

SPORT DATA REVOLUTION

Andy Hyeans

SPORT DATA REVOLUTION

L'analyse des données au service
de la performance sportive

Préfaces de

Alexandre Marles

Docteur en sciences du sport, ancien membre de la Direction technique nationale de la Fédération française de football et ancien responsable de la performance du Paris Saint-Germain et de l'Olympique Lyonnais.

Bill Gerrard

Docteur en économie et enseignant en analyse des données sportives à la Leeds Business School, ancien analyste de la performance pour le club de rugby de Saracens et la Premier League anglaise.

DUNOD

Illustration de couverture : Dmitry Fetisov - 123rf.com
Relecture : Danielle Roque
Mise en page : Nord Compo

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements



d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.

Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du Centre français d'exploitation du

droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).

© Dunod, 2016

5 rue Laromiguière, 75005 Paris
www.dunod.com

ISBN 978-2-10-074716-0

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

« It's about getting things down to one number. Using stats to reread them, we'll find the value of players that nobody else can see. People are over looked for a variety of biased reasons and perceived flaws. Age, appearance, personality. Bill James and mathematics cuts straight through that. »

Peter Brand (interprété par Jonah Hill) à Billy Bean (Brad Pitt)
dans Le Stratège (Moneyball, B. Miller, 2011)

Remerciements

Merci aux sociétés qui ont accepté de partager leur savoir-faire
et ont ainsi contribué à fournir des illustrations concrètes
dans cet ouvrage :

Dartfish

FootoVision

Mac-Lloyd

Tech'4'Team

Merci à toutes celles et ceux qui ont apporté une contribution à cet
ouvrage et notamment :

Gilles Celeux

Mehdi Ennaciri

Anne Germain

Bill Gerrard

Sandrine Godier-Michnik

Alexandre Marles

Ludovic Potrich

Valérie Robert

Julien Roussel

Kévin Vitoz

À Lina-Mazarine

À propos de l'auteur



Andy Hyeans est né en 1980 dans le Nord de la France. Comme souvent dans le monde sportif, c'est son père entraîneur qui lui a transmis sa passion et lui a permis « *d'apprendre beaucoup sur les arcanes du monde sportif sans même le savoir* ».

Diplômé de l'enseignement supérieur en sciences politiques, en

management et en gestion du sport, il est aussi titulaire d'un diplôme d'État de la jeunesse, de l'éducation populaire et du sport ainsi que d'un certificat de qualification professionnelle de technicien sportif.

Il conjugue son poste de cadre du secteur public dans le domaine de la formation avec ses fonctions d'expert sportif. D'abord joueur de basket-ball, il embrasse tôt le statut d'entraîneur assistant en championnat régional, avant de devenir entraîneur en championnat de France. Particulièrement investi dans le sport-entreprise, il y obtient un titre de champion de France et une quatrième place européenne avant de devenir conseiller technique pour le sport dans son administration. En 2011, il diversifie son expérience en occupant les fonctions de manager dans un club de basket-ball professionnel, tout en conduisant des projets de développement dans divers clubs

amateurs. En quête d'une approche globale du système de la performance sportive, il se forme en préparation mentale et en préparation physique. Il suit alors individuellement de jeunes sportifs de haut-niveau, dans le basket-ball et le tennis notamment.

Durant ses activités, il a conduit des recherches pour différentes organisations sur les thèmes de la performance physiologique (« La phosphocréatine : stockage et synthèse chez le basketteur », ressource interne à la société CFPD, 2014), de la performance psychologique (« Tests psychométriques et protocoles d'entretiens », ressource interne à la société H SPORTS, 2011), de la performance des clubs (« La professionnalisation d'un club amateur de basket-ball : Approches élémentaire et systémique », FSSEP Lille, 2012) et des équipes (« Coacher dans la complexité », *Mixt Offense*, 2014). Il a aussi accompagné de nombreuses entreprises qui, en quête de performance, se tournaient vers le monde sportif pour y sensibiliser leurs managers ou y consolider leurs équipes. Il a suivi le chemin inverse en analysant ce que l'ingénierie d'entreprise pouvait apporter au sport, ce qui lui a permis de transposer, d'adapter, de transformer, d'enrichir des savoirs, savoir-faire et savoir-être au profit du monde sportif.

Fondateur de Data Win, service d'ingénierie du sport qui contribue à améliorer les performances des athlètes, des équipes, des clubs et des enceintes sportives, il souhaite développer l'ingénierie du sport en France.

Convaincu qu'une maîtrise approfondie des sciences analytiques et qu'une spécialisation avancée dans certains domaines d'expertise pouvaient améliorer la performance, il a choisi de présenter dans ce livre comment et pourquoi les acteurs sportifs doivent développer leur expertise en analyse des données. Selon lui, l'analyse des données offre des opportunités sans précédents pour fiabiliser la performance sportive sous toutes ses formes.

Préface d'Alexandre Marles



Alexandre Marles est un expert de l'entraînement, de la préparation physique et de la performance sportive. Docteur en sciences du sport, il s'est spécialisé dans les facteurs limitants de la performance afin d'accompagner les sportifs de haut niveau en collaborant à l'amélioration de leur performance. Ancien athlète et footballeur, il a notamment

créé l'entreprise PERF in Sport en 2007, avec laquelle il est intervenu dans de nombreuses disciplines (football, rugby, basket-ball, handball, cyclisme, hockey sur gazon...) auprès de clubs et de fédérations. En 2011, il rejoint la Fédération française de football comme membre de la cellule Recherche de la direction technique nationale, formateur des entraîneurs et préparateur physique des sélections nationales. Il prendra ensuite les fonctions de responsable de la performance au Paris Saint-Germain Football en 2013 puis à l'Olympique Lyonnais Football en juillet 2014. Il est également consultant auprès de la NBA depuis 2008 et de la MLS depuis 2012.

Alors que l'on cherche tous à rationaliser les choses en réduisant au mieux l'incertitude, le big bang de la data, avec un flux de données gigantesque, est en train de marquer l'histoire du sport professionnel. L'idée de recueillir, traiter, combiner, filtrer et comprendre des données ne date pas d'hier, mais l'optimisation des datas afin de mieux connaître et de mieux comprendre l'individu ou les individus ensemble ou par catégorie est devenu un enjeu colossal pour rendre toute chose plus performante au quotidien. Les mathématiques et les statistiques ont toujours existé dans de nombreux domaines, y compris celui du sport. Mais l'utilisation des données de terrain est déjà beaucoup plus récente ; la compréhension des efforts d'un athlète ou d'un joueur lors d'un entraînement ou d'un match est aujourd'hui un sujet d'actualité.

En 1976, Thomas Reilly est un des premiers à utiliser un outil d'analyse manuel pour décrire les performances en match en fonction des postes de jeu. Cela passe par un découpage du terrain, un travail manuel et visuel couplé avec un enregistrement audio et vidéo, ce qui permet d'analyser les joueurs un par un. Le travail est long, compliqué et pas toujours très précis, mais c'est un premier pas vers une analyse plus détaillée des efforts en match. Afin de gagner du temps dans le traitement, des systèmes informatisés seront ensuite développés en combinant les statistiques et la vidéo, limitant les erreurs humaines mais n'améliorant que peu la fiabilité à cause de la faiblesse d'acquisition des caméras vidéo et le peu de données recueillies, ce qui entraîne des pertes de données, des estimations moyennes et donc des approximations. En 1998, l'Australian Institute of Sports s'intéresse à la mesure des facteurs de la performance chez l'athlète de haut niveau mais avec pour objectif principal de pouvoir mesurer des paramètres à chaque instant quel que soit l'exercice réalisé. Même si l'idée de l'outil de *tracking* embarqué sur le joueur n'est certainement pas de cette année, la croissance industrielle et informatique de l'époque va aider à faire émerger un outil GPS. D'abord peu précis et fiable, le *tracking* va évoluer et intégrer un certain nombre de composants permettant d'affiner la mesure. On retrouve actuellement dans les outils les plus modernes, comme le petit dernier Sport Tracking Pro, de la taille d'une gomme et d'un poids inférieur

à 40 grammes, une puce GPS, un accéléromètre 9 axes, un gyroscope, un magnétomètre, un capteur cardiaque... Tous ces capteurs permettent de suivre en temps réel l'ensemble des joueurs en récupérant la distance parcourue par zone d'intensité, le nombre et l'intensité des accélérations, la fréquence cardiaque, une estimation de la dépense énergétique et de nombreux autres paramètres, notamment ceux concernant le suivi médical du joueur. En effet, des nouvelles données sur l'évolution du temps de contact au sol lors de l'entraînement permettent de juger d'une fatigue ou non des membres inférieurs ou encore du déséquilibre jambe droite/jambe gauche, ce qui peut mettre en évidence une compensation ou une problématique médicale chez un joueur. La modernisation et surtout la miniaturisation de l'outil de *tracking* vont permettre petit à petit de pouvoir le porter en compétition. Le port du système de *tracking* en compétition n'a jamais été interdit ni réglementé puisque de nombreux athlètes en sport individuel et de nombreuses équipes en sport collectif l'ont utilisé depuis le début des années 2000, avant que l'International Rugby Board ne l'officialise en 2007 pour le rugby, et la Fédération internationale de Football Association en 2015 pour le football. Si on ne prend par exemple qu'un capteur du système, comme le GPS de 50 hertz soit, 50 données par seconde, cela représente 180 000 données pour une heure d'entraînement pour un joueur. Étant donnée cette quantité de données, des algorithmes utilisés et des systèmes puissants de traitement du signal permettent d'obtenir des données utilisables mais encore faut-il pouvoir les archiver dans une base de données avant d'essayer de les faire parler.

La base de données a pour objectif d'interfacer toutes les informations concernant un joueur sur plusieurs années. Outre la sécurité et la confidentialité des données, cryptées et stockées sous le contrôle du club, cela permet tout d'abord une rapidité d'utilisation par tous en fonction des droits d'accès attribués, mais permet aussi de capitaliser d'une année sur l'autre, d'avoir également un système centralisé et commun à toutes les entités d'un club, de mettre en place des alertes pour les situations clés, ou encore des signes de fatigue ou des blessures, etc., et ainsi de donner de la visibilité en partageant l'information à l'ensemble du staff et à la direction. Tout cela est indispensable pour mieux cibler et rentabiliser

un recrutement, pour suivre l'évolution d'un joueur du centre de formation au groupe professionnel, pour mieux gérer les charges de travail, la récupération, les blessures et donc pour gagner ces 1 à 10 % qui font que l'on peut être à haut niveau ou à très haut niveau, et s'y maintenir en gardant une certaine stabilité. Dans certains pays d'Europe, aux États-Unis, en Australie, les analystes de performances sont entre 5 et 10 par club en fonction des sports. Ils forment de véritables staffs à part entière qui vont apporter des compléments d'informations scientifiques et statistiques au coach principal pour aider à la décision. En France en 2016, il n'existe aucune structure de la sorte, mais juste quelques embryons, avec des personnes bien seules pour tenter de faire évoluer les mentalités sur le sujet. En effet, presque aucun club ne possède d'analyste de performances ni de réelle base de données, ce qui est déjà le minimum pour bien commencer.

Le sport de haut niveau a cependant encore besoin de temps pour mieux appréhender les datas. L'avenir des datas dans le sport passera par le couplage de plusieurs systèmes afin d'être de plus en plus précis et de plus en plus évocateurs et révélateurs d'une situation. Par exemple, le couplage en direct de la vidéo et de l'outil de *tracking* embarqué sur le joueur, permettant au joueur et au coach de mieux comprendre avec l'image les datas qui peuvent être fournies ; la possibilité de sélectionner et d'isoler seulement les accélérations ou actions intenses à l'entraînement ou en match qui sont significatives des actions décisives et de voir la vidéo associée ; ou encore la capacité de voir quels sont les exercices les plus sollicitants pour certains joueurs et parfois les moins sollicitants pour d'autres en sélectionnant les datas, alors qu'auparavant, on n'avait que l'image sans données précises qui pouvait justifier d'un état de fraîcheur ou de fatigue d'un joueur lors d'une action de jeu.

Le joueur moderne est aujourd'hui un joueur multiconnecté, multi-observé, et il doit prendre conscience de bien entretenir et de bien protéger son corps qui est aussi son « outil de travail ». Le coach moderne est celui qui va s'intéresser à ce qui est possible d'en tirer pour mieux manager. Chaque outil permettant d'améliorer un peu la performance dans l'optique d'atteindre le haut niveau et/ou d'y rester, se doit d'être considéré. Face à l'ampleur et à une certaine peur

de toutes ces données, le livre d'Andy sur la révolution des données sportives est le premier dans le genre à recenser les tenants et les aboutissants de la data et à rendre accessible l'importance de celle-ci dans le sport professionnel. Tout manager, coach, staff médical, athlète ou joueur se doit de s'intéresser à ce livre qui est le reflet du monde sportif d'aujourd'hui. La seule difficulté est de ne pas se perdre dans cet infini de données mais de réfléchir et d'échanger autour des informations intéressantes pour augmenter la performance du système entraîneur/joueurs. L'analyse des données est aujourd'hui une aide à la décision et peut permettre de diminuer le risque d'erreur ; elle ne livre toujours pas toutes les vérités, mais elle permet de gagner en performance.

Alexandre Marles
le 25 février 2016

Préface de Bill Gerrard



Bill Gerrard est un pionnier en Europe. Docteur en économie et professeur en analyse des données sportives à la Leeds Business School, il a conduit de nombreuses recherches sur cette thématique. De l'analyse du marché des transferts à celle de la qualité d'une équipe, de l'efficacité du coaching à la performance des sociétés

sportives, il a une expertise de toutes les dimensions sportives. L'ancien rédacteur en chef de la revue européenne Sport Management est cependant loin de n'être qu'un théoricien. Il est titulaire d'une licence d'entraîneur de football de l'UEFA et a travaillé sur le développement de l'analyse des données comme aide à la décision au plus haut niveau. Il est notamment passé par plusieurs clubs de rugby, en Angleterre, en Australie et en Afrique du Sud, par plusieurs clubs de football anglais et néerlandais, par le club de baseball d'Oakland aux États-Unis et a été consultant pour des athlètes olympiques ainsi que pour la Sky Sports Super League. Rien que dans le rugby, il a aidé le club anglais de Saracens à gagner deux titres de champion d'Angleterre et à disputer 14 demi-finales et finales nationales et européennes.

Une des caractéristiques déterminantes de la première partie du XXI^e siècle est la révolution des données (Data Révolution) qui transforme chaque aspect de notre vie. Les exemples sont multiples, des médecins qui utilisent de plus en plus de données pour diagnostiquer et traiter nos problèmes de santé aux entreprises qui utilisent les données pour comprendre nos préférences d'achats et pour mieux cibler leurs offres de produits en passant par les agences de renseignement gouvernementales qui utilisent les données pour détecter les menaces de sécurité potentielles.

Le sport de haut niveau a également été impacté très fortement par la révolution des données. Cependant l'impact a été assez inégal, à la pointe dans certains domaines mais à la traîne dans d'autres. Je dirais que le sport d'élite a connu de profonds changements dans la collecte de données sportives, mais que cela ne s'est pas encore traduit par un véritable changement dans l'usage qui est fait des données.

La collecte de données de performance se limitait, jusqu'à ces dernières années, principalement à la méthode du papier-crayon pour enregistrer les informations clés pendant les matches. Cela fonctionne bien dans les sports comme le baseball où la structure et le rythme du jeu permettent l'enregistrement instantané de la contribution de chaque joueur. En revanche, pour d'autres sports d'équipe, où les joueurs évoluent sur la totalité du terrain et se déplacent rapidement, où les divers codes de jeu sont complexes comme au football, rugby, hockey ou basket-ball par exemple, les informations sont difficiles à enregistrer sans assistance technologique. Traditionnellement dans ces types de sport, les seules données de performance enregistrées systématiquement sont les compositions d'équipes, les scores et les fautes. Même l'avènement des magnétoscopes dans les années 1970 montre peu de changement dans la collecte de données de performance pour le football et les autres sports d'équipe. L'extraction des données était particulièrement chronophage. Au milieu des années 1990, lorsque j'ai entrepris mon premier travail d'analyste dans le football afin de développer un système d'évaluation des joueurs permettant de déterminer les frais de transfert, les seules données de performances disponibles pour les footballeurs professionnels étaient les apparitions publiques en club, les buts marqués et les sélections internationales.