

Installer avec la *data* un environnement de haute performance sportive dans un club professionnel de football ?

Par Arnaud MAZIER

Ingénieur général des Mines

Et Damien COMOLLI

Président du Toulouse Football Club

Dans le monde du football professionnel, les clubs sont au cœur d'une économie complexe qui dépasse l'aspect sportif de la performance.

Avec l'essor des capteurs physiques et vidéo, des *brokers* de données et des pratiques utilisant des modélisations mathématiques, la collecte et l'utilisation de la *data* se sont amplifiées depuis les années 2000-2010 pour développer la performance sportive. La *data* est un outil précieux pour aider les clubs de football professionnels à créer un environnement propice à la haute performance sportive, en optimisant le recrutement ou en suivant la performance individuelle des joueurs au service d'un style de jeu identifiable par ses *supporters*.

Le Toulouse Football Club (Téfécé) a connu de grands moments depuis sa création et a retrouvé l'élite en 2022, remportant une nouvelle Coupe de France en 2023. Le Téfécé en est un exemple concret, ayant réussi à utiliser la *data* de manière efficace sportivement et économiquement, et attirer un public grandissant.

Dans un monde où le football est bien plus qu'un simple sport, les clubs professionnels se retrouvent au cœur d'une économie complexe. Leur gestion ne se limite pas à l'aspect sportif emblématique de la performance de l'équipe ou de chaque joueur, mais comprend de nombreux aspects relevant de la formation, d'un spectacle, en tenant compte du territoire dans lequel il est implanté.



Figure 1 : Logo du Toulouse Football Club (Source : Club).

Le Toulouse Football Club (Téfécé) a connu de grands moments depuis sa création en 1937 avec 51 saisons en ligue 1, une Coupe de France et deux titres de champion de ligue 2. Deux années après un changement d'actionnaire, il a retrouvé l'élite en 2022, a gagné une nouvelle Coupe de France en 2023 pour jouer la ligue Europa. L'affluence au stade est passée de 5 000 à 26 000 spectateurs en moyenne par match, de 4 000 abonnés à 17 500 permettant de stabiliser l'équilibre économique du club.

La donnée et les mathématiques ont accompagné l'analyse de la performance sportive d'un club professionnel.

De données nourrissant les paris sportifs dans les journaux au cours du XX^e siècle, relatant le passé pour prédire l'avenir, la collecte et l'usage de la *data* se sont amplifiés depuis les années 2000-2010, au gré de la constitution d'un écosystème rendu possible par l'essor des capteurs physiques et vidéo, de *brokers* de données, et l'émergence de pratiques utilisant massivement les modélisations mathématiques.

Quels sont les leviers mis en œuvre pour développer la performance d'un club avec la dix-neuvième masse salariale de ligue 1 ? En quoi la donnée peut-elle aider le dirigeant ?

CRÉER L'ENVIRONNEMENT PROPICE À LA HAUTE PERFORMANCE SPORTIVE

Avec 87 % de victoires depuis 1907, l'équipe des All Blacks est la plus performante du monde, tout sport collectif confondu. Dans son ouvrage, véritable bible pour les managers sportifs anglo-saxons, James Kerr décortique les raisons de ce succès : créer un environnement de haute performance sportive bâti sur une culture et des valeurs reconnues.

En 2020, la défaite s'était profondément ancrée dans le club avec 16 mois sans aucune victoire, à tel point que le mot "To Loose" transpirait dans tous les secteurs, des plus jeunes au groupe professionnel, jusqu'à la gestion du club. Il a fallu recréer une culture du succès en prenant appui sur quelques fondamentaux : être identifiable par un style de jeu, installer un environnement de haute performance sportive, retisser les liens avec la communauté Viola¹ (*supporters*, collectivités, *sponsors*).

Le style de jeu est la pierre angulaire du projet sportif et entrepreneurial. Il détermine ce que le secteur sportif incarne (flamboyant, porté sur l'attaque et la possession du ballon), et projette comme marqueur de l'expérience sportive attendue au stade par la communauté Viola autant que d'enracinement du club dans son territoire.

Il reste construit sur une analyse des facteurs objectifs de succès sportifs passés. L'analyse des styles de jeu des 25 équipes montées en Ligue 1 en 18 saisons a ainsi permis de préciser les choix, en déconstruisant le succès, en les objectivant pour éviter de tomber dans des croyances installées dans le football². Au quotidien, la donnée guide la mise en œuvre pratique et précise du style de jeu (exemple : possession haute du ballon dans certaines zones du terrain, récupérer le ballon après 5 secondes, optimiser le PPDA³).

Le style de jeu se décline par chaque éducateur dans toutes les sections, professionnelle, jeune, féminine, lors des exercices à l'entraînement et lors des compétitions.

Le style de jeu se mue en plaisir attendu par ses *supporters* au stade ou à la télévision. La compétition est en effet faite pour attirer et fidéliser les spectateurs, avec une riche offre locale de loisir (Stade Toulousain, Méditerranée, ski en Andorre).

Au service du style de jeu et des performances, la haute performance sportive installe des processus d'amélioration continue basés sur des métriques objectives, questionnant les pratiques, et installant l'excellence dans tous les gestes individuels et collectifs. Elle est par construction holistique et systématique dans tout le club, en utilisant massivement la *data*.

¹ Violet est la couleur du club et l'un des marqueurs de Toulouse.

² Pour monter en Ligue 1, la capacité défensive est déterminante.

³ *Passes Per Defensive Action* (PPDA) : métrique qui met en regard les passes de l'adversaire dans son camp et les interventions défensives (tacles, interceptions, duels et fautes). Plus l'indicateur est bas, plus une équipe est active et intense défensivement dans le camp adverse.

LE RECRUTEMENT

Traditionnellement, le recrutement est le fruit d'un travail de scout sur le terrain. Un grand club disposant de 15 recruteurs ne pourra pas analyser des milliers de joueurs. Or la *data* est omniprésente, car chaque match consolide 3,5 millions de données brutes, et offre une opportunité inégalable pour outiller le processus de recrutement avec une économie de moyens.

Avec 3 recruteurs, le Téfécé utilise cette masse d'information⁴ en couvrant 70 championnats dans le monde, allant de la première à la cinquième division en Angleterre, pour 42 000 athlètes suivis. Dans un contexte compétitif, l'agilité du recrutement est accrue grâce à l'éventail décuplé de prospects en maintenant constantes les exigences sur les qualités des joueurs servant le style de jeu.

Les données sont analysées selon des métriques propres (100 par poste)⁵. Les modèles permettent de faire de la translation en projetant en Ligue 1 les performances d'un athlète jouant en deuxième division du Brésil. Le processus de recrutement établit des listes sur chacun des 11 postes. Les critères financiers (budget de transfert, masse salariale, montant du transfert) complètent l'analyse pour discriminer les choix de recrutement selon la valeur économique subjective du joueur sur le marché.

La méthode mise en œuvre vise à identifier les déficiences du marché en contournant l'irrationalité du marché des transferts, pour surperformer sportivement par rapport à la masse salariale du club.

L'analyse *data* est complétée par des visionnages de séquences sportives par les recruteurs grâce à des outils tels Wyscout ou VidéoProfile.

Le processus se finit par des échanges directs pour mieux cerner la personnalité et l'environnement du joueur. Cette approche *data* permet également d'être rationnel en exposant les raisons qui ont amené le club à s'intéresser à lui, comme aux facteurs d'épanouissement sportif proposés.

LA PERFORMANCE INDIVIDUELLE DU JOUEUR

À leur arrivée, les joueurs font l'objet de tests physiques (capacité athlétique) mesurés par des métriques ; leur passé athlétique est en général transmis par leur ancien club. Chaque joueur s'entraîne et joue avec des capteurs sur les vêtements⁶. Les *data* sont collectées, et accessibles autant au joueur qu'au *staff* médical et technique après chaque entraînement et match. La capacité d'endurance à une charge physique est mesurée (course à haute vitesse, durée de repos, nombre de matchs successifs) selon les contraintes affichées sur un poste (exposition à la vitesse). Le travail, autant à l'entraînement qu'en compétition, est alors individualisé pour optimiser les phases de jeu comme de récupération dans l'objectif de minimiser les blessures notamment musculaires. Cette analyse permet également l'échange avec le joueur pour lui permettre d'améliorer un secteur de jeu⁷.

⁴ *Data providers* StatsBomb ou Opta.

⁵ Partenariat industriel et scientifique avec Zellus Analytics.

⁶ Cardiofréquence-mètre, accéléromètre, GPS.

⁷ Alors qu'elle est phénoménale, la charge cognitive des sportifs de haut niveau est peu comprise par le grand public. Il serait contre-productif de communiquer des données brutes de 60 matchs sans accompagnement.

INDIVIDUAL PERFORMANCES

Good performances by Frank, Aron, Moussa, Kevin, and Rasmus for starters.

Gabi, Thijs and Mikkel created value off the bench.

Name	Role	Mins	GDI	OPV-M Total	OPV-P Total
★ Frank Magri	ST-1	90	49 0.11	47 0.05	89 0.45
★ Gabriel Suazo	LB	33	83 0.07	81 0.05	88 0.16
★ Aron Donnum	RW	75	16 -0.09	59 0.08	74 0.15
★ Moussa Diarra	LB	68	26 -0.11	48 -0.04	73 0.07
★ Thijs Dallinga	ST-1	33	68 0.21	75 0.14	62 0.05
★ Kévin Kebe	RB	75	23 -0.12	62 -0.01	62 0.04
★ Stijn Spierings	M-DP	102	44 -0.04	21 -0.09	57 0.02
★ Rasmus Nicolaisen	LCB-2	102	26 -0.11	83 0.01	57 0.01
★ Mikkel Desler	RB	26	57 -0.03	51 -0.03	44 -0.02
★ César Gelabert	LW	11	31 -0.02	24 -0.03	38 -0.04
★ Cristian Cásseres Jr.	M-DP	102	4 -0.20	28 -0.07	36 -0.03
★ Logan Costa	RCB-2	102	32 -0.09	47 -0.06	33 -0.06
★ Vincent Siero	N. 10	26	36 -0.00	16 -0.06	27 -0.09
★ Denis Genreau	N. 10	102	2 -0.22	10 -0.09	23 -0.11
★ Yanis Bagraoui	LW	68	4 -0.18	23 -0.03	20 -0.12

Figure 2 : L'analyse data de chaque joueur (Source : Club).

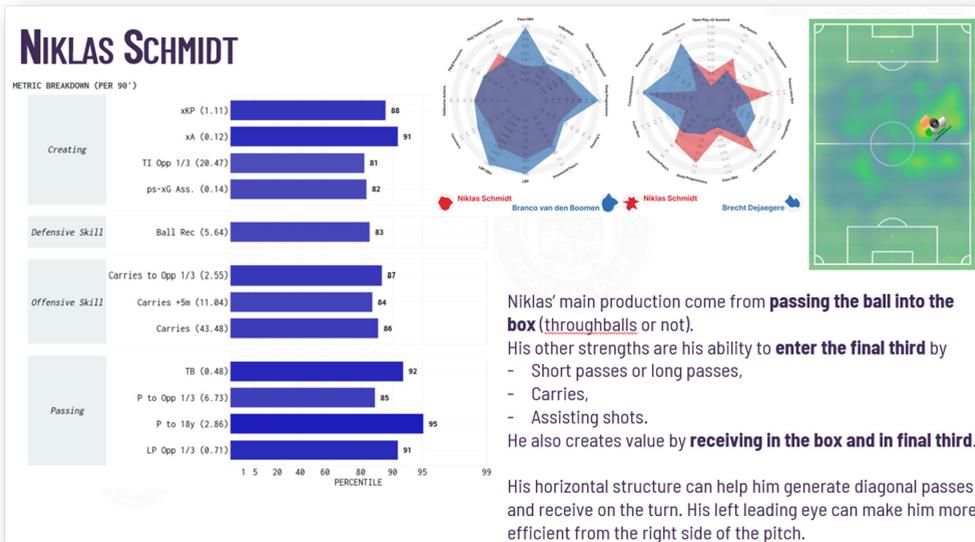


Figure 3 : Exemple d'analyse data d'un joueur (Source : Club).

La communication fluide, entre les secteurs sportifs et médicaux, est le facteur principal de prévention des blessures non traumatiques des athlètes⁸. Des indicateurs quotidiens de bien-être sont ainsi capturés pour aider à l'analyse de la fatigue des athlètes (qualité

⁸ Ekstrand J., Lundqvist D., Davison M. *et al.* (2019), "Communication quality between the medical team and the head coach/manager is associated with injury burden and player availability in elite football clubs", *British Journal of Sports Medicine*, 53(5), pp. 304-308.

du sommeil) par l'équipe médicale, composée d'un médecin, d'un kiné et d'un préparateur mental.

Le taux de disponibilité des joueurs au Téfécé, marquant le faible nombre de blessures, est ainsi de 94 %, le plus haut de Ligue 1 avec l'AS Monaco. Il tient autant à la caractéristique d'un banc composé de jeunes joueurs qu'à l'existence d'un environnement de performance intégrant sportif, physique, mental et médical.

LE MANAGEMENT SPORTIF

Le style de jeu est cardinal, et tout le *staff* technique (entraîneur, préparateur) doit y contribuer. L'aptitude à appuyer cette démarche est un critère de recrutement⁹. Le Téfécé dispose d'une base de données de 350 entraîneurs dans le monde, contenant les diplômes, le style de jeu des équipes dirigées, avec une profondeur historique importante. Elle permet d'analyser les caractéristiques identifiantes¹⁰.

L'entraîneur suit les processus de travail du Téfécé, et y contribue directement. Dès la phase de recrutement, l'entraîneur est confronté à une analyse faite sur sa capacité à adhérer à la démarche collective, et l'approche *data driven* en fait partie.

L'entraîneur dispose d'un retour *data* après chaque match. Cette analyse repose sur 12 métriques fixées au mois de juin pour la saison à venir. Si sur les 34 matchs de la saison, elles sont tenues à 50 % alors les objectifs du club sont atteints. Des réunions collectives d'analyse sont faites tous les 6 matchs pour analyser les performances passées et mesurer l'atteinte des objectifs. Ces métriques sont projetées sur les adversaires pour établir une analyse comparative.

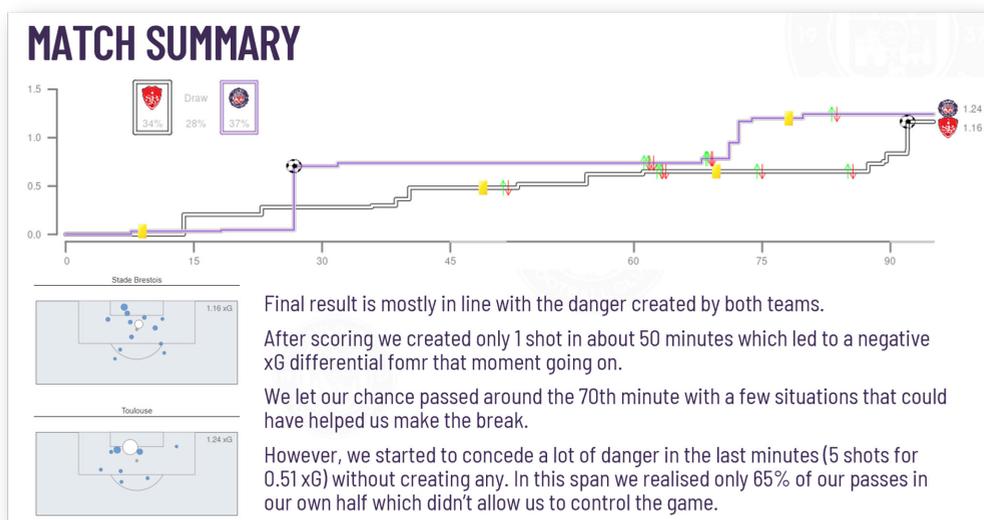


Figure 4 : Exemple de tableau de bord de performance (Source : Club).

⁹ Périodisation tactique, jeu de position.

¹⁰ Pep Guardiola présente le même style, débutant à la réserve du Barça ou entraîneur titré de Manchester City.

La préparation tactique des matchs s'organise dès le lundi. Elle débute par un débrief du match de la veille, puis par l'analyse de l'adversaire à venir, pour fixer des schémas tactiques simulant leur effet. La donnée permet d'aller plus loin que l'analyse vidéo. En analysant 100 000 coups de pied arrêtés passés, la préparation à la défense ou l'attaque est plus fine, avec une finesse d'exécution très importante dans les gestes individuels et collectifs répétés dans la semaine à l'entraînement.

Les données tactiques^a et techniques^b sont produites à partir de vidéos étudiées par des sociétés spécialisées aidées par des algorithmes. Les données sont fiables 24 heures après le match (30 % d'erreur après 1 heure). L'usage de données et d'algorithmique restent encore un sujet non résolu pour établir des schémas tactiques adaptés au temps réel d'un match. Les algorithmes n'ont pas encore prouvé leur efficacité pour l'analyse de schéma tactique défensif.

^a Évaluation globale de l'équipe et des joueurs.

^b Que fait le joueur avec le ballon.

LE CENTRE DE FORMATION

La formation des jeunes joueurs est un continuum partant de l'école de football (U8), à l'académie (U13) jusqu'au centre de formation (16 à 20 ans). La formation est sous la tutelle de la direction technique nationale (DTN) de la fédération française de football. Dès l'académie, le parcours sportif se conjugue avec l'éducatif¹¹.

Les centres de formation sont classés tous les ans par la DTN, leur modèle économique est évalué (retour sur investissement à l'euro investi, joueurs devenus professionnels, contributions aux équipes nationales, scolarité)¹². Un classement mondial évalue leurs performances s'agissant d'un marché économique en tant que tel¹³.

La capacité pour un centre de formation d'attirer et de former des jeunes joueurs est fondamental. Elle dépend tout autant de son ancrage local, de sa capacité à former, et à projeter les joueurs dans le groupe professionnel ou dans un club de l'élite¹⁴.

Dès les U14, les joueurs s'entraînent avec les mêmes capteurs que les professionnels. Le club dispose ainsi d'une connaissance fine sur leur potentiel. Sans être filmés, les analystes suivent les matchs selon des métriques permettant de mesurer l'application du style de jeu et de corriger des effets non anticipés pour mieux ajuster les entraînements¹⁵.

¹¹ Les élèves sont scolarisés à proximité, ceux en contrat professionnel le sont dans une école privée propre au club avec des professeurs en lien étroit avec les entraîneurs. Les jeunes sont associés à des projets citoyens proposés par la fondation du club (exemple : lutte contre les discriminations).

¹² <https://www.fff.fr/article/10523-1-efficacite-2022-2023-des-centres-de-formation.html>

¹³ <https://football-observatory.com/LettreHebdomadaire436>

¹⁴ Le centre de formation du Téfécé est ainsi classé cinquième au niveau national et quinzième pour les 5 grands championnats européens, tout en ayant un budget plus faible que les meilleurs (ROI supérieur).

¹⁵ Équipes belles à regarder jouer mais faciles à déstabiliser par l'adversaire. Les exercices ont été adaptés avec des résultats mesurés en compétition.

Au Téfécé, entre 40 et 50 % de l'effectif professionnel composé de 22 joueurs, est issu du centre de formation. Ce sont des joueurs aisément intégrables sportivement comme financièrement dans le groupe professionnel, ou en capacité de contribuer par leurs transferts au modèle économique du club.

L'analyse de cohortes depuis 20 ans pose le constat que dans une génération d'une vingtaine de jeunes joueurs, seuls 1 ou 2 parviennent à décrocher un contrat professionnel, quand en 2005 ce sont 5 jeunes passés professionnels. Le club s'est questionné sur les raisons. L'analyse des données des joueurs couplée avec l'analyse sociologique du « groupe parents » a marqué une forte corrélation, et a lancé des actions pour reproduire le succès (*mentoring* collectif).

L'EXPÉRIENCE DE MATCH

Le football professionnel est un spectacle vivant. Il doit offrir l'expérience attendue par ses *fans* ou aux *sponsors* pour véhiculer l'image souhaitée. L'expérience doit correspondre à tous les publics, d'un club ancré dans ses racines populaires, aux ultras comme aux invités dans les loges. Elle incarne les valeurs véhiculées par le club. Si la performance est cardinale, l'expérience vécue reste prédominante pour attirer et fidéliser un public qui vient rechercher l'expression du style de jeu du Téfécé.

Le club développe ses services¹⁶ et la communication autour du match pour maximiser le "match experience"¹⁷ et l'adhésion aux valeurs du club. La *data* permet une meilleure connaissance des *supporters*, de leurs attentes pour personnaliser et améliorer les services rendus à chaque public, à l'extérieur comme à l'intérieur du stade.

Sa rentabilité économique s'inscrit dans une volonté de garder le stade populaire, de développer le club avec les collectivités locales¹⁸, mais aussi d'accueillir d'autres types de compétition sportive ou de spectacle.

¹⁶ Partenariat avec le studio d'animation toulousain TAT et un chef mondialement reconnu.

¹⁷ Les *supporters* viennent au stade maintenant 56 mn avant, alors qu'ils arrivaient 30 mn avant.

¹⁸ Centre de formation, football féminin, loges.