 3 Ground challenges won - defensive
5 Jorginho Frello	1 Successful tackles 2 Picking-ups 3 Passes
6 Bruno Fernandes	1 Key passes accurate 2 Crosses 3 Shots
7 Hakim Ziyech	1 Dribbles successful 2 Shots 3 Key passes accurate
8 Raheem Sterling	1 Air challenges won - attacking 2 Shots 3 Key passes accurate
9 Romelu Lukaku	1 Duels aériens gagnés - attaque 2 Tirs 3 Passes pour occasion

For each of the eight classes, we focused on the three variables where the players stand out more than their teammates of all the classes taken together (see Figure 7), and established hierarchies based on these three variables. For example, players from the Van Dijk class were ranked by taking into account the differences in comparison to teammates for the following three variables: aerial defensive duels, ground defensive duels and interceptions.

Figures 8a and 8b present the top 10 footballers from all leagues surveyed for each of the eight classes. Nevertheless, it is also possible to take into consideration the league level at which the footballers play, so as to only compare players in championships of relatively similar strengths, the big-5 for example.

**Figure 8: most productive players with respect to teammates**

**Class 1** Virgil Van Dijk / [AirChalDef], [GroundChalDef], [Intercept]

1	Haiderson Hurtado	26.3 yrs	ŠKF Sereď (SVK)
2	Maximilian Wöber	24.1 yrs	RB Salzburg (AUT)
3	Leonardo Balerdi	23.1 yrs	Olympique Marseille (FRA)
4	Simone Canestrelli	21.5 yrs	FC Crotone (ITA/2)
5	Juuso Hämäläinen	28.2 yrs	Inter Turku (FIN)
6	Frank Boya	25.7 yrs	Zulte Waregem (BEL)
7	Oliver Abildgaard	25.7 yrs	Rubin Kazan (RUS)
8	Frederik Sörensen	29.9 yrs	Ternana Calcio (ITA/2)
9	Nicolas Wimmer	27.0 yrs	Austria Klagenfurt (AUT)
10	Gabriele Angella	32.8 yrs	AC Perugia (ITA/2)

**Class 2** Nemanja Matić / [AirChalDef], [AirChalAtt] and [Intercept]

1	Andrian Kraev	23.0 yrs	Levski Sofia (BUL)
2	Saša Tomanović	32.4 yrs	TSC Bačka Topola (SRB)
3	Sergiy Petko	28.1 yrs	Veres Rivne (UKR)
4	Benjamin André	31.6 yrs	LOSC Lille (FRA)
5	Marco van Ginkel	29.2 yrs	PSV Eindhoven (NED)
6	Krzysztof Kubica	21.8 yrs	Górnik Zabrze (POL)
7	Tomáš Souček	27.0 yrs	West Ham United (ENG)
8	Marshall Munetsi	25.7 yrs	Stade de Reims (FRA)
9	Boubacar Fofana	32.3 yrs	Sepsi OSK (ROM)
10	Oleksiy Dovhyi	32.3 yrs	FK Lviv (UKR)

**Class 3** Marcos Alonso / [Crosses], [Tackles] and [KeyPass]

1	Alfonso Pedraza	25.9 yrs	Villarreal CF (ESP)
2	Iasmin Latovlevici	35.8 yrs	FC Argeş (ROM)
3	Albert Adomah	34.2 yrs	QPR (ENG/2)
4	Illia Putria	23.8 yrs	FK Chornomorets (UKR)
5	Leonardo Lukačević	23.1 yrs	Admira Wacker (AUT)
6	Krum Stoyanov	30.6 yrs	CSKA 1948 (BUL)
7	Rasmus Carstensen	22.2 yrs	Silkeborg IF (DEN)
8	Damián Suárez	33.8 yrs	Getafe CF (ESP)
9	Yanis Hamache	22.6 yrs	Boavista FC (POR)
10	Pedro Porro	22.5 yrs	Sporting CP (POR)

Figure 8b : joueurs les plus actifs par rapport aux co-équipiers

**Classe 4** Kieran Trippier / [Crosses], [Tackles] and [GroundChalDef]

1	João Aurélio	33.5 yrs	Pafos FC (CYP)
2	Ahmet Oğuz	29.1 yrs	Sivasspor (TUR)
3	Óscar de Marcos	32.9 yrs	Athletic Club (ESP)
4	Ignacio Guerrico	23.6 yrs	NK Maribor (SVN)
5	Luis Carlos Murillo	31.4 yrs	HJK Helsinki (FIN)
6	Yvan Dibango	20.0 yrs	FC Isloch (BLR)
7	Daniilo Soares	30.3 yrs	VfL Bochum (GER)
8	Ilya Samoshnikov	24.3 yrs	Rubin Kazan (RUS)
9	Rubén Duarte	26.4 yrs	CD Alavés (ESP)
10	Justas Lasickas	24.4 yrs	FK Voždovac (SRB)

**Classe 5** Jorginho Frello / [Tackles], [PickUp] and [Passes]

1	Antonio Vacca	31.8 yrs	Venezia FC (ITA)
2	Theofanis Tzandaris	28.7 yrs	PAS Lamia (GRE)
3	Ali Mohamed	26.4 yrs	Maccabi Haifa (ISR)
4	Obinna Nwobodo	25.2 yrs	Göztepe SK (TUR)
5	Cameron Devlin	23.7 yrs	Heart of Midlothian (SCO)
6	Mikhail Bashilov	29.1 yrs	Energetik-BGU (BLR)
7	Nikita Korzun	27.0 yrs	Shakhtyor Soligorsk (BLR)
8	Miguel Mellado	29.0 yrs	OFI Crete (GRE)
9	Juan Munafo	33.9 yrs	Asteras Tripolis (GRE)
10	Óscar Valentín	27.5 yrs	Rayo Vallecano (ESP)

**Classe 6** Bruno Fernandes / [KeyPass], [Crosses] and [Shots]

1	Ronaldo Deaconu	24.8 yrs	Gaz Metan Mediaş (ROM)
2	Shahboz Umarov	23.0 yrs	Energetik-BGU (BLR)
3	Ricardo Quaresma	38.4 yrs	Vitória Guimarães (POR)
4	Quentin Cornette	28.1 yrs	Le Havre AC (FRA/2)
5	Alexis Flips	22.1 yrs	Stade de Reims (FRA)
6	Filip Kostić	29.3 yrs	Eintracht Frankfurt (GER)
7	Yuliyen Nenov	27.3 yrs	Botev Vratsa (BUL)
8	Fabian Reese	24.2 yrs	Holstein Kiel (GER/2)
9	Kristijan Lovrić	26.2 yrs	HNK Gorica (CRO)
10	Yannick Bolasie	32.8 yrs	Çaykur Rizespor (TUR)

**Classe 7** Raheem Sterling / [Dribbles], [Shots] and [KeyPass]

1	Adama Traoré	26.1 yrs	Wolverhampton FC (ENG)
2	Heorhii Tsitaishvili	21.3 yrs	FK Chornomorets (UKR)
3	Marko Livaja	28.5 yrs	Hajduk Split (CRO)
4	Gaëtan Laura	26.6 yrs	Paris FC (FRA/2)
5	Rabbi Matondo	21.5 yrs	Cercle Brugge (BEL)
6	Antoine Semenyo	22.1 yrs	Bristol City (ENG/2)
7	Joseph Hungbo	22.1 yrs	Ross County (SCO)
8	Chidera Ejuke	24.2 yrs	CSKA Moskva (RUS)
9	Khvicha Kvaratskhelia	21.0 yrs	Rubin Kazan (RUS)
10	Jens Odgaard	22.9 yrs	RKC Waalwijk (NED)

**Classe 8** Romelu Lukaku / [AirChalAtt], [Shots] and [KeyPass]

1	Milan Djurić	31.8 yrs	US Salernitana (ITA)
2	Tobias Lauritsen	24.5 yrs	Odds BK (NOR)
3	Matt Smith	32.7 yrs	Millwall FC (ENG/2)
4	Mostafa Mohamed	24.2 yrs	Galatasaray SK (TUR)
5	Roberts Uldriķis	23.9 yrs	SC Cambuur (NED)
6	Joselu Mato	31.9 yrs	CD Alavés (ESP)
7	Rok Kidrič	26.8 yrs	NK Aluminij (SVN)
8	El Bilal Touré	20.4 yrs	Stade de Reims (FRA)
9	Andy Delort	30.4 yrs	OGC Nice (FRA)
10	Aliou Badji	24.4 yrs	Amiens SC (FRA/2)

## 6. Conclusion

The game indicators collected by InStat are a treasure from which it is possible to develop multiple innovative research with very practical applications. This report was conceived with a view to linking science and industry, an approach that we value dearly and that constitutes the *raison d'être* of the CIES Football Observatory research group.

The method of players' profiling and classification detailed in this study is particularly useful not only from a descriptive perspective to understand the different roles played by footballers within a team, but also from the point of view of scouting. The calculation of statistical distances between players is indeed particularly useful when targeting potential recruits to replace departing players.

The choice of profiling and classifying players not based on raw statistics, but in comparison to teammates, is also particularly fruitful when it comes to talent spotting. Indeed, it mitigates the recurring problem of the impact of the collective force of a team on individual performances. This approach notably allows us to identify players who do not stand out in absolute terms, but whose productivity is well above that of their teammates.

This report constitutes just another step in the direction of taking full advantage of the possibilities available in terms of research and development through the analysis of technical game data such as those carefully gathered by InStat. We look forward to pursuing this field more fully and to making available new breakthroughs for all those passionate about the game.